The limits of consumption taxation: A theory of optimal capital taxation*

Rafael Berriel† Ravi Jagadeesan‡

March 20, 2025

Abstract

In canonical models with wealth inequality, it is undesirable to tax capital: redistribution can be achieved through consumption taxation without any distortions. These models assume that consumption is observable to the government. However, in practice, business owners can disguise their consumption as a business expense to avoid or evade consumption taxes. How do such imperfections in the enforceability of consumption taxes affect optimal tax policy? We approach this question from a dynamic mechanism design perspective in general equilibrium. The key trade-off that the planner faces is that capital taxation can provide more revenue than consumption taxation, but lowers wages over time. We analytically solve for the long run state of the economy under the optimal dynamic mechanism. The optimal allocation requires sustained intertemporal distortions—as entailed, for example, by sustained capital taxation. The optimal policy can be implemented by a combination of time-varying linear taxes/subsidies on wealth, capital income, and consumption.

Keywords: capital taxation; consumption taxation; dynamic mechanism design

*We would like to thank Daniel Chen, Wioletta Dziuda, Patrick Kehoe, Scott Kominers, Spencer Kwon, Omeed Maghzian, Stefanie Stantcheva, Frank Yang, and, especially, Emmanuel Farhi, for helpful comments. Jagadeesan was supported by a National Science Foundation Graduate Research Fellowship under Grant No. DGE-1745303, and by the Washington Center for Equitable Growth.

†Department of Economics, Stanford University. Email address: rberriel@stanford.edu

‡Department of Economics, Stanford University. Email address: ravi.jagadeesan@gmail.com.

1 Introduction

As wealth and capital income have concentrated in the hands of a small fraction of the population (Saez and Zucman, 2016; Piketty et al., 2018), capital taxation has become increasingly salient in the public debate regarding taxation. However, as capital taxation causes intertemporal distortions, simple models provide arguments against its use to reduce wealth inequality (Kaldor, 1955). Rather, taxing consumption, and using the revenue to lower labor income tax rates and/or fund a demogrant, would allow everyone to share in the fruits of wealth, thereby effectively expropriating initial wealth (Kaldor, 1955; Chamley, 1980).

In practice, however, consumption taxes are imperfect: business owners can disguise their consumption as business expenses to avoid consumption taxes.[1] To understand why this issue arises, recall that consumption taxes are best implemented via value-added taxes (VAT), whereby businesses remit taxes on their sales to the government, but deduct taxes already paid on business inputs from the tax bill. But there is an issue of exactly which inputs should be deductible. For a VAT to be non-distortive, it is critical that all production inputs be fully deductible. Conversely, if any business expense that is not a true production input were deductible for VAT purposes, there would be a channel for tax avoidance: business owners could disguise parts of their own consumption as business expenses to avoid the VAT.

Hence, the structure of a VAT relies on a sharp dichotomy between productive and unproductive business expenses. But there is no such black-and-white distinction in practice. For example, business class plane tickets, corporate cars, parties for employees, and expensive haircuts for TV stars all lie in the "grey area" of expenses that share qualities of both productive spending and disguised consumption.

This paper studies how the imperfection of consumption taxes affects optimal taxation when wealth is concentrated in the hands of capitalists. Our main results characterize the state of the economy in the long run under the optimal dynamic tax policy, as well as the optimal long-run distortions. Thus, our analysis characterizes the optimal balance between consumption and capital taxation in the long run.

To highlight the role of disguised consumption in disciplining the role of consumption taxation, we abstract from agency issues within businesses and model capitalists

[1] A growing empirical literature documents that individuals shift income between individual and corporate tax bases in order to minimize tax payments (see, e.g., Gordon and Slemrod (2000), Devereux et al. (2014), and Alstadsæter and Jacob (2016)).

as the owner-managers of businesses that hold their wealth. We suppose that capitalists experience idiosyncratic risk, which is nonverifiable to the planner. Capitalists can invest to increase their expected profit. However, they can also disguise their consumption as business expenses, which, unlike investment, does not improve profit. We suppose that the social planner cannot distinguish productive spending from disguised consumption for the owner-manager. This imperfect observability limits the planner's ability to redistribute via consumption taxation: business owners can avoid pure consumption taxes by disguising consumption as a business expense.[2]

The planner seeks to maximize the (discounted) utility of the workers who consume hand-to-mouth. However, the planner is constrained to provide a minimum level of utility to the capitalists—representing, e.g., political constraints. As this minimum utility level varies, the optimal allocation traces out the constrained Pareto frontier.

Rather than optimizing within a fixed set of tax instruments, we take a dynamic mechanism design approach. Thus, the model determines not only the trade-off between consumption and capital taxes, but also which tax instruments should be considered in the first place. For example, the model allows for arbitrary forms of capital taxation, including both capital income taxes and wealth taxes.

We focus on understanding the long-run state of the economy under the optimal dynamic tax policy. By a steady state, we mean a dynamic allocation in which the capital stock and the consumption of the workers are constant, though the capitalists' consumption levels and capital holdings may each vary idiosyncratically over time. To ensure the existence of a steady state with positive consumption for the workers, we suppose that the planner is more patient than the worker.[3] Our main technical contribution is to prove that if the planner is sufficiently more patient than the capitalists, a steady state exists, and is in fact unique and locally saddle-path stable.

Despite the complexity of the model, we can fully characterize the steady state in closed form. Three key features emerge. First, the capitalists' total consumption

[2]Under the standard assumption that the planner could perfectly distinguish productive business spending from unproductive business expenses that are in fact disguised owner-manager consumption (Kaldor, 1955; Chamley, 1980), the planner could implement the first best.

[3]When capitalists are as patient as the planner (or more patient than the planner), we show that the optimal policy immiserates the worker in the long run. A similar issue arises in settings where there is asymetric information due to repeated taste shocks (Atkeson and Lucas, 1992), where a more patient planner also would not induce immiseration (Phelan, 2006; Farhi and Werning, 2007); as well as in in models of business cycles with financial frictions in which "experts" or "financial intermediaries" are subject to moral hazard problems (see, e.g., Bernanke et al. (1999)).

as a fraction of the capital stock is proportional to how efficiently they can disguise consumption as a business expense. Intuitively, while the planner would prefer to allocate all consumption to workers in the long run, the capitalists must retain some consumption to deter them from disguising consumption as a business expense.

Second, the intertemporal distortion for the capitalists—i.e., their steady state *savings wedge*—is positive and takes a simple form that depends only on the discount rates and the capitalists' elasticity of intertemporal substitution. Thus, the imperfection of consumption taxes makes long-run intertemporal distortions optimal.

Third, the capital stock is strictly below the modified golden rule level. Indeed, we show that the marginal product of capital exceeds the modified golden rule level by an amount that is increasing in both the savings wedge and the fraction of the capital stock devoted to the capitalists' total steady state consumption. Thus, the capitalists' consumption and the savings wedge mediate how much the moral hazard problem causes intertemporal inefficiencies for the planner.

We also develop an implementation of the optimal policy via linear taxes/subsidies. We show that time-varying linear taxes/subsidies on wealth, unexpected capital income, and consumption can implement that optimal dynamic allocation. In steady state, the taxes/subsidies are constant, and we characterize them in closed form.

The optimal capital income and consumption taxes turn out to bear a simple relation that can be seen from the moral hazard problem. Specifically, the optimal taxes keep the capitalist indifferent (on the margin) between two methods to consume. First, the capitalist can consume in the open, thereby paying the consumption tax. Second, the capitalist can disguise consumption as a business expense, incurring efficiency costs but avoiding both the consumption tax and the capital income tax. To maintain this indifference, the planner faces a trade-off between taxing consumption and taxing unexpected capital income: the latter is desirable as it helps spread risk across capitalists, but limits the ability to raise revenue through the former. For example, when capitalists can disguise consumption as a business expense with relatively low costs, the optimal consumption tax is lower.

The optimal wealth tax balances two forces. First, as the planner is more patient than the capitalists, there is a mechanical motive to subsidize capital, which benefits workers in the long run by raising wages. Second, there is a motive to tax capital for redistributive purposes. How these two forces balance depends on the strength of the moral hazard problem. When it is difficult for capitalists to disguise consumption as

a business expense, consumption taxes are more effective as a redistributive tool, and the motive to subsidize capital dominates, and leads to a negative wealth tax. When it is relatively easy for capitalists to disguise consumption as a business expense, consumption taxes are less effective, so it is optimal instead to tax wealth.

One of our main contributions is to develop a tractable framework for studying capital taxation in which a practical limit to consumption taxes leads to a non-trivial trade-off between capital and consumption taxes. From a methodological perspective, we obtain tractability by using continuous time methods. Our formulation of the dynamic agency problem faced by the planner builds on DeMarzo and Sannikov's (2006) continuous-time model of asset management contracts, as extended by Di Tella (2019) to settings with income effects (see also Di Tella and Sannikov (2021)).[4] We incorporate a labor sector to ensure the existence of a steady state, as in the neoclassical growth model. This allows us to characterize the long run state of the economy under the optimal dynamic allocation in closed form, and to derive simple expressions for the long-run wedges and long-run implementing taxes.

Related literature. This paper relates to four strands of the literature on capital taxation. One influential strand of the literature, initiated by Chamley (1986) and Judd (1985, 1999), considers taxation of existing wealth and restricts the set of instruments available to the planner—ruling out, for example, consumption taxation, to prevent the planner from expropriating capitalists' initial wealth. In such settings, the optimal long-run tax rates have been shown to be zero. However, for low elasticities of intertemporal substitution, the optimal long-run tax rate can be so high that the capital stock falls to its lowest feasible level (Straub and Werning, 2020); for Judd's (1985) model, moderate sustained capital taxation is optimal only in the knife-edge case in which capitalists have log utility (Lansing, 1999). By contrast, this paper does not restrict the set of tax instruments, and shows that moderate sustained intertemporal distortions are optimal in steady state.

In this vein, two more recent papers incorporate effects of taxes on entrepreneurial activity and invesment. Chari et al. (2020) develop a model in which taxes affect the investment and risk-taking decisions of entrepreneur, and show it is consistent with the dynamics of US wealth inequality. They then quantitatively study the impact of taxes in dynamic general equilibrium. Amol (2024) extends Judd (1985) by allowing

[4]Note that Di Tella (2019) incorporated hidden trade, and Di Tella and Sannikov (2021) incorporated hidden savings, which are beyond the scope of our analysis.

(as we do) capitalists to consume through their business and show that the model can match the current levels of US taxes. While we do not consider the dynamics of wealth inequality, we take a mechanism design approach that allows the planner access to all potential tax instruments and solve for the optimal dynamic tax policy.

A second strand of literature generates positive capital income taxes at optimum by supposing that agents acquire capital for precautionary savings (Aiyagari, 1995), or due to bequest motives (Piketty and Saez, 2013) or utility for wealth (Saez and Stantcheva, 2018). By contrast, this paper focuses on capital held by the rich, for which precautionary motives are unlikely to apply, and considers more standard preferences to isolate the interaction between consumption and capital taxation.

A third strand of literature considers the savings wedges entailed by optimal dynamic allocations when agents' skill levels evolve stochastically over time. In such settings, when utility is separable between consumption and leisure, savings wedges are positive (Rogerson, 1985; Golosov, Kocherlakota, and Tsyvinski, 2003). This result is driven by the fact that savings tighten future incentive constraints for labor income taxation (see, e.g., Golosov, Tsyvinski, and Werning (2006)). For example, savings wedges in such models decline to 0 over the life cycle (Farhi and Werning, 2013; Stantcheva, 2017). By contrast, the moral hazard problem in this paper arises due to the riskiness of returns on capital, and the positivity of the savings wedge does not rely on an interaction between capital taxation and labor income taxation.

A fourth strand of the literature studies how differences in tax bases affect individual incentives to shift income in order to avoid taxes. Gordon and Slemrod (2000) exploit time series from the US and find substantial income shifting for high-income people in response to differences between corporate and personal taxes. Using data from the UK, Devereux et al. (2014) estimate the elasticity of corporate income into real responses and income shifting and find a rather modest income shifting elasticity. Alstadsæter and Jacob (2016) use rich Swedish administrative microdata and find robust evidence of extensive income shifting across tax bases in response to a 2006 dividend tax cut. Our paper incorporates these tax avoidance concerns in a canonical neo-classical growth model to understand its impact on the optimal tax structure.

Paper outline. We proceed as follows. Section 2 sets up the model. Section 3 characterizes the optimal dynamic allocation. Section 4 develops an implementation via linear taxes. Section 5 summarizes our solution strategy. Section 6 is a conclusion.

2 Model

2.1 Environment

Time is continuous. There is a representative worker and a continuum of capitalists j, indexed by $\mathbb{J} = [0, 1]$. There is an initial capital stock K_0; capital depreciates at rate δ.

As in Judd (1985), the worker lives hand to mouth and inelastically supply 1 unit of labor to a competitive labor market. Letting $C_{w,t}$ denote the consumption of the representative worker at time t, the worker's utility is given by

$$\int_0^\infty e^{-\zeta t} v(C_{w,t})\, dt, \tag{1}$$

where $\zeta > 0$ is a social discount rate, and $v : \mathbb{R}_{>0} \to \mathbb{R}$ is a Bernoulli utility function.[5] We also allow for the possibility that $\zeta = 0$, in which case the worker's utility is

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T v(C_{w,t})\, dt.$$

Each capitalist j has Epstein–Zin preferences over consumption processes $C_{j,t}$, with discount rate ρ, coefficient of relative risk aversion γ, and elasticity of intertemporal substitution $\psi > 1$. For ease of exposition in the main text, we occasionally write formulae assuming CRRA preferences—i.e, $\gamma = 1/\psi$, but all results are stated for the general case. In the CRRA case, capitalist j's utility is

$$U_{j,0} \equiv \mathbb{E}_0 \int_0^\infty e^{-\rho t} C_{j,t}^{1-\frac{1}{\psi}}\, dt.$$

Appendix A describes Epstein–Zin case, which is based on Duffie and Epstein (1992).

Businesses. Each capitalist is the owner-manager of a business. Without investment, if the capitalist has $K_{j,t}$ units of capital and employs $L_{j,t}$ units of labor, then their expected production at time t is $F(K_{j,t}, L_{j,t}) = K_{j,t}^\theta L_{j,t}^{1-\theta}$, where $0 < \theta < 1$ is the capital share. Writing $K_t = \int_{\mathbb{J}} K_{j,t} dj$ for the aggregate capital stock at time t, perfect competition in the labor market implies that the equilibrium wage rate at time t is $w_t = F_L(K_t, 1)$ and capitalist j's equilibrium labor demand is $L_{j,t} = K_{j,t}/K_t$.

[5] The worker's Bernoulli utility function v is strictly increasing, continuously differentiable, strictly concave, and satisfies Inada conditions.

At each point in time, a capitalist can invest the final good to improve their expected earnings. Letting $I_{j,t}$ denote capitalist j's invest in their business at time t, the capitalist accrues an additional expected earnings $\alpha(I_{j,t}, K_{j,t})$, where α is strictly increasing, strictly concave, homogeneous of degree 1, and satisfies Inada conditions.

Last, each capitalist j experiences idiosyncratic risk $\sigma dZ_{j,t}$ per unit of capital, where $Z_{j,t}$ is a standard Brownian motion that is i.i.d. across j. Thus, capitalist j's profits at time t are

$$d\Pi_{j,t} \equiv \Big[\underbrace{F(K_{j,t}, L_{j,t})}_{\text{production}} - \underbrace{w_t L_{j,t}}_{\text{wages}} - \underbrace{\delta K_{j,t}}_{\text{depreciation}} + \underbrace{\alpha(I_{j,t}, K_{j,t})}_{\text{addt'l earnings}} - \underbrace{I_{j,t}}_{\text{investment}}\Big] dt + \underbrace{\sigma K_{j,t} dZ_{j,t}}_{\text{risk}}. \quad (2)$$

Disguised consumption. The capitalists can disguise consumption as business expenses, which are observably indistinguishable from productive investment. Letting $E_{j,t} \geq 0$ be the amount that capitalist j spends on consumption that is disguised as business expenses and $\widetilde{C}_{j,t}$ their observed consumption, their (effective) consumption level is $C_{j,t} = \widetilde{C}_{j,t} + \phi E_{j,t}$, where $0 < \phi < 1$ is represents value of disguised consumption relative to consumption that is directly observable by the planner.

A capitalist's business expense level is $\widetilde{I}_{j,t} \equiv I_{j,t} + E_{j,t}$. Thus, rather than observing the true profits $d\Pi_{j,t}$, the planner observes (and can verify) reported profits

$$d\widetilde{\Pi}_{j,t} \equiv \Big[\underbrace{F(K_{j,t}, L_{j,t})}_{\text{production}} - \underbrace{w_t L_{j,t}}_{\text{wages}} - \underbrace{\delta K_{j,t}}_{\text{depreciation}} + \underbrace{\alpha(I_{j,t}, K_{j,t})}_{\text{addt'l earnings}} - \underbrace{\widetilde{I}_{j,t}}_{\text{bus. expenses}}\Big] dt + \underbrace{\sigma K_{j,t} dZ_{j,t}}_{\text{risk}}. \quad (3)$$

For tractability, we assume that business expenses $\widetilde{I}_{j,t}$ are non-verifiable. As we show in Lemma 1, this assumption implies that, in every period, the investment level $I_{j,t}$ is the solution of a simple static problem that only depends on $K_{j,t}$. In turn, this assumption allows us to solve the steady state of the economy analytically, greatly simplifying the analysis. Nevertheless, this assumption is not crucial for the main messages of our results. In the Appendix, we consider an extension where business expenses are verifiable, and derive analogues of many of our main results.

Resource constraint. The resource constraint for the economy is

$$\underbrace{C_{w,t} + \int_{\mathbb{J}} \widetilde{C}_{j,t}\, dj}_{\text{observed consumption}} + \underbrace{\int_{\mathbb{J}} \widetilde{I}_{j,t}\, dj}_{\text{bus. expenses}} + \underbrace{\dot{K}_t + \delta K_t}_{\text{gross savings}} \leq \underbrace{\int_{\mathbb{J}} F(K_{j,t}, L_{j,t}) dj}_{\text{production}} + \underbrace{\int_{\mathbb{J}} \alpha(I_{j,t}, K_{j,t}) dj}_{\text{additional earnings}}. \quad (4)$$

2.2 The planning problem

We now set up the social planner's problem. An *individual plan* for capitalist j consists of trajectories $(C_{j,t}, K_{j,t})$ for their consumption and capital holdings—which can depend on the trajectory K_t of the aggregate capital stock and is adapted to (cumulative) observed idiosyncratic profits $\widetilde{\Pi}_{j,t}$.

Given an individual plan, the capitalist chooses processes $E_{j,t}$ for disguised consumption and $I_{j,t}$ for investment (both adapted to the idiosyncratic shock process $Z_{j,t}$) to maximize their utility. To simplify the planning problem, we first show that the capitalist's best responses will always involve efficient choices of investment.

Lemma 1. *For any individual plan, the capitalist will (almost) always choose an efficient level of investment: i.e., all best responses almost surely satisfy*

$$I_{j,t} = \operatorname*{arg\,max}_{\widehat{I}_{j,t} \geq 0} \left\{ \alpha(\widehat{I}_{j,t}, K_{j,t}) - \widehat{I}_{j,t} \right\}. \tag{5}$$

To understand the intuition behind Lemma 1, suppose for sake of deriving a contradiction that a capitalist chose an inefficient investment level. The capitalist could instead choose an efficient investment level and appropriate the excess earnings as disguised consumption, Since business expenses are not verifiable, this strategy would not change any verifiable information, so the capitalist would strict benefit from it. We can therefore assume that $I_{j,t}$ is defined by (5) for the remainder of this section.

Moreover, since disguising consumption as a business expense is inefficient ($\phi < 1$) and the consumption value of disguised consumption is linear, it is without loss of generality to assume that there is no disguised consumption in equilibrium—i.e., $E_{j,t} = 0$ (DeMarzo and Sannikov, 2006; DeMarzo and Fishman, 2007). The following condition requires that $E_{j,t} = 0$ is indeed a best response for the capitalist.

Definition 1. An individual plan $(C_{j,t}, K_{j,t})$ is *incentive-compatible* given a wage trajectory w_t if for all processes $E_{j,t}$ of disguised consumption (adapted to the idiosyncratic shock process $Z_{j,t}$), we have that

$$\mathbb{E}_0 \int_0^\infty e^{-\rho t} C_{j,t}^{1-\frac{1}{\psi}} dt \geq \hat{\mathbb{E}}_0 \int_0^\infty e^{-\rho t} \big(C_{j,t} + \phi E_{j,t}\big)^{1-\frac{1}{\psi}} dt,$$

where the expectation $\mathbb{E}_0$ (resp. $\hat{\mathbb{E}}_0$) is taken assuming that observed idiosyncratic profits are given by (2) (resp. by (3) with $\widetilde{I}_{j,t} = I_{j,t} + E_{j,t}$).

To introduce the space of plans, we aggregate and impose the resource constraint.

Definition 2. A *plan* consists of: a consumption stream $(C_{w,t})$ for the worker, and an incentive-compatible individual plan $(C_{j,t}, K_{j,t})$ for each capitalist j. Such a plan is *feasible* if the resource constraint (4) holds.

The planning problem is to choose a feasible plan to maximize the representative worker's utility (1) for a given initial capital stock K_0, subject to satisfying a promise to provide each capitalist j a utility level of at least $u_0 > 0$.[6] To ensure that the planning problem is well-defined, we make the following technical assumptions.

Assumption 0. There exists a feasible plan with $C_{w,t} > 0$ for all t that delivers utility at least u_0 to each capitalist j.

Assumption 0 holds as long as u_0 is sufficiently small. For example, a sufficient condition is that u_0 is at most the utility level obtained in autarky if capitalists each pay for consumption, investment, and depreciation out of their own wealth. so that obtaining unbounded welfare is impossible.

3 The constrained optimum

In this section, we define the steady state of the constrained optimum, provide conditions under which it exists, and analytically characterize it.

3.1 Steady state

In a steady state, the capital stock and workers' consumption are constant, but each capitalists' consumption and capital holdings are stochastic. We therefore introduce consider the capitalists' aggregate consumption $C_t \equiv \int_{\mathbb{J}} C_{j,t}\, dj$.

Definition 3. A *steady state* consists of

- a capital stock K^*,
- an aggregate promised capitalist utility U^*,
- a worker consumption level $C_w^* > 0$, and

[6] This promise could represent, for example, a political economy constraint. The results would be essentially unchanged if the promised utility varied across capitalists.

- an aggregate capitalist consumption $C^* > 0$

such that given the initial conditions $(K_0, u_0) = (K^*, U^*)$, there exists an optimal plan $(C_{w,t}, (C_{j,t}, K_{j,t})_{j\in\mathbb{J}})$ with $(K_t, C_{w,t}, C_t) = (K^*, C_w^*, C^*)$ for all t

Our definition of steady state requires that worker consumption and capitalist aggregate consumption each be positive in steady state, so the workers are not immiserated and the capitalists are not immiserated in aggregate.

We have three key conditions for the existence of steady state. The first condition ensures that long-run growth is infeasible. To introduce this condition, consider the maximum excess return on capital that can be accrued via investment, which is

$$\alpha^* \equiv \max_{i\geq 0} \big\{\alpha(i,1) - i\big\}.$$

As capitalists can achieve an excess return of α^*, they can achieve a return of α^* even under a high capital stock. To ensure that long-run growth is infeasible, we require that $\delta > \alpha^*$. It will be convenient to consider an *investment-adjusted depreciation rate* $\delta^* \equiv \delta - \alpha^*$, that adjusts the depreciation rate for the excess return that can be accrued from investment; we then require that $\delta^* \geq 0$.

Assumption 1. The investment-adjusted depreciation rate satisfies $\delta^* > 0$.

The second assumption is needed to ensure that it is not optimal to give all the capital to a vanishingly small mass of capitalists to manage. As shown by Di Tella (2019), the key condition is that the capitalists' EIS is large enough.

Assumption 2. The elasticity of intertemporal substitution satisfies $\psi \geq 2$.

Despite the large EIS required by Assumption 2, Epstein–Zin preferences allow us to incorporate realistic levels of capitalist risk aversion in our analysis.

The third assumption is most critical. If the capitalist is at least as patient as the planner, then the optimal allocation immiserates the worker, as in many models with infinitely repeated moral hazard (see, e.g., Atkeson and Lucas (1992)). A standard approach to restoring the existence of a steady state, based on analyses of models with a single type of agent (see, e.g., Phelan (2006) and Farhi and Werning (2007)), is to suppose that the planner is strictly more patient than the capitalist. However, in contrast to those models, even if the planner is more patient than the capitalist in our model, the optimal allocation may still immiserate the worker.

Proposition 1 (Worker immiseration if $\zeta \geq \rho$ or $\zeta \approx \rho$)**.** *Suppose that Assumptions 1 and 2 hold. Fixing all parameters other than ζ, there exists $\zeta^{**} < \rho$ such that for $\zeta \geq \zeta^{**}$ and all optimal plans under planning discount rate ζ, if the quantities $K_t, C_{w,t}, C_t$ each converge to limits in $[0, \infty]$ as $t \to \infty$, then $C_{w,t} \to 0$ as $t \to \infty$.*

Intuitively, if the planner is not patient enough, it is efficient for incentive-provision purposes to frontload the worker's consumption and backload the capitalists' consumption. This backloading in turn leads to the workers being immiserated in the long run. We therefore instead assume that the planner is sufficiently more patient than the worker. We formulate this condition by requiring that the ratio between the planning discount rate and the capitalists' discount rate be at most the labor share.[7]

Assumption 3. The planner's discount rate satisfies $\zeta \leq (1-\theta)\rho$.

We can then show that there exists a unique steady state. Despite the complexity of the model, we can also completely characterize the steady state in closed form.

Theorem 1. *Under Assumptions 1–3, there exists a unique steady-state, where*[8]

$$\frac{C^*}{K^*} = \phi\sigma \frac{2\rho - \zeta}{\sqrt{2(\rho-\zeta)}} \sqrt{\gamma\left(1 - \tfrac{1}{\psi}\right)} \tag{6}$$

$$f'(K^*) = \zeta + \delta^* + \frac{1}{\frac{\zeta(1-\frac{1}{\psi})}{2(\rho-\zeta)} + 1} \cdot \frac{C^*}{K^*} \tag{7}$$

$$\frac{C_w^*}{K^*} = (1-\theta)\frac{f(K^*)}{K^*} + \zeta - \frac{1}{\frac{2(\rho-\zeta)}{\zeta(1-\frac{1}{\psi})} + 1} \cdot \frac{C^*}{K^*}. \tag{8}$$

The steady state has a particularly simple characterization in the case in which the planner is very patient ($\zeta = 0$). Indeed, in this case, (6)–(8) become

$$\frac{C^*}{K^*} = \phi\sigma \sqrt{2\rho\gamma\left(1 - \tfrac{1}{\psi}\right)} \tag{9}$$

$$f'(K^*) = \delta^* + \frac{C^*}{K^*} \tag{10}$$

$$C_w^* = (1-\theta) f(K^*). \tag{11}$$

[7] An alternative approach would be to require that the planner be strictly more patient than the worker and that the investment-adjusted depreciation rate be sufficiently large. This alternative assumption would deliver identical other results.

[8] In the steady state, we also have $U^* = (C^*)^{1-\frac{1}{\psi}}/(2\rho - \zeta)$.

To better understanding the steady state, we study how it depends on the efficiency ϕ of disguising consumption as a business expense. We can see from (6) and (9) that C^*/K^* increases with ϕ, while both K^* and C_w^* decreases with ϕ. Intuitively, the higher ϕ, the harder it is for the planner to provide incentives for the capitalists not to disguise consumption. Thus, to provide such incentives, the planner offers the capitalists a higher share of the economy, so that C^*/K^* increases with ϕ.

Equations (7) and (10) are a version of the (modified) golden rule, adjusted for the fact that capitalists also consume in the steady state. Indeed, in the first best ($\phi = 0$), we would have that the steady state level of capital satisfies the modified golden rule $f'(K^*) = \zeta + \delta^*$. In the case of $\zeta = 0$ (see (10)), the ratio $\frac{C^*}{K^*}$ quantifies how much the steady state rate of return deviates from the golden rule level; in the case of $\zeta > 0$ (see (7)), the deviation of the steady state rate of return from the modified golden rule level also depends on preference parameters. In either case, as a higher ϕ means devoting a larger fraction C^*/K^* of resources to capitalist consumption, it also induces a lower steady-state level of capital K^*. In turn, a lower level of capital generates lower wages and lower consumption for the workers, which is the cost of redistributing through capital taxation.

3.2 The savings wedge

To characterize the distortions entailed by optimal plans, we first characterize the wedges they entail for capitalists' savings. Formally, capitalist j's *savings wedge* $\omega_{j,t}^s$ *at time* t is defined as the difference between the risk-free rate entailed by capitalist j's consumption process and the intertemporal marginal rate of substitution (IMRS) of the social planner, which, under CRRA preferences, is given by

$$\omega_{j,t}^s \, dt = \underbrace{\left(\zeta \, dt - \frac{d(v'(C_{w,t}))}{v'(C_{w,t})} \right)}_{\text{planner's IMRS}} - \underbrace{\left(\rho \, dt - \mathbb{E}_t \frac{d(C_{j,t}^{-\frac{1}{\psi}})}{C_{j,t}^{-\frac{1}{\psi}}} \right)}_{\text{capitalists' risk-free rate}}.$$

Appendix A extends the definition of the savings wedge to the Epstein–Zin case.

It turns out that capitalists share the same savings wedges at each point in time in optimal plans, even outside the steady state. Moreover, these wedges are positive.

Proposition 2. *Under Assumptions 0–2, for all optimal plans, at almost all times*

t, almost all capitalists almost surely share the same positive savings wedge ω_t^s.

Here, the positivity of the savings wedge at each point in time follows from the inverse Euler equation (Rogerson, 1985; Golosov et al., 2003).

In steady state, the savings wedge is constant and has a very simple formula.

Theorem 2. *Under Assumptions 1–3, under any optimal plan corresponding to a steady state, almost everywhere, we have*

$$\omega_{j,t}^s = \frac{2(\rho - \zeta)}{\psi - 1} \equiv \omega_{\text{ss}}^s.$$

Theorem 2 states that the steady state savings wedge is proportional to the difference between the private and social discount rates; the constant of proportionality depends on the capitalists' elasticity of intertemporal substitution. In particular, the optimal dynamic tax policy features sustained intertemporal distortions in steady state. These distortions are not mechanically driven by a motivation to increase savings when the social planner is more patient than the capitalist: indeed, the steady state savings wedge is higher the more patient the planner is relative to the capitalist.

3.3 Stability

While our steady state analysis is tractable, one may be concerned about whether they characterize the long run state of the economy under the optimal dynamic tax policy.[9] Thus, to support our analytical results about steady states, we next provide a stability analysis of the steady state.

We show that if the planner is sufficiently patient (ζ is small), the steady state (K^*, U^*, C_w^*, C^*) is *locally stable* in that the optimal plan $(C_{w,t}, (C_{j,t}, K_{j,t})_{j\in\mathbb{J}})$ converges to (K^*, U^*, C_w^*, C^*), provided that the initial level of capital K_0 and the promised capitalist utility $U_{j,0}$ are sufficiently close to their steady state levels K^*, U^*. To that end, we first formally define our notion of local stability.

Definition 4. A steady state (K^*, U^*, C_w^*, C^*) is *locally saddle-path stable* if there exists a neighborhood $\mathcal{N}$ of (K^*, U^*) such that if the $(K_0, u_0) \in \mathcal{N}$, then given any optimal plan $(C_{w,t}, (C_{j,t}, K_{j,t})_{j\in\mathbb{J}})$, we have $\lim_{t\to\infty}(K_t, C_{w,t}, C_t) = (K^*, C_w^*, C^*)$.

We then establish the steady state's local stability for a sufficiently patient planner.

[9] This concern may be particularly relevant to capital taxation given Straub and Werning (2020).

Theorem 3. *Under Assumptions 1 and 2, if the planner's discount rate ζ is sufficiently small, then the steady state is locally saddle-path stable.*

In light of Theorem 3, we can interpret our results about the steady state as characterizing the long run behavior under the optimal dynamic tax policy, provided that the initial conditions are not too far from steady state.

4 Implementation

We next develop of implementation of the optimal dynamic allocation via linear taxes/subsidies on the capitalists. We show that the constrained optimum can be implemented by a combination of time-varying linear taxes/subsidies on wealth, capital income, and consumption. We also derive the steady-state taxes in closed form.

Formally, the tax system we consider is as follows:

- a tax on observed consumption $\widetilde{C}_{j,t}$ with rate τ_t^c
- a tax on unexpected reported profits $d\widetilde{\Pi}_{j,t} - \mathbb{E}_t[d\Pi_{j,t}]$ with rate τ_t^k
- a wealth tax on capital stocks $K_{j,t}$ with rate $\tau_t^w\, dt$.

We use a tax on unexpected reported profits, rather than all reported profits, to isolate the wealth taxation and risk-sharing roles of capital taxation. (Taxing all reported profits instead would only require adjusting the wealth tax.)

Given these taxes, capitalists are allowed to freely save, invest, and consume. Thus, the law of motion of the capital j's wealth is

$$\begin{aligned} dK_{j,t} &= d\Pi_{j,t} - \tau_t^k\big(d\widetilde{\Pi}_{j,t} - \mathbb{E}_t[d\Pi_{j,t}]\big) - \tau_t^w K_{j,t}\, dt - (1+\tau_t^c)\widetilde{C}_{j,t}\, dt - E_{j,t}\, dt \\ &= d\widetilde{\Pi}_{j,t} - \tau_t^k\big(d\widetilde{\Pi}_{j,t} - \mathbb{E}_t[d\Pi_{j,t}]\big) - \tau_t^w K_{j,t}\, dt - (1+\tau_t^c)\widetilde{C}_{j,t}\, dt. \end{aligned} \tag{12}$$

Given a path of taxes $(\tau_t^c, \tau_t^k, \tau_t^w)$ and an initial capital stock $K_{j,0} = K_0$ each capitalist j chooses $(C_{j,t}, L_{j,t}, I_{j,t}, E_{j,t})$ to maximize their lifetime utility subject to the law of motion for capital and the usual no Ponzi condition.

We first establish that we can implement the optimal dynamic mechanism with only these three taxes.

Proposition 3. *Under Assumptions 0–2, there exist a unique trajectory of taxes $(\tau_t^c, \tau_t^k, \tau_t^w)_t$ that implements the constrained optimum plan.*

To understand the logic behind Proposition 3, we first summarize how to solve for the capital income tax τ_t^k. Intuitively, the capital income tax is set so that the capitalist receives the right amount of skin-in-the-game: $\tau_t^k = 1$ would imply perfect risk-sharing of the idiosyncratic risk $dZ_{j,t}$, while $\tau_t^k = 0$ would imply the capitalist j bears all the risk of the technology associated with capital K_{jt}. More precisely, since the capital holdings $K_{j,t}$ under the optimal allocation follows an Itō process adapted to $Z_{j,t}$, it can be written as

$$\frac{dK_{j,t}}{K_{j,t}} = \mu_{K_j} dt + \sigma_{K_j} dZ_{j,t},$$

where σ_{K_j} is the idiosyncratic proportional volatility of capital. Then, the capital income tax τ_t^k must be set so that the idiosyncratic proportional volatility of capital reaches its level under the optimal dynamic allocation. Substituting the formulae for true and reported profits $d\Pi_{j,t}$, $d\widetilde{\Pi}_{j,t}$ from (2) and (3) into (12) yields that

$$\begin{aligned} dK_{j,t} = \left[F(K_{j,t}, L_{j,t}) - w_t L_{j,t} - \delta K_{j,t} + \alpha(I_{j,t}, K_{j,t}) - I_{j,t}\right] dt \\ + (1-\tau_t^k)\sigma K_{j,t} dZ_{j,t} - \tau_t^w K_{j,t} - (1+\tau_t^c) C_{j,t} dt \\ + \left[\phi(1+\tau_t^c) - (1-\tau_t^k)\right] E_{j,t} dt, \quad (13) \end{aligned}$$

so the proportional volatility of capital is $(1-\tau_t^k)\sigma$. It follows that we must have

$$(1-\tau_t^k)\sigma = \sigma_{K_j}$$

Next, we describe how the presence of the possibility of consumption disguise through business expenses ($E_{j,t}$) imposes a tight link between the consumption and the capital income tax, which pins down the consumption tax rate. Since $\phi < 1$, the optimal allocation implies that the capitalists do not disguise consumption through business expenses ($E_{j,t} = 0$). An inspection in the last term of (13) implies that the taxes must satisfy $(1-\tau_t^k) \geq \phi(1+\tau_t^c)$. A strict inequality would require the capitalist to strictly prefer consuming observably to engaging in disguised consumption, which we show to be inconsistent with the optimal allocation, as it would involve providing excessive risk to the capitalist. We can therefore establish the following result.

Proposition 4 (Limits of Consumption Taxation)**.** *Under Assumptions 0–2, the implementing taxes satisfy*

$$(1-\tau_t^k) = \phi(1+\tau_t^c). \tag{14}$$

Equation (14) captures a key force of our model: to make the agent indifferent between disguising consumption as a business expense and consuming, the marginal capital income retention ratio must equal the product of the gross consumption tax with the efficiency of disguising consumption (ϕ). Indeed, obtaining an effective consumption of ϕ through disguised consumption requires incurring business expenses of 1, which costs $1-\tau_t^k$ after capital income taxes, while obtaining an effective consumption of ϕ by observable consumption costs $\phi(1+\tau_t^c)$. This equation also highlights the trade-off that the planner faces between risk-sharing and raising revenue. Taxing unexpected capital income improves risk-sharing, and therefore allows the planner to provide less consumption to capitalists, but limits the planner's ability to raise revenue through consumption taxation—a limit to consumption taxation.

Equation (14) also allows us to calculate consumption taxes from capital income taxes. Given capital income and consumption taxes, imposing the law of motion of the capitalists' wealth on the optimal dynamic allocation pins down the wealth tax.

4.1 Steady State Taxes

While the implementing taxes are not analytically tractable along the transition path to steady state, we can characterize the the steady state taxes in closed form.

Theorem 4. *Under Assumptions 1–3, any optimal plan corresponding to the steady state can be implemented with constant linear taxes* $(\tau_{SS}^c, \tau_{SS}^k, \tau_{SS}^w)$ *given by*

$$1+\tau_{SS}^c = \frac{1}{\phi\sigma}\sqrt{\frac{2(\rho-\zeta)}{\gamma\left(1-\frac{1}{\psi}\right)}} \tag{15}$$

$$1-\tau_{SS}^k = \frac{1}{\sigma}\sqrt{\frac{2(\rho-\zeta)}{\gamma\left(1-\frac{1}{\psi}\right)}} \tag{16}$$

$$\tau_{SS}^w = -2(\rho-\zeta) + \phi\sigma\sqrt{\frac{\gamma\left(1-\frac{1}{\psi}\right)}{2(\rho-\zeta)}}\cdot\frac{2\rho-\zeta}{\frac{\zeta\left(1-\frac{1}{\psi}\right)}{2(\rho-\zeta)}+1}. \tag{17}$$

As with the steady state allocation, the steady state taxes have a simpler characterization in the case in which the planner is very patient ($\zeta = 0$). Indeed, in this case, equations (15)-(17) become

$$1 + \tau^c_{SS} = \frac{1}{\phi\sigma}\sqrt{\frac{2\rho}{\gamma\left(1 - \frac{1}{\psi}\right)}}$$

$$1 - \tau^k_{SS} = \frac{1}{\sigma}\sqrt{\frac{2\rho}{\gamma\left(1 - \frac{1}{\psi}\right)}}$$

$$\tau^w_{SS} = -2\rho + \phi\sigma\sqrt{2\rho\gamma\left(1 - \frac{1}{\psi}\right)}.$$

To better understand the steady-state taxes, we consider the comparative statics of the taxes in the value ϕ of disguised consumption. The consumption tax rate is inverse to ϕ and the wealth tax rate is linear in ϕ. Indeed, with very low ϕ, the moral hazard problem is not significant, and consumption taxation is a powerful redistributive tool. The planner can then set a very high consumption tax rate without incentivizing capitalists to consume through their business, as in Kaldor's (1955) ideal consumption tax. As a result, the planner, being more patient than the capitalists, subsidizes wealth to encourage capital accumulation and higher long run wages.

Thus, as ϕ increases, consumption taxes become less effective. The steady state capital income tax rate is unchanged, so the consumption tax rate must decrease. The wealth subsidy then decreases—and eventually becomes a wealth tax. Intuitively, the ineffectiveness of the consumption taxes due to the possibility of disguising consumption as a business expense forces the planner to turn to a wealth tax to effect redistribution, despite the intertemporal distortions that it entails.

Next, we consider comparative statics in the profit volatility σ. When profits are more volatile, the moral hazard problem is also worse, as it is harder for the planner to make inferences about the level of disguised consumption from reported profits. Analogously to raising ϕ, raising σ lowers the consumption tax and raises the wealth tax. However, raising σ also makes it more desirable for the planner to share risk across capitalists, and therefore raises the consumption tax rate as well. Similar comparative statics apply in the capitalists' risk aversion γ, as more risk aversion makes risk-sharing more valuable and therefore worsens the moral hazard problem.

4.2 Numerical Example

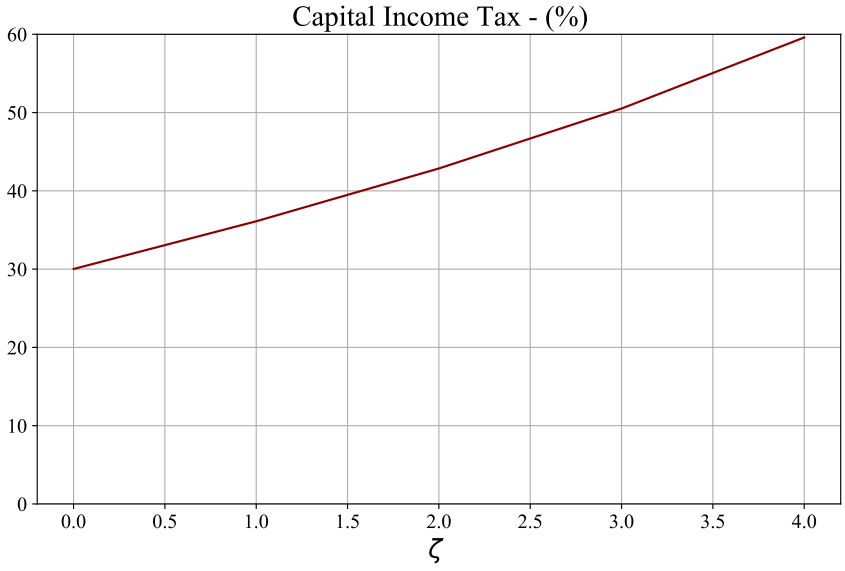
We next quantitatively illustrate the steady state taxes. To do so, we use standard value for parameters $(\rho, \gamma, \psi, \sigma)$, reported in Table 1. We then consider ranges or the value ϕ of disguising consumption and the planner's discount rate ζ, which are important and harder to pin down.

We consider a range of planner's discount rates $\zeta \in (0, 4\%)$. This implies a steady state capital income tax in the range $\tau^k_{SS} \in (30\%, 60\%)$, as depicted in Figure 1. The upper bound for ζ is the maximum value for which Assumption 3 is satisfied when the capitalists' discount rate is $\rho = 6\%$ and the capital share in production is $\theta = 1/3$ For a baseline case of $\zeta = 2\%$, we obtain a capital income tax rate of $\tau^k_{SS} = 42.85\%$.

parameter	symbol	value	source
capitalist discount rate	ρ	6%	standard
capitalist risk aversion	γ	3	standard
capitalist EIS	ψ	3	standard
volatility of profits	σ	35%	Moskowitz and Vissing-Jørgensen (2002)

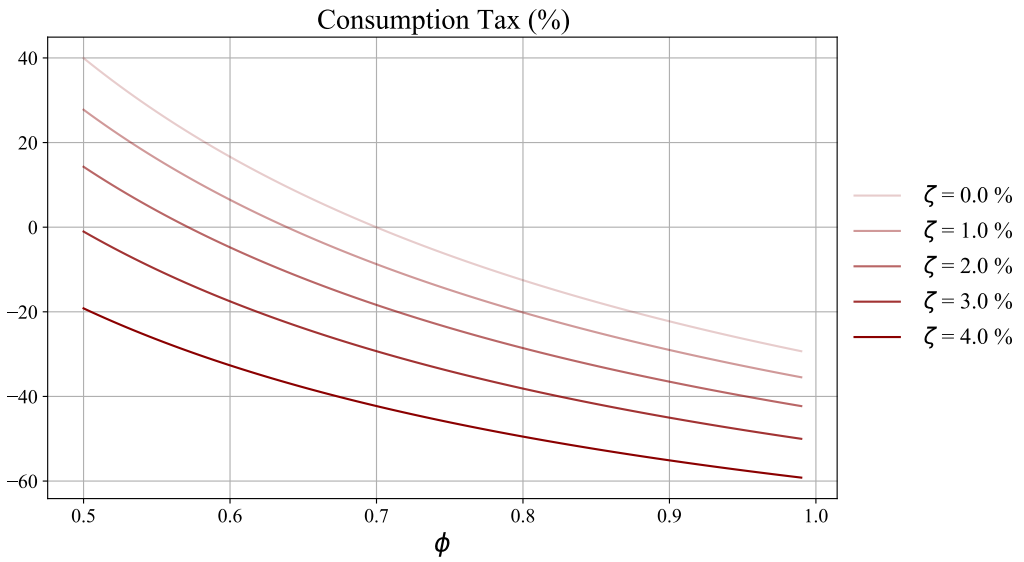
Table 1: Parameters for numerical example.

Figure 1



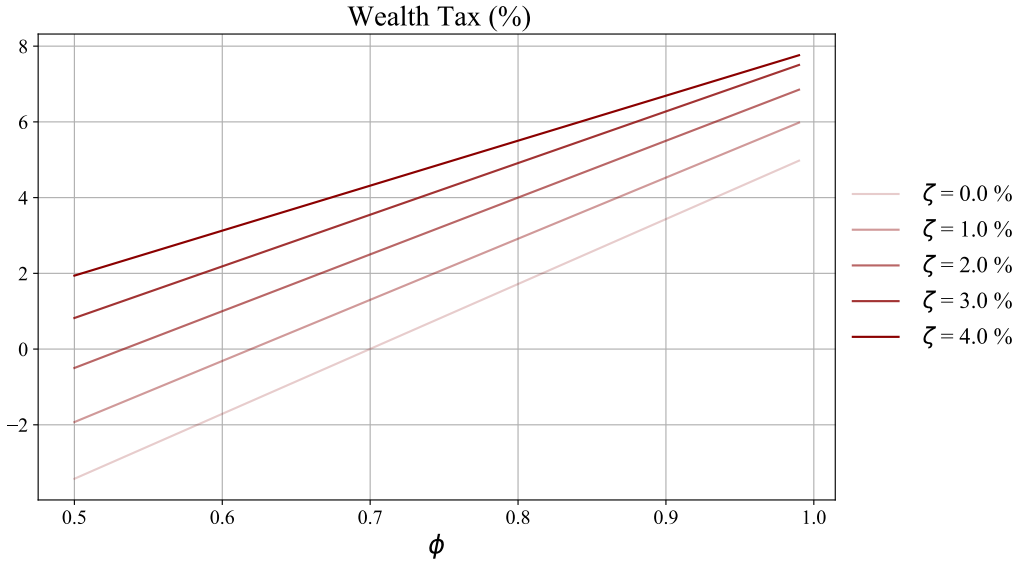
Notes: the y-axis is the capital income tax τ^k_{SS}, reported in %, whose formula is given in equation (16). The x-axis is the planner's intertemporal discount rate ζ, reported in %. For the remaining parameters, we assume the values reported in Table 1.

Figure 2



Notes: the y-axis is the consumption tax τ^c_{SS}, reported in %, whose formula is given in equation (15). The x-axis is the value ϕ of disguised consumption. The different curves consider different intertemporal discount rate ζ for the planner, with more transparent curves associated with lower values of ζ. For the remaining parameters, we assume the values reported in Table 1.

Figure 3



Notes: the y-axis is the wealth tax τ^w_{SS}, reported in %, whose formula is given in equation (17). The x-axis is the value ϕ of disguised consumption. The different curves consider different intertemporal discount rate ζ for the planner, with more transparent curves associated with lower values of ζ. For the remaining parameters, we assume the values reported in Table 1.

We consider a range for the value of disguised consumption $\phi \in (0.5, 1)$. We then use equations (15) and (17) to plot the corresponding steady-state consumption tax τ^c_{SS} and wealth tax τ^w_{SS}, depicted in Figures 2 and 3. The optimal levels of these taxes are quite sensitive to ϕ, as that parameter controls the strength of the moral hazard problem. In a baseline case where $\zeta = 2\%$ and $\phi = 0.5$, the consumption tax is $\tau^c_{SS} = 14.28\%$ and the wealth tax is $\tau^w_{SS} = -0.5\%$. An increase of ϕ from 0.5 to 0.6 would decrease the consumption tax τ^k_{SS} from 14.28% to -4.76% and increase the wealth tax τ^w_{SS} from -0.5% to 1%. As ϕ increases further, the consumption tax rate continues to fall and the wealth tax rate continues to increase.

With a very patient planner ($\zeta = 0$), the consumption tax are higher: they range from from 40% (with $\phi = 0.5$) to -30% (with $\phi = 1$). Wealth taxes are in turn lower, which reflects that a patient planner is more willing to defer consumption to accumulate capital and raise wages in the long run.

5 Analysis

In this technical section, we summarize our overall strategy to solve the model. We show that the homotheticity of preferences and technologies can be used to the planner's problem as a control problem. We use the HJB equation for this control problem to analytically characterize the steady state, and provide conditions for its existence and stability. Detailed arguments are in the Appendix.

5.1 Incentive compatibility

To begin, we need to characterize the incentive compatible individual plans. It will be useful to consider the the evolution of capitalists' *continuation utilities*, which under CRRA preferences (i.e, $\gamma = 1/\psi$), satisfies

$$U_{j,t} = \mathbb{E}_t \int_t^\infty e^{-\rho(s-t)} C_{j,s}^{1-\frac{1}{\psi}} ds.$$

The following standard lemma (see, e.g., Propositions 1 and 2 in Sannikov (2008) or Lemmata 1 and 2 in Di Tella (2019)) characterizes the law of motion of continuation utility, and expresses incentive compatibility as a condition on the law of motion.

Lemma 2. *For any individual plan $(C_{j,t}, K_{j,t})$ for a capitalist j for which $U_{j,0}$ is well-defined, capitalist j's continuation utility process $U_{j,t}$ satisfies*

$$dU_{j,t} = \left(\rho U_{j,t} - C_{j,t}^{1-\frac{1}{\psi}} + \frac{\left(\gamma - \frac{1}{\psi}\right)\sigma_{j,t}^2}{2\left(1-\frac{1}{\psi}\right)U_{j,t}} \right) dt + \sigma_{j,t} dZ_{j,t}$$

for some stochastic process $\sigma_{j,t}$. Such an individual plan is incentive compatible if and only if the inequality

$$\sigma_{j,t} \geq \frac{\left(1-\frac{1}{\psi}\right)\phi\sigma K_{j,t}}{C_{j,t}^{\frac{1}{\psi}}}, \tag{18}$$

holds almost surely for almost all t.

Intuitively, the inequality (18) corresponds to a skin-in-the-game constraint (Sannikov, 2008). This constraint is tighter the higher the value of disguised consumption—i.e., the worse the moral hazard problem is.

5.2 Reduction to a control problem

We now reduce the planner's problem to a control problem by exploiting the homotheticity of preferences and technologies. The state variables for this control problem are the capital stock K_t, and a particular aggregate of capitalists' continuation utilities[10] given by

$$X_t = \int_{\mathbb{J}} X_{j,t}\, dj,$$

where $X_{j,t} \equiv U_{j,t}^{\frac{1}{1-\frac{1}{\psi}}}$ is capitalist $j's$ continuation utility measured in consumption units. The control variables are aggregate capitalist consumption $C_t = \int_{\mathbb{J}} C_{j,t}\, dj$ and worker consumption $C_{w,t}$.

The reduction to a control problem of this type relies on Assumption 2, which requires that the capitalists' EIS ψ be at least 2. This condition is needed to ensure that the planner does not desire to implement plans in which a vanishing mass of the capitalists hold all the capital (Di Tella, 2019). The key consequence of this assumption is the following lemma, which states roughly that in optimal plans, individual plans must be homogeneous in $X_{j,t}$.

[10]The use of continuation utilities as a state variable is motivated by Spear and Srivastava (1987).

Proposition 5. *Under Assumptions 2, for each optimal plan, we have that*

$$(C_{j,t}, K_{j,t}) = \left(\frac{X_{j,t}}{X_t}\right)(C_t, K_t) \tag{19}$$

almost surely for almost all capitalists j.

Proposition 5 leads to a reduction of the planning problem to a control problem with state variables K_t and X_t. Restricting to plans that satisfy the homogeneity condition (19), the investment efficiency condition (5) implies that $I_t = \iota K_t$, where

$$\iota = \arg\max_{i \geq 0} \left\{\alpha(i, 1) - i\right\} \tag{20}$$

For homogeneous plans, the resource constraint (4) can be written as

$$\dot{K}_t \leq f(K_t) - \delta^* K_t - C_t - C_{w,t}, \tag{21}$$

where recall that $\delta^* = \delta - (\alpha(\iota, 1) - 1)$ and $f(K) \equiv F(K_t, 1)$.

Moreover, combining Lemma 2 and Itō's Lemma also shows that the inequality

$$\dot{X}_t \geq \frac{\rho X_t - C_t^{1-\frac{1}{\psi}} X_t^{\frac{1}{\psi}}}{1 - \frac{1}{\psi}} + \frac{\gamma \phi^2 \sigma^2 K_t^2}{2 C_t^{\frac{2}{\psi}} X_t^{1-\frac{2}{\psi}}} \tag{22}$$

must hold. We therefore consider a deterministic control problem for each K_0, X_0:

$$V(K_0, X_0) \equiv \max_{(C_t, C_{w,t}, \dot{K}_t, \dot{X}_t)} \int_0^\infty e^{-\zeta t} v(C_{w,t})\, dt$$
$$\text{subject to (21), (22), and } \lim_{t \to \infty} e^{-\rho t} X_t = 0. \tag{23}$$

Building on Lemma 2 and Proposition 5, the following lemma relates the social planner's problem to the control problem (23)—showing how to construct optimal plans from solutions to (23).

Lemma 3. *Fix an initial capital stock K_0. Under Assumptions 1 and 2:*

(a) The control problem (23) *has a unique solution $(C_t, C_{w,t}, K_t, X_t)$ for $X_0 = u_0^{\frac{1}{1-\frac{1}{\psi}}}$.*

(b) There exist continuation utility trajectories $X_{j,t}$ such that the trajectory $C_{w,t}$ and the individual plans defined by (19) *comprise an optimal plan.*

Proposition 5 and Lemma 3 show that the control problem (23) entirely captures the planning problem. The key to our analysis is then to exploit the tractability of the control problem (23).

We can then obtain a Hamilton–Jacobi–Bellman (HJB) equation: $V(K, X)$ solves

$$\zeta V(K,X) = \max_{C, C_w} \left\{ \begin{array}{l} v(C_w) + V_K \left(f(K) - \delta^* K - C - C_w \right) \\ \qquad + V_X \left(\dfrac{\rho X_t - C^{1-\frac{1}{\psi}} X^{\frac{1}{\psi}}}{1-\frac{1}{\psi}} + \dfrac{\gamma \phi^2 \sigma^2 K^2}{2 C^{\frac{2}{\psi}} X^{1-\frac{2}{\psi}}} \right) \end{array} \right\} \tag{24}$$

Under Assumption 2, we can show that (24) is a concave maximization problem. We proceed by using the HJB formulation (24) to analytically characterize the steady state and establish its (local) stability.

5.3 Steady state

A steady state is defined by level of capital and aggregate promised utility to the capitalists (K^*, X^*) such that for the optimal control trajectories $C_t = C(K_t, X_t)$ and $C_{w,t} = C_w(K_t, X_t)$ that solves (24), we have that $\dot{X}_t = \dot{K}_t = 0$ for almost all t.

To that end, under Assumptions 1–3, we use the first-order conditions with respect to (C_w, C), together with the Envelope Theorem to establish the steady state described in Theorem 1, where the associated steady-state aggregate capitalists promised utility is $X^* = C^*/(2\rho - \zeta)^{\frac{1}{1-\frac{1}{\psi}}}$.

In light of Lemma 3, taking the initial capital stock to be $K_0 = K^*$ and the promised utility to capitalists to be $u_0 = (X^*)^{1-\frac{1}{\psi}}$, there are optimal plans corresponding to the steady state. In such plans, each capitalists' individual plan features stochastic consumption and capital holdings. However, the consumption of the capitalists and the (aggregate) capital stock are constant, as is the worker's consumption.

The assumption that the planner is sufficiently patient (Assumption 3) is critical in the existence of a steady state. Indeed, if the planner's discount rate ζ is close enough, or above, the capitalist's discount rate ρ, a steady state does not exist. More precisely, in this case, if the state and control variables of (23) each converge to limits under an optimal plan, then the worker must be immiserated—i.e., $C_{w,t}$ must approach 0 as $t \to \infty$, a result we formally establish in Proposition 1.

5.4 Stability

We briefly discuss how to establish local saddle-path stability of the steady state (Theorem 3). The argument relies on using the HJB equation (24) to derive a non-linear dynamical system for $(C_t, C_{w,t}, K_t, X_t)$. We then linearize this system around the steady state and use a standard linear differential equations argument to establish our local stability result. Specifically, the linearized system is of the form

$$\dot{y}_t = M_{4\times 4} y_t$$

where $y_t \equiv (C_t - C^*, C_{w,t} - C_w^*, K_t - K^*, X_t - X^*)$. A sufficient condition for local saddle-path stability is that $M_{4\times 4}$ has two strictly positive and two strictly negative eigenvalues, which holds if the planner's discount rate ζ is sufficiently small.

6 Conclusion

A standard argument against capital taxation, due to Kaldor (1955), is that redistribution can be done more efficiently via consumption taxes than via capital taxes. This argument assumes that consumption is perfectly observable, which is not true in practice: business owners can disguise consumption taxes as a business expense. This paper studies the optimal form of taxation when the tax authority cannot distinguish business expenses that are disguised consumption from business expenses that are in fact productive investments by the business.

Our main analytical results characterize the steady state of the economy under the optimal dynamic tax policy, and show that the optimal allocation can be implemented by linear taxes/subsidies on capital income, consumption, and wealth. Our framework highlights that the ineffectiveness of consumption taxation creates a role for capital taxation. Thus, the planner faces a trade-off between raising revenue from capital taxation and lowering the capital stock (and hence wages). The optimal policy balances this trade-off, and involves both consumption and capital taxes.

Our analysis provides a step toward a more comprehensive investigation of optimal capital taxation. Despite the complexity of the model, our model can be studied tractably by using continuous time methods. An interesting direction for future work would be to study other trade-offs in capital taxation by building on our framework.

A Epstein–Zin Preferences

This appendix develops the case in which the capitalists have Epstein–Zin preferences with discount rate ρ, relative risk aversion γ, and elasticity of intertemporal substitution ψ. Formally, given an effective consumption process $C_{j,t}$, capitalist j's utility process $U_{j,t}$ is defined by the stochastic integral equation

$$U_{j,t} = \mathbb{E}_t \left[\int_t^\infty e^{-\rho(s-t)} \left(\rho U_{j,s} - C_{j,s}^{1-\frac{1}{\psi}} + \frac{\left(\gamma - \frac{1}{\psi}\right)}{2\left(1-\frac{1}{\psi}\right) U_{j,s}} \frac{d}{ds}[U_j]_s \right) ds \right],$$

where $[U_j]_t$ denotes the quadratic variation process of $U_{j,t}$ (Duffie and Epstein, 1992).[11]

The definition of the savings wedge needs to be adapted because the capitalists' risk-free rate is different with Epstein–Zin preferences. For individual plans that are Markov and homogenous of degree 1 in X_t, capitalist j's *savings wedge* $\omega_{j,t}^s$ *at time* t is defined as

$$\omega_{j,t}^s \, dt = \underbrace{\left(\zeta \, dt - \frac{d(v'(C_{w,t}))}{v'(C_{w,t})} \right)}_{\text{planner's IMRS}} - \underbrace{\left(\rho \, dt - \mathbb{E}_t \frac{d(C_{j,t}^{-\frac{1}{\psi}})}{C_{j,t}^{-\frac{1}{\psi}}} - \left(\gamma - \frac{1}{\psi}\right) \frac{1+\frac{1}{\psi}}{2} \frac{\mathrm{Var}_t \, dC_{j,t}}{C_{j,t}^2} \right)}_{\text{capitalists' risk-free rate}}.$$

B Proofs

To be added.

C Extension: Verifiable Business Expenses

To be added.

[11] Theorem C1 in Duffie and Epstein (1992) guarantees the existence and uniqueness of an integrable semi-martingale solution $U_{j,t}$ to this integral equation provided that $\mathbb{E}_0 \int_0^\infty e^{-\rho t} C_{j,t}^2 < \infty$.

References

Aiyagari, S. R. (1995). Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting. *Journal of Political Economy 103*(6), 1158–1175.

Alstadsæter, A. and M. Jacob (2016). Dividend taxes and income shifting. *Scandinavian Journal of Economics 118*(4), 693–717.

Amol, A. (2024). Ramsey taxation with endogenous financial frictions. Working paper.

Atkeson, A. and R. E. Lucas (1992). On efficient distribution with private information. *Review of Economic Studies 59*(3), 427–453.

Bernanke, B. S., M. Gertler, and S. Gilchrist (1999). The financial accelerator in a quantitative business cycle framework. *Handbook of Macroeconomics 1*, 1341–1393.

Chamley, C. (1980). Optimal intertemporal taxation and the public debt. Cowles Foundation Discussion Paper No. 554, Yale University.

Chamley, C. (1986). Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica 54*(3), 607–622.

Chari, V. V., P. J. Kehoe, E. Pastorino, and S. Salgado (2020). Taxing the rich. Working paper.

DeMarzo, P. M. and M. J. Fishman (2007). Optimal long-term financial contracting. *Review of Financial Studies 20*(6), 2079–2128.

DeMarzo, P. M. and Y. Sannikov (2006). Optimal security design and dynamic capital structure in a continuous-time agency model. *Journal of Finance 61*(6), 2681–2724.

Devereux, M. P., L. Liu, and S. Loretz (2014). The elasticity of corporate taxable income: New evidence from UK tax records. *American Economic Journal: Economic Policy 6*(2), 19–53.

Di Tella, S. (2019). Optimal regulation of financial intermediaries. *American Economic Review 109*(1), 271–313.

Di Tella, S. and Y. Sannikov (2021). Optimal asset management contracts with hidden savings. *Econometrica 89*(3), 1099–1139.

Duffie, D. and L. G. Epstein (1992). Stochastic differential utility. *Econometrica 60*(2), 353–394.

Epstein, L. G. and S. E. Zin (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica 57*(4), 937–969.

Farhi, E. and I. Werning (2007). Inequality and social discounting. *Journal of Political Economy 115*(3), 365–402.

Farhi, E. and I. Werning (2013). Insurance and taxation over the life cycle. *Review of Economic Studies 80*(2), 596–635.

Golosov, M., N. Kocherlakota, and A. Tsyvinski (2003). Optimal indirect and capital taxation. *Review of Economic Studies 70*(3), 569–587.

Golosov, M., A. Tsyvinski, and I. Werning (2006). New dynamic public finance: A user's guide [with comments and discussion]. *NBER Macroeconomics Annual 21*, 317–387.

Gordon, R. H. and J. Slemrod (2000). Are "real" responses to taxes simply income shifting between corporate and personal tax bases? In J. Slemrod (Ed.), *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*. Cambridge University Press.

Judd, K. L. (1985). Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics 28*(1), 59–83.

Judd, K. L. (1999). Optimal taxation and spending in general competitive growth models. *Journal of Public Economics 71*(1), 1–26.

Kaldor, N. (1955). *An Expenditure Tax*. London: Allen & Unwin.

Lansing, K. J. (1999). Optimal redistributive capital taxation in a neoclassical growth model. *Journal of Public Economics 73*(3), 423–453.

Moskowitz, T. J. and A. Vissing-Jørgensen (2002). The returns to entrepreneurial investment: A private equity premium puzzle? *American Economic Review 92*(4), 745–778.

Phelan, C. (2006). Opportunity and social mobility. *Review of Economic Studies 73*(2), 487–504.

Piketty, T. and E. Saez (2013). A theory of optimal inheritance taxation. *Econometrica 81*(5), 1851–1886.

Piketty, T., E. Saez, and G. Zucman (2018). Distributional national accounts: Methods and estimates for the United States. *Quarterly Journal of Economics 133*(2), 553–609.

Rogerson, W. P. (1985). Repeated moral hazard. *Econometrica 53*(1), 69–76.

Saez, E. and S. Stantcheva (2018). A simpler theory of optimal capital taxation. *Journal of Public Economics 162*, 120–142.

Saez, E. and G. Zucman (2016). Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *Quarterly Journal of Economics 131*(2), 519–578.

Sannikov, Y. (2008). A continuous-time version of the principal-agent problem. *Review of Economic Studies 75*(3), 957–984.

Spear, S. E. and S. Srivastava (1987). On repeated moral hazard with discounting. *Review of Economic Studies 54*(4), 599–617.

Stantcheva, S. (2017). Optimal taxation and human capital policies over the life cycle. *Journal of Political Economy 125*(6), 1931–1990.

Straub, L. and I. Werning (2020). Positive long run capital taxation: Chamley-Judd revisited. *American Economic Review 110*(1), 86–119.