Optimal Taxation with Heterogeneous Rates of Return

Nicholas Hoffman   Ali Shourideh

July 1, 2021

Abstract

We study optimal capital taxation in a model in which individuals are heterogeneous in the rates of return that they face. By offering individuals the choice between risk-free saving and risky investing, we capture the role that the government plays in ensuring that its most able entrepreneurs invest capital into their businesses. We allow the government to levy separate taxes on interest income from risk-free savings and on capital income earned by investing into private entrepreneurial enterprises. We begin in a static, two-period setting, and demonstrate properties of constrained-efficient allocations both analytically, and in a numerical model. We then turn to a fully dynamic model, and demonstrate that the optimal allocations and wedges in this setting are independent of an agent's history of capital income. Finally, we discuss the implementation of these allocations in a decentralized economy with taxes and transfers.

1   Introduction

Recent advances in data collection and economic modelling have given economists a strong understanding of the way in which the distribution of wealth in the United States has evolved to its current shape.[1] We know that the fortunes of those on the top rung of the economic ladder are primarily composed of risky assets such as ownership of businesses and shares of stock. These fortunes are held by entrepreneurs, investors, and business owners—talented individuals who face a range of risks and rates of return in accumulating their wealth. Given these forces shaping the wealth distribution, however, it is less clear what sort of tax and transfer system is optimal to both satisfy the redistributive motives of the government and encourage investors to undertake risky projects with benefits that are beneficial to society. In order to study this question, we construct a model in which agents face differing rates of return, and a benevolent government levies taxes in order to satisfy utilitarian redistributive motives.

The best way in which to tax the wealthy is a central focus of the current political debate. Saez and Zucman (2019) argue that, if one accounts for state, local, and sales taxes, the richest one hundredth of Americans pay a disproportionately small share of their income in taxes. Saez and Zucman attribute this regressive feature of the ex-post tax system to the treatment of capital income: owners of capital, they argue, are able to control the timing, amount, and even location of

---

[1]See for instance Benhabib *et al.* (2011) or Benhabib *et al.* (2019).

their income in a way that laborers are not. As a solution, they propose a tax on wealth directly, which they argue is more difficult to move or conceal. Critics of this approach note that such a tax discourages the type of risky investment that has long been the engine of economic growth.

The Dynamic Public Finance literature, which began with the seminal work of Mirrlees (1971), offers us a way in which to weigh these concerns in implementing optimal capital taxation. What distinguishes this literature from prior work in optimal taxation ("Ramsey" taxation) is that no exogenous restrictions are imposed on the tax schedule. Instead, the government can implement any type of tax that it wishes, subject to revenue requirements and endogenous informational frictions. Given the nonlinear tax schedules and informational frictions present in reality, we view this as an intuitively appealing setting in which to study capital taxation.

A canonical result in the Ramsey literature, in which the government seeks to raise its revenue using the most non-distortionary linear taxes possible, is that the tax on capital should be set to zero.[2] Perhaps the simplest reasoning behind this result is that the supply of capital is highly elastic, and any distortions introduced to the savings decisions of agents in the economy will decrease the future capital stock. In the presence of informational frictions, by contrast, it is often the case that capital taxes are nonnegative at the optimum.[3] In this case, the government taxes capital in order to prevent agents from self-insuring against idiosyncratic income shocks, ensuring that they will continue to exert labor effort. In most cases, however, these results arise in environments wherein even the wealthiest agents earn the bulk of their income in the labor market, and everyone saves at a common, risk-free rate. We investigate whether the optimality of nonnegative intertemporal wedges remains when individuals face risk to their rates of return, and when the primary way in which wealthy agents earn income is through returns to capital–as is the case in reality.

Given these wedges, we then turn to implementation in a competitive, dynamic economy. Studying implementations provides further motivation for allowing rates of return to be heterogeneous across the population. If all assets earn a constant rate, then there is an equivalence between taxing wealth and taxing capital income, and to a certain extent, either can be used to appropriately discourage savings. With different rates of return, however, this equivalence is lost, and taxes on capital income and wealth have fundamentally different impacts on behavior, and consequently, on government revenues. For this reason, a setting with heterogeneous and stochastic rates of return is an appropriate context for studying the different impacts of these tax schemes.

We also consider this work a step in studying optimal taxation in the face of plausible processes for wealth accumulation. A number of different augmentations have been introduced to the workhorse heterogeneous models of Aiyagari (1994) and Huggett (1996) in order to generate long-run distributions of wealth that match their empirical counterparts. Examples of such augmentations include Quadrini (2000), Hubmer *et al.* (2016), and Benhabib *et al.* (2019), all of which produce long-run distributions of wealth more in line with the data. The widespread use of augmentations such as these demands a more thorough understanding of optimal taxation in a setting

---

[2]See for instance Atkinson and Stiglitz (1976), Chamley (1986), and Judd (1982).

[3]Golosov *et al.* (2003) show this result in a general setting.

where agents face differing returns to saving. Our goal is to begin to build such an understanding: though at times we refer to agents who save capital in risky vehicles as "entrepreneurs," our results apply to any setting in which agents can choose a portfolio of investments of varying risk and return.

We proceed as follows. Section 2 reviews the progress made in the literature around this topic to date, and highlights precisely how we aim to expand on these findings. In section 3, we present a two-period version of our model, in which we establish results on optimal intertemporal (savings) wedges. In section 4 we to the fully dynamic model. Here, we demonstrate that optimal wedges are independent of history, and discuss how the constrained-efficient allocations might be implemented in a decentralized market economy. Section 5 concludes.

2   Literature Review

Our study builds primarily upon a long line of research on optimal taxation. Mirrlees (1971) introduced the canonical model, a static economy in which agents are privately informed of their own ability to turn labor effort into output. In contrast to prior literature, Mirrlees allowed the government to choose nonlinear tax schedules, provided that they incentivized each type to exert its allocated effort. Nevertheless, in his numerical examples, he found that the optimal labor income tax schedule was approximately linear over most of the distribution of productivities, and zero at the top and bottom. Following Mirrlees (1971), Diamond (1998) and Saez (2001) demonstrated, using empirically plausible preferences and distributions of ability, that while the optimal tax schedule may in fact be quite nonlinear, it is increasing in income, or progressive, over most of the distribution of ability.

Golosov *et al.* (2003), Kocherlakota (2005), and Albanesi and Sleet (2006) extend this approach to dynamic economies, in which agents' skills evolve stochastically over time and risk-free savings is the only vehicle for self-insurance. Golosov *et al.* (2003) derive one of the key results in this literature: they demonstrate that for a wide class of specifications for the risk that agents face, the optimal tax system distorts their intertemporal consumption-savings decision, discouraging them from carrying too many assets into each period. Kocherlakota (2005) shows that this result holds in an economy subject to aggregate shocks. He also shows that the intertemporal wedge can be implemented by a tax system that is nonlinear in capital gains and linear in current wealth, and that in such a system, the tax on wealth is zero in the aggregate and raises no revenue. Albanesi and Sleet (2006) construct a dynamic economy in which agents are subject to idiosyncratic shocks to their disutility of labor, and show that optimal allocations can be implemented in a market economy using a simple tax schedule conditioned on wealth and current labor income. In this decentralization, the tax on capital may be nonzero, depending on how the tax system incorporates an agent's labor income history.

One common feature of these three papers is that the intertemporal rate of return on savings is constant across individuals, even if it may vary over time. Such a process of accumulation,

however, fails to capture the way in which wealth is built in reality. Benhabib *et al.* (2011) show that the main force behind the thick tails in the wealth distribution– such as those in the US data–is variation in individual rates of return. To see intuitively why this is the case, consider two favorable shocks to an agent: one affects his income, and the other affects the rate of return. The shock to income allows him to save more, increasing his wealth additively. The shock to his rate of return, however, *multiplies* his wealth. It is this second type of shock that fills in the thick upper tail in the stationary distribution of wealth–particularly if an agent has the potential to realize several such favorable shocks consecutively.

Because variable rates of return can help models capture the thick tail in the empirical wealth distribution, a more recent strand of literature has attempted to characterize optimal taxation in settings where agents earn heterogeneous returns. Phelan (2019a) considers a framework in which agents can hold a number of different assets, and the government is able to tax income from these assets according to different schedules. He finds that the tax on income from risky capital is constant throughout the population, and dependent only on the degree to which information is private, rather than individual productivities. In our model, the optimal tax on risky investment income does depend on the individual's idiosyncratic rate of return, or type. Phelan (2019b), meanwhile, studies optimal wealth, labor, and capital income taxation in a context similar to ours, in which business owners' returns evolve according to a law of motion that depends upon their history of effort, as well as a random drift. Where our approaches differ is in the forces that generate disparate returns: agents in Phelan (2019b) have utility over consumption and leisure, and earn higher returns by investing more *effort* over time. Our model, meanwhile, abstracts away from this mechanism, instead treating agent's average returns as an intrinsic characteristic, and allocating to entrepreneurs the investment of capital rather than effort.

Our model is most closely related to that in Shourideh (2014), in which agents are subject to both ex-ante and ex-post heterogeneity in the returns to their entrepreneurial investments. In this environment, where type, investment, and ex-post shocks are all unobservable, the optimal wedges for both risky investment and risk-free savings are increasing in type θ. We build upon this work in two ways. To begin, we reduce the ex-post heterogeneity in returns to two states, rather than a continuum. This change both makes the model more tractable, and allows us to consider a distribution for type θ that is unbounded. In addition, we explicitly compare implementing the tax system with wealth and capital income taxes separately, and in doing so, aim to shed light on whether a tax on capital ought to depend on an agent's entire accumulated stock of assets, or just his most recent realized return.

3   Two-Period Economy

Our model is most closely related to that in Shourideh (2014). Although the full model is of infinite length, for expositional purposes, we begin with a two-period version of the model, in which a number of results on optimal taxation in the full model can be shown.

3.1 Model

The economy is populated by a continuum of agents, indexed by $i \in [0, 1]$. Each agent has a privately-known type $\theta \in \Theta = [\underline{\theta}, \bar{\theta}]$, distributed according to the C.D.F $F(\theta)$. Time is discrete, and is given by $t \in \{0, 1\}$. Agents derive utility from consumption and discount the future at rate β. All agents are endowed with initial wealth w, which they allocate at $t = 0$ between consumption and savings. There are two assets in which agents may save for future consumption: risk-free savings bonds b, and investment into their private entrepreneurial production technology k. The price of the risk-free bond is normalized to one, and its return given by $R > 0$. If an agent of type θ invests k into his entrepreneurial technology, his production at $t = 1$ is as follows:

$$y = \begin{cases} \theta k & \text{with probability } \alpha \\ 0 & \text{with probability } 1 - \alpha \end{cases} \tag{1}$$

where α is an exogenous parameter.

The government levies a tax on capital income, and is able to differentiate between income from risk-free saving and risky investment. Thus, the tax function has the form $T(y, Rb)$, and the partial derivatives T_1 and T_2 give the marginal tax rates on the two types of investment income. The government is unable, however, to observe type θ, investment k, or shock ε. As noted in Mirrlees (1971), the government can be equivalently characterized as a benevolent social planner, who collects reports from agents on their type θ and income y, and allocates first-period consumption and investment $(c_0(\theta), k(\theta))$ and second-period consumption $c_1(\theta, y)$ and $c_1(\theta, 0)$ upon realization of a successful and unsuccessful investment, respectively. Henceforth, we use the terms "government" and "planner" interchangeably.

The planner's objective, then, is to maximize

$$\int_{\underline{\theta}}^{\bar{\theta}} \mathcal{U}(\theta) f(\theta) d\theta \tag{2}$$

where

$$\mathcal{U}(\theta) = u\left(c_0\left(\theta\right)\right) + \beta\left[\alpha u\left(c_1\left(\theta, y\right)\right) + (1 - \alpha) u\left(c_1\left(\theta, 0\right)\right)\right] \tag{3}$$

by choosing allocations $(c_0(\theta), k(\theta), c_1(\theta, y))$ for $\theta \in \Theta$ and $y \in \mathbb{R}_+$. Throughout this paper, we assume that $u(c) = \log(c)$. An allocation is said to be *feasible* if it satisfies the resource constraints for $t = 0, 1$:

$$\int \left[c_0\left(\theta\right) + k\left(\theta\right)\right] dF\left(\theta\right) \leq w \tag{4}$$

$$\int \left[\alpha c_1\left(\theta, y\right) + (1 - \alpha) c_1\left(\theta, 0\right)\right] dF\left(\theta\right) \leq \int \alpha \theta k\left(\theta\right) dF\left(\theta\right) \tag{5}$$

Additionally, because the planner cannot observe θ or k, she faces a moral hazard problem, and thus she must choose allocations that are *incentive compatible*. There are two strategies that

agents can choose from that deviate from the planner's recommendations. First, an agent of type θ can claim to be of type $\hat{\theta}$. If he selects this strategy, he must invest $\frac{\hat{\theta}}{\theta} k\left(\hat{\theta}\right)$, in order to mimic the output of type $\hat{\theta}$ in the event that his return is positive. This formulation of the constraints embeds the assumption that if the planner observes capital income y inconsistent with reported type θ, she infers that the agent has misreported his type and punishes him with zero consumption and infinite disutility. Formally, these incentive constraints are formulated as follows: $\forall \theta \in \Theta$,

$$\theta \in \arg\max_{\hat{\theta}} u\left(c_0\left(\hat{\theta}\right) + k\left(\hat{\theta}\right) - \frac{\hat{\theta}}{\theta} k\left(\hat{\theta}\right)\right) + \beta\left[\alpha u\left(c_1\left(\hat{\theta}, y\right)\right) + (1-\alpha) u\left(c_1\left(\hat{\theta}, 0\right)\right)\right] \quad (6)$$

Note that because both type θ and investment k are private information, (6) contains a continuum of constraints. In order to impose these constraints in the planner's problem we employ the first-order approach, following Jewitt (1988). The constraints in (6) can be equivalently stated as follows:

$$\mathcal{U}'(\theta) = \frac{k(\theta)}{\theta} u'\left(c_0(\theta)\right) \quad (7)$$

Justification for this formulation can be found in the appendix.

The stochastic nature of output affords agents another deviation strategy, which further constrains the planner's solution. An agent has the ability to claim to be of type $\hat{\theta}$ and invest nothing, thus producing nothing in the second period. Due to the possibility for an investment to fail, the planner cannot distinguish between agents who choose this strategy, and agents who comply with her recommendations but are simply unlucky. In order to discourage this strategy, we require that $\forall \theta \in \Theta$,

$$\mathcal{U}(\theta) \geq \max_{\hat{\theta}} u\left(c_0(\hat{\theta}) + k(\hat{\theta})\right) + \beta u\left(c_1(\hat{\theta}, 0)\right) \quad (8)$$

In order to simplify the double continuum of constraints in (8), we note from (7) that $\mathcal{U}$ is weakly increasing in θ, and thus it is sufficient to require that

$$\mathcal{U}(\underline{\theta}) \geq u\left(c_0\left(\underline{\theta}\right) + k\left(\underline{\theta}\right)\right) + \beta u\left(c_1^0\left(\underline{\theta}\right)\right) \quad (9)$$

Thus, the planner's problem is to maximize (2) subject to (4), (5), (7), and (9).

3.2 Characterizing Optimal Allocations

Recall that the government levies taxes on risky investment income and risk-free saving via the (possibly nonlinear) tax function $T(y, Rb)$, with partial derivatives $T_1 \equiv \tau_k$ and $T_2 \equiv \tau_b$. As we will show, while these derivatives are somewhat analogous to marginal tax rates, they are not exactly identical. Nevertheless, our first step in deriving optimal taxes will be to derive the optimal

intertemporal wedges for investment and saving, defined respectively as follows:

$$\tau_k(\theta) = 1 - \frac{u'(c_0(\theta))}{\alpha\beta\theta u'(c_1(\theta,y))} \tag{10}$$

$$\tau_b(\theta) = 1 - \frac{u'(c_0)}{\beta R\left[\alpha u'(c_1(\theta,y)) + (1-\alpha)u'(c_1(\theta,0))\right]} \tag{11}$$

While a full solution to the planner's problem is not available, we show a few characteristics of optimal allocations that shed light on the behavior of taxes in this framework. The first demonstrates the selection of entrepreneurs in this economy, and the means by which these entrepreneurs are incentivized to take on risk:

Proposition 1. *If $k(\theta) > 0$, then it must be the case that $\alpha\theta > R$. Furthermore, $k(\theta) > 0$ also implies that the multiplier on (9), $\phi(\theta)$, is strictly positive, and*

$$c_1(\theta,y) = c_1(\theta,0) + \frac{\beta\phi(\theta)}{\lambda_1(1-\alpha)}$$

where λ_1 is the multiplier on the second period feasibility constraint (5).

The proof is left to the appendix. Proposition 1 demonstrates that if the planner finds it optimal for an agent to invest in their project, it must be the case that their expected return $\alpha\theta$ is greater than the risk-free rate R. In addition, because these agents take on risk, clearly they must be compensated with higher consumption in the second period. The identity in proposition 1 describes how these incentives are provided: the difference between $c_1(\theta,y)$ and $c_1(\theta,0)$ is a linear function of $\phi(\theta)$, and thus this can be thought of as a measure of risk and reward.

We now turn to the characterization of optimal wedges in the planner's problem:

Proposition 2. *If $k(\theta) > 0$, $\tau_k(\theta) \leq 0$, with equality if $\theta \in \{\underline{\theta}, \bar{\theta}\}$, and $\tau_b(\theta) > 0$. If $k(\theta) = 0$, meanwhile, $\tau_b(\theta) = 0$.*

The proof can again be found in the appendix. Entrepreneurs in this model face a positive intertemporal wedge risk-free savings, and a negative wedge on risky capital. In the absence of distortions, agents wish to insure against the possibility of an unsuccessful investment by saving in the risk-free instrument. The planner, however, finds it optimal to discourage such insurance: she incentivizes agents to invest at a higher rate than they would absent distortions, and provides insurance through a transfer if the project is unsuccessful. The wedge on risky investment, meanwhile, gives entrepreneurs an incentive to invest more than they would in the non-distorted environment. Note that proposition 2 implies that the optimal tax schedule is U-shaped, similar to that in Diamond (1998) or Saez (2001). In addition, the optimality of a marginal tax of zero on the highest type is consistent with Mirrlees (1971). The savings decisions of those who do not invest, meanwhile, are not distorted. These agents lend their capital to be used by those who are designated to be entrepreneurs, and earn rate R with certainty.

3.3 Numerical Example

To shed further light on allocations in this setting, we solve a numerical version of the model. Utility remains log, and we assume that θ is uniformly distributed over the interval $\underline{\theta}, \bar{\theta}$. We parameterize the problem in the following way:

$$\alpha = 0.5 \qquad\qquad \underline{\theta} = 3.2$$
$$\beta = 0.95 \qquad\qquad \bar{\theta} = 4.6$$
$$w = 1.2 \qquad\qquad a = 1.5$$
$$R = 1.5 \qquad\qquad \mathcal{U}^* = 0.2$$

As shown in proposition 2, agents who do not invest face no distortions. As such, in this exercise we focus on values of $\theta > R/\alpha$. In the illustrations that follow, we solve the dual problem, wherein the planner minimizes the cost of delivering a fixed level of aggregate utility $\mathcal{U}^*$, subject to the incentive and feasibility constraints. The planner's problem can be reduced to a system of differential equations in $\mathcal{U}$ and μ, where μ is the multiplier on (7):

$$\mathcal{U}'(\theta) = \frac{k(\theta)}{\theta c_0(\theta)}$$

$$\mu'(\theta) = \begin{cases} \gamma - \frac{c_1(\theta,y)}{\beta R} - \mu(\theta)\frac{f'(\theta)}{f(\theta)} & \theta > \underline{\theta} \\ \gamma - \frac{c_1(\theta,y)}{\beta R} - \mu(\theta)\frac{f'(\theta)}{f(\theta)} + \phi(\theta) & \theta = \underline{\theta} \end{cases}$$

Here, γ and $\phi(\theta)$ are multipliers in the planner's problem. We use a numerical ODE solver in Julia to solve the above system for the unknowns $c_0(\theta), k(\theta), c_1(\theta,y), c_1(\theta,0)$, We require that the solution satisfies one of two sets of first-order conditions in the planner's problem: if $k(\theta) > 0$,

$$1 - \frac{c_1(\theta,y)}{\beta R c_0(\theta)} - \frac{\mu(\theta)k(\theta)}{\theta c_0(\theta)^2} + \frac{\phi(\theta)}{c_0(\theta) + k(\theta)} = 0$$

$$1 - \frac{\alpha\theta}{R} + \frac{\mu(\theta)}{\theta c_0(\theta)} + + \frac{\phi(\theta)}{c_0(\theta) + k(\theta)} = 0$$

$$1 - \frac{c_1(\theta,y)}{c_1(\theta,0)} + \frac{\beta R\phi(\theta)}{(1-\alpha)c_1(\theta,0)} = 0$$

$$\log c_0(\theta) + \beta\left[\alpha\log c_1(\theta,y) + (1-\alpha)\log c_1(\theta,0)\right] = \mathcal{U}(\theta)$$

$$\log(c_0(\theta) + k(\theta)) + \beta\log c_1(\theta,0) = \mathcal{U}(\underline{\theta})$$

If $k = 0$, then $c_1(\theta,y) = c_1(\theta,0) \equiv c_1(\theta)$, meanwhile, the solutions must satisfy

$$\log c_0(\theta) + \beta\log c_1(\theta) = \mathcal{U}(\theta)$$

$$1 - \frac{c_1(\theta,y)}{\beta R c_0(\theta)} = 0$$

In practice, we impose the boundary conditions $\mu(\underline{\theta}) = \mu(\bar{\theta}) = 0$. The choice of $\mathcal{U}(\underline{\theta})$ ensures that the upper boundary condition on μ holds, while γ is chosen such that the constraint $\int_{\Theta} \mathcal{U}(\theta)\, dF(\theta) = \mathcal{U}^*$ binds.

Figure 1 shows allocations and wedges in the numerical model. Proceeding counterclockwise, the top left panel shows the allocations in the first period. One counterintuitive results is that $k(\theta)$ is decreasing in θ, which runs counter to the shape that allocations in this model usually take. In addition, while these allocations satisfy the local incentive constraints in (7), they violate some global incentive constraints: for instance, agents of type $\bar{\theta}$ under these allocations can benefit by claiming to be of type $\hat{\theta}$ some distance below $\bar{\theta}$, and investing the deviation amount $(\hat{\theta}/\bar{\theta})k(\hat{\theta})$. More work is needed to determine why this is the case; because this approach relies on first-order conditions, it may be that this solution method locates the wrong stationary point. Nevertheless, these allocations are somewhat helpful in visualizing the ways in which incentive constraints shape the solution to the planner's problem. As shown in the top right panel, $\phi(\theta)$, and thus the gap between $c_1(\theta, y)$ and $c_1(\theta, 0)$, is increasing in k. In the bottom-left panel, τ_k displays the parabolic shape implied by proposition 2. The wedge on risk-free savings, meanwhile, is indeed positive, and is increasing in $k(\theta)$.

Figure 1: Allocations and Wedges in the Planner's Problem
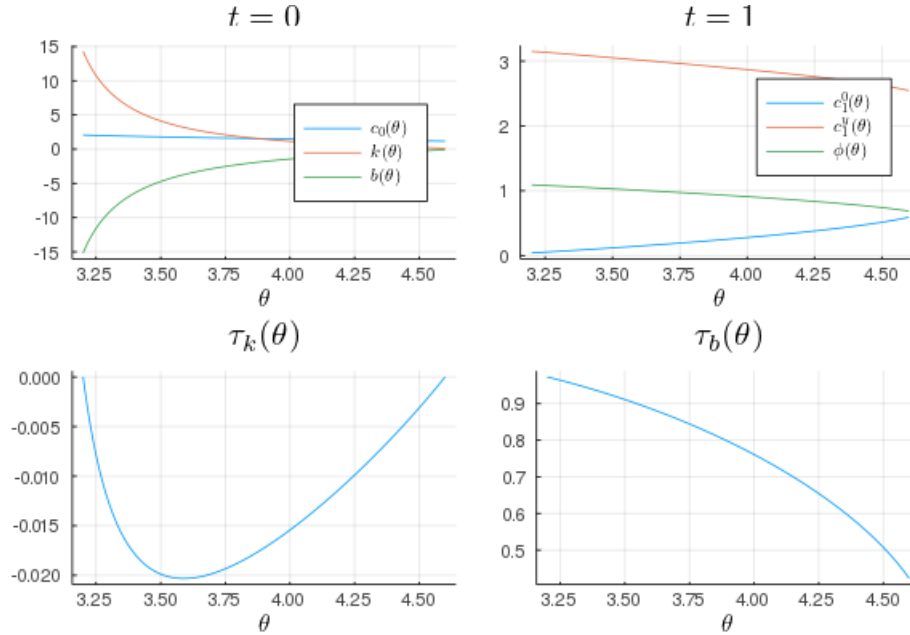
These results demonstrate the forces exerted by the moral hazard problem on optimal allocations. Although returns to scale are constant at the individual level, in the aggregate, the model exhibits decreasing returns, as in Shourideh (2014). Specifically, the planner finds it optimal to have a range of θ-types invest, rather than just the most productive type. The amount of investment k allocated to any type is limited by the incentive constraints; in particular, if too much capital is allocated to any one type, then other agents can benefit by reporting this type, eating the entire

first-period allocation $c_0 + k$, and investing and producing nothing. Thus, the role for types of $\theta < \bar{\theta}$ to become entrepreneurs arises endogenously from the moral hazard problem, rather than from an assumption of decreasing returns to scale.

4   Full Dynamic Model

While the two-period model of section 3.1 provided a number of insights on optimal taxation in our framework, it does not allow us to differentiate between capital income and wealth. In order to make this distinction, we turn our attention to an infinite horizon model where, with the right assumptions, only slight modifications to the two-period model are needed.

Time is again discrete and indexed by t. Each period plays out in the following order: an agent begins the period by realizing his capital income y_t, given his prior type θ_{t-1} and investment k_t, according to the following process:

$$
y_t = \begin{cases} \theta_{t-1} k_t & \text{with probability } \alpha \\ 0 & \text{with probability } 1 - \alpha \end{cases} \tag{12}
$$

The agent then draws a new type θ_t from the same distribution with CDF $F(\theta)$; these draws are independent and identically distributed across agents and time. Denoting his history of draws θ^t and capital incomes y^t, he then makes consumption and savings choices $c_t\left(\theta^t, y^t\right)$, $k_{t+1}\left(\theta^t, y^t\right)$, and $b_{t+1}\left(\theta^t, y^t\right)$. Let $\mu_t\left(\theta^t, y^t\right)$ be the measure of period-t histories induced by the stochastic process governing θ_t and the random nature in y_t described in (12).

The planner chooses allocations $\left\{c_t\left(\theta^t, y^t\right), k_{t+1}\left(\theta^t, y^t\right), b_{t+1}\left(\theta^t, y^t\right)\right\}_{t=0}^{\infty}$ in order to solve

$$
\max \sum_{t=0}^{\infty} \beta^t \int u\left(c_t\left(\theta^t, y^t\right)\right) d\mu_t\left(\theta^t, y^t\right) \tag{13}
$$

Subject to a series of feasibility constraints:

$$
\int \left[c_t\left(\theta^t, y^t\right) + k_{t+1}\left(\theta^t, y^t\right)\right] d\mu_t\left(\theta^t, y^t\right) = \alpha \int \theta_{t-1} k_t\left(\theta^{t-1}, y^{t-1}\right) d\mu_{t-1}\left(\theta^{t-1}, y^{t-1}\right) \tag{14}
$$

As in the static case, the planner's allocations must be incentive compatible. To begin, the promise utility allocated to an agent with history θ^t conditional on the realization of y^{t+1} is given by

$$
w_{t+1}\left(\theta^t, y^{t+1}\right) = \sum_{s=t+1}^{\infty} \beta^{s-t-1} \int u\left(c_s\left(\theta^s, y^s\right)\right) d\mu_s\left(\theta^s, y^s | \theta^t, y^{t+1}\right) \tag{15}
$$

Thus, upon realization of capital income y_t and report θ^t, the planner allocates consumption and investment to each agent, along with promised utility as given above. Promised utility w_t is also a sufficient statistic to summarize joint histories (θ^t, y^t), because types θ are not autocorrelated. There are two deviation strategies available to the agents. First, they can claim to be type $\hat{\theta}^t \neq \theta^t$,

and then mimic this type by investing $\left(\hat{\theta}_t/\theta_t\right) k\left(\theta^{t-1}, \hat{\theta}^t, y^t\right)$. The second available strategy is to claim to be type $\hat{\theta}_t$, eat the entire endowment $c_t + k_{t+1}$, and claim in the next period to have been unlucky. Thus, the allocations must satisfy the following constraints: $\forall \theta^t, \hat{\theta}_t, y^t, t,$

$$u\left(c_t\left(\theta^t, y^t\right)\right) + \beta\left[\alpha w_{t+1}\left(\theta^t, \{y^t, y\}\right) + (1-\alpha) w_{t+1}\left(\theta^t, \{y^t, 0\}\right)\right] \geq$$
$$u\left(c_t\left(\theta^{t-1}, \hat{\theta}_t, y^t\right) + k\left(\theta^{t-1}, \hat{\theta}_t, y^t\right) - \frac{\hat{\theta}_t k\left(\theta^{t-1}, \hat{\theta}_t, y^t\right)}{\theta_t}\right) +$$
$$\beta\left[\alpha w_{t+1}\left(\theta^{t-1}, \hat{\theta}_t, \{y^t, y\}\right) + (1-\alpha) w_{t+1}\left(\theta^{t-1}, \hat{\theta}_t, \{y^t, 0\}\right)\right] \quad (16)$$

$$u\left(c_t\left(\theta^t, y^t\right)\right) + \beta\left[\alpha w_{t+1}\left(\theta^t, \{y^t, y\}\right) + (1-\alpha) w_{t+1}\left(\theta^t, \{y^t, 0\}\right)\right] \geq$$
$$u\left(c_t\left(\theta^{t-1}, \hat{\theta}_t, y^t\right) + k\left(\theta^{t-1}, \hat{\theta}_t, y^t\right)\right) + \beta w_{t+1}\left(\theta^{t-1}, \hat{\theta}_t, \{y^t, 0\}\right) \quad (17)$$

In formulating these constraints, we rely on the one-shot deviation principle in mechanism design, which allows us to limit our attention to strategies wherein an agent deviates in only one period, rather than multiple periods. Note that the left-hand side of (16) and (17) both denote the time-t lifetime utility allocated to an agent of type θ who follows the planner's recommendations; we again denote this object $\mathcal{U}_t\left(\theta^t, y^t\right)$. With this formulation, the problem is a direct analogue of the two-period version. Accordingly, we can employ the same arguments as in section 3.1 in order to simplify the incentive constraints in (16) and (17). The constraints that we impose, then, are

$$\frac{\partial}{\partial \theta_t}\mathcal{U}\left(\theta^t, y^t\right) = u'\left(c_t\left(\theta^t, y^t\right)\right) \frac{k_{t+1}\left(\theta_t, y^t\right)}{\theta_t} \quad (18)$$

$$\mathcal{U}\left(\underline{\theta}_t\right) \geq u\left(c_t\left(\theta^t, y^t\right) + k_{t+1}\left(\theta^t, y^t\right)\right) + \beta w_{t+1}\left(\theta^t, \{y^t, 0\}\right) \quad (19)$$

To solve for allocations in the dynamic setting, we formulate the dual to the planner's problem in (13). Here, the planner seeks to minimize the cost of providing a prespecified level of utility $\mathcal{U}^*$:

$$\min \sum_{t=0}^{\infty} \frac{1}{\Pi_{s\leq t-1}R_s} \int \left[c_t\left(\theta^t, y^t\right) - \alpha\theta_{t-1}k_t\left(\theta^{t-1}, y^{t-1}\right)\right] d\mu_t\left(\theta^t, y^t\right) \quad (20)$$

subject to

$$\sum_{t=0}^{\infty} \beta^t \int u\left(c_t\left(\theta^t, y^t\right)\right) d\mu_t\left(\theta^t, y^t\right) \geq \mathcal{U}^* \quad (21)$$

and the incentive constraints in (18) and (19). In order to further simplify the problem, we restrict attention to a component planner, who solves the dual problem for an agent of a particular history $\left(\theta^t, y^t\right)$ assuming that interest rates are fixed over time. The recursive formulation of the component planner's problem is as follows (suppressing dependence of allocations on current promise utility

w):

$$C\left(w\right) = \min_{c,k',w'_y,w'_0,\mathcal{U}} \int \left\{ c\left(\theta\right) - \frac{\alpha\theta k\left(\theta\right)}{R} + R^{-1}\left[\alpha C\left(w'_y\left(\theta\right)\right) + \left(1-\alpha\right)C\left(w'_0\left(\theta\right)\right)\right]\right\} dF\left(\theta\right) \quad (22)$$

$$\text{s.t.}$$

$$\int \mathcal{U}\left(\theta\right) dF\left(\theta\right) \geq w$$

$$\mathcal{U}\left(\theta\right) = u\left(c\right) + \beta\left[\alpha w'_y\left(\theta\right) + \left(1-\alpha\right)w'_0\left(\theta\right)\right]$$

$$\mathcal{U}'\left(\theta\right) = u'\left(c\right)\frac{k}{\theta}$$

$$\mathcal{U}\left(\underline{\theta}\right) \geq u\left(c+k\right) + \beta w'_0$$

4.1 Characterizing Dynamic Allocations

As was the case in the similar model of Shourideh (2014), the value function $C(w)$ in (22), along with the attendant policy functions, admit closed-form representations. We assume that utility is given as follows:

$$u\left(c\right) = \begin{cases} \frac{c^{1-\sigma}}{1-\sigma} & \sigma \neq 1 \\ \log c & \sigma = 1 \end{cases} \quad (23)$$

Proposition 3. *If $\sigma = 1$:*

$$C(w) = A\exp\left(\left(1-\beta\right)w\right) \qquad\qquad w'\left(\theta,0,w\right) = w'\left(\theta,0\right) + w$$
$$k\left(\theta,w\right) = k\left(\theta\right)\exp\left(\left(1-\beta\right)w\right) \qquad\qquad w'\left(\theta,y,w\right) = w'\left(\theta,y\right) + w$$
$$c\left(\theta,w\right) = c\left(\theta\right)\exp\left(\left(1-\beta\right)w\right) \qquad\qquad \mathcal{U}\left(\theta,w\right) = \mathcal{U}\left(\theta\right) + w$$

for some $A, c\left(\theta\right), k\left(\theta\right), w'\left(\theta,y\right), w'\left(\theta,0\right),$ and $\mathcal{U}\left(\theta\right)$. If $\sigma \neq 1$, meanwhile,

$$C(w) = A\left(\left(1-\sigma\right)w\right)^{\frac{1}{1-\sigma}} \qquad\qquad w'\left(\theta,0,w\right) = w'\left(\theta,0\right)\left(1-\sigma\right)w$$
$$k\left(\theta,w\right) = k\left(\theta\right)\left(\left(1-\sigma\right)w\right)^{\frac{1}{1-\sigma}} \qquad\qquad w'\left(\theta,y,w\right) = w'\left(\theta,y\right)\left(1-\sigma\right)w$$
$$c\left(\theta,w\right) = c\left(\theta\right)\left(\left(1-\sigma\right)w\right)^{\frac{1}{1-\sigma}} \qquad\qquad \mathcal{U}\left(\theta,w\right) = \mathcal{U}\left(\theta\right)\left(1-\sigma\right)w$$

The proof can be found in the appendix. The key insight is that the planner's problem in (22) is homogeneous in w. Proposition 3 allows for a much simpler characterization of allocations in the dynamic model than there otherwise may have been. In particular, it gives us the following result:

Proposition 4. *In the dynamic model, the intertemoporal wedges are independent of history $\left(\theta^t, y^t\right)$. Instead, the wedges $\tau_{t,k}$ and $\tau_{t,b}$ depend only on the pair $\left(\theta_t, \theta_{t+1}\right)$.*

The proof is left to the appendix. While proposition 4 makes this model much more tractable, it does make the task of implementing the optimal allocations in a decentralized economy somewhat more difficult. Proposition 4 states that two individuals with different capital income histories but

the same current and next-period types should face the same distortions. As we will demonstrate, however, constructing a tax schedule which induces agents to choose these constrained-efficient allocations will require some care, as it must ensure that two agents of different levels of wealth– but the same type–make identical choices. Note also that a key assumption that led to proposition 4 is that the shocks to type θ are i.i.d.; while it may seem more natural for entrepreneurial talent to be somewhat persistent, for now, we have not made that the case, in order to gain tractability.

4.2 Decentralization

To this point, we have demonstrated properties of optimal *wedges*, distortions placed on allocations in a mechanism design problem solved by a benevolent social planner. It remains to show how these wedges translate into optimal *taxes* in a dynamic, competitive economy. In order to do so, we consider a decentralized economy, following Benhabib *et al.* (2011). In this economy, households begin with initial wealth w_0, and again have the option of intertemporally moving resources by either investing in their own risky entrepreneurial venture, or saving in a risk-free bond. The government observes income from both of these instruments, and levies taxes according to a function T, depending on some combination of capital income y, risk-free interest income Rb, and wealth w. Our goal is to describe the features of the tax function T that implements the constrained-efficient allocations described in section 4.1.

We chose this decentralization in order to most closely match the mechanism design problem that we have studied thus far. Nevertheless, in the mechanism design problem we abstracted away from the nature of private contracts, which we must address in formulating the competitive economy. In particular, in the mechanism design problem we assumed that the agents were able to borrow and lend to one another at risk-free rate R, and that through her allocations, the planner imposed a borrowing constraint and enforced repayment. In the decentralized economy, meanwhile, we must specify how the market for risk-free funds operates, and the degree to which agents can insure themselves against shocks to their type θ. As a starting point, we make the following assumptions, in order to make the government's role as analogous as possible to that of the planner.

The market for risk-free funds is operated by an impartial intermediary, who facilitates the sale and purchase of risk-free bonds in whatever amounts the agents choose. Aside from these risk-free savings and transfers from the government, the agents are uninsured from θ shocks. The government collects taxes and distributes transfers according to the function T, and in doing so, raises revenue for redistribution, and insures against adverse income shocks ($y = 0$). Another choice that must be made in constructing a decentralization is the enforcement of a borrowing constraint, which prevents agents from running Ponzi schemes. Again, we assume that this responsibility falls to the government, which uses the tax function T to ensure that agents do not choose to borrow more than an exogenous amount $\underline{b}$.

Even given this specification for the decentralization, there are still a number of choices that we must make regarding the tax function. As discussed in section 4.1, if we assume that θ shocks are i.i.d., we need to choose a tax function that ensures that individuals of the same type, but with

different levels of wealth, face the same distortions.

5 Conclusion

The prominent role played by capital income risk in generating the thick tail in the distribution of wealth in the US necessitates a better understanding of optimal taxation when agents face risky, heterogeneous returns to investment. To this end, we have presented results on the optimal taxation of capital income in an environment where agents can choose to save in risk-free bonds, or invest capital into an entrepreneurial venture with heterogeneous and uncertain returns. Following prior studies on optimal taxation following Mirrlees (1971), we begin by studying optimal allocations and wedges in a mechanism design problem. We show that in a static two-period model, the optimal tax on risky capital is U-shaped and negative, with a marginal tax of zero on the least and most productive entrepreneurs and a subsidy to everyone else. The optimal tax on risk-free interest, meanwhile, is positive for entrepreneurs, steering them away from savings and towards productive–but risky–investment. In the dynamic model, meanwhile, we demonstrate that optimal intertemoporal wedges are independent of history, and instead depend only on current and next-period type.

We then turn to a decentralized economy, in which agents can borrow and lend at the risk-free rate in a sequence of spot markets, and the government uses taxes and transfers to provide partial insurance to entrepreneurs and enforce a limit on borrowing. Our next step in this exercise is to construct a tax function that implements the constrained-efficient allocations in the decentralized economy. In addition, we plan to examine how the nature of allocations in the dynamic model changes when θ_t exhibits persistence over time.

References

S Rao Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics*, 109(3):659–684, 1994.

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

Anthony Barnes Atkinson and Joseph E Stiglitz. The design of tax structure: direct versus indirect taxation. *Journal of public Economics*, 6(1-2):55–75, 1976.

Jess Benhabib, Alberto Bisin, and Shenghao Zhu. The distribution of wealth and fiscal policy in economies with finitely lived agents. *Econometrica*, 79(1):123–157, 2011.

Jess Benhabib, Alberto Bisin, and Mi Luo. Wealth distribution and social mobility in the us: A quantitative approach. *American Economic Review*, 109(5):1623–47, 2019.

Christophe Chamley. Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica: Journal of the Econometric Society*, pages 607–622, 1986.

Peter A Diamond. Optimal income taxation: an example with a u-shaped pattern of optimal marginal tax rates. *American Economic Review*, pages 83–95, 1998.

Mikhail Golosov, Narayana Kocherlakota, and Aleh Tsyvinski. Optimal indirect and capital taxation. *The Review of Economic Studies*, 70(3):569–587, 2003.

Joachim Hubmer, Per Krusell, and Anthony A Smith Jr. The historical evolution of the wealth distribution: A quantitative-theoretic investigation. Technical report, National Bureau of Economic Research, 2016.

Mark Huggett. Wealth distribution in life-cycle economies. *Journal of Monetary Economics*, 38(3):469–494, 1996.

Ian Jewitt. Justifying the first-order approach to principal-agent problems. *Econometrica: Journal of the Econometric Society*, pages 1177–1190, 1988.

Kenneth L Judd. Redistributive taxation in a simple perfect foresight model. Technical report, Discussion Paper, 1982.

Narayana R Kocherlakota. Zero expected wealth taxes: A mirrlees approach to dynamic optimal taxation. *Econometrica*, 73(5):1587–1621, 2005.

James A Mirrlees. An exploration in the theory of optimum income taxation. *The review of economic studies*, 38(2):175–208, 1971.

Thomas Phelan. On the optimality of differential asset taxation. 2019.

Thomas Phelan. The optimal taxation of business owners. *Federal Reserve Bank of Cleveland Working Papers*, (WP 19-26), 2019.

Vincenzo Quadrini. Entrepreneurship, saving, and social mobility. *Review of economic dynamics*, 3(1):1–40, 2000.

Emmanuel Saez and Gabriel Zucman. *The triumph of injustice: How the rich dodge taxes and how to make them pay*. WW Norton & Company, 2019.

Emmanuel Saez. Using elasticities to derive optimal income tax rates. *The review of economic studies*, 68(1):205–229, 2001.

Ali Shourideh. Optimal taxation of wealthy individuals. Technical report, Working paper, 2014.

Appendix

In the arguments that follow, we suppress the dependence of the allocations on θ; for instance, we substitute $c_0 = c_0(\theta)$ and $c_1^y = c_1(\theta, y)$.

First-Order Approach An equivalent formulation of the incentive constraints in (6) is:

$$\mathcal{U}(\theta) = \max_{\hat{\theta}} u\left(c_0\left(\hat{\theta}\right)\right) + \beta\left(\alpha u\left(\theta k\left(\hat{\theta}\right) + Rb\left(\hat{\theta}\right) - T\right) + (1 - \alpha) u\left(Rb - T\right)\right) \tag{A.1}$$

The envelope condition gives

$$\mathcal{U}'(\theta) = \beta\alpha\left(1 - T_1\right) k u'\left(c_1^y\right) \tag{A.2}$$

The household optimality condition for k, meanwhile, gives

$$u'(c_0) = \beta\alpha\theta\left(1 - T_1\right) u'\left(c_1^y\right) \tag{A.3}$$

Combining (A.2) and (A.3) gives

$$\mathcal{U}'(\theta) = u'(c_0)\frac{k}{\theta}$$

the form in (7).

Proof of Proposition 1

Proof. The proof proceeds by contradiction. Assume that for some θ, $k(\theta) > 0$ at the optimum of the planner's problem, but $\alpha\theta < R$. Then, from the first-order condition for k in the planner's problem,

$$R - \alpha\theta = \frac{\mu}{\lambda_1\theta c_0} - \frac{\phi}{\lambda_1\left(c_0 + k\right)} > 0$$

which implies that

$$\frac{\mu}{\lambda_1\theta c_0} > \frac{\phi}{\lambda_1\left(c_0 + k\right)} > 0$$

Thus, μ–the multiplier on the incentive constraint–is positive. Then, rearranging the first-order conditions for c_0 and k gives

$$R + \frac{\phi}{\lambda_1\left(c_0 + k\right)} = \frac{c_1^y}{\beta c_0} - \frac{\mu k}{\lambda_1\theta c_0^2}$$

$$R + \frac{\phi}{\lambda_1\left(c_0 + k\right)} = \alpha\theta + \frac{\mu}{\lambda_1\theta c_0}$$

which implies

$$R + \frac{\phi}{\lambda_1\left(c_0 + k\right)} = \frac{c_1^y}{\beta c_0} - \frac{k}{c_0}\left(\underbrace{R + \frac{\phi}{\lambda_1\left(c_0 + k\right)} - \alpha\theta}_{>0}\right)$$

16

The term in parentheses on the right-hand side is equal to $\frac{\mu}{\lambda_1 \theta c_0}$, which by the above, is positive. This implies that

$$\frac{c_1^y}{\beta c_0} > R + \frac{\phi}{\lambda_1 (c_0 + k)}$$

This, however, implies that $\mu < 0$, contradicting the assumption. Thus, $k(\theta) > 0 \implies \alpha\theta > R$.

The first-order condition for $c_1(\theta, 0)$, meanwhile, can be rearranged to give

$$c_1^y = c_1^0 + \frac{\beta\phi}{\lambda_1 (1 - \alpha)} \tag{A.4}$$

By definition, $\phi \geq 0$, with equality if the constraint in (9) does not hold. Equation (A.4) shows that $k > 0 \implies \phi > 0$; otherwise, the agents would have no incentive to bear the risk of investing. □

Proof of Proposition 2

Proof. Recall that the intertemporal wedges are given by

$$\tau_k(\theta) = 1 - \frac{c_1^y}{\alpha\beta\theta c_0}$$

$$\tau_b(\theta) = 1 - \frac{\frac{1}{c_0}}{\beta R \left(\frac{\alpha}{c_1^y} + \frac{1-\alpha}{c_1^0} \right)}$$

Begin with the case where $k > 0$. Combining the first-order conditions in the planner's problem for c_0 and k gives

$$\frac{c_1^y}{\beta c_0} - \frac{\mu k}{\lambda_1 \theta c_0^2} = \alpha\theta + \frac{\mu}{\lambda_1 \theta c_0}$$

which implies

$$\alpha\theta - \frac{c_1^y}{\beta c_0} = \frac{\mu}{\lambda_1 \theta c_0} \left(-\frac{k}{c_0} - 1 \right)$$

$$\tau_k = 1 - \frac{c_1^y}{\alpha\beta\theta c_0} = \frac{\mu}{\lambda_1 \alpha\theta^2 c_0} \left(-\frac{k}{c_0} - 1 \right)$$

In practice, $\mu \geq 0$, and so because all other variables on the right-hand side of the final equality are nonnegative, the above shows that $\tau_k \leq 0$. The boundary conditions in the planner's problem are $\mu(\underline{\theta}) = \mu(\bar{\theta}) = 0$, and thus $\tau_k(\underline{\theta}) = \tau_k(\bar{\theta}) = 0$ by the above.

Turning to the wedge on risk-free savings, note that if $k(\theta) > 0$, then period-1 consumption

$c_1(\theta)$ becomes a random variable, so by Jensen's inequality,

$$\frac{\alpha}{c_1^y} + \frac{1-\alpha}{c_1^0} > \frac{1}{\alpha c_1^y + (1-\alpha) c_1^0} \implies$$

$$\alpha c_1^y + (1-\alpha) c_1^0 > \frac{1}{\frac{\alpha}{c_1^y} + \frac{1-\alpha}{c_1^0}} \implies$$

$$\frac{\alpha c_1^y + (1-\alpha) c_1^0}{\beta R c_0} > \frac{\frac{1}{c_0}}{\beta R \left(\frac{\alpha}{c_1^y} + \frac{1-\alpha}{c_1^0}\right)}$$

So in order to establish a positive wedge τ_b, it is sufficient to show that $\beta R c_0 > \alpha c_1^y + (1-\alpha) c_1^0$. From the first-order conditions in the planner's problem,

$$\alpha c_1^y + (1-\alpha) c_1^0 = c_1^y - \frac{\beta \phi}{\lambda_1}$$

and

$$\beta R c_0 = c_1^y - \frac{\beta \mu k}{\lambda_1 \theta c_0} - \frac{\beta \phi c_0}{\lambda_1 (c_0 + k)}$$

Combining and multiplying through by $(-\beta/\lambda_1)$, it suffices to show that

$$\phi \left(1 - \frac{c_0}{c_0 + k}\right) \geq \frac{\mu k}{\theta c_0}$$

Note, however, that this inequality holds trivially when $k(\theta) > 0$: by definition, the left hand side is positive, while the right is negative.

In the case where $k(\theta) = 0$, $c_1(\theta, y) = c_1(\theta, 0) \equiv c_1(\theta)$; the planner does not find it optimal for these types to invest, and so she has no need to incentivize them to do so. By Proposition 1, then, $\phi(\theta) = 0$. Then, the first-order condition for $c_0(\theta)$ in the planner's problem gives

$$1 = \frac{c_1}{\beta R c_0}$$

and thus $\tau_b(\theta) = 0$ for types who do not invest. $\qquad\square$

Proof of Proposition 3

Proof. Begin with the case where $\sigma = 1$. Under the posited form for the value and policy functions, the component planner's problem becomes

$$A = \min$$

$$\int \left\{ c(\theta) + k(\theta) - \frac{\alpha \theta k(\theta)}{R} + \frac{1}{R} \left[\alpha A \exp\left((1-\beta) w_y'(\theta)\right) + (1-\alpha) A \exp\left((1-\beta) w_0'(\theta)\right)\right] \right\} dF(\theta)$$

$$\tag{A.5}$$

18

Subject to

$$0 = \int \mathcal{U}(\theta)\, dF(\theta)$$

$$\mathcal{U}(\theta) = \log c(\theta) + \beta \left[\alpha w_1'(\theta) + (1-\alpha) w_0'(\theta)\right]$$

$$\mathcal{U}'(\theta) = \frac{k(\theta)}{\theta c(\theta)}$$

$$\mathcal{U}(\underline{\theta}) = \log(c(\theta) + k(\theta)) + \beta w_0'(\theta)$$

Note that now the problem is independent of w. Thus, the existence of A as the solution to the problem in (A.5) proves the proposition. We use Blackwell's sufficient conditions to verify that the mapping on the right-hand side of (A.5), denoted $(TA)(w)$, is a contraction mapping, with a fixed point of A. First, this mapping satisfies monotonicity: consider two sets of policy functions $\{c(\theta), k(\theta), w_y'(\theta), w_0'(\theta), \mathcal{U}(\theta)\}$, one uniformly higher than the other. Clearly, the total cost of delivering utility under the greater of the policy functions will have a higher cost, and thus this mapping is monotonic. In addition,

$$T(A+a)(w) = \min$$

$$\int \left[c(\theta) + k(\theta) - \frac{\alpha\theta k(\theta)}{R} + \frac{1}{R}\alpha\left\{A\exp\left((1-\beta) w_y'(\theta)\right) + a\right\} + \frac{1}{R}(1-\alpha)\left\{A\exp\left((1-\beta) w_y'(\theta)\right) + a\right\}\right]$$

$$= T(A) + a\frac{1}{R}$$

As long as $R > 1$, this mapping is a contraction, and has a fixed point $A(w)$. However, since the problem was independent of w, $A(w)$ is a constant A.

The case where $\sigma \neq 1$ is similarly shown. Under the posited forms, when $w = (1-\sigma)^{-1}$, the planner solves

$$A = \min$$

$$\int \left\{c(\theta) + k(\theta) - \frac{\alpha\theta k(\theta)}{R} + \frac{1}{R}\left[\alpha A\left((1-\sigma) w_y'\right)^{\frac{1}{1-\sigma}} + (1-\alpha) A\left((1-\sigma) w_0'\right)^{\frac{1}{1-\sigma}}\right]\right\} dF(\theta) \quad (A.6)$$

By the same arguments as in the case where $\sigma = 1$, A exists as the fixed point of the contraction map defined by the right-hand side of (A.6). $\qquad \square$

Proof of Proposition 4

Proof. In the dynamic model, the intertemporal wedge on risky capital is given by

$$\tau_{t,k}\left(\theta^t, y^t\right) = 1 - \frac{u'\left(c_t\left(\theta^{t-1}, y^t\right)\right)}{\alpha\beta\theta_t u'\left(c_{t+1}\left(\theta^{t+1}, \{y^t, y\}\right)\right)} \quad (A.7)$$

Consider the case $\sigma = 1$. Under proposition 3,

$$c_t \left(\theta^t, y^t \right) = c \left(\theta_t \right) \exp \left((1 - \beta) \, w_t \left(\theta^{t-1}, y^t \right) \right)$$

$$w_{t+1} \left(\theta^t, \{y^t, y\} \right) = w_t \left(\theta^{t-1}, y^t \right) + w_{t+1}^y \left(\theta_t \right)$$

and furthermore,

$$c_{t+1} \left(\theta^{t+1}, \{y^t, y\} \right) = c \left(\theta_{t+1} \right) \exp \left[(1 - \beta) \, w_{t+1} \left(\theta^t, \{y^t, y\} \right) \right]$$

$$= c \left(\theta_{t+1} \right) \exp \left[(1 - \beta) \left\{ w_t \left(\theta^{t-1}, y^t \right) + w_{t+1}^y \left(\theta_t \right) \right\} \right]$$

Thus, the wedge in (A.7) can be written as

$$\tau_{t,k} \left(\theta^t, y^t \right) = 1 - \frac{1/c_t \left(\theta^t, y^t \right)}{\alpha \beta \theta_t \left\{ 1/c_{t+1} \left(\theta^{t+1}, \{y^t, y\} \right) \right\}}$$

$$= 1 - \frac{1/c \left(\theta_t \right) \exp \left((1 - \beta) \, w_t \left(\theta^{t-1}, y^t \right) \right)}{\alpha \beta \theta_t \left\{ 1/c \left(\theta_{t+1} \right) \exp \left[(1 - \beta) \left\{ w_t \left(\theta^{t-1}, y^t \right) + w_{t+1}^y \left(\theta_t \right) \right\} \right] \right\}}$$

$$= 1 - \frac{1/c \left(\theta_t \right) \exp \left((1 - \beta) \, w_t \left(\theta^{t-1}, y^t \right) \right)}{\alpha \beta \theta_t \left\{ 1/c \left(\theta_{t+1} \right) \exp \left[(1 - \beta) \, w_t \left(\theta^{t-1}, y^t \right) \right] \exp \left[(1 - \beta) \, w_{t+1}^y \left(\theta_t \right) \right] \right\}}$$

$$= 1 - \frac{1/c \left(\theta_t \right)}{\alpha \beta \theta_t \left\{ 1/c \left(\theta_{t+1} \right) \exp \left[(1 - \beta) \, w_{t+1}^y \left(\theta_t \right) \right] \right\}}$$

The final line gives the desired result: $\tau_{t,k} \left(\theta^t, y^t \right)$ only depends on θ_t and θ_{t+1}. The wedge on risk-free savings is similar:

$$\tau_{t,b} \left(\theta^t, y^t \right) = 1 - \frac{1/c_t \left(\theta^t, y^t \right)}{\beta R \left[\alpha / c_{t+1} \left(\theta^{t+1}, \{y^t, y\} \right) + (1 - \alpha) / c_{t+1} \left(\theta^{t+1}, \{y^t, 0\} \right) \right]}$$

$$= 1 - \frac{1/c \left(\theta_t \right)}{\beta R \left\{ \alpha / c \left(\theta_{t+1} \right) \exp \left[(1 - \beta) \, w_{t+1}^y \left(\theta_t \right) \right] + (1 - \alpha) / c \left(\theta_{t+1} \right) \exp \left[(1 - \beta) \, w_{t+1}^0 \left(\theta_t \right) \right] \right\}}$$

Thus, $\tau_{t,b}$ also depends only on θ_t and θ_{t+1}. The same can easily be shown for the case where $\sigma \neq 1$. $\square$