

Optimal Taxation with Heterogeneous Rates of Return

Nicholas Hoffman

Ali Shourideh

July 31, 2020

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 derive properties of constrained-efficient allocations, and use a numerical model to illustrate the types of intertemporal wedges that may arise due to incentive constraints. The aim of this project is ultimately to compare the implementation of these wedges with a tax on capital income, and a tax on accumulated capital stock, or wealth.

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. 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 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.¹ 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.² 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 where agents face differing returns to saving. Our goal is to begin to build such an understanding:

¹See for instance Atkinson and Stiglitz (1976), Chamley (1986), and Judd (1982).

²Golosov *et al.* (2003) shows this result in a general setting.

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 establish results on optimal intertemporal (savings) wedges. Section 4 extends our model to be fully dynamic, lasting for $T > 2$ periods. In this context, we aim to study how the optimal *wedges* derived in section 3 translate into optimal *taxes* in a competitive equilibrium under various tax regimes. 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 using a simple tax schedule conditioned on wealth and current labor income.

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 driver the thick tails in the wealth distribution—such as those in the US data—is variation in individual rates of return. To see 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.

This paper begins by expanding the model of 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. We add new sources of uncertainty in returns, and by doing so, impose additional constraints that the government must satisfy in order to incentivize truthful reporting. While the added informational frictions complicate the model, we show that they constrain the government in such a way as to allow us to incorporate a distribution for ability with wider support. 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 contains $T \leq \infty$ periods, 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 . where c_t denotes consumption at time t . 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:

$$y = \begin{cases} \theta k \varepsilon & \text{with probability } \alpha \\ 0 & \text{with probability } 1 - \alpha \end{cases} \quad (1)$$

ε is an idiosyncratic shock to productivity, drawn from a distribution with C.D.F. $H(\varepsilon)$. In particular, we assume that $\varepsilon \sim \Gamma(\gamma, \gamma^{-1})$. We define the distribution of second-period output, conditional on type θ and investment k as

$$G(y|\theta, k) = H\left(\frac{y}{\theta k}\right) \quad (2)$$

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 $c_1(\theta, y)$. Henceforth, we use the terms “government” and “planner” interchangeably. For ease of exposition, we redefine the second-period allocations in terms of the shock ε , rather than income y : we define $c_1(\theta, y) = c_1(\theta, \theta k(\theta) \varepsilon) \equiv c_1(\theta, \varepsilon)$.

The planner’s objective, then, is to maximize

$$\int_{\underline{\theta}}^{\bar{\theta}} \mathcal{U}(\theta) f(\theta) d\theta \quad (3)$$

where

$$\mathcal{U}(\theta) = u(c_0(\theta)) + \beta \left(\alpha \int_0^\infty u(c_1(\theta, \varepsilon)) dH(\varepsilon) + (1 - \alpha) u(c_1(\theta, 0)) \right) \quad (4)$$

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 $t = 0, 1$:

$$\int_{\underline{\theta}}^{\bar{\theta}} [c_0(\theta) + k(\theta)] f(\theta) d\theta \leq w \quad (5)$$

$$\alpha \int_{\underline{\theta}}^{\bar{\theta}} \int_0^\infty [\theta k(\theta) \varepsilon - c_1(\theta, \varepsilon)] h(\varepsilon) d\varepsilon f(\theta) d\theta + (1 - \alpha) \int_{\underline{\theta}}^{\bar{\theta}} c_1(\theta, 0) f(\theta) d\theta \leq 0 \quad (6)$$

Additionally, because the planner cannot observe θ or k , she must choose allocations that are *incentive compatible*. Formally, the incentive constraints are formulated as follows: $\forall \theta \in \Theta$,

$$\theta, k(\theta) \in \arg \max_{\hat{\theta}, \hat{k}} u \left(c_0(\hat{\theta}) + k(\hat{\theta}) - \hat{k} \right) + \beta \left(\alpha \int_0^\infty c_1(\hat{\theta}, y) dH \left(\frac{y}{\theta \hat{k}} \right) + (1 - \alpha) u(c_1(\theta, 0)) \right) \quad (7)$$

Note that because both type θ and investment k are private information, (7) contains a double 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 (7) can be equivalently stated as follows:

$$u'(c_0(\theta)) = \beta \alpha \int_0^\infty u(c_1(\theta, \varepsilon)) \frac{1}{k(\theta)} (-\varepsilon h'(\varepsilon) - h(\varepsilon)) d\varepsilon \quad (8)$$

$$\mathcal{U}'(\theta) = \frac{k(\theta)}{\theta} u'(c_0(\theta)) \quad (9)$$

Justification for this formulation can be found in the appendix. The crucial step is to think of the planner as observing y and collecting reports on type θ . Then, under the presumption that the agent has truthfully reported his type and invested the recommended $k(\theta)$, the planner imputes a value for the agent's shock. Thus, if an agent of type θ reports type $\hat{\theta}$ and invests \hat{k} , the shock that the planner will impute is given by

$$\tilde{\varepsilon} = \frac{\varepsilon \theta \hat{k}}{\hat{\theta} k(\hat{\theta})}$$

The agent will then be allocated consumption $c_1(\theta, \tilde{\varepsilon})$. This transformation allows us to calculate the marginal benefit of investing additional \hat{k} in terms of its effect on the distribution of imputed $\tilde{\varepsilon}$ values.

The planner faces one additional constraint in formulating optimal allocations. As we will show, if R is the risk-free rate, then the planner finds it optimal to incentivize all types $\theta > R$ to invest $k(\theta) > 0$. The agents who do so will face risk, including the nonzero probability that their capital investment return will be zero. The planner must ensure that these types are not better off eating their entire first-period allotment $c_0(\theta) + k(\theta)$, investing nothing (and thus producing $y = 0$ with certainty), and then claiming to have been unlucky. Formally, we require that $\forall \theta \in \Theta$,

$$\mathcal{U}(\theta) \geq \max_{\hat{\theta}} u \left(c_0(\hat{\theta}) + k(\hat{\theta}) \right) + \beta u \left(c_1(\hat{\theta}, 0) \right) \quad (10)$$

In order to simplify the double continuum of constraints in (10) we make the following two assumptions, which will prove to be innocuous. The first is that if $\theta > R$, $\mathcal{U}'(\theta) > 0$; that is, the utility from optimal allocations is increasing in θ . The second is that

$$\bar{\theta} = \arg \max_{\hat{\theta}} u \left(c_0(\hat{\theta}) + k(\hat{\theta}) \right) + \beta u \left(c_1(\hat{\theta}, 0) \right) \quad (11)$$

that is, if an agent plans on absconding with all of his allocated capital and claiming to have been unlucky, it is optimal for him to claim to be type $\bar{\theta}$. With these assumptions in place, (10) can be reduced to a single constraint:

$$\mathcal{U}(R) \geq u \left(c_0(\bar{\theta}) + k(\bar{\theta}) \right) + \beta u \left(c_1(\bar{\theta}, 0) \right) \quad (12)$$

Thus, the planner's problem is to maximize (3) subject to (5), (6), (8), (9), and (12).

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:

$$\tau_k(\theta) = 1 - \frac{u'(c_0(\theta))}{\beta \theta \int u'(c_1(\theta, \varepsilon)) \varepsilon h(\varepsilon) d\varepsilon} \quad (13)$$

$$\tau_b(\theta) = 1 - \frac{u'(c_0(\theta))}{\beta R \int u'(c_1(\theta, \varepsilon)) h(\varepsilon) d\varepsilon} \quad (14)$$

A full, closed-form characterization of optimal allocations in this context cannot be solved for. Nevertheless, we can prove a few properties that optimal allocations must satisfy, which will be helpful in determining the above wedges at the optimum. The first is that for the agents who produce $y > 0$, consumption in the second period is linear in the shock ε :

Proposition 1. *Second-period consumption $c_1(\theta, \varepsilon) = \phi(\theta) + \vartheta(\theta)\varepsilon$; that is, $c_1(\theta, \varepsilon)$ is linear in ε .*

The proof, which closely follows Shourideh (2014), can be found in the appendix. The linearity of $c_1(\theta, \varepsilon)$ is a consequence of our assumptions on utility and the distribution of ε , but the fact that consumption is increasing in income is invariant to these assumptions. Agents who invest capital into their private projects face risk, and because they are assumed to be risk-averse, they must be compensated with increased consumption in order to incentivize them to take on this risk.

3.2.1 A Discrete Analogue

In order to gain further insights into the optimal allocations in this model, we formulate a discrete analogue to the continuous model of section 3.1. We allow the distribution for ε to have

continuous support over the real line, and we assume that $\varepsilon \sim \Gamma(\gamma, \gamma^{-1})$, where $\Gamma(a, b)$ indicates the Gamma distribution with shape a and scale b . We assume log utility, and parametrize the problem as follows:

$$\begin{array}{lll} \beta = 0.95 & R = 1.0 & w = 1.2 \\ \alpha = 0.9 & \gamma = 2.0 & \end{array}$$

We assume a discrete grid for income $y \in Y = \{y_1, \dots, y_{N_y}\}$. Agents and the planner can choose any $k \geq 0$, which determines their probability mass function over Y , denoted $\Pr(y|\theta, k)$ ³. Agents are of type $\theta \in \Theta = \{\theta_1, \dots, \theta_{N_\theta}\}$, each with equal probability N_θ^{-1} . The economy is once again populated by a continuum of agents, and so the law of large numbers guarantees that the various probabilities in our model translate into proportions of the population; for example, a proportion N_θ^{-1} of the population has a given type i , and of those, α will produce $y > 0$, while $1 - \alpha$ will produce zero. Because k can take any nonnegative value, we use a two-step procedure to compute the incentive constraints. First, for each pair $(\theta, \hat{\theta}) \in \Theta$, we compute

$$k^*(\theta, \hat{\theta}) = \arg \max_{\hat{k}} u \left(c_0(\hat{\theta}) + k(\hat{\theta}) - \hat{k} \right) + \beta \sum_{i=1}^{N_y} u \left(c_1(\hat{\theta}, y) \right) \Pr(y|\theta, \hat{k})$$

Thus, $k^*(\theta, \hat{\theta})$ gives the optimal investment for an agent of type θ reporting type $\hat{\theta}$. Then, we impose the N_θ^2 incentive constraints, which require that $\forall \theta, \hat{\theta} \in \Theta$,

$$\mathcal{U}(\theta) \geq \left(c_0(\hat{\theta}) + k(\hat{\theta}) - k^*(\theta, \hat{\theta}) \right) + \beta \sum_{i=1}^{N_y} u \left(c_1(\hat{\theta}, y) \right) \Pr(y|\theta, k^*(\theta, \hat{\theta})) \quad (15)$$

given the k^* values computed above. Additionally, we require that $k^*(\theta, \theta) = k(\theta)$; that is, if an agent truthfully reports his type, he must find it optimal to invest the amount recommended to him by the planner. Finally, we impose (12): we assume that $R < \theta_1$, and thus all types invest $k > 0$, and require that

$$U(\theta_1) \geq u(c_0(\theta_{N_\theta}) + k(\theta_{N_\theta})) + \beta u(c_1(\theta_{N_\theta}, 0)) \quad (16)$$

We solve for allocations $\{c_0(\theta), k(\theta), c_1(\theta, y), c_1(\theta, 0)\}$ for each $\theta \in \Theta$ and $y \in Y$. From the household's problem, we define the wedges as follows:

$$\begin{aligned} \tau_k(\theta) &= \frac{u'(c_0(\theta))}{\beta \theta \alpha \sum_{i=1}^{N_y} u'(c_1(\theta, y)) \frac{y}{k(\theta)} \Pr(y|\theta, k(\theta), y > 0)} \\ \tau_b(\theta) &= \frac{u'(c_0(\theta))}{\beta \theta \sum_{i=1}^{N_y} u'(c_1(\theta, y)) \Pr(y|\theta, k(\theta))} \end{aligned}$$

³ $\Pr(y|\theta, k)$ incorporates the fact that $y = 0$ with probability $1 - \alpha$.

Figure 1 shows the optimal allocations for a grid Θ with ten equally spaced points between 0.8 and 3.0, and a grid Y of forty points between zero and fifteen. We construct the grid Y such that lower y values are clustered closer together; most income values will be towards the bottom of this range, so we construct the grid in this way in order to make the solution more accurate at these lower y values. Beginning with the upper left-hand panel, higher types are unsurprisingly allocated more investment in the first period. Accordingly, the upper right-hand panel shows that these types can expect higher capital incomes in the second period, with more weight in the right tail of the distribution. As demonstrated in the bottom panel, consumption in the second period $c_1(\theta, y)$ is increasing in y , as expected. Interestingly, though, the ordering of $c_1(\theta, y)$ is not constant over the grid Y . Instead, for each θ_i , there is an interval over the space of incomes Y where θ_i is allocated the most consumption among all θ values. This nonlinear feature of the consumption schedule is a result of the incentive constraints: given a reported type, the planner knows the range of y values into which the agent's income will most likely fall. Knowing this, she sets the consumption schedule for this θ such that the agent is best off investing the allocated amount, and ending up in their designated range. If the agent invests less than $k(\theta)$, they are likely to fall below this range of incomes, where their consumption schedule decreases rapidly. Meanwhile, if they invest more, their income will likely be higher, but the consumption schedule flattens out at these values such that the loss in first-period consumption necessitated by this higher investment leaves the agents worse off on net.

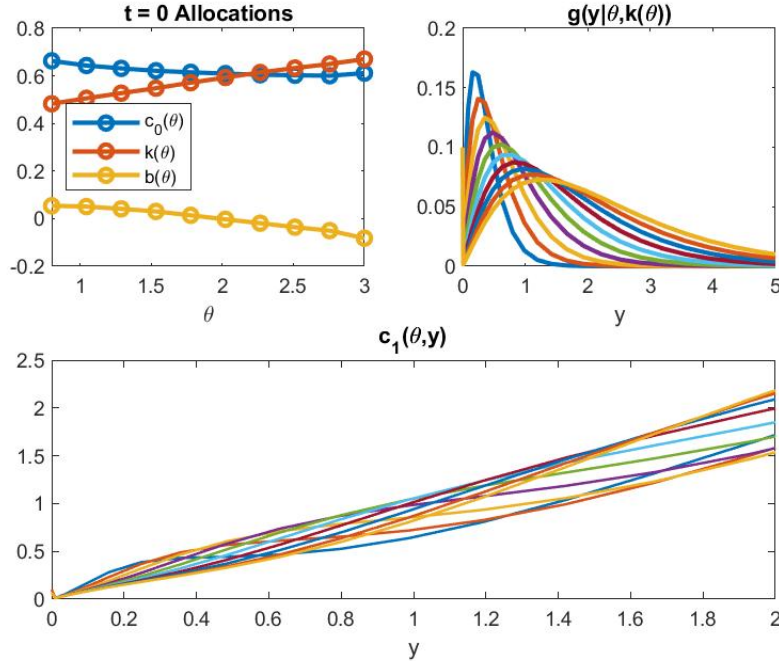


Figure 1: Allocations in the Discrete Model

Figure 2, meanwhile, shows the wedges at the optimum. In this example, capital income is actually subsidized at all levels, with $\tau_k(\theta) < 0$ for all $\theta \in \Theta$. These results primarily demonstrate

the influence of the incentive constraints on the planner's problem. The lower θ types, being less productive on average, would like to invest a greater proportion of their portfolio in a long position in the risk-free bond, and less into their risky capital, as compared to higher types. From the planner's perspective, however, allowing the lower productivity types to have such a portfolio is not compatible with her incentive constraints. Instead, the planner would rather the lower types have a higher proportion of risky assets, in order to discourage nearby types from misreporting their type. Accordingly, the planner sets the wedge on risk-free savings τ_b at a high level, inversely proportional to θ .

While the results in the discrete formulation are illustrative of the forces governing allocations in our model, these results are by no means definitive. Due to the maximization problem in the constraints, the algorithm is slow to converge even for low values of θ . As such, we are limited in the degree of ex-ante heterogeneity that we can capture. Furthermore, there is no guarantee that this problem converges to a global optimum. Indeed, the solution to this algorithm is a vector of allocations of dimension $2N_\theta + N_\theta N_y = 420$, which makes searching over the entire space infeasible. As such, we are only guaranteed a local minimum, and further analytical results are needed to determine whether the allocations in Figure 1 can be improved upon.

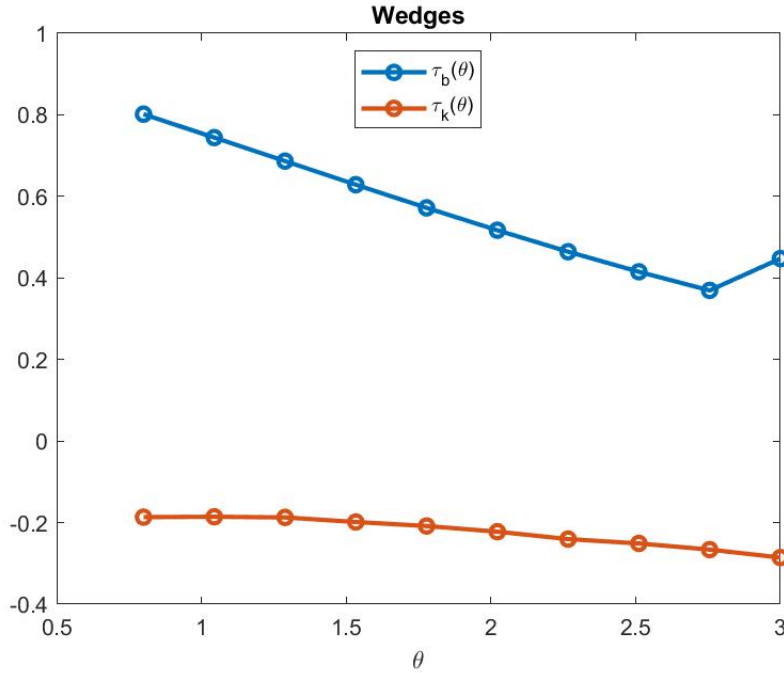


Figure 2: Wedges in the Discrete Model

4 Full Dynamic Model

5 Conclusion

We have studied optimal capital taxation a model in which agents can save in two assets: a risk-free instrument with common return, and a private entrepreneurial project, the return to which is subject to risk and heterogeneous throughout the population. We have shown that at the constrained efficient solution, consumption allocated to agents who invest capital is a linear function of the multiplicative shock ε to their return, with type-dependent slope and intercept. We also formulate a semi-discrete version of our model, and shown results for a specific parametrization. Under this specification, the planner finds it optimal to subsidize investment, and to set a high, positive wedge on interest income from risk-free saving.

What remains is twofold: to carry out the analysis of optimal wedges in the two-period model, and to demonstrate how those wedges translate into taxes in the dynamic competitive version of the model. While our discrete formulation is useful for illustrating the ways in which the incentive constraints shape optimal allocations, the numerical results from this exercise do not hold in general. Instead, we will show that more features of optimal allocations can be derived analytically.

Turning to the full dynamic model, we will study how the wedges derived in section 3 translate into taxes in a competitive equilibrium. We will construct a model in discrete time, in which agents make a (possibly infinite) series of consumption and savings decisions. In this economy, the government will tax either capital income or wealth according to a tax-and-transfer function $T(\cdot)$. This function *implements* the optimal distortions if it induces the agents to choose the constrained-efficient allocations. Finding these implementations will allow us to study whether the government is better off levying a tax on capital income, or on wealth.

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

First-Order Approach The incentive constraints are

$$\theta, k(\theta) \in \arg \max_{\hat{\theta}, \hat{k}} u \left(c_0(\hat{\theta}) + k(\hat{\theta}) - \hat{k} \right) + \beta \left(\alpha \int_0^\infty c_1(\hat{\theta}, y) dH \left(\frac{y}{\theta \hat{k}} \right) + (1 - \alpha) u(c_1(\theta, 0)) \right) \quad (\text{A.17})$$

These constraints can be stated equivalently as follows:

$$\theta, k(\theta) \in \arg \max_{\hat{\theta}, \hat{k}} u \left(c_0(\hat{\theta}) + k(\hat{\theta}) - \hat{k} \right) + \beta \left(\alpha \int_0^\infty c_1 \left(\hat{\theta}, \frac{\varepsilon \theta \hat{k}}{\hat{\theta} \hat{k}(\hat{\theta})} \right) dH(\varepsilon) + (1 - \alpha) u(c_1(\theta, 0)) \right) \quad (\text{A.18})$$

In the second formulation, the planner provides a schedule for consumption $c_1(\theta, \varepsilon)$, and imputes the value of ε based on the agent's type and observed income, presuming that the agent has reported truthfully and invested the recommended amount. The first-order approach requires that the first-order conditions for $\hat{\theta}$ and \hat{k} of (A.18) hold when evaluated at $\hat{\theta} = \theta$ and $\hat{k} = k(\theta)$. The first-order condition for (A.18) with respect to \hat{k} is

$$u'(c_0(\theta)) = \beta \alpha \int_0^\infty u(c_1(\theta, \varepsilon)) \frac{1}{k(\theta)} (-\varepsilon h'(\varepsilon) - h(\varepsilon)) \quad (\text{A.19})$$

as in (8). For $\hat{\theta}$, we utilize the Envelope condition:

$$\mathcal{U}(\theta) = \max_{\hat{\theta}, \hat{k}} u \left(c_0(\hat{\theta}) + k(\hat{\theta}) - \hat{k} \right) + \beta \left(\alpha \int_0^\infty c_1(\hat{\theta}, y) dH \left(\frac{y}{\theta \hat{k}} \right) + (1 - \alpha) u(c_1(\theta, 0)) \right) \quad (\text{A.20})$$

and so by the Envelope theorem,

$$\mathcal{U}'(\theta) = \frac{\partial}{\partial \hat{\theta}} \beta \alpha \int_0^\infty c_1(\hat{\theta}, y) dH \left(\frac{y}{\theta \hat{k}} \right) \quad (\text{A.21})$$

Because the right-hand side of (A.21) is a function of the product $\theta \hat{k}$, we make use of the following identity:

$$\frac{\partial}{\partial \theta} f(\theta k) = \frac{k}{\theta} \frac{\partial}{\partial k} f(\theta k)$$

Thus,

$$\begin{aligned} \mathcal{U}'(\theta) &= \frac{k}{\hat{\theta}} \frac{\partial}{\partial k} \beta \alpha \int_0^\infty c_1(\hat{\theta}, y) dH \left(\frac{y}{\theta \hat{k}} \right) \\ &= \frac{k}{\theta} u'(c_0) \end{aligned}$$

where the final equality follows from the right-hand side of (A.19). This is the form in (9).

Proof of Proposition 1

Proof. Consider the planner's problem in (3). The first-order conditions for $c_1(\theta, \varepsilon)$ and $c_1(\theta, 0)$ give:

$$\begin{aligned}\lambda_1 &= \beta u'(c_1(\theta, 0)) \eta \\ \lambda_1 &= \beta u'(c_1(\theta, \varepsilon)) \eta + \beta u'(c_1(\theta, \varepsilon)) \frac{1}{k(\theta)} (-\varepsilon h'(\varepsilon) - h(\varepsilon)) \kappa f\end{aligned}$$

where λ_1 is the multiplier on the feasibility constraint in the first period (6), and $\eta(\theta)$ and $\kappa(\theta)$ are the multipliers on (4) and (8), respectively. Rearranging and combining, along with the assumption of log utility, gives

$$\begin{aligned}c_1(\theta, \varepsilon) &= c_1(\theta, 0) + \kappa \beta \frac{1}{\lambda_1 k} \frac{(-\varepsilon h'(\varepsilon) - h(\varepsilon))}{h(\varepsilon)} \\ &= c_1(\theta, 0) + \kappa \beta \frac{1}{\lambda_1 \gamma k} (\varepsilon - 1) \\ &\equiv \phi(\theta) + \vartheta(\theta) \varepsilon\end{aligned}$$

where the penultimate inequality follows from the assumption that ε is drawn from a $\Gamma(\gamma, \gamma^{-1})$ distribution. \square