Labor and Capital Taxation in Stochastic Economies*

Mikhail Golosov
University of Chicago

Ilia Krasikov
Arizona State University

Ken Miyahara
University of Chicago

September 2026

Abstract

We develop a general approach to the analysis of labor and capital taxation in incomplete market models in the spirit of Aiyagari (1994). Taxes on labor and capital can be arbitrary, potentially non-linear functions. We derive explicit expressions for the equilibrium responses of invidiual behavior and macroeconomic aggregates to perturbations of such taxes in terms of two objects: the transition probability of wages and assets and the Frisch elasticities of earnings and savings. Using these expressions, we characterize optimal labor and capital taxes and develop numerical techniques to compute them. In the calibrated economy, optimal labor taxes are substantially lower than in a comparable Mirrleesian model because of the adverse impact that labor taxes have on wealth accumulation. Capital taxes may be positive or negative depending on the prevailing interest rates but generally feature subsidies for individuals with low wealth.

*We think Anmol Bhandari, Roozbeh Hosseini, Virgiliu Midrigan, Nicolas Werquin, and participants at EIEF and SED for comments. Golosov and Krasikov thank the NSF for support (award number 2416505). Golosov thanks the EIEF for hospitality.

1 Introduction

Individuals live long lives and face uncertainty about their future earnings. They save against that uncertainty, but can borrow only to a limited extent and often face restrictions on the assets they can hold. They eventually die and leave their assets to their heirs. Yet, most of what we know about optimal taxation comes from static settings that abstract from these features. Which objects determine how labor supply, consumption, and savings respond to changes in taxes? How do tax reforms affect different households, the distribution of wealth, and government tax revenues? What do optimal taxes look like when they shape not only how much people work but also how much wealth they accumulate and pass on? The aim of this paper is to make progress on these questions.

We consider labor and capital taxation in a canonical continuous-time incomplete-markets economy with borrowing constraints. The government imposes arbitrary, potentially non-linear taxes on labor earnings and capital. Individuals are mortal but altruistic towards their children. Our model is isomorphic to the "infinitely lived agents" specification of Aiyagari (1994) but the dynastic structure allows us to be explicit about the impact of taxes on different generations.

We develop a general approach to characterizing the effects of tax policy in such economies. Its central ingredient is the following thought experiment. Take an economy that has been operating under a given tax policy long enough to converge to the invariant distribution, and introduce a tax reform, i.e., a small perturbation of taxes. We allow arbitrary reforms in both how they depend on earnings and assets and how they change over time: tax reforms can be linear or non-linear, affect labor or capital, be permanent or transitory, sudden or anticipated. We derive equilibrium responses to such policy changes, both for individual households and for macroeconomic aggregates.

At the individual level, our key result describes the responses of earnings and savings (and, therefore, consumption) to any tax reform. We show that these responses admit a simple and intuitive representation in terms of two objects: the Frisch elasticities of earnings and savings, and the transition probability of wages and assets. Both the transition probability and the Frisch elasticities are independent of the tax reform, have direct empirical counterparts, and can be easily constructed numerically from pre-reform allocations.

Our representation applies uniformly to all agents, whether borrowing constrained or not. This is in contrast to the statistics commonly used in static tax analysis, such as Hicks and Marshall elasticities, which remain valid only for borrowing constrained agents. For those agents, our formulas can equivalently be written in terms of Hicks and Marshall elasticities, recovering the classical characterization of (Saez, 2001). For unconstrained agents, however, no such representation is available: their earnings and savings responses are shaped by uncertainty and by the possibility of hitting the borrowing constraint in the future, which Hicks and Marshall elasticities cannot capture.

At the macroeconomic level, we show how these individual responses can be aggregated to characterize the impact of the tax reform on the asset distribution, tax revenues, welfare, and other macroeconomic aggregates. We also show that the same techniques extend to other types of taxation, such as U.S.-style taxation of total income of individuals.

We use our characterization of the equilibrium responses to derive formulas for optimal labor and capital taxes. We adopt the Harsanyi-Rawls' "behind the veil of ignorance" criterion of optimality under which the planner sets taxes to maximize the lifetime utility of a person born into economy. This criterion is equivalent to assigning equal Pareto weights to all agents, irrespective of their assets, productivity, date of birth, and heritage. We derive optimal taxes for the most common types of economies studied in the macroeconomic literature: the small open economy, the closed economy with a constant returns to scale technology, and the economy with a fixed amount of assets a-la Huggett (1993).

Our tax formulas have the flavor of the "ABC" formula of Diamond (1998) and Saez (2001): optimal taxes balance the benefits of redistribution against the costs of behavioral responses of labor and savings to taxes. In contrast to static settings, however, both responses have important forward-looking components that capture how the expectation of future taxes affect decisions today. Because decisions are forward-looking, a change in taxes for any given level of earnings or capital affects not only agents who currently have that amount of labor income or wealth but also many other agents. Our characterization nests the famous optimal tax formula of Mirrlees (1971) as a special case, which emerges in the small open economy in which the interest rate is sufficiently low so that all agents are borrowing constrained.

Finally, we design a numerical algorithm to compute optimal labor and capital taxes. Non-linear taxes are, in principle, an infinite-dimensional object, which puts direct search methods out of reach: each candidate tax schedule requires solving for the equilibrium anew, and the space of candidate schedules is too vast to explore by brute force. Our analytical characterization restores tractability by identifying the most efficient direction in searching for the candidate tax schedule. The algorithm is applicable to both environments in which the planner can tax labor and capital non-linearly and when she faces further restrictions on allowable taxes, such as being able to tax capital only linearly.

We apply the algorithm to a calibrated small open economy version of Aiyagari economy. We find thatthe optimal labor taxes are substantially lower and less progressive than in comparable Mirrleesian settings. The reason for is twofold. First, the ability to borrow and lend allows agents to partially self-insure against income shocks, which reduces the social benefits of redistribution. Second, and quantitatively more importantly, taxation of labor income, especially that of high-earners, has an adverse effect on aggregate savings. Lower savings decrease welfare of future generations, who, all things being equal, are better off to be born into economy which large and more equally distributed wealth. The same

trade-off between redistribution and wealth accumulation shapes optimal capital taxes. The optimal non-linear capital taxes are progressive. featuring negative rates at low levels of wealth and positive at high wealth levels, which allows the planner to compress the wealth distribution. Under the restriction to linear capital taxes, the optimal rate is positive when interest rates are low and negative when they are high.

1.1 Related literature

A substantial literature in public finance, starting with the seminal work of Mirrlees (1971), studied optimal labor taxation in static environments. Saez (2001) developed perturbational techniques that allowed to represent the effects of Mirrleesian taxes in terms of Hicksian and Marshallian elasticities in models with rich heterogeneity, and Golosov et al. (2014) further extended this approach to models with many goods and multiple sets of tax instruments. A substantial follow-up literature used those techniques to study effects of taxation in variety of settings.[1] Direct application of those techniques to dynamic settings with incomplete markets is challenging: uncertainty imposes measurability restrictions on choices, while borrowing constraints introduce kinks in individual behavior. Both frictions are hard to incorporate into classical consumer theory.[2] Our characterization of the effects of taxation in dynamic settings overcomes these difficulties by using Frisch rather than Hicks and Marshall elasticities, and by exploiting the information embedded in the transition probability about the recursive structure of the consumer problem.

Chang and Park (2021) is the closest to our paper. Like us, those authors study taxation in incomplete market settings with borrowing constraints. Their main focus is on labor responses and their tax implications in the case when the government can use only labor taxes and there are no income effects on labor supply. We characterize labor and savings responses in a model with general preferences and both capital and labor taxes. Ignoring savings responses is important even when preferences have no income effects. Changes in labor taxes alter savings, and through them, the asset distribution. The change of the asset distribution has has a welfare impact if assets are taxed, if aggregate prices respond to changes in capital stock, or if the planner assigns positive weight on future generations - all irrespective of whether income effects are present or not. In addition, if income effects are not zero, which is in line with recent empirical work (e.g., (Golosov et al., 2024)), changes in savings directly prompt labor supply responses.

[1] See, e.g., Saez (2002), Jacquet et al. (2013), Allcott et al. (2019), Sachs et al. (2020), Bierbrauer et al. (2023), Ferey et al. (2024), Becko and Sztutman (2026).

[2] For example, Findeisen and Sachs (2017) consider nonlinear labor and linear capital taxes in a stochastic model in which consumers face no borrowing constraints and there are no income effects in labor supply. Even under these simplifying assumptions, their characterization of optimal taxes depends on a stochastic sequence of the Lagrange multipliers that are hard to related to economically-interpretable objects.

Piketty and Saez (2013) consider bequest taxation in a stochastic dynamic economy. They express the equilibrium responses to bequest tax changes in terms of the so-called "policy elasticities" (see also Hendren (2016)) that are reduced-form objects that depend on the tax reform itself. Such elasticities are hard to use to study non-linear taxes, which can be perturbed in infinitely many ways, eachassociated a separate and unknown policy elasticity. In contrast, we characterize tax responses in terms of the Frisch elasticities, which do not depend on the tax reform and are known explicitly.

Our approach is complementary to the New Dynamic Public Finance literature[3] that uses the mechanism design techniques to study optimal informationally-constrained taxes. Taxes that emerge from such approach can be very complex and history-dependent, which makes them hard to relate to policies that are used in practice. In contrast, our perturbational approach maps naturally into observed policies and is easy to adapt to different types of taxes that a policy-marker may consider.

A vast Ramsey literature studies optimal taxes numerically.[4] Papers in this tradition posit a parametric family of taxes that the government can use, and numerically optimize with respect to those parameters. In contrast, we do not need to restrict our taxes to any parametric family to gain tractability. Our analytical characterization of the effects of taxation is central both for designing efficient numerical algorithms for computing such taxes and for decomposing the optimal taxes into components that highlight how different economic trade-offs faced by the planner shape qualitative and quantitative properties of the optimal tax schedules.

Our characterization of labor and savings responses to tax reforms is related to the work of Farhi et al. (2022), who describe how consumption of individuals responds to changes of interest rates and income in incomplete market settings. Their approach relies on measures of prudence and stochastic stopping times associated with hitting the borrowing constraint. Our approach provides an alternative representation of consumption, labor, and savings responses to shocks (which, in our settings are more complex due to tax non-linearities) in terms of agents' marginal utilities of consumption and the transition probability. This representation maps more naturally and directly into the analysis of optimal taxation. It is also easier to use in empirical work because the transition probability, unlike stopping times, has a direct empirical analogue in the data that is readily available.

In working with our continuous time model we build on the insights of Achdou et al. (2022), which we further extend to settings with endogenous labor supply and non-linear taxes. While continuous time is not crucial for any part of our analysis, it is helpful in two ways. First, it simplifies analytical characterization of consumer behavior because the set of borrowing constrained agents takes a particularly simple form. Second, it streamlines numerical analysis allowing us to find the equilibrium, the invariant

[3]See, e.g., Golosov et al. (2003), Albanesi and Sleet (2006), Farhi and Werning (2013), Hosseini et al. (2013), Golosov et al. (2016), Ndiaye (2018), Hellwig (2021), Hellwig and Werquin (2022).

[4]See, e.g., Conesa et al. (2009), Dyrda and Pedroni (2023), Ferriere et al. (2023).

distribution, and all objects that appear in the optimal tax formulas in a fast and computationally efficient way.

The rest of the paper is organized as follows. To present our techniques in the most transparent way possible, we first start by considering a small open economy version of the Aiyagari model. We describe this environment in Section 2. In Section 3, we present two key objects that we use repeatedly in our analysis: the transition probability and the Frisch elasticities. In Section 4, we describe consumer responses to arbitrary tax reforms, and show how to aggregate these responses to find effects of tax reforms on macroeconomic aggregates. In Section 5, we extend our approach to environments in which aggregate prices respond to taxes, derive formulas for optimal labor and capital taxes, and show how to use our techniques to study other types of taxes, such as U.S.-style taxes on the total income of individuals. In Section 6, we describe a numerical algorithm to compute optimal tax formulas and study optimal taxes quantitatively in a calibrated economy. Section 7 concludes.

2 Environment

We consider a continuous time, small open economy version of the economy of Aiyagari (1994) with explicitly specified dynastic structure. At every instant, the economy is populated by a unit measure of agents who are heterogeneous in wealth and labor productivity. Each agent works, pays labor and capital taxes, consumes and saves. The flow budget constraint is

$$c_t + s_t = y_t - T(y_t) + ra_t - P(a_t), \tag{1}$$

where c_t is consumption, s_t is the flow of savings, y_t is labor earnings, a_t is asset holdings, r is an exogenous interest rate, T and P are labor and capital taxes. Apart from some regularity conditions that we discuss later, we keep functions T and P general and do not impose on them any specific functional form. We normalize $P(0) = 0$. This is without loss of generality since any intercept of P can be absorbed into $T(0)$ without changing the budget constraint.

Assets evolve according to $da_t = s_t dt$. Agents save subject to the borrowing constraint $a_t \geq 0$.[5] For technical reasons, we require a_t be bounded above by some a_{max}, where a_{max} is large enough so that it is immaterial for the solution.

Each agent has intratemporal preferences $u(c_t, l_t)$ over consumption c_t and labor supply l_t, where u satisfies standard concavity and differentiability assumptions. Individual earnings are given by $y_t = e^{\theta_t} l_t$, where θ_t is idiosyncratic labor productivity. We assume that θ_t follows the OU process

$$d\theta_t = \kappa(\bar{\theta} - \theta_t)dt + \sigma dB_t$$

[5] We set the borrowing limit at zero to simplify exposition. The analysis extends directly to any borrowing constraint of the form $a_t \geq -a_{min}$ with $a_{min} \geq 0$.

with reflecting boundaries at θ_{min} and θ_{max}, where $\bar{\theta} \in (\theta_{min}, \theta_{max})$. In this specification, B_t is a standard Brownian motion, $\sigma > 0$ controls the volatility of idiosyncratic shocks, $\bar{\theta}$ governs the mean of the process, and $\kappa > 0$ controls the speed of the mean reversion. This process is the continuous time version of the AR(1) process commonly used in discrete time models.

Each person dies at some constant Poisson rate and at death is replaced by her child who inherits parent's wealth and productivity. A parent discounts the future at rate $\rho > r$ and values her child's utility exactly as she values her own continuation utility. Because of this equal weighting, the death rate does not affect individual decisions, and each person's lifetime utility is given by the standard infinite-horizon problem with discount rate ρ.

Consider a person with assets a and productivity θ. We refer to $x = (a, \theta)$ as that person's state. Her lifetime utility $\mathrm{v}(x)$ is the value of the maximization problem

$$\mathrm{v}(x) := \max_{(c_t, y_t, s_t, a_t)} \mathbb{E}_0 \left[\int_0^\infty e^{-\rho t} u\left(c_t, y_t e^{-\theta_t}\right) dt \right] \tag{2}$$

subject to (1), $da_t = s_t dt$, $a_t \geq 0$, and $x_0 = x$.

Maximization problem (2) is the same as the maximization problem of an infinitely lived agent in the economy considered by Aiyagari (1994). While Aiyagari's framework is the workhorse framework to model earnings, savings and consumption dynamics, its assumption that people never die and no new agents are born is inadequate for a comprehensive analysis of tax policy. We therefore follow Becker and Barro (1988) and Barro and Becker (1989) and adopt their dynastic microfoundation of the infinitely lived maximization problem.

The problem (2) can be written recursively as the following Hamilton–Jacobi–Bellman (HJB) equation:

$$\rho \mathrm{v}(x) = \max_{c,y} u(c, ye^{-\theta}) + \partial_a \mathrm{v}(x) \underbrace{(y + ra - T(y) - P(a) - c)}_{=s} + \kappa(\bar{\theta} - \theta)\partial_\theta \mathrm{v}(x) + \frac{\sigma^2}{2} \partial_{\theta\theta} \mathrm{v}(x), \tag{3}$$

where ∂ denotes partial derivatives. Equation (3) holds everywhere on the interior of the state space $X = [0, a_{max}] \times [\theta_{min}, \theta_{max}]$ and is complemented by the boundary conditions that we detail in the appendix. In particular, the reflecting boundaries of the productivity process imply that $\partial \mathrm{v}_\theta = 0$ at $\theta = \theta_{\min}, \theta_{\max}$.

Let $\mathrm{c}, \mathrm{y} : X \to \mathbb{R}$ be policy functions, so that $\mathrm{c}(x)$ and $\mathrm{y}(x)$ denote the optimal consumption and earnings choices in (3) in state x. We use this font to distinguish functions c, y that map from the state space X from scalars c, y that appear in (3). We use the same font to denote other functions with domain of X, for example, the value function v. Let $\mathrm{a} : X \to \mathbb{R}$ be the function that returns the level of assets in state x, i.e., $\mathrm{a}(a, \theta) = a$. We use s to denote agents' optimal savings:

$$\mathrm{s} := \mathrm{y} + r\mathrm{a} - T(\mathrm{y}) - P(\mathrm{a}) - \mathrm{c}.$$

Finally, $\mathrm{u} := u(\mathrm{c}, \mathrm{y}e^{-\theta})$ and $\mathrm{u}_c := \partial_c u(\mathrm{c}, \mathrm{y}e^{-\theta})$ denote agents' instantaneousand marginal utilities of consumption, respectively.

We assume that tax and utility functions are sufficiently regular to ensure continuity and appropriate differentiability of the value and all policy functions. Those assumptions largely mirror assumptions in Achdou et al. (2022), that we extend to our economy with endogenous labor supply. In terms of economically meaningful restrictions, we assume that taxes T and P are such that v remains strictly concave in assets, and that, for each x, the maximization problem (3) has a unique maximum that satisfies the second order conditions with strict inequality. We state and discuss all of these regularity assumptions and technical details in the appendix.

3 Preliminaries

To understand what determines the effects of taxes on agents' utilities and choices, welfare and tax revenue, we consider the following thought experiment. Take an economy that has been operating under a given tax policy (T, P) long enough to reach the invariant distribution. Introduce a tax reform – a small perturbation of the tax functions – and consider the impact of this reform on equilibrium variables. We allow tax reforms that are essentially arbitrary: tax changes can be permanent or transitory, sudden or delayed, apply to labor or capital, and be linear or non-linear. We derive equilibrium responses to such reforms analytically. Specifically, we show that the responses of all equilibrium objects – continuation utilities, earnings, the asset distribution, tax revenues, etc – can be characterized in terms of the reform itself and two sets of objects: the transition probability, which describes the probability distribution over the state in which an agent may find herself t units of time ahead, and a set of earnings and savings Frisch elasticities. Both objects are independent of the reform, pinned down in the pre-reform economy, and have empirical counterparts in the data. The transition probability can be measured directly from the joint dynamics of assets and wages of individuals. The Frisch elasticities can be estimated econometrically, for example, along the lines of Blundell et al. (2016).

Our characterization describes equilibrium responses to tax reforms in a parsimonious and economically intuitive way. It is also central both to deriving analytical formulas for optimal labor and capital taxes and to designing efficient computational algorithms for finding such taxes numerically. The same characterization extends with fairly minor modifications to other environments, such as settings in which taxes affect aggregate prices or the government uses other forms of taxation.

Since the transition probability and the Frisch elasticities play an important role in our characterization, we now define them formally and discuss some of their properties.

3.1 Transition probability and the invariant distribution

In equilibrium, state x_t follows a Markov process. We use the notation $\Pi_t(\cdot|x)$ to denote the probability distribution over future states that a person who is currently in state x can reach after t units of time. We refer to Π_t as the transition probability. The transition probability allows us to write the expectation of any function φ that depends on x_t as $\mathbb{E}_0[\varphi(x_t)] = \int \varphi(\tilde{x})\Pi_t(d\tilde{x}|x)$. We use G to denote the invariant distribution of the Markov process x_t. If the economy is in the invariant distribution, the cross-sectional distribution of states remains unchanged over time, which implies that $\int \Pi_t(\tilde{x}|x)\mathrm{G}(dx) = \mathrm{G}(\tilde{x})$ for any $\tilde{x}$ and t.

Borrowing constraints impose a certain structure on both G and Π_t. Below, we discuss this structure heuristically; formal details can be found in the appendix and in Achdou et al. (2022). In continuous time settings, agents with strictly positive assets are not borrowing constrained because any choice of savings over an infinitesimal period dt is feasible. Whether agents with zero assets are borrowing constrained depends on their productivity. Under our regularity assumptions, agents with zero assets are not borrowing constrained if their productivity is above a certain cut-off θ_c, in which case they choose $\mathrm{s} > 0$. Conversely, agents with zero assets are borrowing constrained if their productivity is below that cut-off and they choose $\mathrm{s} = 0$. Let $\Theta_c = [\theta_{min}, \theta_c)$ be the set of productivities of the borrowing constrained agents, and $\bar{\Theta}_c$ be its closure, $\bar{\Theta}_c = [\theta_{min}, \theta_c]$. We use the indicator variable $\mathbb{I} : X \to \mathbb{R}$ to denote the set of states in which the borrowing constraints is slack. This indicator is equal to 0 for $x = (0, \theta)$ with $\theta \in \bar{\Theta}_c$ and to 1 otherwise.

Due to the borrowing constraint, both G and Π_t may have mass points at $a = 0$. Specifically, G can be written as

$$\mathrm{G}(dx) = [m(\theta)\delta(a) + g(x)]dx,$$

where δ is the Dirac delta function, m is the density of productivities of agents at the borrowing constraint, and g is the density of the borrowing unconstrained types. Similarly, Π_t can be expressed as

$$\Pi_t(d\tilde{x}|x) = [\mu_t(\tilde{\theta}|x)\delta(\tilde{a}) + \pi_t(\tilde{x}|x)]d\tilde{x}.$$

The invariant distribution G and the transition probability Π_t are related to policy functions through the generator of the Markov process (x_t), which we denote by $\mathcal{L}$. The generator is the linear operator that for any sufficiently differentiable function[6] $\varphi : X \to \mathbb{R}$ returns another function $\mathcal{L}\varphi : X \to \mathbb{R}$ defined by the equation

$$\mathcal{L}\varphi := \mathrm{s}\partial_a\varphi + \kappa(\bar{\theta} - \theta)\partial_\theta\varphi + \frac{\sigma^2}{2}\partial_{\theta\theta}\varphi.$$

[6]By "sufficient differentiability" of φ we mean that φ has all derivatives and that those derivatives are continuous for the expressions to make sense. Henceforth, all φ that appear in text will be assumed to be sufficiently differentiable.

The generator is constructed from combining the law of motion for productivity θ and the savings policy function s. The HJB equation (3) can be written in terms of the generator as

$$\rho \mathrm{v} = \mathrm{u} + \mathcal{L}\mathrm{v}. \tag{4}$$

One can show that for any sufficiently smooth function φ, the transition probability Π_t and the generator $\mathcal{L}$ are related by

$$\frac{d}{dt}\int \varphi(\tilde{x})\Pi_t(d\tilde{x}|x) = \int (\mathcal{L}\varphi)(\tilde{x})\Pi_t(d\tilde{x}|x) \text{ for all } x. \tag{5}$$

This equation shows that $\mathcal{L}$ captures the instantaneous rate of change of the expected value of $\varphi(x_t)$, $(\mathcal{L}\varphi)(x_t) = \lim_{dt\to 0} \frac{\mathbb{E}_t[\varphi(x_{t+dt})-\varphi(x_t)]}{dt}$. Since φ is arbitrary, one can think of (5) as defining Π_t from the savings policy function and stochastic process for θ.

Similarly, G and $\mathcal{L}$ are related through the equation

$$\int (\mathcal{L}\varphi)(x)\mathrm{G}(dx) = 0, \tag{6}$$

which must hold for any sufficiently differentiable φ with $\partial_\theta \varphi = 0$ at $\theta = \theta_{\min}, \theta_{\max}$. Equation (6) is known as the weak form of the Kolmogorov forward equation (KFE). This weak formulation will be more convenient for us to work with than the more frequently used strong form of the KFE, which states that the invariant distribution must satisfy $\mathcal{L}^*\mathrm{G} = 0$, where $\mathcal{L}^*$ is the adjoint of $\mathcal{L}$.[7]

As we analyze effects of taxation, we often need to know the value of the derivative $\partial_a \mathbb{E}_0[\varphi(x_t)]$ of expectation $\mathbb{E}_0[\varphi(x_t)]$ for a given function φ. Note that this derivative is entirely characterized by the transition probability Π_t and can be expressed as

$$\partial_a \mathbb{E}_0[\varphi(x_t)] = \int \varphi(\tilde{x})\partial_a \Pi_t(d\tilde{x}|x_0), \tag{7}$$

where $\partial_a \Pi_t(\cdot|x)$ is the derivative of $\Pi_t(\cdot|x)$ with respect to a.[8] One can show that $\partial_a \Pi_t(\cdot|x) = \mathbf{0}$ for the borrowing constrained agents, so that giving a dollar of assets to an agent with $a = 0$ and $\theta \in \Theta_c$ results in that agent consuming the extra wealth without changing her savings behavior.

An alternative but mathematically equivalent way to write $\partial_a \mathbb{E}_0[\varphi(x_t)]$ is as follows. Consider an agent in state x and let $t^* := \inf\{t|a_t = 0, \theta_t \in \overline{\Theta}_c\}$ be the first time the agent hits the borrowing

[7] The intuition for the two ways of defining the invariant distribution is easiest to see in a finite state, discrete time Markov chain. Let Π be its transition probability matrix. A vector g is the invariant distribution if it satisfies $g = \Pi^T g$, or equivalently, $(I - \Pi)^T g = 0$. The analogue of this equation in continuous time is the strong KFE, with the adjoint playing the role of the transponse and the generator $\mathcal{L}$ describing transition probabilities over infinitesimal time intervals. Note that if we multiply $(I - \Pi)^T g = 0$ by any vector φ and take the transpose, we obtain $g^T(I - \Pi)\varphi = 0$. The continuous time analogue of this equation is the weak KFE, our equation ((6)).

[8] Since $\Pi_t(\cdot|x)$ is not a function in the classical sense, $\partial_a \Pi_t(\cdot|x)$ should be understood as the weak-* derivative. The weak-* derivative is a linear functional such that $\partial_a \mathbb{E}_0[\varphi(x_t)] = \int \varphi(\tilde{x})\partial_a \Pi_t(d\tilde{x}|x_0)$ for any continuous φ. When $\Pi_t(\cdot|x)$ is sufficiently smooth, we can write $\partial_a \Pi_t(d\tilde{x}|x) = [\partial_a \mu_t(\tilde{\theta}|x)\delta(\tilde{a}) + \partial_a \pi_t(\tilde{x}|x)]d\tilde{x}$, but the smoothness assumption is not needed for our analysis.

constraint; t^* is a random variable that depends not only on x but also on the realizations of (θ_t). One can represent derivative $\partial_a \mathbb{E}_0[\varphi(x_t)]$ as

$$\partial_a \mathbb{E}_0[\varphi(x_t)] = \mathbb{E}_0 \left[\partial_a \varphi(x_t) e^{\int_0^t \partial_a \mathrm{s}(x_{\tilde{t}}) d\tilde{t}} \iota_{t<t^*} \right]. \tag{8}$$

To understand this equation, consider any sample path of productivity realizations. First, suppose that the agent does not hit the borrowing constraint by time t along this path, so that $\iota_{t<t^*} = 1$. Perturbing the initial assets a_0 in the law of motion, $da_{\tilde{t}} = \mathrm{s}(a_{\tilde{t}}, \theta_{\tilde{t}}) d\tilde{t}$, and differentiating this law of motion gives $\frac{d}{d\tilde{t}} \left(\frac{\partial a_{\tilde{t}}}{\partial a_0} \right) = \partial_a \mathrm{s}(x_{\tilde{t}}) \frac{\partial a_{\tilde{t}}}{\partial a_0}$, which solves to $\frac{\partial a_t}{\partial a_0} = e^{\int_0^t \partial_a \mathrm{s}(x_{\tilde{t}}) d\tilde{t}}$ when paired with the initial condition $\frac{\partial a_0}{\partial a_0} = 1$. That is, an extra dollar of assets in period 0 results in $e^{\int_0^t \partial_a \mathrm{s}(x_{\tilde{t}}) d\tilde{t}}$ extra dollars of assets at time t. Each extra dollar increases the value of $\varphi(x_t)$ by $\partial_a \varphi(x_t)$, which results in the payoff $\partial_a \varphi(x_t) e^{\int_0^t \partial_a \mathrm{s}(x_{\tilde{t}}) d\tilde{t}}$. On the other hand, if the agent hits the borrowing constraint along the sample path, then $\iota_{t<t^*} = 0$ and an extra dollar of assets at $t = 0$ does not affect $\varphi(x_t)$ along that path. The right hand side of equation (8) integrates across all sample paths along which agents do not hit their borrowing constraint prior to time t.

The two representations of $\partial_a \mathbb{E}_0$ given in equations (7) and (8) are equivalent but have different advantages and disadvantages. Equation (8) captures the economic essence of how extra assets affect outcomes in the future. This representation, however, is difficult to work with because it requires sorting through many sample paths and determining if and when the borrowing constraint binds along each of them. Representation (7) is more convenient. The transition probability Π_t is readily available in numerical analysis and has a direct empirical counterpart in the data. Moreover, equation (7) does not require differentiating φ, which often results in simpler and more intuitive expressions.

3.2 Elasticities

In static environments, the effects of taxation are determined by the Hicksian and Marshallian income and substitution effects. As we shall show later in this paper, in dynamic economies this role is played by their Frisch counterparts. In this section, we define Frisch elasticities most relevant for tax analysis.

Consider the maximization problem

$$\max_{c,l} u(c,l) + q[(1+\Delta)wl + ra - c - T(wl) - P(a)], \tag{9}$$

parameterized by (Δ, w, q). The solution to this problem satisfies the following optimality conditions:

$$MRS = (1 + \Delta - T')w, \quad \partial_c u = q, \tag{10}$$

where $MRS := -\partial_l u / \partial_c u$. These equations define Frischerian demands c^F and l^F that are functions of (Δ, w, q). The derivatives $\frac{\partial c^F}{\partial \Delta}$ and $\frac{\partial l^F}{\partial \Delta}$, evaluated at $\Delta = 0$, capture the effects of increasing the

after-tax retention rate $1-T'$ on consumption and labor holding the marginal utility of consumption fixed. The derivatives $\frac{\partial c^F}{\partial w}$, $\frac{\partial l^F}{\partial w}$ capture the wage effects, and $\frac{\partial c^F}{\partial q}$, $\frac{\partial l^F}{\partial q}$ capture the Frisch income effects.

Our approach to defining Frischerian responses in the economies with non-linear taxation closely mirrors the definitions of Hicksian and Marshallian responses used by Jacquet et al. (2013) and Scheuer and Werning (2017) to study the effects of taxation in static models. Hicksian demand functions are defined by replacing the second condition in equation (10) with $u = \bar{u}$; Marshallian demand functions are defined by replacing it with $(1+\Delta)wl + ra - c - T(wl) - P(a) = m$. Then, variations in $\bar{u}$ and m capture then Hicksian and Marshallian income effects, respectively. Frischerian, Hicksian and Marshallian income and substitution effects all summarize local properties of the utility function around a given point (c, l) by expanding u along the corresponding demand curves. In measuring the effect of tax changes on consumption and labor, Frisch holds the marginal utility of consumption fixed, Hicks holds the utility level fixed, and Marshall holds the value of resources fixed. All three types of responses can be computed analytically by differentiating equation (10) and its analogues.

The values of these responses depend on the point (c, l) at which responses are evaluated. Note that the optimal consumption and labor supply decisions of an agent in state x in the dynamic economy, equation (3), correspond to problem (9) in which $(\Delta, w, q) = (0, e^{\theta}, \partial_a \mathrm{v}(x))$. Thus, for any x we can construct appropriate Frischerian, well as Hicksian and Marshallian, demand functions and responses.

In the analysis of taxation, it would be more convenient to work not with responses of labor and consumption but with responses of earnings and savings. Let $y^F = wl^F$ and $s^F = (1+\Delta)wl^F + ra - c^F - T(wl^F) - P(a)$ be Frischerian earnings and savings demands. We use the notation $\eta_y = \frac{\partial y^F}{\partial \ln q}$ and $\eta_s = \frac{\partial s^F}{\partial \ln q}$ to denote their income effects. The Frisch elasticity of labor supply with respect to the after-tax retention rate is defined by $\xi_y = (1-T')\frac{\partial \ln y^F}{\partial \Delta}$,where the normalization by $1-T'$ converts the derivative into the elasticity. It will be convenient to define the Frisch (semi)-elasticity of savings with respect to the after-tax retention rate as $\xi_s = (1-T')\left(MRS\frac{\partial l^F}{\partial \Delta} - \frac{\partial c^F}{\partial \Delta}\right)$.[9] These objects can be expressed analytically as

$$\begin{pmatrix} \xi_y & \eta_y/y \\ \xi_s & \eta_s \end{pmatrix} = \begin{pmatrix} \frac{l\partial_l MRS}{MRS} + \frac{\partial_c MRS}{MRS} y(1-T') + \frac{yT''}{1-T'} & -\frac{\partial_c MRS}{MRS} \\ -\frac{\partial_c MRS}{MRS} y(1-T') & -\frac{\partial_{cc} u}{\partial_c u} \end{pmatrix}^{-1}. \tag{11}$$

To build intuition, consider the GHH preferences, introduced by Greenwood et al. (1988), for which u has the form $u(c - \upsilon(l))$ for some convex function υ. Under the GHH preferences, labor supply is determined solely by the marginal return on labor and is independent of q. For such preferences, $\partial_c MRS = 0$, and it follows from equation (11) that η_y and ξ_s are equal to zero. For more general

[9] Note that the response of savings is $\frac{\partial s^F}{\partial \Delta} = \left(MRS\frac{\partial l^F}{\partial \Delta} - \frac{\partial c^F}{\partial \Delta}\right) + wl^F$, where the last term is mechanical because $1+\Delta$ appears directly in the definition of s^F. Our convention uses ξ_s to denotes changes in savings that emerge due to changes in labor and consumption. This convention is immaterial for the analysis but it emphasizes the behavioral component of savings responses and simplifies tax formulas.

preferences, η_y captures the dependence of earnings on income, and ξ_s captures how that dependence affects savings holding the marginal utility of consumption fixed.

The Frisch, Hicks, and Marshall elasticities all summarize properties of the same utility functions and therefore related. Just as the Hicks and Marshall elasticities are related by the Slutsky equation, the Hicks tax elasticity of earnings $\xi_y^H = (1-T')\frac{\partial \ln y^H}{\partial \Delta}$ and the Marshall income effect $\eta_y^M = \frac{\partial y^M}{\partial m}$ can be expressed in terms of the Frisch elasticities:

$$\xi_y^H = \xi_y - \frac{\xi_s \eta_y}{y \eta_s}, \quad \eta_y^M = \frac{\eta_y}{\eta_s}. \tag{12}$$

Using equation (10) one can also define Frischerian wage responses. The Frisch wage elasticities coincide with tax elasticities when labor taxes are locally linear and differ by a term proportional to T'' under non-linear taxes. We present these formulas in the appendix. Wage elasticities will not appear in our analysis until we turn to economies with general equilibrium effects in Section 5.

4 Tax reforms and their effects

Our main thought experiment considers the effect of introducing a tax reform in the economy that has been living in the stationary distribution corresponding to some taxes (T, P). We first make several observations about stationary equilibrium. In such an economy, the joint distribution of assets and productivities is time-invariant and is given by G. Similarly, all macroeconomic aggregates, such as total earnings $\int \mathrm{y} d\mathrm{G}$ and the capital stock $\int \mathrm{a} d\mathrm{G}$ are constant over time. The government tax revenues $\mathscr{R}_t$ are the same at each time instant and are given by

$$\mathscr{R}_t = \int (T(\mathrm{y}) + P(\mathrm{a})) d\mathrm{G}.$$

Furthermore, a fixed fraction of agents is dying and being replaced by newborns per unit of time. Since a newborn inherits her parent's state and the probability of dying is the same for all agents, she draw her initial state from the distribution G. Therefore, the lifetime expected utility of a person born at time t, that we denote by $\mathscr{W}_t$, is given by

$$\mathscr{W}_t = \int \mathrm{v} d\mathrm{G}.$$

We refer to $\mathscr{W}_t$ as the welfare of the generation born at t.

We now introduce the concept of a tax reform.

4.1 Tax reforms

Without loss of generality, we normalize the time of the reform announcement to 0, and use the variable τ to denote time elapsed since the announcement, or *the start*, of the reform. A tax reform is a tuple

of functions $(\dot{T}_\tau, \dot{P}_\tau)_{\tau \geq 0}$ such that the new labor and capital taxes taxes at time τ are

$$T(y) + \varepsilon \dot{T}_\tau(y), \quad P(a) + \varepsilon \dot{P}_\tau(a), \tag{13}$$

where ε is a scalar. Our focus is on understanding the impact of this reform on the economy as $\varepsilon \to 0$.

We assume that functions $\dot{T}_\tau(y)$, $\dot{P}_\tau(a)$ are continuously differentiable in y, a and τ and satisfy mild regularity conditions discussed in the appendix. Without loss of generality, we assume that $\dot{P}_\tau(0) = 0$. Apart from that, we keep tax reforms general, both in how they depend on time τ and on (y, a). For example, tax reforms may be time-invariant, so that $\dot{T}_\tau$ and $\dot{P}_\tau$ are the same for all τ, orthey may induce gradual changes in taxes over time. Tax reforms may transitory, so that $\dot{T}_\tau$ and $\dot{P}_\tau$ converge to zero as $\tau \to \infty$, or they may be permanent. Tax changes may be unanticipated, so that $\dot{T}_0$ or $\dot{P}_0$ are not identically zero, or anticipated, so that $\dot{T}_\tau, \dot{P}_\tau$ are zero for some time after the reform announcement. Time-invariant reforms will be important both for illustrations of some of our results and in several applications we consider. We drop the time subscript and denote them by $(\dot{T}, \dot{P})$.

The tax reform (13) is defined by functions whose domains are earnings and assets. It will be convenient to represent it by functions on the state space X by defining $\dot{\mathrm{T}}_\tau(x) := \dot{T}_\tau(\mathrm{y}(x))$ and $\dot{\mathrm{P}}_\tau(x) := \dot{P}_\tau(\mathrm{a}(x))$. The functions $\dot{\mathrm{T}}_\tau$ and $\dot{\mathrm{P}}_\tau$ have economic interpretation of the statutory tax changes, i.e., they capture how tax liabilities of a person in state x would change if she kept her earnings at the pre-reform level. Similarly, we use $\dot{\mathrm{T}}'_\tau(x) := \dot{T}'_\tau(\mathrm{y}(x))$, $\dot{\mathrm{P}}'_\tau(x) := \dot{P}'_\tau(\mathrm{a}(x))$ to denote the statutory changes in marginal taxes. Let $\mathrm{R}(x) := T(\mathrm{y}(x)) + P(\mathrm{a}(x))$ be the total taxes collected from in stage x before the reform, and let $\dot{\mathrm{R}}_\tau(x) := \dot{T}_\tau(\mathrm{y}(x)) + \dot{P}_\tau(\mathrm{a}(x))$ be the statutory change in total taxes.

In the rest of this section, we characterize equilibrium responses to a given tax reform. We use dots to denote these responses.[10] In Section 4.2, we describe how the reform affects agents' continuation utilities. These responses depend on the time τ elapsed since the start of the reform and are denoted by $(\dot{\mathrm{v}}_\tau)_{\tau \geq 0}$. In Section 4.3, we describe earnings and savings responses $(\dot{\mathrm{y}}_\tau, \dot{\mathrm{s}}_\tau)_{\tau \geq 0}$. In Section 4.4, we show how the asset distribution changes with the reform, $(\dot{\mathrm{G}}_\tau)_{\tau \geq 0}$, and use this to construct equilibrium responses of various macroeconomic aggregates.

4.2 Continuation utility responses

We start by characterizing equilibrium responses of the continuation utilities to tax reforms. We use $\dot{\mathrm{v}}_\tau$ to denote the change in the function v at time τ after the reform announcement.

Lemma 1. *Response $\dot{\mathrm{v}}_\tau$ satisfies for all x*

$$\dot{\mathrm{v}}_\tau(x) = -\int_0^\infty \int e^{-\rho t} \mathrm{u}_c(\tilde{x}) \dot{\mathrm{R}}_{\tau+t}(\tilde{x}) \Pi_t(d\tilde{x}|x) dt. \tag{14}$$

[10] For any equilibrium object z, we use $\dot{\mathrm{z}}$, to denote its Gateaux derivative with respect to ε at $\varepsilon = 0$, i.e., $\dot{\mathrm{z}} = \lim_{\varepsilon \to 0} \frac{\mathrm{z}^\varepsilon - \mathrm{z}}{\varepsilon}$.

Equation (14) has a simple economic interpretation. Re-write it as

$$\dot{\mathrm{v}}_\tau(x) = -\mathrm{u}_c(x) \times \int_0^\infty \int \frac{e^{-\rho t}\mathrm{u}_c(\tilde{x})}{\mathrm{u}_c(x)} \dot{\mathrm{R}}_{\tau+t}(\tilde{x}) \Pi_t(d\tilde{x}|x) dt.$$

The term $\frac{e^{-\rho t}\mathrm{u}_c(\tilde{x})}{\mathrm{u}_c(x)}\Pi_t(d\tilde{x}|x)$ is the pre-reform stochastic discount factor (SDF) of an agent in state x. Therefore, the double integral in this equation is the present discounted value of statutory tax changes from time τ onward, discounted with that SDF. Thus, Lemma 1 shows that the change in utility $\dot{\mathrm{v}}_\tau(x)$ equals the present value of future taxes the agent expects to pay, converted into utils by the marginal utility of consumption $\mathrm{u}_c(x)$.

Lemma 1 is a consequence of the envelope theorem. It might appear surprising that the envelope theorem applies in our incomplete market economy with borrowing constraints, or that the formula for $\dot{\mathrm{v}}_\tau$ is identical to the one that would arise in a complete market model without asset trading frictions. However, borrowing constraints per se do not invalidate applicability of the envelope theorem: irrespective of the structure of asset markets or asset trading frictions, the shadow value of a dollar of income in any future state is given by the marginal utility of consumption adjusted by the probability of reaching that state.

An important object that will appear in our analysis is the marginal value of assets, $\partial_a \mathrm{v}$, and we need to characterize how it is affected by the tax reform. Since $\partial_a \mathrm{v} = \mathrm{u}_c$ from the first order conditions for (3), this characterization also describes the effect of the tax reform on agents' marginal utilities of consumption.

Lemma 2. *The response $\partial_a \dot{\mathrm{v}}_\tau$ of the borrowing unconstrained agents satisfies*

$$\partial_a \dot{\mathrm{v}}_\tau(x) = -\int_0^\infty \int e^{-\rho t}\mathrm{u}_c(\tilde{x}) \dot{\mathrm{R}}_{\tau+t}(\tilde{x}) \partial_a \Pi_t(d\tilde{x}|x) dt. \tag{15}$$

The response $\partial_a \dot{\mathrm{v}}_\tau$ of the borrowing constrained agents is the limit of this expression as $a \to 0$. In particular, for all $\theta \in \Theta_c$,

$$\partial_a \dot{\mathrm{v}}_\tau(0,\theta) = \lim_{a\to 0} \partial_a \dot{\mathrm{v}}_\tau(a,\theta) = \frac{1}{\eta_s(0,\theta)} \left(\dot{\mathrm{R}}_\tau(0,\theta) + \frac{\xi_s(0,\theta)}{1-\mathrm{T}'(0,\theta)} \dot{\mathrm{T}}'_\tau(0,\theta) \right) \partial_a \mathrm{v}(0,\theta). \tag{16}$$

Equation (15) shows that the structure of $\partial_a \dot{\mathrm{v}}_\tau$ for the borrowing unconstrained agents is very similar to that of $\dot{\mathrm{v}}_\tau$ except that future statutory tax changes are discounted with $\frac{e^{-\rho t}\mathrm{u}_c(\tilde{x})}{\mathrm{u}_c(x)}\partial_a\Pi_t(d\tilde{x}|x)$ rather than with $\frac{e^{-\rho t}\mathrm{u}_c(\tilde{x})}{\mathrm{u}_c(x)}\Pi_t(d\tilde{x}|x)$. The limit of this expression as $a \to 0$ gives the formula for the borrowing constrained agents.

It is tempting to compute $\partial_a \dot{\mathrm{v}}_\tau(0,\theta)$ for the borrowing constrained agents by simply substituting $x = (0,\theta)$ into equation (15). This substitution is not mathematically warranted[11] and would, in fact, produce an incorrect answer. For each fixed t, $\lim_{a\to 0} \int \mathrm{u}_c(\tilde{x}) \dot{\mathrm{R}}_{\tau+t}(\tilde{x}) \partial_a \Pi_t(d\tilde{x}|a,\theta) = 0$ because $\partial_a \Pi_t(\cdot|x) =$

[11] The problem is the following. Let $\mathrm{f}_t(x) := \int e^{-\rho t}\mathrm{u}_c(\tilde{x}) \dot{\mathrm{R}}_{\tau+t}(\tilde{x}) \partial_a \Pi_t(d\tilde{x}|x)$ for the borrowing unconstrained agents, so that equation (15) reads $\partial_a \dot{\mathrm{v}}_\tau(a,\theta) = -\int_0^\infty \mathrm{f}_t(a,\theta) dt$. Functions $(\mathrm{f}_t(a,\theta))_{t,a}$ are not uniformly bounded in (t,a) around $(0,0)$. For this reason, the order of integration and limits cannot be exchanged, and $\lim_{a\to 0} \int_0^\infty \mathrm{f}_t(a,\theta) dt \neq \int_0^\infty \lim_{a\to 0} \mathrm{f}_t(a,\theta) dt$.

0 for borrowing constrained agents. Yet, the limit of the double integral, $\lim_{a\to 0}\int_0^\infty e^{-\rho t}\int \mathrm{u}_c(\tilde{x})\dot{\mathrm{R}}_{\tau+t}(\tilde{x})\partial_a\Pi_t(d\tilde{x}|a,\theta)dt$, is not zero and is instead given by the expression in equation (16). To obtained this expression, one needs to consider the limit of the double integral on the right hand side of (15) explicitly.

The expression for the limiting value $\partial_a\dot{\mathrm{v}}_\tau(0,\theta)$ in equation (16) appears surprising in light of equation (15). Equation (15) has no behavioral elasticities, and reforms of capital and labor taxes appear there as a sum $\dot{\mathrm{R}}_{\tau+t}=\dot{\mathrm{T}}_{\tau+t}+\dot{\mathrm{P}}_{\tau+t}$. In contrast, the limit of this equation contains the Frisch elasticities ξ_s and η_s; it depends only on the current statutory changes $\dot{\mathrm{R}}_\tau$ but not on the future ones; it involves the statutory changes in marginal labor taxes$\dot{\mathrm{T}}'_\tau$ but not in marginal capital taxes $\dot{\mathrm{P}}'_\tau$. We provide a sketch of the derivation here leaving details to the appendix.

For simplicity, consider a time-invariant reform so that the time subscript τ is redundant and we can write equation (14) using the generator $\mathcal{L}$ as

$$\rho\dot{\mathrm{v}}=-(\dot{T}(\mathrm{y})+\dot{P}(\mathrm{a}))\partial_a\mathrm{v}+\mathcal{L}\dot{\mathrm{v}}, \tag{17}$$

where we substituted $\partial_a\mathrm{v}$ for u_c. Differentiate this expression with respect to a, remembering that the generator $\mathcal{L}$ also depends on a via s, to obtain

$$\rho\partial_a\dot{\mathrm{v}}=-\dot{\mathrm{R}}\partial_{aa}\mathrm{v}-(\dot{\mathrm{T}}'\partial_a\mathrm{y}+\dot{\mathrm{P}}')\partial_a\mathrm{v}+\mathcal{L}\partial_a\dot{\mathrm{v}}+\partial_a\mathrm{s}\partial_a\dot{\mathrm{v}}. \tag{18}$$

Using definitions of the Frisch elasticities, it is easy to verify from the first order conditions to (3) that $\partial_a\mathrm{y}$ and $\partial_a\mathrm{s}$ satisfy $\partial_a\mathrm{y}=\eta_y\frac{\partial_{aa}\mathrm{v}}{\partial_a\mathrm{v}}=\frac{\xi_s}{1-\mathrm{T}'}\frac{\partial_{aa}\mathrm{v}}{\partial_a\mathrm{v}}$ and $\partial_a\mathrm{s}=(r-\mathrm{P}')+\eta_s\frac{\partial_{aa}\mathrm{v}}{\partial_a\mathrm{v}}$. Achdou et al. (2022) were the first to point while while the marginal utility $\partial_a\mathrm{v}$ remains finite, the second derivative $\partial_{aa}\mathrm{v}$ explodes to $-\infty$ as agents' assets approach the borrowing constraint - a result that extends to our economy with endogenous labor and taxes. Substitute the expressions for $\partial_a\mathrm{y}$ and $\partial_a\mathrm{s}$, divide by $\partial_{aa}\mathrm{v}$ and take limits as $a\to 0$ to obtain

$$0=-\dot{\mathrm{R}}-\dot{\mathrm{T}}'\frac{\xi_s}{1-\mathrm{T}'}+\eta_s\frac{\partial_a\dot{\mathrm{v}}}{\partial_a\mathrm{v}},$$

which is equation (16) for time-invariant tax reforms.

The key mathematical property behind equation (16) is that $\partial_{aa}\mathrm{v}$ explodes near the borrowing constraint. This property has important economic content. Since $\partial_a\mathrm{v}=\mathrm{u}_c$, the second derivative $\partial_{aa}\mathrm{v}$ captures how agent's marginal utility of consumption is affected by having an extra dollar of assets. As agent's assets approach the borrowing constraint, small changes in savings start to have an increasingly large effect on her marginal utility of consumption. For such an agent, this becomes the dominant consideration for how she works and saves.

4.3 Earnings and savings responses

We now turn our attention to the responses of earnings and savings to tax reforms.

Lemma 3. *The earnings and savings responses* $\dot{\mathrm{y}}_\tau$ *and* $\dot{\mathrm{s}}_\tau$ *satisfy, for all* x

$$\dot{\mathrm{y}}_\tau = -\frac{\mathrm{y}\xi_y}{1-\mathrm{T}'}\dot{\mathrm{T}}'_\tau + \eta_y \frac{\partial_a \dot{\mathrm{v}}_\tau}{\partial_a \mathrm{v}}, \qquad \dot{\mathrm{s}}_\tau = -\dot{\mathrm{R}}_\tau - \frac{\xi_s}{1-\mathrm{T}'}\dot{\mathrm{T}}'_\tau + \eta_s \frac{\partial_a \dot{\mathrm{v}}_\tau}{\partial_a \mathrm{v}}.$$

Lemma 3, when combined with Lemma 2, fully characterizes the earnings and savings responses to an arbitrary tax reform. In particular, these responses can be written as

$$\dot{\mathrm{y}}_\tau(x) = -\frac{\mathrm{y}(x)\xi_y(x)}{1-\mathrm{T}'(x)}\dot{\mathrm{T}}'_\tau(x) - \eta_y(x)\int_0^\infty \int \frac{e^{-\rho t}\mathrm{u}_c(\tilde{x})}{\mathrm{u}_c(x)}\dot{\mathrm{R}}_{\tau+t}(\tilde{x})\partial_a \Pi_t(d\tilde{x}|x)dt, \tag{19}$$

$$\dot{\mathrm{s}}_\tau(x) = -\dot{\mathrm{R}}_\tau(x) - \frac{\xi_s(x)}{1-\mathrm{T}'(x)}\dot{\mathrm{T}}'_\tau(x) - \eta_s(x)\int_0^\infty \int \frac{e^{-\rho t}\mathrm{u}_c(\tilde{x})}{\mathrm{u}_c(x)}\dot{\mathrm{R}}_{\tau+t}(\tilde{x})\partial_a \Pi_t(d\tilde{x}|x)dt, \tag{20}$$

with an understanding that the responses of the borrowing constrained agents are obtained as the limits of these equations as $a \to 0$.

Equation (19) shows that the earnings response $\dot{\mathrm{y}}_\tau$ can be represented as the sum of two terms. The first term on the right hand side of (19) is the earnings response at time τ to the contemporaneous statutory change in marginal labor taxes $\dot{\mathrm{T}}'_\tau$. This response is governed by the Frisch price elasticity of earnings ξ_y.[12] The second term on the right hand side of (19) is the present value of future statutory tax changes discounted with the $\partial_a \Pi_t$-adjusted SDF. The earnings response to this present value change in captured by the Frisch income effect η_y.

Equation (20) shows that the response of savings $\dot{\mathrm{s}}_\tau$ is described by a similar formula. Savings respond to the contemporaneous change in marginal tax $\dot{\mathrm{T}}'_\tau$ and to the present value of future tax changes proportionally to the corresponding Frisch price[13] and income effects, as described by ξ_s and η_s. In addition, savings respond mechanically to the contemporaneous statutory changes in tax levels, which is captured by the term $-\dot{\mathrm{R}}_\tau$ in equation (20).

While the labor and capital taxes T and P appear in the same way in the household budget constraint, their effects on earnings and savings responses are qualitatively different. In particular, neither $\dot{\mathrm{y}}_\tau$ nor $\dot{\mathrm{s}}_\tau$ respond directly to $\dot{\mathrm{P}}'_\tau$. This is driven by the fact that from the point of view of time-τ agent, the reform $\dot{\mathrm{P}}_\tau$ is equivalent to a lump-sum tax.

The responses $\dot{\mathrm{y}}_\tau$ and $\dot{\mathrm{s}}_\tau$ also admit an alternative representation. From Lemma 1, $\dot{\mathrm{v}}_\tau$ is minus the expected present value of statutory tax changes, and $\partial_a \dot{\mathrm{v}}_\tau$ is simply the derivative of that expectation, that is

$$\partial_a \dot{\mathrm{v}}_\tau(x_0) = -\partial_a \mathbb{E}_0\left[\int_0^\infty e^{-\rho t}\partial_c u(c_t, l_t)\dot{\mathrm{R}}_{\tau+t}dt\right].$$

Equations (19) and (20) are obtained by substituting for $\partial_a \mathbb{E}_0$ the representation given in equation (7). As we discussed in Section 3.1, one can equivalently represent $\partial_a \mathbb{E}_0$ via equation (8). By going this route,

[12] The term $\frac{\mathrm{y}}{1-\mathrm{T}'}$ appears merely because of the definition of the elasticity. The tax reform changes the log after-tax retention rate $1-\mathrm{T}'$ by $-\frac{\dot{\mathrm{T}}'_\tau}{1-\mathrm{T}'}$, and the Frisch elasticity ξ_y captures the response of log earnings $\frac{\dot{\mathrm{y}}_\tau}{\mathrm{y}}$ to this change. Consequently, the response of the level of earnings $\dot{\mathrm{y}}_\tau$ to changes in the marginal tax rate $\dot{\mathrm{T}}'_\tau$ is proportional to $\frac{\mathrm{y}\xi_y}{1-\mathrm{T}'}$.

[13] Since we defined ξ_s as a semi-elasticity, there is no s in the $\frac{\xi_s}{1-\mathrm{T}'}\dot{\mathrm{T}}'_\tau$ term.

one obtains a characterization of the savings and earnings responses in terms of the hitting times of the borrowing constraint, the marginal propensity to save $\partial_a \mathrm{s}$, and the derivative $\frac{d}{da}\left(\partial_c u(c_t, l_t)\dot{\mathrm{R}}_{\tau+t}\right)$. We present these formulas in the appendix and show that they are related to the results in Farhi et al. (2022), who examine the responses of consumption to income and interest rate shocks in an incomplete market settings. While this approach provides an alternative perspective on the economics behind the earnings and savings responses to tax reforms, it is cumbersome to use for our tax analysis.

Equations (19) and (20) are the dynamic analogue of the "sufficient statistics" formulas pioneered by Saez (2001). To understand the connection with his work and a large literature that followed, it will be helpful to first characterize explicitly the earnings and savings responses of the borrowing constrained agents by computing limits of equations (19) and (20) as $a \to 0$.

Corollary 1. *The responses of the borrowing constrained agents satisfy* $\dot{\mathrm{s}}_\tau = 0$ *and*

$$\dot{\mathrm{y}}_\tau = -\left(\xi_y - \frac{\xi_s \eta_y}{\mathrm{y}\eta_s}\right)\frac{\mathrm{y}}{1-\mathrm{T}'}\dot{\mathrm{T}}'_\tau + \frac{\eta_y}{\eta_s}\dot{\mathrm{R}}_\tau = -\frac{\mathrm{y}\xi_y^H}{1-\mathrm{T}'}\dot{\mathrm{T}}'_\tau + \eta_y^M \dot{\mathrm{R}}_\tau, \tag{21}$$

where the second equality follows from (12).

Corollary 1 shows that the borrowing constrained agents do not change their savings behavior in response to a tax reform so that $\dot{\mathrm{s}}_\tau = 0$ for such agents. Equation (21) provides two equivalent characterizations of their earnings responses $\dot{\mathrm{y}}_\tau$, one in terms of the Frisch elasticities that arise from taking the limit of equation (19), and the other in terms of the Hicks price elasticity and the Marshallian income effect that follow from the relationship between the Frisch, Hicks and Marshall responses described in Section 3.2. The second representation is also equivalent to the characterization of Saez (2001) in his study of earnings taxation in static settings. We recover his formula in our dynamic setting because the borrowing constrained agents do not adjust their savings to small tax changes and, consequently, they act as if they solve a static problem.

Motivated by the characterization of earning responses in static economies, a large literature in empirical public finance estimates Hicksian and Marshallian elasticities to construct policy counterfactuals and to evaluate the deadweight losses of taxation. Our results caution against this practice. The static formula (21) applies only to agents at, or at least very near, the borrowing constraint. The universal formula that applies to all agents is described by equations (19) and (20). These equations highlight that key objects are the Frisch rather than the Hicks or Marshall elasticities, and that responses of earnings and savings also include a forward-looking component, which can be accounted for by the transition probability Π_t.

4.4 Responses of asset distribution and macroeconomic aggregates

In this section we describe how the asset distribution responds to tax reforms. This response is an important ingredient for understanding the effect of taxes on macroeconomic aggregates.

To motivate our discussion, consider the impact of the reform on government tax revenues. The pre-reform revenues are $\mathscr{R} = \int [T(\mathrm{y}) + P(\mathrm{a})] d\mathrm{G}$, where both y and G depend on taxes. Perturbing this expression in the direction of the tax reform, we obtain that the response of tax revenues at time τ is given by

$$\dot{\mathscr{R}}_\tau = \int \dot{\mathrm{R}}_\tau d\mathrm{G} + \int \mathrm{T}' \dot{\mathrm{y}}_\tau d\mathrm{G} + \int \mathrm{R} d\dot{\mathrm{G}}_\tau. \tag{22}$$

The first term on the right hand side of this expression captures how revenues would be affected by the reform if agents did not change their earnings and savings behavior. It integrates the statutory tax changes against the pre-reform distribution G. The second term captures the change in revenues due to the changes in individual earnings. Finally, the last term captures the change in revenues due to the changes in savings behavior. The savings response to the tax reform alters the joint distribution of asset and productivites, which is captured by $\dot{\mathrm{G}}_\tau$. Since we solved for $\dot{\mathrm{y}}_\tau$ in equation (19), we know the value of the first two terms for any reform. It remains to find $\int \mathrm{R} d\dot{\mathrm{G}}_\tau$.

Lemma 4. *For any φ satisfying the regularity conditions stated in the appendix,*

$$\int \varphi(x) \dot{\mathrm{G}}_\tau(dx) = \int \int_0^\tau \dot{\mathrm{s}}_{\tilde{\tau}}(x) \left[\int \varphi(\tilde{x}) \partial_a \Pi_{\tau - \tilde{\tau}}(d\tilde{x}|x) \right] d\tilde{\tau} \mathrm{G}(dx). \tag{23}$$

Since we have already solved for $\dot{\mathrm{s}}_{\tilde{\tau}}$ in equation (20), this lemma explicitly describes $\int \varphi d\dot{\mathrm{G}}_\tau$ for any function φ. In particular, to find the expression for the last term in equation (22), we substitute $\varphi = \mathrm{R}$ in equation (23).

Equation (23) has a natural economic interpretation. The term $\dot{\mathrm{s}}_{\tilde{\tau}}$ is a forward-looking component. Its value at state x, $\dot{\mathrm{s}}_{\tilde{\tau}}(x)$, captures how a person in that state changes savings at time $\tilde{\tau}$ after the reform and, as we showed in equation (20), depends on the tax changes from time $\tilde{\tau}$ onward. The term $\partial_a \mathbb{E}_0 \varphi(x_{\tau - \tilde{\tau}}) = \int \varphi(\tilde{x}) \partial_a \Pi_{\tau - \tilde{\tau}}(d\tilde{x}|x_0)$ is a backward-looking component. It captures how much an extra dollar of assets saved at time $\tilde{\tau}$in state x affects the value of variable φ at time τ, that is $\tau - \tilde{\tau}$ units of time later. Together, the forward- and backward-looking components summarize all the effects of the tax reform on asset distribution.

Since φ is arbitrary, Lemma 4 allows us to describe the responses of any macroeconomic aggregate. Consider, for example, the welfare effects of the tax reform. The reform changes the continuation utility of each person alive at $\tau = 0$ by $\dot{\mathrm{v}}_0$, which we characterized in Lemma 1. A newborn person in period $\tau = 0$ draws her initial state from the distribution G. Since the lifetime utility in state x changes by $\dot{\mathrm{v}}_0(x)$, the welfare of generation $\tau = 0$ changes by $\dot{\mathscr{W}}_0 = \int \dot{\mathrm{v}}_0 d\mathrm{G}$. As time passes since the introduction

of the reform, the distribution of assets changes. This affects the distribution from which newborns draw their initial state. The impact of the tax reform on welfare of the generation born at time τ is therefore

$$\dot{\mathscr{W}}_\tau = \int \dot{\mathrm{v}}_\tau d\mathrm{G} + \int \mathrm{v} d\dot{\mathrm{G}}_\tau, \tag{24}$$

where $\dot{\mathrm{G}}_0 = \mathbf{0}$. Lemmas 1 and 4 allow us to find these two integrals explicitly for any tax reform: the first one is obtained by substituting $\dot{\mathrm{v}}_\tau$ from equation (14) and the second one by taking $\varphi = \mathrm{v}$ in equation (23).

Long run effects and time invariant reforms

We now describe the long-run effects of tax reforms. For simplicity, we restrict attention to time-invariant reforms, although the same arguments apply to any reform that eventually settles to some permanent tax change.

Let $(\dot{T}, \dot{P})$ be a time-invariant reform. Our results in Lemma 3 show that the earnings and savings responses are time invariant and are given by

$$\dot{\mathrm{y}}(x) = -\frac{\mathrm{y}(x)\xi_y(x)}{1-\mathrm{T}'(x)}\dot{\mathrm{T}}'(x) - \eta_y(x)\int \dot{\mathrm{R}}(\tilde{x})\partial_a\bar{\Pi}(d\tilde{x}|x), \tag{25}$$

$$\dot{\mathrm{s}}(x) = -\dot{\mathrm{R}}(x) - \frac{\xi_s(x)}{1-\mathrm{T}'(x)}\dot{\mathrm{T}}'(x) - \eta_s(x)\int \dot{\mathrm{R}}(\tilde{x})\partial_a\bar{\Pi}(d\tilde{x}|x), \tag{26}$$

where $\partial_a\bar{\Pi}$ is defined by

$$\partial_a\bar{\Pi}(d\tilde{x}|x) := \int_0^\infty \frac{e^{-\rho t}\mathrm{u}_c(\tilde{x})}{\mathrm{u}_c(x)}\partial_a\Pi_t(d\tilde{x}|x)dt$$

for the borrowing unconstrained agents and as the limit of this expression for the borrowing-constraints individuals. While the earnings and savings responses do not depend on the time since the start of the reform, the asset distribution may exhibit a non-trivial transition dynamics. The long run response of the asset distribution is captured by $\dot{\mathrm{G}}$, and its characterization closely parallels that $\dot{\mathrm{G}}_\tau$ in Lemma 4.

Lemma 5. *Let* $\dot{\mathrm{s}}$ *and* $\dot{\mathrm{G}}$ *be the responses of savings and of the invariant distribution to a time-invariant reform. For any* φ *satisfying the regularity conditions stated in the appendix,*

$$\int \varphi(x)\dot{\mathrm{G}}(dx) = \int\int_0^\infty \dot{\mathrm{s}}(x)\left[\int \varphi(\tilde{x})\partial_a\Pi_{\tilde{\tau}}(d\tilde{x}|x)\right]d\tilde{\tau}\mathrm{G}(dx). \tag{27}$$

Equation (27) highlights the economic mechanism behind the distributional effect $\int \varphi d\dot{\mathrm{G}}$. While it can be used directly in computational work, it is often simpler and faster to use an alternative characterization of $\int \varphi d\dot{\mathrm{G}}$ in numerical analysis. For any given φ, let $\overline{\varphi} = \int \varphi d\mathrm{G}$ be the average value of $\overline{\varphi}$ under the invariant measure. Let function $\phi : X \to \mathbb{R}$ be a solution to

$$\mathcal{L}\phi = \varphi - \overline{\varphi}. \tag{28}$$

The key property of this function is that the relationship between φ and ϕ is given by[14]

$$\int \varphi d\dot{\mathrm{G}} = -\int \dot{\mathrm{s}} \partial_a \phi d\mathrm{G}. \tag{29}$$

The usefulness of function ϕ defined in equation (28) and using it to find $\int \varphi d\dot{\mathrm{G}}$ via equation (29) stems from the following observations. In numerical analysis, the state space X is discretized and (28) is a sparse system of linear equations. Sparsity allows one to find ϕ corresponding to any φ quickly and numerically efficiently. Once ϕ is found, its derivative $\partial_a \phi$ is obtained by straightforward differencing operation. This provides a fast method of computing $\int \varphi d\dot{\mathrm{G}}$ for any φ even on a very fine discretization. We use this observation in our computation of optimal taxes in Section 6.

5 Applications and extensions

In this section, we apply our analysis to characterize optimal labor and capital taxes. We also extend our approach to other types of taxation and environments with general equilibrium responses.

5.1 Optimal labor and capital taxation

To study optimal labor and capital taxation one needs to take a stance on the welfare criterion used by the planner. We consider the problem of choosing taxes that maximize the expected utility of a consumer "behind the veil of ignorance", i.e., welfare of the consumer who does not know the time or state into which she will be borne to. The original foundations for maximization of the expected utility of consumers were developed in Harsanyi (1953, 1955). Rawls, in his influential philosophical treatise *A Theory of Justice* (Rawls (1971)) coined the term "the veil of ignorance". He argued the principle of justice should be "chosen behind a veil of ignorance" because "no one knows his place in society, his class position or social status; nor does he know his fortune in the distribution of natural assets and abilities, his intelligence and strength, and the like" (p. 137 in Rawls (1971)). While Rawls famously thought that the "veil of ignorance" principle implied that society should focus on maximizing well-being of the least-fortunate groups, Harsanyi (1975) pointed that its mathematical formalization leads to maximization of the expected lifetime utility.

The Harsanyi-Rawls' "veil of ignorance" criterion is a natural starting point to study normative questions in public economics as it imposes equal Pareto weights on all agents irrespective of their birth time.[15] Mathematically, we formulate our problem as choosing time-invariant, balanced-budget taxes,

[14] Equation (28) is known in the theory of Markov processes as the Poisson equation. Its solution has the form $\phi(x) = -\int_0^\infty \int \varphi(\tilde{x}) \left(\Pi_{\tilde{\tau}}(d\tilde{x}|x) - G(d\tilde{x})\right) d\tilde{\tau}$. For the borrowing unconstrained agents, one can switch the order of differentiating and integration so that $\partial_a \phi(x) = -\int_0^\infty \int \varphi(\tilde{x}) \partial_a \Pi_{\tilde{\tau}}(d\tilde{x}|x) d\tilde{\tau}$ for such agents. This is not necessarily the case for the borrowing constrained agents but, since $\dot{\mathrm{s}} = 0$ for such agents, equations (27) and (29) provide equivalent characterizations of $\int \varphi d\dot{\mathrm{G}}$.

[15] More generally, consider the planner in period $t = 0$ who discounts utilities of future generations at some rate $\alpha > 0$. We can write this objective as $\bar{\alpha} \int \exp(-\alpha t) \mathscr{W}_t dt$, where $\bar{\alpha}$ is the normalization constant $\bar{\alpha} := \int \exp(-\alpha t) dt$ that ensures

that maximize ergodic welfare $\mathscr{W}$. Such taxes maximize

$$\mathscr{W} + \lambda \mathscr{R},$$

where λ is the Lagrange multiplier on the budget-balanced constraint $\mathscr{R} \geq 0$.

Up to this point, we were agnostic about which types of taxes we consider: both taxes and tax reforms could be essentially arbitrary. In practice, a policy-maker may face additional restrictions on what types of taxes she can use. In our analysis, we restrict attention to three types of tax systems. For all three systems we allow labor taxes to be arbitrary non-linear taxes and consider three cases for capital taxation: (i) no capital taxes; (ii) linear capital taxes; (iii) non-linear capital taxes. Case (i) is a convenient starting point as it will allow us to contrast most clearly how earnings taxation in dynamic environments differ from the canonical analysis of Mirrlees (1971). The assumption of linear capital taxes in case (ii) is in the spirit of Aiyagari (1995); such taxes emerge as the only feasible form of capital taxation if it is difficult for tax authorities to attribute capital ownership to individual taxpayers (see, e.g., Golosov and Tsyvinski (2007)). Finally, case (iii) puts capital and labor taxation on an equal footing.

In all three cases we can write the optimality condition in the same way,

$$\dot{\mathscr{W}} + \lambda \dot{\mathscr{R}} = 0 \text{ for all admissible } (\dot{T}, \dot{P}), \tag{30}$$

with the only restriction on which taxes tax reforms are admissible. In case (i) we have $P = \dot{P} = 0$, in case (ii) P is linear and $\dot{P} = \pm a$, and in case (iii) P and $\dot{P}$ are unrestricted.

Regardless of the class of taxes we consider, equation (30) can be written as

$$0 = \lambda^{-1}\dot{\mathscr{W}} + \dot{\mathscr{R}} = \underbrace{\int \left(\lambda^{-1}\dot{\mathrm{v}} + \dot{\mathrm{R}}\right) d\mathrm{G}}_{\text{mechanical}} + \underbrace{\int \mathrm{T}'\dot{\mathrm{y}} d\mathrm{G}}_{\text{labor}} + \underbrace{\int \left(\lambda^{-1}\mathrm{v} + \mathrm{R}\right) d\dot{\mathrm{G}}}_{\text{savings}}. \tag{31}$$

This formula represents the tax optimality condition as a sum of three effects. The first one is mechanical. Terms $\int \dot{\mathrm{R}} d\mathrm{G}$ and $\int \dot{\mathrm{v}} d\mathrm{G}$ capture changes in revenues and welfare that would prevail if agents did not adjust their earnings and savings to tax reforms. The multiplier λ captures the shadow cost of transferring a dollar from agents to the government, so that λ^{-1} converts utils to dollars.

The two other effects on the right hand side of (31) capture behavioral responses: how much revenues and utilities of agents are effects by the changes in labor and savings behavior induced by the reform.

that generational Pareto weights sum to one. The limit $\alpha \to 0$ of this objective assigns equal weights to all generations and corresponds to the Harsanyi-Rawls "veil of ignorance" objective. The limit $\alpha \to \infty$ corresponds to the objective of the planner that maximizes utility of the initial generation, $\mathscr{W}_0$. This is implicitly the welfare criterion considered by Aiyagari (1995) in his study of optimal capital taxation as well as of other authors such as Chamley (1986) and Judd (1985). Applying such an extreme welfare criterion is problematic not only from the ethical but also from the practical point of view. Taxes chosen to maximize welfare of the initial old generation often leave later generations in extreme poverty (see Straub and Werning (2020) and Auclert et al. (2024)), and it is unrealistic to expect later generations to be willing to go along with such policies.

Changes in labor supply affect tax revenues through changes in earnings, $\dot{\mathrm{y}}$, which carry the fiscal cost of $\int \mathrm{T}'\dot{\mathrm{y}}d\mathrm{G}$. Changes in savings induce changes in the asset distribution, $\dot{\mathrm{G}}$, which impacts welfare through two distinct channels. First, changes in assets affect tax revenues that the government collects through labor and capital taxes. This effect is captured by the term $\int \mathrm{R}d\dot{\mathrm{G}}$. Second, changes in assets affect utility of the newborn agents as, all things being equal, they are better off to be born in an economy with greater and more evenly distributed wealth. This effect is captured by term $\int \mathrm{v}d\dot{\mathrm{G}}$.

We know all the equilibrium responses that appear in equation (31) from our results in Section 4.1. To derive formulas for optimal taxes, we want to substitute those expressions and change the order of integration to explicitly find how $\lambda^{-1}\dot{\mathscr{W}} + \dot{\mathscr{R}}$ is affected by $\dot{\mathrm{T}}(x)$ and $\dot{\mathrm{P}}(x)$ for any x. We discuss each of the effects separately.

To re-write the mechanical effect, substitute $\dot{\mathrm{v}}$ from Lemma 1 to obtain

$$\int \dot{\mathrm{v}}d\mathrm{G} = -\int \int_0^\infty \int e^{-\rho t}\mathrm{u}_c(\tilde{x})\dot{\mathrm{R}}(\tilde{x})\Pi_t(d\tilde{x}|x)dt\mathrm{G}(dx).$$

Change the order of integrals and integrate first with respect to x, noting that the invariance property of G implies that $\int \Pi_t(d\tilde{x}|x)\mathrm{G}(dx) = \mathrm{G}(d\tilde{x})$. This gives[16]

$$\int \dot{\mathrm{v}}d\mathrm{G} = -\int \left(\int_0^\infty e^{-\rho t}dt\right)\mathrm{u}_c(\tilde{x})\dot{\mathrm{R}}(\tilde{x})\mathrm{G}(d\tilde{x}) = -\rho^{-1}\int \dot{\mathrm{R}}\mathrm{u}_c d\mathrm{G}.$$

Therefore, the mechanical effect can be written as

$$\text{mechanical} = \int \dot{\mathrm{R}}\left(1 - \rho^{-1}\lambda^{-1}\mathrm{u}_c\right)d\mathrm{G}.$$

This expression shows that statutory tax change $\dot{\mathrm{R}}(x)$ in each stage x is weighted by $1 - \rho^{-1}\lambda^{-1}\mathrm{u}_c(x)$. This weight captures the benefits of redistribution, and we discuss it in more details below.

Consider now the labor effect in equation (31). As we showed in (25), $\dot{\mathrm{y}}$ depends not only on changes in the tax levels, $\dot{\mathrm{R}}$, but also in marginal tax rates $\dot{\mathrm{T}}'$. Thus, the labor effect separates into two components, corresponding to $\dot{\mathrm{R}}$ and $\dot{\mathrm{T}}'$.

The dependence of the labor effect on $\dot{\mathrm{T}}'$ is straightforward. Let $\bar{\xi}_y$ be the effective substitution elasticity of earnings, defined as $\bar{\xi}_y(x) = \xi_y(x)$ if x is borrowing unconstrained and $\bar{\xi}_y(x) = \xi_y(x) - \frac{\xi_s(x)\eta_y(x)}{\mathrm{y}(x)\eta_s(x)} = \xi_y^H(x)$ otherwise. Then $\dot{\mathrm{T}}'(x)\frac{\mathrm{T}'(x)\mathrm{y}(x)\bar{\xi}_y(x)}{1-\mathrm{T}'(x)}$ captures the response of labor tax revenues to $\dot{\mathrm{T}}'(x)$ for both constrained and unconstrained agents.

The dependence of the labor effect on $\dot{\mathrm{R}}$ can be written as

$$-\int \mathrm{T}'(x)\left[\eta_y(x)\int_0^\infty \dot{\mathrm{R}}(\tilde{x})\partial_a\bar{\Pi}(d\tilde{x}|x)\right]\mathbb{I}(x)\mathrm{G}(dx) + \int \mathrm{T}'(x)\eta_y^M(x)\dot{\mathrm{R}}(x)(1-\mathbb{I}(x))\mathrm{G}(dx),$$

where we explicitly separated the unconstrained and the constrained agents. As with the mechanical effect, we change the order of integration in the first term in this expression. To rewrite the resulting

[16] Another way to see that $\int \dot{\mathrm{v}}d\mathrm{G} = -\rho^{-1}\int \dot{\mathrm{R}}\mathrm{u}_c d\mathrm{G}$ is to integrate (4) with respect to G and apply the weak KFE (6).

formula succinctly, we introduce a linear operator $\mathcal{H}_y$ that for any function $\varphi : X \to \mathbb{R}$ returns the function $\mathcal{H}_y\varphi(x)$ defined by

$$(\mathcal{H}_y\varphi)(x) := \begin{cases} -\int \frac{\varphi(\tilde{x})\eta_y(\tilde{x})}{g(x)} \partial_a \bar{\pi}(x|\tilde{x}) \mathbb{I}(\tilde{x}) G(d\tilde{x}) & \text{if } \mathbb{I}(x) = 1, \\ \varphi(x)\eta_y^M(x) - \int \frac{\varphi(\tilde{x})\eta_y(\tilde{x})}{g(x)} \partial_a \bar{\mu}(\theta|\tilde{x}) \mathbb{I}(\tilde{x}) G(d\tilde{x}) & \text{if } \mathbb{I}(x) = 0, \end{cases}$$

where $\bar{\mu}$ are $\bar{\pi}$ the components of the decomposition $\bar{\Pi}(d\tilde{x}|x) = [\bar{\mu}(\tilde{\theta}|x)\delta(\tilde{a}) + \bar{\pi}(\tilde{x}|x)]d\tilde{x}$, in line with Section 3.1. Using this operator, we can write the above formula as $\int \dot{\mathrm{R}}(x)(\mathcal{H}_y \mathrm{T}')(x)\mathrm{G}(dx)$, so that the whole labor effect in equation (31) can be written as

$$\text{labor} = -\int \dot{\mathrm{T}}' \frac{\mathrm{T}' \mathrm{y} \bar{\xi}_y}{1 - \mathrm{T}'} d\mathrm{G} + \int \dot{\mathrm{R}} \mathcal{H}_y \mathrm{T}' d\mathrm{G}.$$

The operator $\mathcal{H}_y$ will play an important role in several applications, and it is useful to build intuition for it. Recall that in dynamic economies a dollar increase in taxes in any state x leads to labor supply adjustments of agents in all states. The operator $\mathcal{H}_y$ uses the income effects and the transition probability to compute those adjustments in all states and aggregate earnings responses with any weighting function φ. Consider first the unitary function $\varphi = \mathbf{1}$. In this case, $(\mathcal{H}_y \mathbf{1})(x)$ captures how much total earnings in the economy would change if taxes for agents in state x increased by 1 dollar. In the formula for the labor effect, we have $\varphi = \mathrm{T}'$ because changes in tax revenues in any state are proportional to product of the changes in earnings $\dot{\mathrm{y}}$ and the pre-reform marginal tax rates T'. Therefore, $(\mathcal{H}_y \mathrm{T}')(x)$ returns the total change in government revenues from labor supply adjustments to a dollar increase in taxes in state x.

The savings effect in equation (31) has a similar operator representation. To define it, it would be helpful to first consider the structure of the expression $\int \varphi d\dot{\mathrm{G}}$ for an arbitrary φ. From Lemma 5, this integral can be written as

$$\int \varphi d\dot{\mathrm{G}} = \int \dot{\mathrm{s}}(x) \int_0^\infty \int \varphi(\tilde{x}) \partial_a \Pi_\tau(d\tilde{x}|x) d\tau \mathrm{G}(dx).$$

The savings response $\dot{\mathrm{s}}$, just like $\dot{\mathrm{y}}$, depends on both $\dot{\mathrm{R}}$ and $\dot{\mathrm{T}}'$. Following the same steps as with labor effect, we can the aggregate savings responses to $\dot{\mathrm{R}}$ and $\dot{\mathrm{T}}'$ using two operators, $\mathcal{H}_s$ and $\mathcal{E}_s$, so that

$$\int \varphi d\dot{\mathrm{G}} = \int \dot{\mathrm{T}}' \mathcal{E}_s \varphi d\mathrm{G} + \int \dot{\mathrm{R}} \mathcal{H}_s \varphi d\mathrm{G}. \tag{32}$$

The operators $\mathcal{H}_s$ and $\mathcal{E}_s$ aggregate income and substitution elasticities of savings, η_s and ξ_s. $(\mathcal{H}_s\varphi)(x)$ captures how the value of any function φ aggregated across states with G would change if a person in state x faced an extra dollar of tax liability. Similarly, $(\mathcal{E}_s\varphi)(x)$ captures the effect of an extra percentage point of the marginal tax rate on labor in state x on the aggregate value of φ. Just like the operator $\mathcal{H}_y$, the operators $\mathcal{H}_s$ and $\mathcal{E}_s$ will play an important role in the other applications that we consider in this section. To obtain the savings effect in equation (31), we substitute $\varphi = \lambda^{-1}\mathrm{v} + \mathrm{R}$ to

write

$$\text{savings} = \int \dot{\mathrm{T}}' \mathcal{E}_s \left(\lambda^{-1}\mathrm{v} + \mathrm{R}\right) d\mathrm{G} + \int \dot{\mathrm{R}} \mathcal{H}_s \left(\lambda^{-1}\mathrm{v} + \mathrm{R}\right) d\mathrm{G}.$$

We now have all the ingredients needed to characterize optimal labor and capital taxes.

5.1.1 Optimal labor taxes

The easiest way to derive the formula that characterizes optimal labor taxes is to consider reforms of the form $\dot{P}(a) = 0$ and $\dot{T}(y) = \iota_{y>y^*}$, i.e., reforms that increase labor taxes by a dollar on all earnings above some y^*. While such functions $\dot{T}(y)$ are not differentiable in the classical sense, they are differentiable in the sense of distributions,[17] and their distributional derivatives are given by $T'(y) = 0$ for $y \neq y^*$ and $T'(y) = \delta(y - y^*)$, where δ is the Dirac delta function. Substituting these derivatives into our expressions for mechanical, labor, and savings effects, we obtain a characterization of optimal tax schedule at any y^*.

Let F_y denote the cumulative distribution of earnings under a given tax system (T, P), and let f_y denote its density.

Proposition 1. *The optimal labor taxes satisfy, at every earnings level y^*,*

$$\underbrace{\mathbb{E}_{\mathrm{y}>y^*}\left(1 - \rho^{-1}\lambda^{-1}\mathrm{u}_c\right)}_{mechanical} - \underbrace{\frac{T'(y^*)}{1-T'(y^*)} \frac{y^* f_y(y^*)}{1-F_y(y^*)} \mathbb{E}_{\mathrm{y}=y^*} \bar{\xi}_y - \mathbb{E}_{\mathrm{y}>y^*} \mathcal{H}_y \mathrm{T}'}_{labor} \tag{33}$$

$$+ \underbrace{\frac{f_y(y^*)}{1-F_y(y^*)} \mathbb{E}_{\mathrm{y}=y^*} \mathcal{E}_s(\lambda^{-1}\mathrm{v} + \mathrm{R}) + \mathbb{E}_{\mathrm{y}>y^*} \mathcal{H}_s(\lambda^{-1}\mathrm{v} + \mathrm{R})}_{savings} = 0.$$

The Lagrange multiplier λ is given by $\lambda^{-1} = \frac{1 + \mathbb{E}\mathcal{H}_y \mathrm{T}' + \mathbb{E}\mathcal{H}_s \mathrm{R}}{\rho^{-1}\mathbb{E}\mathrm{u}_c - \mathbb{E}\mathcal{H}_s \mathrm{v}}$.

To understand the intuition for equation (33), consider a reform that slightly increases marginal taxes in a small neighborhood of some level of earnings y^*. The mechanical effect captures the welfare impact of this reform if agents did not adjust their labor and savings behavior. The term $\rho^{-1}\mathbb{E}_{\mathrm{y}>y^*}\mathrm{u}_c$ captures welfare losses of all agents with earnings greater than y^*, who now face higher labor taxes, and λ captures the gains from lump-sum redistribution of these extra revenues. Therefore, the mechanical effect compares the welfare benefits of taking money from agents earning more than y^* and giving the proceeds to all agents lump sum. As the marginal utility of consumption is generally lower for agents with higher earnings, the mechanical benefits of redistribution are increasing in y^*.

The other terms capture behavioral effects. Agents with earnings $\mathrm{y} = y^*$ change their labor supply and savings. The labor supply response is governed by $\bar{\xi}_y$, and the resulting reduction in labor tax revenues is $\frac{T'(y^*)}{1-T'(y^*)} \frac{y^* f_y(y^*)}{1-F_y(y^*)} \mathbb{E}_{\mathrm{y}=y^*} \bar{\xi}_y$. The response of savings is governed by ξ_s. This elasticity appears

[17] See Kanwal (1998) for an introduction to the theory of distributions and generalized functions.

in $\mathcal{E}_s$, which describes the resulting response of the asset distribution. As the asset distribution changes, so do tax revenues and welfare of the newborn. These two effects are given by $\frac{f_y(y^*)}{1-F_y(y^*)}\mathbb{E}_{\mathrm{y}=y^*}\mathcal{E}_s\mathrm{R}$ and $\frac{f_y(y^*)}{1-F_y(y^*)}\mathbb{E}_{\mathrm{y}=y^*}\mathcal{E}_s\frac{\mathrm{v}}{\lambda}$, respectively.

The remaining terms in formula (33) capture effects from the labor and savings adjustments to the tax increase for earnings above the y^* threshold. In dynamic economies, agents both above and below threshold y^* adjust their labor and savings to such a tax increase. The operators $\mathcal{H}_y$ and $\mathcal{H}_s$ aggregate the responses of all agents through their income their income effects on earnings and savings, η_y and η_s.

The general lesson of equation (33) is that optimal non-linear labor taxes balance the mechanical benefits of redistribution with the behavioral effects of earnings and savings responses at every earnings level y^*. The same formula (33) characterizes optimal labor taxes irrespective of how capital taxes are set.

Our Proposition 1 nests the optimal tax formulas derived by Mirrlees (1971) and Saez (2001) in static settings. Consider a special case of our economy in which the interest rate r is sufficiently low so that all agents are borrowing constrained. In such economy there are no assets, and the distribution of earnings is determined by the exogenous cross-sectional distribution of productivities. Since $\dot{\mathrm{s}}=0$ for the borrowing constrained agents by Corollary 1, the savings term in equation (33) drops out and the optimal labor tax formula becomes

$$\mathbb{E}_{\mathrm{y}>y^*}\left(1-\rho^{-1}\lambda^{-1}\mathrm{u}_c\right)-\frac{T'(y^*)}{1-T'(y^*)}\frac{y^*f_y(y^*)}{1-F_y(y^*)}\xi_y^H(y^*)+\mathbb{E}_{\mathrm{y}>y^*}\eta_y^M\mathrm{T}'=0, \tag{34}$$

where $\lambda^{-1}=\frac{1+\mathbb{E}\eta_y^M\mathrm{T}'}{\rho^{-1}\mathbb{E}\mathrm{u}_c}$. This equation is isomorphic to the optimal tax formula obtained by Saez (2001), which, as he showed in his paper, is also equivalent to the optimal tax formula of Mirrlees (1971). Comparison of equations (34) and (33) reveals several important differences that emerge in dynamic settings. First, one needs to take into account responses of savings to labor taxes. The labor responses are primarily governed by the Frisch rather than Hicks-Marshal elasticities. Finally, the Mirrlees-Saez formula implicitly assumes that only agents above the y^* cut-off respond to the tax reform, while in general one needs to take into account responses of all agents.

Two hallmark results about properties of optimal labor taxes in the Mirrlees model are widely known: that the optimal marginal taxes are positive and that they converge to zero for the highest earners. The top marginal tax rate can be seen immediately from (34) if we multiply it by $1-F_y(y^*)$ and take the limit $y^*\to y^{max}$ to obtain

$$-\frac{T'(y^{max})}{1-T'(y^{max})}y^{max}f_y(y^{max})\xi_y^H(y^{max})=0,$$

which can only be equal to zero if $T'(y^{max})=0$. The easiest way to see that marginal taxes are positive for all agents is by considering preferences with no income effects, $\eta_y^M=0$. Since the marginal utility

of consumption is decreasing in earnings, the expression $(1 - F_y(y^*))\mathbb{E}_{\mathrm{y}>y^*}\left(1 - \rho^{-1}\lambda^{-1}\mathrm{u}_c\right)$ must be positive for all y^* above the lowest earnings level, and it must converge to zero for the highest earners. Since the compensated elasticity ξ_y^H is always positive, optimal marginal taxes $\frac{T'(y^*)}{1-T'(y^*)}$ must inherit this pattern.

Neither of these two results extends to dynamic economies. The marginal tax rate for the top earners satisfies

$$-\frac{T'(y^{max})}{1-T'(y^{max})}y^{max}f_y(y^{max})\bar{\xi}_y(y^{max}) + \mathbb{E}_{y=y^{max}}\mathcal{E}_s(\lambda^{-1}v + R) = 0,$$

so that it balances the cost of labor distortions with the benefits induces by changing savings behavior of the top earners. Similarly, the saving saving may be sufficiently strong to make it desirable to subside labor supply of some agents. We illustrate in our quantitative section that such effects may arise in calibrated economies for agents with relatively low earnings, providing a rationale for labor subsidy programs in the spirit of the Earned Income Tax Credit in the U.S.

5.1.2 Optimal capital taxes

The characterization of optimal capital taxes is very similar to that of labor taxes. To derive optimal non-linear capital taxes we consider a family of tax reforms $\dot{T} = 0$ and $\dot{P}(a) = \iota_{a>a^*}$ for any asset level a^*; for linear taxes it is sufficient to consider the reform $\dot{P}(a) = a$.

Proposition 2. *(a) The optimal non-linear capital taxes satisfy, at every asset level a^*,*

$$\underbrace{\mathbb{E}_{\mathrm{a}>a^*}\left(1 - \rho^{-1}\lambda^{-1}u_c\right)}_{mechanical} + \underbrace{\mathbb{E}_{\mathrm{a}>a^*}\mathcal{H}_y\mathrm{T}'}_{labor} + \underbrace{\mathbb{E}_{\mathrm{a}>a^*}\mathcal{H}_s(\lambda^{-1}\mathrm{v} + \mathrm{R})}_{savings} = 0. \tag{35}$$

(b). The optimal linear capital taxes satisfy

$$\underbrace{\mathbb{E}a\left(1 - \rho^{-1}\lambda^{-1}u_c\right)}_{mechanical} + \underbrace{\mathbb{E}\mathrm{a}\mathcal{H}_y\mathrm{T}'}_{labor} + \underbrace{\mathbb{E}\mathrm{a}\mathcal{H}_s(\lambda^{-1}\mathrm{v} + \mathrm{R})}_{savings} = 0. \tag{36}$$

The optimal tax formula for non-linear capital taxes, equation (35), is very similar to the optimal non-linear labor taxes, equation (33). It is also simpler since, as we showed in Lemma 3, the marginal tax on savings $\dot{P}'(a)$ does not directly impact earnings and savings. Consequently, there is no analogue of the substitution effects in the formulas for the optimal capital tax.

When the planner uses unrestricted non-linear savings taxes, she can equalize the mechanical benefits of redistribution to the behavioral costs of taxation separately for each asset level a^*. When the planner is restricted to linear taxation, such a precise allocation of distortions is impossible. Instead, as shown in equation (36), the planner equates the benefits and the costs of taxation "on average", in proportion to asset holdings of each agent.

5.2 Taxation of gross income

The techniques that we developed to study optimal labor and capital taxes are easy to extend to other types of taxation. Suppose, for example, that the government taxes gross income $y+ra$ rather earnings and capital separately. Such a tax resembles the one used in the U.S., where the base for individual taxation consists of earnings and interest income.

Let $Q(y+ra)$ be the tax on the gross income $y+ra$. As before, we use y to denote earnings under such taxes and z = y + ra to denote the gross income. A tax reform of Q induces changes in the statutory tax levels $\dot{\mathrm{Q}}(x) = \dot{Q}(\mathrm{z}(x))$ and marginal tax rates $\dot{\mathrm{Q}}'(x) = \dot{Q}'(\mathrm{z}(x))$. Consequently, all our previous analysis applies if we replace $\dot{\mathrm{R}}(x)$ with $\dot{\mathrm{Q}}(x)$ and $\dot{\mathrm{T}}'(x)$ with $\dot{\mathrm{Q}}'(x)$. This immediately gives the following characterization of the optimal tax on gross income

Proposition 3. *The optimal non-linear tax on the gross income satisfies, at every level of gross income* z^*,

$$\underbrace{\mathbb{E}_{\mathrm{z}>z^*}\left(1-\rho^{-1}\lambda^{-1}\mathrm{u}_c\right)}_{mechanical} - \underbrace{\frac{Q'(z^*)}{1-Q'(z^*)}\frac{z^* f_z(z^*)}{1-F_z(z^*)}\mathbb{E}_{\mathrm{z}=z^*}\bar{\xi} - \mathbb{E}_{\mathrm{z}>z^*}\mathcal{H}_y\mathrm{Q}'}_{labor} \tag{37}$$
$$+\underbrace{\frac{f_z(z^*)}{1-F_z(z^*)}\mathbb{E}_{\mathrm{z}=z^*}\mathcal{E}_s(\lambda^{-1}\mathrm{v}+\mathrm{Q}) + \mathbb{E}_{\mathrm{z}>z^*}\mathcal{H}_s(\lambda^{-1}\mathrm{v}+\mathrm{Q})}_{savings} = 0,$$

where F_z *is the cumulative distribution of* z *and* f_z *is its density.*

Comparison of Propositions 1 and 3 reveals that the optimal taxes on labor earnings and on gross income are characterized by essentially the same formula. The only difference is that labor taxes use individual earnings as the base, and so all expectations are with respect to y, while gross income taxes use the gross income as a base and so all expectations are with respect to z.

5.3 Taxation with general equilibrium effects

So far, we have considered taxation in a small open economy where taxes do not affect aggregate wages and interest rates. In this section we extend our more general environments. Let w and r be aggregate wages and interest rates in the invariant distribution associated with some taxes (T, P). The agent's budget constraint is still given by (1) but the relationship between labor supply l and earnings y is now $y = we^{\theta}l$.

In general, tax reforms would affect both interest rates and wages. Let $(\dot{r}_\tau, \dot{w}_\tau)_{\tau\geq 0}$ be responses of these variables. Unlike the savings and earnings responses $\dot{\mathrm{s}}_\tau$, $\dot{\mathrm{y}}_\tau$, price responses $\dot{r}_\tau, \dot{w}_\tau$ are scalars. This makes them much easier to work with.

We organize our discussion in several steps. First, we treat the price responses $(\dot{r}_\tau, \dot{w}_\tau)_{\tau\geq 0}$ as given and extend our results from Section 4.1 to characterize the responses of individual policies to arbitrary

changes in taxes and aggregate prices. After that, we describe how to use the market clearing conditions to find $(\dot{r}_\tau, \dot{w}_\tau)_{\tau \geq 0}$. Finally, we extend our optimal tax formulas to economies with general equilibrium effects.

Our starting point is the following lemma that extends our main results from Section 4.1.

Lemma 6. *(a) The eesponses $\dot{\mathrm{v}}_\tau$ and $\partial_a \dot{\mathrm{v}}_\tau$ are characterized by equations (14) and (15) except that $\dot{\mathrm{R}}_{\tau+t}$ is replaced by $\dot{\mathrm{R}}_{\tau+t} - \dot{r}_{\tau+t}\mathrm{a} - (1-\mathrm{T}')\mathrm{y}\frac{\dot{w}_{\tau+t}}{w}$ in both formulas.*

(b). The earnings and savings responses $\dot{\mathrm{y}}_\tau$ and $\dot{\mathrm{s}}_\tau$ satisfy

$$\begin{aligned} \dot{\mathrm{y}}_\tau &= -\frac{\mathrm{y}\xi_y}{1-\mathrm{T}'}\dot{\mathrm{T}}'_\tau + \mathrm{y}\zeta_y \frac{\dot{w}_\tau}{w} + \eta_y \frac{\partial_a \dot{\mathrm{v}}_\tau}{\partial_a \mathrm{v}}, \\ \dot{\mathrm{s}}_\tau &= -\left(\dot{\mathrm{R}}_\tau - \dot{r}_\tau \mathrm{a} - (1-T')\mathrm{y}\frac{\dot{w}_\tau}{w}\right) - \frac{\xi_s}{1-\mathrm{T}'}\dot{\mathrm{T}}'_\tau + \zeta_s \frac{\dot{w}_\tau}{w} + \eta_s \frac{\partial_a \dot{\mathrm{v}}_\tau}{\partial_a \mathrm{v}}, \end{aligned} \tag{38}$$

where ζ_y and ζ_s are the Frisch elasticities of labor and savings with respect to wages.

(c). Lemmas 4 and 5 are unchanged.

Part (a) generalizes Lemmas 1 and 2. This extension is a straightforward application of the envelope theorem. Part (b) extends Lemma 3. The main new objects that appear are the pair of Frisch wage elasticities, ζ_y and ζ_s. As we explained in Section 3.2, these elasticities emerge because of non-linearities induced by labor taxes. When these taxes are locally linear, labor and savings respond to a wage change $\frac{\dot{w}_\tau}{w}$ in the same way as they respond to a change in the after-tax retention rate $\frac{\dot{\mathrm{T}}'_\tau}{1-\mathrm{T}'}$. However, when taxes are non-linear, the two responses differ by a term proportional to T''. Part (c) shows that our characterization of the response of asset distribution to tax reforms in Lemmas 4 and 5 also applies to economies with general equilibrium effects.

Lemma 6 describes how to find individual responses for a given tax reform and a sequence of price changes. Price changes, however, are not arbitrary but are determined by the market clearing conditions. We now describe how to find these changes. We first present key ideas using the market-clearing condition of Huggett (1993). This economy allows us to convey main insights in a particularly simple and transparent way. After that, we extend our analysis to a general constant returns to scale technology, such as the one used by Aiyagari (1994; 1995), and explain how general equilibrium responses modify our optimal tax formulas.

5.3.1 Taxation under Huggett price determination

Huggett (1993) assumes that wages are exogenous, which we normalize to one, but the interest rates are determined endogenously. In particular, he assumes that there is an exogenous fixed supply of assets $\bar{K}$ so that the market clear condition in his economy can be written as $\int \mathrm{a} d\mathrm{G} = \bar{K}$. This equation implies that the distributional changes must satisfy $\int \mathrm{a} d\dot{\mathrm{G}}_\tau = 0$ for all τ, which, in combination with Lemma

4 gives

$$\int \int_0^\tau \dot{\mathrm{s}}_{\tilde{\tau}}(x) \left[\int \mathrm{a}(\tilde{x}) \partial_a \Pi_{\tau - \tilde{\tau}}(d\tilde{x}|x) \right] d\tilde{\tau} \mathrm{G}(dx) = 0 \text{ for all } \tau. \tag{39}$$

Since wages are fixed, $\dot{w}_\tau = 0$ in the Huggett economy. Therefore, $\dot{\mathrm{s}}_{\tilde{\tau}}$ in equation (39) depends solely on the tax reform and future path of interest rates $(\dot{r}_\tau)_{\tau \geq \tilde{\tau}}$. Since (39) holds for every τ, this defines a system of linear equations that determine the path of interest rates $(\dot{r}_\tau)_{\tau \geq 0}$ for any tax reform.

To illustrate how this works, consider the long-run effect of time-invariant reforms on interest rates. Applying Lemma 4 and equation (32), we obtain the following result

Lemma 7. *The long-run responses of interest rates to tax reforms in the Huggett economy are given by*

$$\dot{r} = \varsigma^r \int \dot{\mathrm{T}}' \mathcal{E}_s \mathrm{a} d\mathrm{G} + \varsigma^r \int \dot{\mathrm{R}} \mathcal{H}_s \mathrm{a} d\mathrm{G}, \tag{40}$$

where the scalar ς^r is given by $\varsigma^r = (\int \mathrm{a} \mathcal{H}_s \mathrm{a} d\mathrm{G})^{-1}$.

Equation (40) characterizes the long-run interest rate response to an arbitrary tax reform. The dependence of interest rates on the marginal and total taxes are captured by the same operators $\mathcal{E}_s$ and $\mathcal{H}_s$ that we constructed in Section 5.1. The scalar ς^r aggregates Frisch income elasticities of savings of different agents and is known from the pre-reform economy.

Using (30), it is easy to extend our analysis of optimal taxation. Proceeding as in Section 5.1, we obtain

$$\lambda^{-1} \dot{\mathscr{W}} + \dot{\mathscr{R}} = \text{mechanical} + \text{labor} \ + \ \text{savings} \ + \varsigma^{opt} \dot{r}, \tag{41}$$

where mechanical, labor, and savings components are the same that defined in that section, and ς^{opt} is a pre-reform scalar parameter that, just like ς^r, is known explicitly. Substituting $\dot{r}$ from (40) to (41) gives the expression for $\lambda^{-1} \dot{\mathscr{W}} + \dot{\mathscr{R}}$ for any tax reform. Setting that expression to zero for all admissible reforms allows us characterize optimal taxes. For example, the expression for the optimal labor taxes is given by

$$\underbrace{\mathbb{E}_{\mathrm{y} > y^*} \left(1 - \rho^{-1} \lambda^{-1} \mathrm{u}_c\right)}_{\text{mechanical}} - \underbrace{\frac{T'(y^*)}{1 - T'(y^*)} \frac{y^* f_y(y^*)}{1 - F_y(y^*)} \mathbb{E}_{\mathrm{y} = y^*} \bar{\xi}_y - \mathbb{E}_{\mathrm{y} > y^*} \mathcal{H}_y \mathrm{T}'}_{\text{labor}} \tag{42}$$

$$+ \underbrace{\frac{f_y(y^*)}{1 - F_y(y^*)} \mathbb{E}_{\mathrm{y} = y^*} \mathcal{E}_s \mathrm{f} + \mathbb{E}_{\mathrm{y} > y^*} \mathcal{H}_s \mathrm{f}}_{\text{savings}} = 0.$$

where $\mathrm{f} := \lambda^{-1} \mathrm{v} + \mathrm{R} + \varsigma^r \varsigma^{opt} \mathrm{a}$. Comparison of equations (33) and (42) reveals that the expression for the optimal labor taxes in the Huggett economy is remarkably similar to the optimal tax formula in the small open economy. The only difference between the two formulas is that the savings effect includes an additional general equilibrium term, whose strength is determined by $\varsigma^r \varsigma^{opt}$. The formulas describing optimal taxes on capital and on gross income extend analogously.

The analysis of optimal taxation in the Huggett economy reveals several patterns that extend to more general production technologies. First, optimal taxes are still determined by the trade-offs between the mechanical gains from the redistribution and the costs of behavioral responses of labor and savings, but the savings effect includes also adjustments for general equilibrium responses. Second, these adjustments are characterized by the same operators that we constructed to study optimal taxation in a small open economy.

5.3.2 Taxation under Aiyagari price determination

Aiyagari (1994; 1995) assumes that output is produces by a constant returns to scale production technology $\mathcal{F}(K,L)$ where $L=\int e^{\theta}\mathrm{l}d\mathrm{G}$ is the efficiency-adjusted total labor supply. The interest rates and wages in this economy are given by marginal products of capital and labor, $r=\mathcal{F}_K(K,L)$ and $w=\mathcal{F}_L(K,L)$.

Since $\mathcal{F}_K$ and $\mathcal{F}_L$ are homogeneous of degree zero, aggregate prices and aggregate quantities of capital and labor are related via

$$\frac{\dot{r}_\tau}{r}=\frac{\mathcal{F}_{KK}K}{\mathcal{F}_K}\left(\frac{\dot{K}_\tau}{K}-\frac{\dot{L}_\tau}{L}\right),\qquad \frac{\dot{w}_\tau}{w}=\frac{\mathcal{F}_{LL}L}{\mathcal{F}_L}\left(\frac{\dot{L}_\tau}{L}-\frac{\dot{K}_\tau}{K}\right).$$

Unpacking $\dot{K}_\tau$ and $\dot{L}_\tau$ using our approach from Section 4.4, we can show that these equations imply

$$\frac{\dot{r}_\tau}{r}=\left[\frac{\mathcal{F}_{KK}K}{\mathcal{F}_K}\int\left(\frac{\mathrm{a}}{K}-\frac{\mathrm{y}}{wL}\right)d\dot{\mathrm{G}}_\tau+\frac{\dot{w}_\tau}{w}-\int\frac{\dot{\mathrm{y}}_\tau}{wL}d\mathrm{G}\right],\qquad \frac{\dot{w}_\tau}{w}=-\left(\frac{\mathcal{F}_{LL}L}{\mathcal{F}_L}/\frac{\mathcal{F}_{KK}K}{\mathcal{F}_K}\right)\frac{\dot{r}_\tau}{r}. \tag{43}$$

Substituting $\int\left(\frac{\mathrm{a}}{K}-\frac{\mathrm{y}}{wL}\right)d\dot{\mathrm{G}}_\tau=\int\int_0^\tau\dot{\mathrm{s}}_{\tilde{\tau}}(x)\left[\int\left(\frac{\mathrm{a}(\tilde{x})}{K}-\frac{\mathrm{y}(\tilde{x})}{wL}\right)\partial_a\Pi_{\tau-\tilde{\tau}}(d\tilde{x}|x)\right]d\tilde{\tau}\mathrm{G}(dx)$ into these equation, and $(\dot{\mathrm{s}}_\tau,\dot{\mathrm{y}}_\tau)_{\tau\geq0}$ from (38), we obtain a sequence of linear equations that determine the price responses $(\dot{r}_\tau,\dot{w}_\tau)_{\tau\geq0}$ for any tax reform. As an illustration of this approach, we derive analytical formulas for the long-run response of the interest rates.

Lemma 8. *The long-run response of interest rates to tax reforms in the Aiyagari economy is given by*

$$\begin{aligned}\dot{r}=&\varsigma_1^r\int\dot{\mathrm{T}}'\mathcal{E}_s\left(\frac{\mathrm{a}}{K}-\frac{\mathrm{y}}{wL}\right)d\mathrm{G}+\varsigma_2^r\int\dot{\mathrm{R}}\mathcal{H}_s\left(\frac{\mathrm{a}}{K}-\frac{\mathrm{y}}{wL}\right)d\mathrm{G}\\&+\varsigma_3^r\int\dot{\mathrm{T}}'\frac{\mathrm{y}\bar{\xi}_y}{1-\mathrm{T}'}d\mathrm{G}+\varsigma_4^r\int\dot{\mathrm{R}}\mathcal{H}_y\mathbf{1}d\mathrm{G},\end{aligned} \tag{44}$$

where the expressions for scalars ς_1^r, ς_2^r, ς_3^r, ς_4^r are given in the appendix. The long-run response of wages is $\dot{w}=-\left(\frac{\mathcal{F}_{LL}L}{\mathcal{F}_L}/\frac{\mathcal{F}_{KK}K}{\mathcal{F}_K}\right)\frac{w}{r}\dot{r}$.

The first line of equation (44) describes the effect of savings on interest rates and is the analogue of the right hand side of (40) in the Huggett economy. The second line describes the effect of changes in labor supply on interest rates, which in the Huggett economy are equal to zero.

It is easy to extend our optimal tax formulas for the Aiyagari economy using Lemma 8. In particular, the expression for $\lambda^{-1}\dot{\mathscr{W}}+\dot{\mathscr{R}}$ is still given by (41) except that the scalar ς^{opt} is different. Let $\varsigma_i := \varsigma_i^r \varsigma^{opt}$. By substituting $\dot{r}$ from (44) into that equation we obtain the following characterization of optimal labor taxes:

$$\underbrace{\mathbb{E}_{\mathrm{y}>y^*}\left(1-\rho^{-1}\lambda^{-1}\mathrm{u}_c\right)}_{\text{mechanical}} - \underbrace{\frac{T'(y^*)}{1-T'(y^*)}\frac{y^* f_y(y^*)}{1-F_y(y^*)}(1+\varsigma_3)\mathbb{E}_{\mathrm{y}=y^*}\bar{\xi}_y - \mathbb{E}_{\mathrm{y}>y^*}\mathcal{H}_y(\mathrm{T}'+\varsigma_4\mathbf{1})}_{\text{labor}} \tag{45}$$
$$+ \underbrace{\frac{f_y(y^*)}{1-F_y(y^*)}\mathbb{E}_{\mathrm{y}=y^*}\mathcal{E}_s\mathrm{f}_1 + \mathbb{E}_{\mathrm{y}>y^*}\mathcal{H}_s\mathrm{f}_2}_{\text{savings}} = 0,$$

where $\mathrm{f}_i = \lambda^{-1}\mathrm{v} + \mathrm{R} + \varsigma_i\left(\frac{\mathrm{a}}{K} - \frac{\mathrm{y}}{wL}\right)$ for $i = 1, 2$. Thus, the optimal labor tax formuls preserves all the insights of Proposition 1 but the labor and savings effects include additional general equilibrium adjustments, whose strength is determined by the four scalars $\varsigma_1, ..., \varsigma_4$. The analysis of other types of taxes is similar.

6 Quantitative application

In this section, we illustrate quantitative applications of our analysis. We consider optimal labor and capital taxation in the small open economy version of Aiyagari economy that we analyzed theoretically in Section 5.1. Our main focus is on understanding how dynamic consideration affects optimal labor and capital taxes.

6.1 Numerical algorithm

Our algorithm for computing optimal taxes proceeds in three main stages:

i. **Stage 1 (Model Solution)**: Given tax policies (T, P), solve the HJB equation and KF equation using a finite difference (FD) method. From the resulting policy functions, compute the key elasticities ξ_y and η_s.

ii. **Stage 2 (Gradient Evaluation)**: Form the first and second derivatives of the objective $\lambda^{-1}\mathscr{W} + \mathscr{R}$ with respect to perturbation scale ϵ under tax reform $T + \epsilon\dot{T}$ and $P + \epsilon\dot{P}$

iii. **Stage 3 (Optimization and Budget Balance)**: For a given Lagrange multiplier λ, iterate on (T, P) until the gradient is driven to zero. Then, update λ until the government budget is balanced.

Stage 1. Solution of HJB and KFE To implement Stage 1, we follow the continuous-time heterogeneous agent method popularized by Achdou et al. (2022). Working in continuous-time is particularly

advantageous here, as it allows for a significantly simpler characterization of borrowing constraints compared to discrete-time setups. Our approach departs from standard methods in two key respects: the treament of the labor supply decision and the formulation of capital returns.

We discretize the state space over a grid of points $x_{ij} = (a_i, \theta_j)$ defined on a $n_a \times n_\theta$ grid. We set the number of grid points for productivity and assets to $n_w = 30$ and $n_a = 200$, respectively. The upper bound $a_{\max}$ is chosen large enough so the points near that boundary carry zero probability up to machine precision.

To solve for optimal labor supply, we compute y in equation (9) using Newton's method. Because this procedure requires evaluating y across *any* positive real value, we parameterize the marginal labor tax $T'(y)$ as a *function*. We implement this using B-splines defined over a wide earnings domain, verifying that the choice of interval bounds does not truncate labor supply choices.

To incorporate non-linear asset taxes, we use a key insight: unlike discrete-time models, continuous-time settings yield no first order condition for asset holdings. Only the overall tax level $P(a)$ enters the household problem.[18] Consequently, for a given a n_a-dimensional vector P_i representing $P(a_i)$, no functional interpolation is required to solve the agent's problem. We simply replace the assets return ra with the net return $ra - P(a)$ and apply standard solution methods.

Once policy functions (c, s, y) are obtained, we compute key objects shaping welfare and revenue responses to tax reforms. Specifically, we define the $n_a \times n_\theta$ vector u_c, the marginal utility of consumption under optimal policies, as well as the elasticity vectors ξ_y and η_s defined in equation (11). Given the policy functions and B-spline specification, these elasticities are available in closed form.[19]

As a byproduct of solving the HJB and KFE, we obtain two matrices: a partial derivative matrix D_a that approximates $\partial_a \phi$ for functions of assets, and a matrix L which approximates the generator $\mathcal{L}$ describing the joint dynamics of assets and wages. Finally, we obtain the joint stationary distribution g.

2. Gradient Evaluation Given a Lagrange multiplier λ, Stage 2 constructs an approximation of the gradient, the discrete counterparts of the Gateaux derivatives of the planner's objective $\lambda^{-1}\mathscr{W} + \mathscr{R}$ with respect to generic reform directions $\left(\dot{T} + \dot{P}\right)$. For example, if preferences are GHH, the derivative with respect to the perturbational parameter decomposes into an impact through $\dot{T} + \dot{P}$ and $\dot{T}'$:

$$\frac{d}{d\epsilon}\left[\lambda^{-1}\mathscr{W} + \mathscr{R}\right] = \left\langle \Gamma \cdot g, \dot{T} + \dot{P} \right\rangle + \left\langle \frac{T'(\mathrm{y})}{1 - T'(\mathrm{y})} \cdot \mathrm{y} \cdot \xi_y \cdot g, \dot{T}' \right\rangle \tag{46}$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product over the state space $n_a \times n_\theta$, g is the stationary distribution vector and vector ξ_y captures earnings elasticities. The vector Γ summarizes social marginal

[18] The savings rate is implied by the consumption rate decision, in equation (10), which only depends on the value function through $q = \partial_a \mathrm{v}$.

[19] Because a B-spline is a function built from polynomials with exact derivatives.

valuations including wealth-distribution effects:

$$\Gamma = 1 - \rho^{-1}\lambda^{-1}\mathrm{u}_c + H_s\left(\lambda^{-1}\mathrm{v} + \mathrm{R}\right).$$

The first two terms represent the standard net social marginal benefit of redistribution at relative valuation λ. The matrix H_s captures the welfare and fiscal effects of a change in the wealth distribution.

$H_s\varphi$ for a generic vector φ is built as follows. First define $\tilde{L}$ as the matrix such that ϕ satisfies the linear system analogous to equation (28)

$$\phi = \tilde{L}\varphi: \qquad L\phi = \varphi - \langle \varphi, g\rangle$$

subject to $\langle \phi, g\rangle = 0$. Then

$$H_s\varphi := D_a\tilde{L}\varphi + \frac{\mathrm{u}_c}{g}\cdot\left\{\left(\rho - L^T\right)^{-1}\left[D_a^T\left(\eta_s\cdot\frac{D_a\tilde{L}\varphi}{\mathrm{u}_c}\cdot g\right)\right]\right\},$$

where L^T denotes the transpose of matrix L.

3. Optimization and Budget Balance Given a fixed Lagrange multiplier λ, Stage 1 allows us to compute the planner's objective $\lambda^{-1}\mathscr{W} + \mathscr{R}$, while Stage 2 provides a gradient with respect to tax reforms $\left(\dot{T}, \dot{P}\right)$. With these objects in hand, one can use gradient ascent or quasi-Newton algorithms to update the tax functions (T, P). Alternatively, when second-order derivatives are available, state-of-the-art trust-region methods can be used for faster and more accurate convergence.[20]

For a given λ, this inner optimization algorithm yields tax schedules that may not satisfy government budget balance. We wrap this routine within an outer loop that applies a root-finding algorithm to the budget imbalance $\langle R(\lambda), g\rangle$, thereby recovering optimal tax schedules that zero the budget.

6.2 Calibration

We focus on the small open economy version of Aiyagari economy. We assume that consumers have GHH preferences $u(c,l) = \ln(c - \frac{1}{1+1/\psi}l^{1+1/\psi})$ and set the labor supply elasticity to $\psi = 0.5$. We set parameters (κ, σ) of the stochastic process for θ so that the annual autocorrelation of wages is 0.9 and the stationary variance of log wages is 0.45, in line with estimates in Heathcote et al. (2023). The reflective boundaries are set at the ± 4.5 standard deviations of the unbounded OU process with these parameters. The mean value $\bar{\theta}$ is a normalization parameter that only affects the scale of the units. We choose the scale so that the average annual earnings are \$70K. The annual discount factor is set to $\rho = 0.04$. The interest rate r will be a free parameter since its comparative statics will be insightful

[20] In practice, we find that the cross-curvature between labor and asset tax schedules---as well as within non-linear asset taxes---is quantitatively significant. Standard gradient ascent methods ignore these off-diagonal Hessian terms, leading to slow convergence or numerical instability relative to trust-region algorithms.

about several economic mechanisms affecting optimal taxes. For our baseline specification we set $r = 0$, consistent with finding in Mehrotra and Sergeyev (2021) that safe interest rates in developed countries are usually close to or even below the rate of economic growth, which is zero in our economy.

Our choices of values of productivity and preference parameters are fairly standard. The main substantive assumption that we make is that labor supply has no income effects. We do it solely for pedagogical reasons as this assumption allows us to focus more clearly on novel mechanisms behind optimal taxation. Recall from Section 3.2 that both η_y and ξ_s are equal to zero under GHH preferences, which implies that operators $\mathcal{H}_y$ and $\mathcal{E}_s$ are identically equal to 0. Since labor supply does not depend on wealth distribution, we also have $\mathbb{E}_{\mathrm{y}>y^*}\mathcal{H}_s\mathrm{R} = \mathbb{E}_{\mathrm{y}>y^*}\mathcal{H}_s\mathrm{P}$. This allows us to simplify the optimal tax formulas (33) and (35) to

$$\frac{T'(y^*)}{1-T'(y^*)} = \underbrace{\frac{1-F_y(y^*)}{y^* f_y(y^*)\psi}\mathbb{E}_{\mathrm{y}>y^*}\left(1-\rho^{-1}\lambda^{-1}\mathrm{u}_c\right)}_{M_y(y^*)} + \underbrace{\frac{1-F_y(y^*)}{y^* f_y(y^*)\psi}\mathbb{E}_{\mathrm{y}>y^*}\mathcal{H}_s\mathrm{P}}_{SF_y(y^*)} + \underbrace{\frac{1-F_y(y^*)}{y^* f_y(y^*)\psi}\mathbb{E}_{\mathrm{y}>y^*}\mathcal{H}_s\frac{\mathrm{v}}{\lambda}}_{SW_y(y^*)} \quad (47)$$

and

$$0 = \underbrace{\frac{1-F_a(a^*)}{a^* f_a(a^*)}\mathbb{E}_{a>a^*}\left(1-\rho^{-1}\lambda^{-1}u_c\right)}_{M_s(a^*)} + \underbrace{\frac{1-F_a(a^*)}{a^* f_a(a^*)}\mathbb{E}_{a>a^*}\mathcal{H}_s\mathrm{P}}_{SF_s(a^*)} + \underbrace{\frac{1-F_a(a^*)}{a^* f_a(a^*)}\mathbb{E}_{a>a^*}\mathcal{H}_s\frac{\mathrm{v}}{\lambda}}_{SW_s(a^*)} = 0, \quad (48)$$

where F_a is the cdf of the distribution of assets and f_a is its density.

Equations (47) and (48) provide a useful decomposition of the trade-offs that shape optimal labor and capital taxes. Consider the labor tax decomposition first, equation (47). Term M_y captures the mechanical redistributory effect of labor taxes adjusted by the elasticity of labor supply. In static settings, this is the only mechanism that determines optimal labor taxes, and Mirrlees' tax formula can be written as $\frac{T'}{1-T'} = M_y$. Terms SF_y and SW_y capture how labor taxes affect government budget and welfare of future generations through their impact on asset distribution. Terms M_s, SF_s, and SW_s in equation (48) have analogous interpretation about impact of capital taxes: M_s captures the mechanical redistributory effect of those taxes, while SF_s and SW_s capture the effect of capital taxes on government revenue and welfare of future generations through the impact on wealth distribution.

6.3 Results

We first consider optimal labor taxes in the absence of capital taxation. This allows us to illustrate clearly some of the important insights about effects of labor taxation in dynamic settings. We then extend our analysis to taxation with both linear and non-linear capital taxes.

6.3.1 Optimal labor taxation without capital taxes

In Figure 1 we plot the optimal labor taxes in the calibrated economy and compare them to the optimal taxes in Mirrlees model. Recall from our discussion in Section 5.1 that we nest Mirrlees settings when

the interest rates are sufficiently low so that all agents are borrowing constrained. In our economy, this occurs when r is about -0.1.

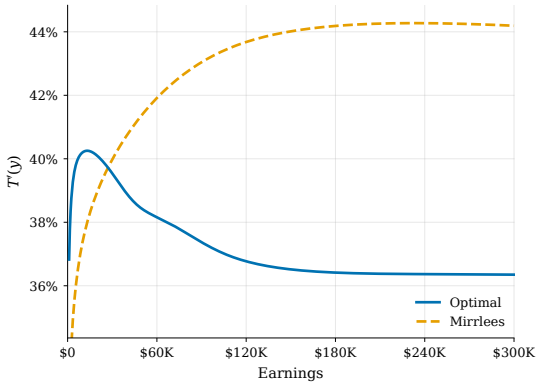


Figure 1: Optimal T' with no capital taxes in the baseline economy. Solid line: baseline line with $r = 0$. Dashed line: the model with $r = -0.1$. In the latter case, all agents are borrowing constrained and the optimal tax schedule coincides with the optimal taxes in static Mirrlees (1971) environment.

It is clear from Figure 1 that the optimal labor taxes in the dynamic economy differ substantially from static Mirrleesian taxes. Mirrleesian taxes are increasing in earnings for most agents and reach 45 percent for high-earners.[21] In contrast, the optimal marginal labor taxes in the dynamic economy decrease in earnings for most agents,[22] and reach 36 percent for high earners.

To understand what drives these differences in Figure 2(a) we decompose optimal labor taxes using equation (47). In Mirrlees economy, only the M_y component is present, which is shown by the dotted line. Since for now we assume that there are no capital taxes, $P = \mathbf{0}$, the decomposition of labor taxes in the dynamic economy consists of two terms, M_y and SW_y. Figure 2(a) shows that optimal labor taxes in the dynamic economy are generally lower for two reasons. First, the redistributive component M_y is uniformly lower in the dynamic economy, calling for lower labor taxes. Second, the SW_y is negative for most earnings levels, with being particular low for high earners. We now explain the intuition for these results.

Term M_y measures mechanical benefits of redistribution relative to the cost of labor distortion. Since the productivity distribution and labor supply elasticity is the same in both economies, the cost of labor supply distortion is the same. Thus, differences in M_y are driven by differences in the benefits of redistribution, which are proportional to $\mathrm{u}_c/\mathbb{E}\mathrm{u}_c$. In Mirrlees economy, differences in earnings translated

[21] Since the cross-sectional distribution of productivities is a truncated log-normal distribution, the optimal marginal tax rates eventually start to decrease for very high earners. We considered an extension of our model in which households occasionally receive large productivity shocks, which allows us to match the Pareto tail of the earnings distribution. Main qualitative and quantitative insights from that extension as the same as in the model in the text. In a companion paper (Golosov et al. (2026)) we systematically explore implications of fat tails in earnings and asset distribution for labor and capital taxation.

[22] The empirical analogue of $T(y)$ consists of all labor taxes and transfers paid or received by households as a function of their earnings y. Phase outs in transfers to the poor create positive effective marginal tax rate on earnings at low earnings levels. These effective marginal tax rates can be quite high in data and exceed top marginal tax rates for high earners. See Maag et al. (2012) for the description of effective marginal tax schedules in the U.S. data.

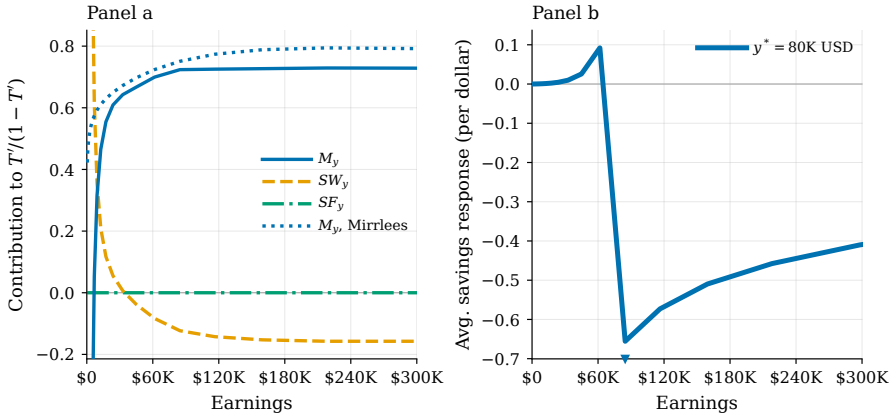


Figure 2: Panel (a): the three terms from decomposition (47). Panel (b): the response $\mathbb{E}[\dot{s}|y = y]$ as a function of y for a reform $\dot{T}(y) = 1_{y>y^*}$ where $y^* = 80K$.

one to one into differences in consumption, leading to a relatively large dispersion in marginal utilities and large gains from redistribution. In contrast, in dynamic economies savings allow agents to smooth productivity shocks. This reduces dispersion in the marginal utilities of consumption and lowers the benefits of redistribution. Consequently, the redistributive component calls for lower marginal taxes.

Term SW_y captures affects of changes in savings induced by labor taxes. Consider a tax reform $\dot{T}(y) = 1_{y>y^*}$ for a given y^*. Panel (b) of Figure 2 plots savings responses to this reform across earnings distribution, $\mathbb{E}[\dot{s}|y = y]$ as a function of y. We want to highlight three features of this graphs: (i) the direction of the savings response (i.e., whether agents increase or decrease savings following the reform) changes discontinuously at $y = y^*$, with high-earners decreasing savings and low-earners increasing savings, (ii) responses are smaller if y is further from y^*, and (iii) the net effect of the reform on savings is negative.

To understand these responses, consider agent's Euler equation, that we can write heuristically as

$$u'(c_t)dt = (1-e^{\rho dt})\{\Pr(y_{t+1} > y^*)\mathbb{E}_t[u'(c_{t+dt})|y_{t+1} > y^*] + \Pr(y_{t+1} \leq y^*)\mathbb{E}_t[u'(c_{t+dt})|y_{t+1} \leq y^*]\}. \quad (49)$$

Take two agents who are almost identical but one agent earners slightly more than y^* while the other agent earns slightly less. The tax reform decreases incomes of both agents in all future states in which they earn above y^*. Since agents are almost identical, the reform raises their right hand sides of (49) by a similar amount. In contrast, the responses of their left of sides of (49) are very different. The below-y^* agent does not experience any change in her current income, so the left hand of her (49) is unaffected. Consumption smoothing then implies that this agent should increase her savings. In contrast, the above-y^* agent sees an immediate drop in the disposable income, which increases the left hand side of her (49) by an amount that exceeds the expected increase in the marginal utility in the future. Therefore, such agent smoothes consumption by decreasing savings. This explains the

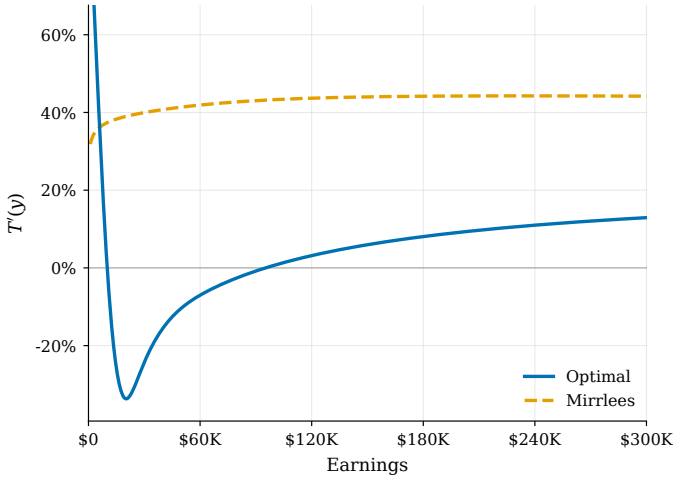


Figure 3: Optimal T' with no capital taxes with high interest rate. Solid line: economy with $r = 0.02$. Dashed line: the model with $r = -0.1$. In the latter case, all agents are borrowing constrained and the optimal tax schedule coincides with the optimal taxes in static Mirrlees (1971) environment.

discontinuity in Figure 2(b) around the y^* threshold. Savings of agents who are far away from y^* are less affected, since the probability that their earnings cross the y^* threshold is remote and both sides are (49) change more equally. Finally, there are two reasons why decreases of savings of the above-the-y^* earners exceed increases of savings of the below-the-y^* earners. One is mechanical. A person above the y^* threshold sees an immediate drop in the after-tax income, which makes less resources available for savings. Second, individuals with low earnings are more likely to have low assets. As we explained in Section 4.2 when we discussed savings responses near the borrowing constraints, agents with low assets are particularly reluctant to adjust their savings, which reduces their savings responses.

To main lesson from Figure 2(b) is that increasing labor taxes on high-earners has negative impact on aggregate wealth. Lower aggregate wealth has adverse welfare effect, which is captured by SW_y line in Figure 2(a). This lowers the optimal labor taxes relative to the static models in which wealth accumulation is absent.

Our discussion highlights the important role that interest rates play in optimal labor taxation. As interest rates increase, more agents move away from borrowing constraints, and savings responses get amplified. In Figure 3 we plot optimal labor taxes in the economy with the interest rate of two percent. Optimal labor taxes in this case fall even further below the Mirrlees benchmark. Moreover, the optimal tax rates are now negative for low-earners, resembling labor subsidy programs such as the Earned Income Tax Credit in the U.S.

6.3.2 Optimal labor and capital taxation

We now introduce capital taxation in addition to non-linear labor taxation. Optimal capital and labor taxation has been considered by number of authors, such as influential papers by Judd (1985), Chamley

(1986) and Aiyagari (1995). Fundamentally, those authors focus on a different notion of optimality than we do. Judd, Chamley, Aiyagari all consider taxes that maximize welfare of the initial generation – Adam and Eve, if one follows the biblical tradition of microfounding that generation – combined with additional restrictions on planner's ability to tax initial capital. The long-run behavior of optimal taxes, which is the central focus of those papers, is driven by that generation's discounting of the future and restrictions on initial capital taxes.[23] Instead, our optimality criterion treats all generations equally and we do not impose ad-hoc restrictions on capital tax restrictions.

Most of papers in the Ramsey tradition focus on linear capital taxation. It will be instructive to consider the trade-offs in setting such taxes before discussing non-linear capital taxation.

Linear capital taxation

Optimal linear capital taxes are characterized by equation (35). Under GHH preferences, there is no labor term in equation (35), the mechanical effect $\mathbb{E}a\left(1-\rho^{-1}\lambda^{-1}u_c\right)$ is positive while the welfare term $\mathbb{E}\mathrm{a}\mathcal{H}_s\mathrm{v}$ is negative. These last two terms highlight the basic trade-off that the planner faces when choosing capital taxes: higher capital taxes improve welfare of current generations by increasing redistribution but hurt future generations by lowering their wealth. The relative strength of these two effects depends on the magnitudes of the response of aggregate savings to capital taxes. As we discussed above, savings responses depend on the level of interest rates. In our baseline calibration with $r = 0$, the welfare term $\mathbb{E}\mathrm{a}\mathcal{H}_s\mathrm{v}$ prevails and the optimal capital taxes are negative, to the tune of -0.44 percent. The optimal labor taxes are roughly the same as in the economy without labor taxation, see Figure 4(a). The only notable difference from labor taxes in Figure 1 is that now high-earners face slightly higher marginal labor tax rates to offset some of the income transfer that they receive through capital subsidies.

If we lower r to -0.5 percent the redistributory motive dominates and the planner chooses to tax capital. In this case, the optimal marginal taxes for high-earners are slightly lower than in Figure 1 to offset some of the losses they incur due to linear capital taxes.[24] We show these taxes in Figure 4(b).

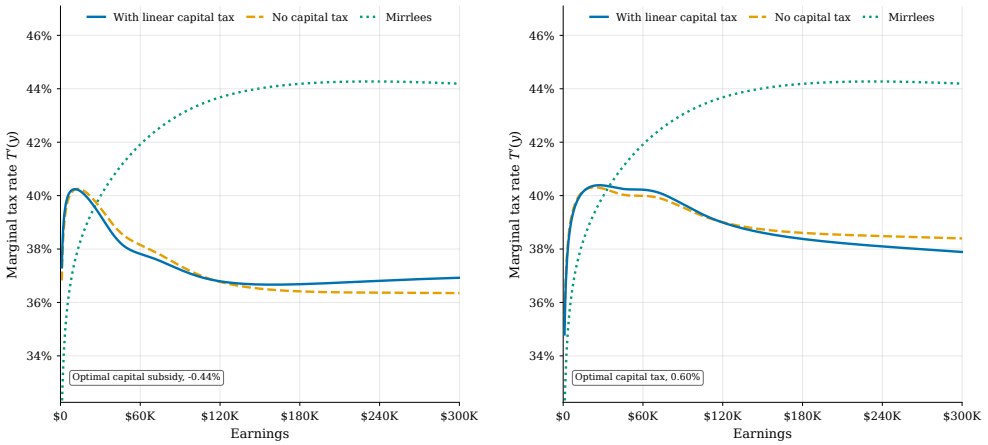
Non-linear capital taxation

We now turn to optimal non-linear labor and capital taxation. Panels (a) and (b) of Figure 5 plot T' and P'. To help visualize the optimal capital tax schedule, we also plot P in panel (c).

The optimal labor tax rates are still much lower than in Mirrlees economy but take somewhat different shape than in the case of no capital taxation. The marginal labor taxes are U-shaped, with

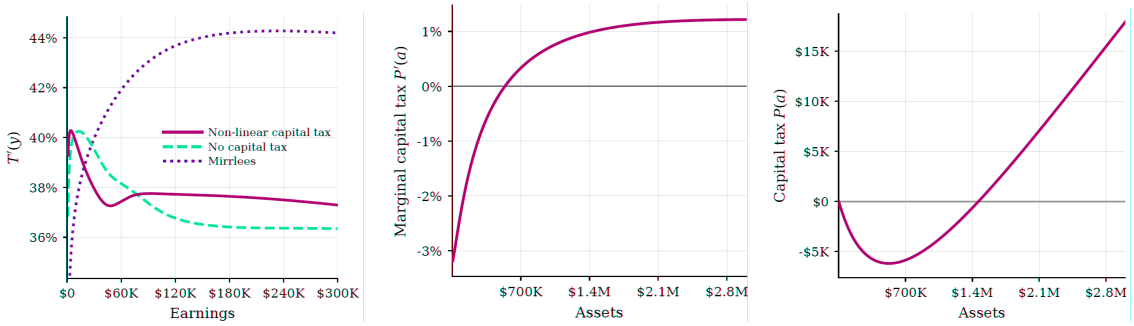
[23] Because initial olds discount the future, they are often happy to set taxes to achieve starvation in the long run. See Straub and Werning (2020) and Auclert et al. (2024). Restrictions on period 0 capital taxes often play an important role in this approach to prevent the planner from attaining her objectives through non-distortionary, period 0 taxation.

[24] Hellwig and Werquin (2022) make a similar point about implications of capital taxes for top marginal labor tax rates in a two period Mirrleesian model.



(a) $r = 0$ (b) $r = -0.005$

Figure 4: Optimal labor taxation under linear capital taxes under $r = 0$ and $r = -0.005$. Both panels also plot taxes from Figure 1 for reference.



(a) Optimal marginal labor tax T' (b) Optimal marginal capital tax P' (c) Optimal capital tax P

Figure 5: Optimal non-linear labor and capital taxes. Panel (a) also gives tax schedules from Figure 1 for references.

low-earners facing lower and high-earners facing higher rates than they were in Figure 1. The marginal capital tax rate P' is increasing in assets. It is negative for low-wealth individuals and positive for high-wealth individuals. The level of capital taxes P is U-shaped, with individuals with wealth of around \$600K receiving the largest subsidy.

To understand the intuition for these results, Figure 6 plots the decomposition of labor and capital taxes that we defined in equations (47) and (48). It will be instructive to consider the capital tax decomposition first. Panel (b) shows that the redistributive component of capital taxes M_a is positive while the welfare component SW_a is negative, mirroring our discussion of the trade-offs with linear taxes. Importantly, one can also see that in the absolute value the SW_a term dominates the M_a term for the asset-poor, while the opposite holds for the asset-rich. This represents the fact that it is particularly costly for a person to be born into a family with little wealth, where the SW_a term is particularly low, while gains from being born into a very wealthy family quickly decline with the marginal utility of

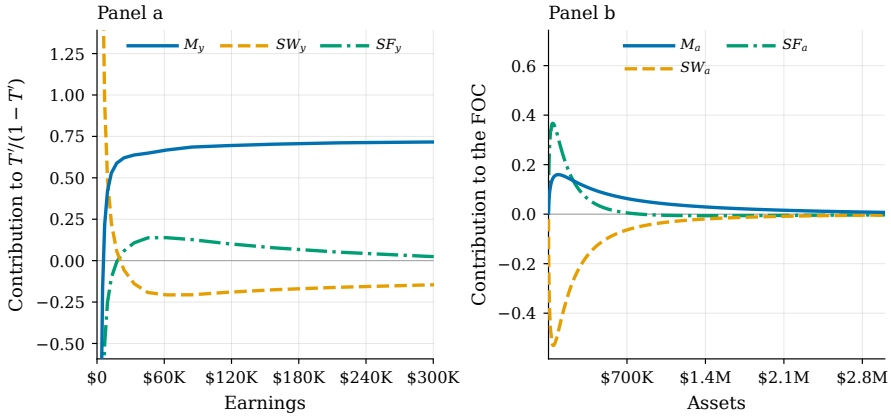


Figure 6: Panel (a): the decomposition (47) of optimal labor taxes. Panel (b): the decomposition (48) of optimal capital taxes.

consumption. The sum of M_a and SW_a gives raise to optimal shape of P' in Figure 5(b): the planner subsidizes wealth accumulation of the asset-poor and taxes the asset-rich. The fiscal impact of such policies depends not on the capital tax rate P' but on the capital tax level P. Since $P(0) = 0$ and the planner subsidizes the asset-poor, the fiscal component SF_a of capital taxes is inversely U-shaped as shown on Figure 6(b).

Panel (a) of Figure 6 plots the decomposition of labor taxes. There are two main differences from Figure 2(a). First, the fiscal component SF_y is no longer zero. SF_y is positive for most agents indicating a force for lower labor taxes. The intuition for this result closely follows from the thought experiment in Figure 2(b): higher labor taxes reduce savings which also reduces capital subsidy spending for the government. Second, the welfare component SW_y is closer to zero for high-earners than it was in Figure 2(a). Optimal capital taxes make the distribution of wealth more concentrated muting saving responses of high-earners.

7 Conclusion

We develop a general approach to studying the effects of taxation in dynamic stochastic incomplete-market environments. While we focused on one particular version of such environments — the economy considered by Aiyagari (1994) — our approach can be extended to models with other sources of risk, such as aggregate uncertainty or stochastic capital returns, to models with additional shocks that match the Pareto tails of the earnings and wealth distributions, and to models with rents and monopoly distortions. A careful quantitative analysis of such settings is a separate endeavor that we leave for future research.

We purposefully modeled altruism so that labor supply, consumption, and savings behavior in our model is identical to that in Aiyagari (1994), the workhorse model of the modern quantitative heterogeneous-agent literature. More generally, the shape that altruism takes and how it affects savings

is an empirical matter. Further work on estimating altruistic behavior in the data and incorporating it into incomplete-market models would be valuable not only for normative tax analysis but also for many quantitative questions in macroeconomics more broadly. It is also natural to consider additional forms of taxation, such as bequest taxes, in these settings. In practice, a significant share of resources is transferred intergenerationally outside of formal bequests, and a systematic analysis of the different channels through which such transfers occur, and of how different taxes affect them, is needed.

References

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. 2022. "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach." *The Review of Economic Studies*, 89(1): 45–86.

Aiyagari, S. Rao. 1994. "Uninsured Idiosyncratic Risk and Aggregate Saving." *The Quarterly Journal of Economics*, 109(3): 659–684.

Aiyagari, S. Rao. 1995. "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting." *Journal of Political Economy*, 103(6): 1158–1175.

Albanesi, Stefania, and Christopher Sleet. 2006. "Dynamic optimal taxation with private information." *Review of Economic Studies*, 73(1): 1–30.

Allcott, Hunt, Benjamin B. Lockwood, and Dmitry Taubinsky. 2019. "Regressive Sin Taxes, with an Application to the Optimal Soda Tax." *The Quarterly Journal of Economics*, 134(3): 1557–1626.

Auclert, Adrien, Michael Cai, Matthew Rognlie, and Ludwig Straub. 2024. "Optimal Long-Run Fiscal Policy with Heterogeneous Agents." Working paper, August 2024.

Barro, Robert J., and Gary S. Becker. 1989. "Fertility Choice in a Model of Economic Growth." *Econometrica*, 57(2): 481–501.

Becker, Gary S., and Robert J. Barro. 1988. "A Reformulation of the Economic Theory of Fertility." *The Quarterly Journal of Economics*, 103(1): 1–25.

Becko, John Sturm, and André Sztutman. 2026. "Efficiency Criteria, Income Taxation, and Heterogeneous Elasticities." *American Economic Review*, Forthcoming.

Bierbrauer, Felix J., Pierre C. Boyer, and Emanuel Hansen. 2023. "Pareto-Improving Tax Reforms and the Earned Income Tax Credit." *Econometrica*, 91(3): 1077–1103.

Blundell, Richard, Luigi Pistaferri, and Itay Saporta-Eksten. 2016. "Consumption Inequality and Family Labor Supply." *American Economic Review*, 106(2): 387–435.

Chamley, Christophe. 1986. "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives." *Econometrica*, 54(3): 607–622.

Chang, Yongsung, and Yena Park. 2021. "Optimal Taxation with Private Insurance." *Review of Economic Studies*, 88(6): 2766–2798.

Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger. 2009. "Taxing Capital? Not a Bad Idea after All!." *American Economic Review*, 99(1): 25–48.

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

Dyrda, Sebastian, and Marcelo Pedroni. 2023. "Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Income Risk." *The Review of Economic Studies*, 90(2): 744–780.

Farhi, Emmanuel, Alan Olivi, and Iván Werning. 2022. "Price Theory for Incomplete Markets." Working Paper 30037, National Bureau of Economic Research.

Farhi, Emmanuel, and Ivan Werning. 2013. "Insurance and Taxation over the Life Cycle." *Review of Economic Studies*, 80(2): 596–635.

Ferey, Antoine, Benjamin B. Lockwood, and Dmitry Taubinsky. 2024. "Sufficient Statistics for Nonlinear Tax Systems with General Across-Income Heterogeneity." *American Economic Review*, 114(10): 3206–3249.

Ferriere, Axelle, Philipp Grubener, Gaston Navarro, and Oliko Vardishvili. 2023. "On the Optimal Design of Transfers and Income Tax Progressivity." *Journal of Political Economy Macroeconomics*, 1(2): .

Findeisen, Sebastian, and Dominik Sachs. 2017. "Redistribution and Insurance with Simple Tax Instruments." *Journal of Public Economics*, 146 58–78.

Golosov, Mikhail, Michael Graber, Magne Mogstad, and David Novgorodsky. 2024. "How Americans Respond to Idiosyncratic and Exogenous Changes in Household Wealth and Unearned Income." *The Quarterly Journal of Economics*, 139(2): 1321–1395.

Golosov, Mikhail, Narayana Kocherlakota, and Aleh Tsyvinski. 2003. "Optimal Indirect and Capital Taxation." *Review of Economic Studies*, 70(3): 569–587.

Golosov, Mikhail, Ilia Krasikov, and Ken Miyahara. 2026. "Risk, markups, and optimal labor and capital taxation."Technical report.

Golosov, Mikhail, Maxim Troshkin, and Aleh Tsyvinski. 2016. "Redistribution and Social Insurance." *American Economic Review*, 106(2): 359–86.

Golosov, Mikhail, and Aleh Tsyvinski. 2007. "Optimal Taxation with Endogenous Insurance Markets." *Quarterly Journal of Economics*, 122(2): 487–534.

Golosov, Mikhail, Aleh Tsyvinski, and Nicolas Werquin. 2014. "A Variational Approach to the Analysis of Tax Systems." Working Paper 20780, National Bureau of Economic Research.

Greenwood, Jeremy, Zvi Hercowitz, and Gregory W. Huffman. 1988. "Investment, Capacity Utilization, and the Real Business Cycle." *American Economic Review*, 78(3): 402–417.

Harsanyi, John C. 1953. "Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking." *Journal of Political Economy*, 61(5): 434–435.

Harsanyi, John C. 1955. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility." *Journal of Political Economy*, 63(4): 309–321.

Harsanyi, John C. 1975. "Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls's Theory." *American Political Science Review*, 69(2): 594–606.

Heathcote, Jonathan, Fabrizio Perri, Giovanni L. Violante, and Lichen Zhang. 2023. "More Unequal We Stand? Inequality Dynamics in the United States, 1967–2021." *Review of Economic Dynamics*, 50 235–266.

Hellwig, Christian. 2021. "Static and Dynamic Mirrleesian Taxation with Non-separable Preferences: A Unified Approach." CEPR Discussion Paper 16254, Centre for Economic Policy Research.

Hellwig, Christian, and Nicolas Werquin. 2022. "A Fair Day's Pay for a Fair Day's Work: Optimal Tax Design as Redistributional Arbitrage." TSE Working Papers 22-1284, Toulouse School of Economics (TSE).

Hendren, Nathaniel. 2016. "The Policy Elasticity." *Tax Policy and the Economy*, 30(1): 51–89.

Hosseini, Roozbeh, Larry E. Jones, and Ali Shourideh. 2013. "Optimal Contracting with Dynastic Altruism: Family Size and Per Capita Consumption." *Journal of Economic Theory*, 148(5): 1806–1840.

Huggett, Mark. 1993. "The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies." *Journal of Economic Dynamics and Control*, 17(5–6): 953–969.

Jacquet, Laurence, Etienne Lehmann, and Bruno Van der Linden. 2013. "Optimal Redistributive Taxation with Both Extensive and Intensive Responses." *Journal of Economic Theory*, 148(5): 1770–1805.

Judd, Kenneth L. 1985. "Redistributive Taxation in a Simple Perfect Foresight Model." *Journal of Public Economics*, 28(1): 59–83.

Kanwal, Ram P. 1998. *Generalized Functions: Theory and Technique.*: Springer Science & Business Media.

Maag, Elaine, C. Eugene Steuerle, Ritadhi Chakravarti, and Caleb Quakenbush. 2012. "How Marginal Tax Rates Affect Families at Various Levels of Poverty." *National Tax Journal*, 65(4): 759–782.

Mehrotra, Neil R., and Dmitriy Sergeyev. 2021. "Debt Sustainability in a Low Interest Rate World." *Journal of Monetary Economics*, 124(S): S1–S18.

Mirrlees, James. 1971. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies*, 38(2): 175–208.

Ndiaye, Abdoulaye. 2018. "Flexible Retirement and Optimal Taxation." Working Paper 2018-18, Federal Reserve Bank of Chicago.

Piketty, Thomas, and Emmanuel Saez. 2013. "A Theory of Optimal Inheritance Taxation." *Econometrica*, 81(5): 1851–1886.

Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Sachs, Dominik, Aleh Tsyvinski, and Nicolas Werquin. 2020. "Nonlinear Tax Incidence and Optimal Taxation in General Equilibrium." *Econometrica*, 88(2): 469–493.

Saez, Emmanuel. 2001. "Using Elasticities to Derive Optimal Income Tax Rates." *Review of Economic Studies*, 68(1): 205–229.

Saez, Emmanuel. 2002. "The Desirability of Commodity Taxation under Non-Linear Income Taxation and Heterogeneous Tastes." *Journal of Public Economics*, 83(2): 217–230.

Scheuer, Florian, and Iván Werning. 2017. "The Taxation of Superstars." *The Quarterly Journal of Economics*, 132(1): 211–270.

Straub, Ludwig, and Ivan Werning. 2020. "Positive Long-Run Capital Taxation: Chamley-Judd Revisited." *American Economic Review*, 110(1): 86–119.