working paper

19 17

On the Optimality of Differential Asset Taxation

Thomas Phelan

On the Optimality of Differential Asset Taxation
Thomas Phelan

How should a utilitarian government balance redistributive concerns with the need to provide incentives for business creation and investment? Should they tax business profits, the (risk-free) savings of owners, or some combination of both? To address this question, this paper presents a model in which the desirability of differential asset taxation emerges endogenously from the presence of agency frictions. I consider an environment in which entrepreneurs hire workers and rent capital to produce output subject to privately observed shocks and have the ability to both divert capital to private consumption and abscond with a fraction of assets. To provide incentives to invest, the wealth of an agent must depend on the performance of her firm, leading to ex-post inequality in all efficient allocations. I show that the efficient stationary distribution of wealth exhibits a thick right (Pareto) tail, with the degree of inequality monotonically increasing in the number of workers per entrepreneur. The efficient allocation is then implemented in a general equilibrium model using history-independent linear taxes on risk-free savings and (reported) business profits. The tax on entrepreneurs' savings may be positive or negative, while the tax on business profits depends solely upon the degree of private information and is independent of all technological and preference parameters.

---

1 Introduction

The appropriate taxation of capital income has long been a contentious issue in both policy debates and the academic literature. Economists have typically addressed this topic through the lens of models that treat capital as a homogeneous entity and view capital income as the returns to savings. Since no distinction is made between different sources of capital income, these models give little guidance to the policymaker as to how (or whether) they ought to discriminate between dividend income, capital gains, or other forms of capital income.

In this paper I present a simple model that highlights the economic forces that may justify the differential treatment of capital income. I consider a perpetual youth environment in which agents may either run their own business (be an entrepreneur) or work for another agent's business (be a worker). Entrepreneurship is subject to two agency frictions: physical capital is subject to (privately observed) stochastic depreciation shocks and may be diverted to (privately observed) consumption; and entrepreneurs may abscond with a fixed fraction of delegated capital. An allocation must specify the occupation of every agent and the amount of capital and labor delegated to each business, all as a function of the history of observable outcomes. The ability of entrepreneurs to divert capital to private consumption implies that consumption must depend upon business performance, while the ability of entrepreneurs to abscond with a fraction of their assets limits the amount of capital that may be delegated to them. I characterize a class of (constrained) efficient allocations in which aggregate quantities and cross-sectional distributions of firm size and income are constant over time. In doing so I do not restrict attention to a fixed set of policy instruments, but allow the government to choose any allocation that respects the incentive-constraints imposed by asymmetric information. The associated distributions of consumption and firm size are characterized in closed-form and exhibit fat (Pareto) tails with thickness determined jointly by the nature of the incentive problem and the technological fundamentals.

I then provide two separate implementations of these allocations that differ in the degree of risk-sharing (i.e. the presence of insurance markets) in the private sector. In the first, agents may write contracts with a competitive sector of risk-neutral intermediaries and their insurance opportunities are inhibited only by the above agency frictions. In this setting, arguments similar to those in Prescott and Townsend (1984) show that efficient allocations require only lump-sum payments to newborn individuals or linear taxes on the profits of intermediaries. Although a natural benchmark, since it assumes the government is no more capable of overcoming the informational asymmetries than private agents, the assumption of such contracting opportunities may be inappropriate if such markets are absent in the real world. For this reason I consider a second implementation in which agents are assumed capable only of saving in a risk-free bond in zero net supply and the government uses fiscal policy to provide social insurance. With this market structure a motive for differential asset taxation emerges. I show that the stationary efficient allocations may be implemented with linear taxes on risk-free savings and (reported) business profits. The decentralization is therefore

quite simple and also illustrates a (to the best of my knowledge) novel observation: when risk-sharing is impeded by a hidden consumption friction, the optimal tax on profits is the highest level consistent with incentive-compatibility and depends solely upon the extent of the agency problem.

In order to analyze optimal taxation in the presence of dynamic agency, the benchmark case largely omits any kind of ex-ante heterogeneity among agents. The presence of adverse selection in addition to moral hazard would greatly complicate the analysis without necessarily providing additional insight. However, extending the analysis to incorporate *observable* heterogeneity in individual productivity poses no difficulty; the efficient allocation may again be implemented with linear taxes on savings and profits *if* the government has the ability to levy type-specific taxes. In this case agents with higher productivity face higher savings taxes, but the prescription for the profits tax is unchanged: it remains the highest level consistent with incentive-compatibility and is again independent of individual productivity. Thus although all taxes are independent of time and wealth, the model does provide a kind of justification for progressive savings taxes as higher types (who dominate the right tail in the stationary distribution) face higher savings taxes.

The tractability of the model also allows for a sharp characterization of the long-run distributions of consumption and firm size. I show that in this model the stationary distributions of both consumption and firm size admit closed-form densities of the "double-Pareto" form,[1] replicating the qualitative features of their empirical counterparts. Furthermore, the degree of inequality in the upper tail may be expressed in closed-form representation and so allows for simple comparative statics. The thickness of the upper tail depends upon the mean and volatility of consumption growth in the efficient allocation. Due to the nature of the agency problem and the implied need to have consumption depend upon the history of firm performance, consumption volatility (and hence inequality) is increasing in the marginal product of capital. Therefore, changes in technology and demographics that increase this latter quantity (such as changes in factor intensities or the number of workers per entrepreneur) will also increase inequality in the *efficient* allocation. Although much is known about the determinants of inequality in economies with exogenously incomplete markets,[2] these comparative statics concerning efficient allocations are new.

Related literature. An extensive literature, surveyed in Chari and Kehoe (1999), has analyzed optimal capital and labor taxation in environments in which agents face no idiosyncratic risk and the government is assumed to have access only to linear taxes on various forms of income. Within this so-called "Ramsey" tradition, all forms of non-labor income are typically grouped together, with no distinction made between risky or non-risky investments. Recent contributions in this vein such as Panousi and Reis (2017) and Evans extend this analysis to economies in which agents are subject to idiosyncratic risk, but do not consider the possibility (and benefits of) taxing different forms of capital income differently.

---

[1]Continuous piecewise polynomials on the positive halfline.
[2]See, e.g., Benhabib and Bisin (2018) for a survey.

The analysis in this paper is closer in spirit to the mechanism design approach of the New Dynamic Public Finance literature emanating from Golosov et al. (2003). Rather than seeking the optimal government policy lying with a pre-specified parametric class, this literature considers all allocations that satisfy incentive-constraints arising from informational asymmetries. However, the majority of this literature has focused on the implications of private-information in labor productivity and has not explicitly accounted for entrepreneurial activity. Therefore, it provides little guidance for the environment in this paper, in which multiple assets with different risk and return coexist. This omission is particularly striking given the observed variation in the taxation of different forms of capital income observed in practice (see, e.g., Gordon and Slemrod (1988) and Auerbach (2002)).

Three notable exceptions are Albanesi (2007), Shourideh (2012) and Phelan (2019). Albanesi (2007) considers a two-period model in which initial wealth is exogenous and common across agents and the returns to entrepreneurial activity depend upon unobserved (and privately-costly) effort. She finds that the optimal intertemporal wedge differs from the case with unobserved labor productivity and may assume either sign, and, as in this paper, considers multiple decentralizations of the efficient allocation. However, the two-period setup precludes an analysis of the efficient long-run distribution of wealth. More closely related with the current paper is Shourideh (2012), who also analyzes an agency model in which entrepreneurs may divert assets to private consumption. I reformulate the agency problem in continuous-time and adopt a welfare notion and life-cycle structure that leads to simpler characterizations of both efficient allocations and their implied taxes.[3] The modeling of the agency problem also qualitatively changes the nature of efficient intertemporal distortions. In contrast to the findings of both Albanesi (2007) and Shourideh (2012), I show that the inverse Euler equation of Rogerson (1985) and Golosov et al. (2003) continues to hold in the presence of production risk for a wide range of parameter values.

Finally, Phelan (2019) considers an environment in which the productivity (or human capital) of entrepreneurs grows randomly over time and depends on unobserved effort. The focus of Phelan (2019) is the characterization of efficient allocations in an environment with a novel agency problem involving human capital (rather than physical capital) and the implied effect on the efficient bearing of risk. Although Phelan (2019) and this paper both characterize efficiency in dynamic environments with agency frictions, the modeling assumptions, scope and results are quite different. Rather that analysing a novel agency problem, this paper instead shows how a variation on a previously explored agency problem leads to both increased tractability and novel results for the decentralization of the efficient allocation in a general equilibrium model. From a methodological point of view, I draw upon the continuous-time contracting literature (due to tractability gained), and in particular the martingale techniques pioneered in Sannikov (2008), for the recursive analysis. The method by

---

[3]I also allow entrepreneurs to abscond with a fraction of assets under their control, a restriction that turns out to be necessary for the problem to be well-defined.

which the principal-agent analysis is embedded within a macroeconomic setting follows Phelan (2019), which in turn is an extension of the techniques of Farhi and Werning (2007).

The outline of this paper is as follows: Section 2 analyzes a principal-agent model in which both the productivity of the agent and the interest rate are exogenous; Section 3 then embeds this into a macroeconomic model and characterizes stationary constrained-efficient allocations when productivity is endogenously determined by aggregate physical and labor resource constraints; Section 4 decentralizes these stationary efficient allocations in a general equilibrium model with exogenously incomplete markets and linear taxes on (risk-free) savings and profits, plots a numerical example and provides some comparative statics; and Section 5 concludes. Technical proofs and a discrete-time version of the environment that relates the findings of the main text to those of the related literature are outlined in the appendix.

2 Principal-agent model

This section characterizes the optimal risk-sharing arrangement between a risk-averse agent (she) and a risk-neutral principal (he) in an environment where the agent may operate a risky production technology, her consumption is private information, and she may abscond with a fraction of the physical assets under her control. Labor is absent from production, and both the marginal product of capital and interest rate are exogenous. This problem will later be embedded into a macroeconomic model in which flow payoffs to the principal are endogenously determined by aggregate resource constraints for both labor and capital.

2.1 Formal setup

The environment is a variation of that considered in Di Tella and Sannikov (2016).[4] Time is continuous and indefinite. The economy consists of a single risk-averse agent and a risk-neutral principal, both of whom live forever. The preferences of the agent over stochastic sequences of consumption $c := (c_t)_{t \geq 0}$ are represented by the utility function

$$U^A(c) := \rho \int_0^\infty e^{-\rho t} \mathbb{E}[\ln c_t] dt.$$

The agent has the ability to operate a constant-returns-to-scale technology subject to random shocks to the capital stock. Only the agent may operate the production function and so the principal must delegate capital to the agent in order for production to take place. When the capital delegated follows the process $K := (K_t)_{t \geq 0}$, output net of depreciation and borrowing costs $Y := (Y_t)_{t \geq 0}$ evolves according to

$$dY_t = [\Pi - r - \tau_I] K_t dt + \sigma K_t dB_t \tag{1}$$

---

[4]The problem considered here is simpler in one respect because savings are observable. However, as Di Tella and Sannikov (2016) observe, the principal's problem often gives infinite profits in the absence of hidden savings. Rather than allowing for hidden savings I instead assume that the agent may abscond with delegated capital and thereafter trade only a risk-free bond.

where $(B_t)_{t\geq 0}$ is distributed according to standard Brownian motion and defined on a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$. The constant Π in (1) may be interpreted as the marginal product of capital (net of depreciation) and will be made endogenous in Section 3, while r is the rate at which the principal discounts and τ_I is a tax on investment. The principal is risk-neutral and so their preferences over allocations $(K, c) := (K_t, c_t)_{t\geq 0}$ are represented by the function

$$U^P(K, c) := \int_0^\infty e^{-rt}\mathbb{E}[dY_t - c_t dt].$$

The agent has the ability to divert a fraction of output to private consumption. If the agent diverts a fraction a_t per unit of time then observed output evolves according to the law

$$dY_t = [\Pi - r - \tau_I - a_t]K_t dt + \sigma K_t dB_t. \tag{2}$$

The agent may only consume a fraction ϕ of the diverted output $a_t K_t$, where $\phi \in (0,1)$ is an exogenous constant. The parameter ϕ may be thought of as a measure of the severity of the agency problem and will play an important role in the decentralization. The specification in (2) may be interpreted as the continuous-time limit of the following discrete-time environments: the principal delegates resources to the agent, investment is *publicly* observed, but the capital stock is subject to idiosyncratic shocks that are *privately* observed.[5] In addition to the unobservability of consumption described above, I will also assume that the agent may, at any time, take a fraction ι of the capital delegated to her and abscond, and after doing so may only trade the same risk-free bond to which the principal has access.

An allocation in this environment must specify the consumption of the agent, the amount of capital delegated by the principal to the agent, and the fraction of capital the principal recommends the agent divert to private consumption, after every history of output. To be formal, let the underlying probability space be $(C[0,\infty), \mathcal{F}_t, P)$, where $\mathcal{F} := (\mathcal{F}_t)_{t\geq 0}$ is the σ-algebra generated by the evaluation maps[6] $(x_t)_{t\geq 0}$ and P is Wiener measure.

Definition 2.1. An allocation chosen by the planner is a pair $(K, c, \tilde{a})$ of $\mathcal{F}$-adapted processes on $C[0,\infty)$. An agent's strategy is a single $\mathcal{F}$-adapted process a defined on $C[0,\infty)$.

When the agent varies a, she alters the law of motion of output and so changes the measure used to evaluate $(K_t, c_t, \tilde{a})_{t\geq 0}$. Denote the measure associated with a by P^a and the corresponding expectation operator by $\mathbb{E}^a$ and note that the utility from adhering to such a strategy is

$$U^A(K, c, \tilde{a}; a) := \rho \int_0^\infty e^{-\rho t}\mathbb{E}^a[\ln(c_t + \phi a_t K_t)]dt.$$

Finally, associated with each allocation $(K, c, \tilde{a})$ and strategy a is the process $W \equiv W(K, c, \tilde{a}, a) = (W_t)_{t\geq 0}$ for continuation utility defined by

$$W_t := \rho \int_t^\infty e^{-\rho(s-t)}\mathbb{E}^a[\ln(c_t + \phi a_t K_t)|\,\mathcal{F}_t]ds. \tag{3}$$

---

[5]Further discussion on the relation with discrete-time models is given in Appendix A.
[6]Defined by $x_t(\omega) := \omega(t)$ for all $\omega \in C[0,\infty)$ and $t \geq 0$.

The following lemma requires only elementary algebra and so the proof is omitted.

Lemma 2.1. *When the agent absconds with K units of capital the utility from having access to a bond market with return rate r is given by $W = \ln(\rho K) + \rho^{-1}(r - \rho)$.*

Lemma 2.1 implies that when an agent may abscond with a fraction ι of the delegated capital and promised utility is given by $(W_t)_{t \geq 0}$, capital assignment $(k_t)_{t \geq 0}$ is subject to the additional constraint for all $t \geq 0$ a.s.

$$\iota k_t \leq \omega \exp W_t \tag{4}$$

where I have abbreviated $\omega := \rho^{-1} \exp(1 - r/\rho)$. An allocation is incentive-compatible if the agent wishes to follow the recommendations of the principal after every history of output. The formal definition is as follows.

Definition 2.2. An allocation $(K, c, \tilde{a})$ is incentive-compatible if

$$U^A(K, c, \tilde{a}; a) \geq U^A(K, c, \tilde{a}; \tilde{a}) \qquad \text{for all } a \in \mathcal{A} \tag{5}$$

and if the no-absconding constraint $K_t \leq \omega \exp(W_t + 1 - r/\rho)$ holds for all $t \geq 0$ almost surely. The set of incentive-compatible allocations is denoted $\mathcal{A}^{IC}$.

Since $\phi < 1$, output is destroyed whenever the agent diverts assets to private consumption. To characterize efficient allocations, it is therefore without loss of generality to restrict attention to allocations with $\tilde{a} = 0$ for all $t \geq 0$ almost surely. For ease of notation I will henceforth omit $\tilde{a}$ from the description of an allocation. Finally, in order to rule out allocations in which the principal runs a Ponzi-like scheme, I will impose the requirement that the present discounted value of transfers and production be finite, or

$$\int_0^\infty e^{-rt} \big| \mathbb{E}[(\Pi - r - \tau_I)K_t + c_t] \big| dt < \infty. \tag{6}$$

An allocation is incentive-feasible if it is both incentive-compatible and satisfies (6). The set of incentive-feasible allocations is denoted $\mathcal{A}^{IF}$. I may now define the principal's problem formally.

Definition 2.3. Given the utility from the agent's outside option W, the marginal product of capital Π, and the interest rate r, the problem of the principal is given by

$$V(W) = \max_{(K,c) \in \mathcal{A}^{IF}} \int_0^\infty e^{-rt} \mathbb{E}[(\Pi - r - \tau_I)K_t - c_t] dt$$

$$W = \int_0^\infty \rho e^{-\rho t} \mathbb{E}[\ln c_t] dt.$$

As is well-known, the principal's problem is recursive in the state variable W. Standard arguments from the continuous-time contracting literature[7] ensure that incentive-compatibility is equivalent to the requirement that promised utility follow a diffusion process with volatility weakly

---

[7]See, e.g., Sannikov (2008) andDi Tella and Sannikov (2016) and the references therein.

exceeding the marginal benefit of diverting output to consumption. Specifically, the requirement that the allocation (K, c) be incentive-compatible may be replaced by the explicit law of motion

$$dW_t = \rho(W_t - \ln c_t)dt + \rho\phi\sigma k_t c_t^{-1}dB_t. \tag{7}$$

Note that the drift term in (7) is the law of motion of W_t that would obtain in the absence of any uncertainty, as can be seen by simply differentiating (3) with respect to time. It will be convenient to measure promised utility in units of consumption and define $u_t := \exp W_t$. Ito's lemma implies

$$\frac{du_t}{u_t} = \left(\rho(\ln u_t - \ln C_t) + \frac{1}{2}(\rho\phi\sigma k_t/C_t)^2 \right)dt + \rho\phi\sigma(k_t/C_t)dB_t. \tag{8}$$

Notice that the elasticity of u_t with respect to output shocks is proportional to the amount of capital delegated and the marginal utility of consumption, since this product is the marginal utility of diverting a unit of output. The solution to the principal's problem admits a simple characterization whenever it is well-defined.

Proposition 2.2. *The value function of the principal is well-defined for sufficiently small Π, in which case it solves the Hamilton-Jacobi-Bellman equation*

$$rv(u) = \max_{\substack{c,k\geq 0 \\ k\leq \omega}}[\Pi - r - \tau_I]k - c + \left(-\rho\ln c/u + \frac{1}{2}(\rho\phi\sigma k/c)^2 \right)uv'(u) + \frac{1}{2}(\rho\phi\sigma k/c)^2 u^2 v''(u)$$

For such Π, both the value function and policy functions are linear in normalized utility. Further, the inverse Euler equation holds whenever the no-absconding inequality is strict, in which case the policy functions are given by

$$c(u) = \exp\left(1 - r/\rho + x(\Pi)^2/2\right)u \qquad\qquad k(u) = \frac{c(u)x(\Pi)}{\sqrt{\rho}\phi\sigma}$$

where $x(\Pi)$ is given by

$$x(\Pi) = \frac{1}{2R(\Pi)}\left[1 - \sqrt{1 - 4R(\Pi)^2}\right]$$

for $R(\Pi) := [\Pi - r - \tau_I]/(\sqrt{\rho}\phi\sigma)$.

Proof. See Appendix B. □

Sufficient conditions for the no-absconding constraint to not hold with equality are given in Lemma B.1. All the numerical examples in the main text satisfy this condition with $\omega = 1/\rho$, which by Lemma 2.1 corresponds to the case in which the agent may abscond with the entire capital stock.

Note that the principal is willing to assign more risk to the agent when the marginal product of capital is high. This hints at the results in Section 4 relating technological parameters with long-run inequality, as the latter is partly determined by the amount of risk to which each entrepreneur is exposed. Forces that increase the marginal value to society of an additional entrepreneur will therefore also tend to increase inequality in the efficient allocation.

8

As mentioned in the introduction, a large literature has extended the static model of Mirrlees (1971) and analyzed constrained-efficient allocations in dynamic environments with privately observed labor productivity shocks. An important result in this literature, established by Rogerson (1985) in a principal-agent setting and extended to a dynamic Mirrleesian environment byi Golosov et al. (2003), shows that intertemporal distortions are characterized by an inverse Euler equation. The characterization of such distortions in much of the literature follows a similar argument: beginning at an (as-yet-unknown) efficient allocation, it cannot be possible to perturb the allocation in such a way that the payoff to the principal is increased and incentive-compatibility is preserved. A similar perturbational argument is applicable here. First, note that the homotheticity of preferences and the log-linearity of the law of motion of wealth together imply that (k, c) satisfies (5) if and only if $(\eta k, \eta c)$ satisfies (5) for any deterministic process η. This observation provides us with a convenient class of incentive-compatible perturbations. Suppose that (k, c) is the efficient allocation and define a class of functions $(\eta_t(u))_{t \geq 0}$ according to

$$
\eta_t(u) = \begin{cases} \exp(u) & \text{if } t \in [t_0, t_0 + dt] \\ \exp(-ue^{\rho(t_1 - t_0)}) & \text{if } [t_1, t_1 + dt] \end{cases}
$$

for some positive t_0, t_1 and dt. The allocation $(\eta k, \eta c)$ continues to satisfy (5) and the implied change in utility to the agent is (up to first-order in dt)

$$
\Delta U = \rho e^{-\rho t_0} u \, dt + \rho e^{-\rho t_1} [-ue^{\rho(t_1 - t_0)}] dt = \rho[e^{-\rho t_0} - e^{-\rho t_1} e^{\rho(t_1 - t_0)}] u \, dt = 0. \tag{9}
$$

Therefore, the choice of η_t preserves promise-keeping and (5). It follows that if the no-absconding constraint is slack, the choice $u = 0$ must locally maximize profits. The change in the present discounted value of profits as one varies u in a neighborhood of the origin is approximately

$$
\Delta \Pi \approx (\Pi - r) \Big[e^{-rt_0} k_{t_0} \exp(u) + e^{-rt_1} \mathbb{E}[k_{t_1}] \exp(-ue^{\rho(t_1 - t_0)}) \Big] dt. \tag{10}
$$

Taking the derivative with respect to u and evaluating at zero gives $0 = e^{-rt_0} k_{t_0} - e^{-rt_1} e^{\rho(t_1 - t_0)} \mathbb{E}[k_{t_1}]$ and hence $k_{t_0} = e^{-r(t_1 - t_0)} e^{\rho(t_1 - t_0)} \mathbb{E}[k_{t_1}]$. Finally, since k/c is independent of history, we have $k_0/c_0 = k_t/c_t$ or

$$
1 = e^{-(r-\rho)(t_2 - t_1)} \mathbb{E}\left[\frac{u'(c_0)}{u'(c_1)} \right] \tag{11}
$$

which is exactly the inverse Euler equation. Note that the assumption that the no-absconding constraint does not hold with equality is necessary for the above argument because otherwise the perturbation $(\underline{c}(u), \underline{k}(u))$ will not be incentive-compatible for $u > 0$. Indeed, the inverse Euler equation may fail to hold in this case.[8] However, note that even when the no-absconding constraint holds with equality, it remains true that the perturbed allocation is incentive-compatible for $u \geq 0$ and so we still have the inequality $\mu_c \leq r - \rho$.

---

[8]This is easiest to see by solving the Hamilton-Jacobi-Bellman equation in the case where $\phi\sigma = 0$.

Section 4 shows how a class of stationary efficient allocations may be decentralized in a general equilibrium model using a particular set of taxes and transfers. Such a characterization is necessarily specific to the choice of Pareto weights attached to different generations, the set of instruments available to the government, and the assumed market structure. To isolate the role of agency frictions independently of general equilibrium effects, it is instructive to first analyze efficient distortions by comparing the solution to the above principal-agent problems with the allocations that arise when the agent may invest in either capital or the risk-free bond available to the principal. To motivate the following analysis, first note that if the return from continually investing in an asset over the interval $[t, t + \Delta]$ is given by $R = R_{t,t+\Delta}$, then intertemporal optimization implies

$$u'(c_t) = \exp(-\rho\Delta)\mathbb{E}[Ru'(c_{t+\Delta})|\mathcal{F}_t]. \tag{12}$$

The wedges defined in Definition 2.4 measure the extent to which (12) fails for an arbitrary return.

Definition 2.4. Given the consumption process $(c_t)_{t\geq0}$ in the principal-agent problem, for each asset A with return process $(R_t^A)_{t\geq0}$ the associated wedge ν^A is defined implicitly by

$$u'(c_0) = \exp(-\rho t)\mathbb{E}\big[\exp(-\nu^R t)R_t^A u'(c_t)\big].$$

Denote by ν^K and ν^B the wedges associated with risky capital and the risk-free bond, respectively, and note that the associated return processes R^K and R^B are given by

$$R_t^K = \exp\left(\big[\Pi - \sigma^2/2\big]t + \sigma B_t\right) \qquad\qquad R_t^B = \exp\left(rt\right)$$

for all $t \geq 0$. It is easy to verify that these wedges are the unique constants such that the solution to the problem

$$\max_{(c_t,k_t)_{t\geq0}} \rho \int_0^\infty e^{-\rho t}\mathbb{E}[\ln c_t]dt$$

$$da_t = \big[(r - \nu^B)(a_t - k_t) - c_t + (\Pi - \nu^K)k_t\big]dt + \sigma k_t dB_t$$

coincides with the solution to the principal-agent problem. As such, they represent the extent to which the presence of private information forces the technological returns on each asset to differ from the returns accruing to the agent. The closed-form expression for consumption allows for a sharp characterization of the intertemporal wedges.

Lemma 2.3. *The intertemporal wedges for risky capital and the risk-free bond are given by*

$$\nu^K = \Pi - r + \rho x(\Pi)^2 - \sqrt{\rho}\sigma x(\Pi) \qquad\qquad \nu^B = \rho x(\Pi)^2.$$

Further, $\nu^B \geq \nu^K$ and $\nu^B \geq 0$, while the risky wedge ν^K may assume either sign.
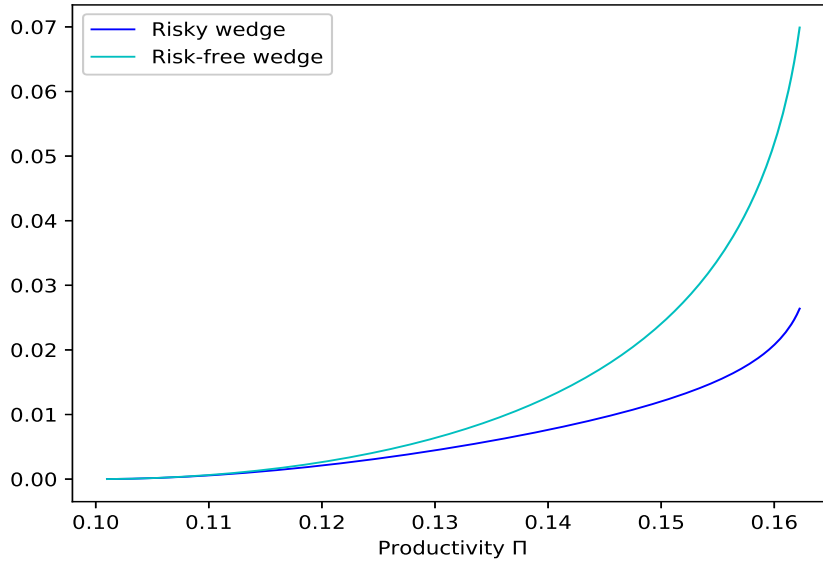
Proof. See Appendix B.3. □

Figure 1: Intertemporal wedges

Figure 1 depicts the intertemporal wedges for both risky capital and the risk-free bond as a function of the marginal product of capital, for the parameters $(\rho, r, \phi, \sigma, \delta) = (0.1, 0.1, 1., 0.3, 0.05)$. As noted in Lemma 2.3, the wedge on risky capital is everywhere below that on the risk-free bond and may in fact be negative. However, I will show in Section 4 that these wedges do not translate immediately into taxes in the decentralization of the efficient allocation. Indeed, although the wedge on the return on the risky asset everywhere exceeds that of the safe return, it does not follow that the tax on savings must exceed the tax on profits.

3   Stationary efficient allocations

The previous section characterized the efficient contract between a single risk-averse agent and risk-neutral principal given an exogenous intertemporal rate and net productivity. This section uses the efficient contract found above to completely characterize a particular stationary constrained efficient allocation in an economy with a continuum of agents and endogenous productivity.

3.1   Formal setup

Time is again indefinite and continuous. At any moment there is a continuum of agents of mass L who do not care for their descendants, discount at rate ρ_S, and die at rate ρ_D. Agents may engage in two activities: the first is identified with entrepreneurial activity (running a business) and the second with wage labor (working for someone else). These activities are not mutually exclusive and so agents may perform both simultaneously. All agents are endowed with one unit of time and have

preferences over sequences $(c_t)_{t \geq 0}$ represented by

$$U(c) := \rho \int_0^\infty e^{-\rho t} \mathbb{E}[\ln c_t] dt$$

where $\rho := \rho_S + \rho_D$ may be interpreted as the effective rate of discount, inclusive of the possibility of death. Entrepreneurs have access to a risky production technology that produces consumption using physical capital and labor. Specifically, if an entrepreneur assigns capital and labor to their technology according to the processes $(K_t, L_t)_{t \geq 0}$, then the law of motion of physical capital is

$$dY_t = \left(A K_t^\alpha L_t^{1-\alpha} - \delta K_t \right) dt + \sigma K_t dB_t$$

where $B := (B_t)_{t \geq 0}$ is a standard Brownian motion, $A > 0$ and $\alpha \in (0, 1)$ are exogenous constants and δ the depreciation rate. An allocation is now indexed by an entire initial distribution Φ over v. The formal definition is the following. Since agents experience no disutility from labor, I will for brevity (and without loss) omit labor supply from the definition.

Definition 3.1. Given a distribution Φ over promised utility, an allocation A consists of consumption, capital assignments, and labor assignments

$$A = \left\{ (c_t^v, k_t^v, l_t^v)_{t \geq 0}, \ (c_t^T, k_t^T, l_t^T)_{t \geq T \geq 0} \ \middle| \ v \in \text{supp}(\Phi) \right\}$$

for the initial generation, and all subsequent generations, respectively.

In contrast with the principal-agent setting of Section 2, incentive-compatibility for an allocation here requires that promises made to the initial generation be satisfied, as well as that no agent have an incentive to divert assets to private consumption.

Definition 3.2. Given a distribution Φ over promised utility v, an allocation A satisfies promise-keeping if $U(c^v) = v$ for all $v \in \text{supp}(\Phi)$. An allocation is incentive-compatible if it satisfies promise-keeping and the incentive-compatibility conditions of the previous section are satisfied.

Denote by $C_t(A), K_t(A), Y_t(A)$ and $L_t(A)$ aggregate consumption, existing capital, output, and labor assigned at date t given the allocation A. Formal definitions are given in Appendix C.1.

Definition 3.3. An allocation A is resource feasible given the capital stock K if $K_0 = K$ and

$$C_t(A) + \dot{K}_t(A) \leq Y_t(A)$$
$$L_t(A) \leq L$$

for all $t \geq 0$. The set of such allocations will be denoted $\mathcal{A}^{RF}$. An allocation is incentive feasible given Φ and K if it is both resource feasible and incentive-compatible given Φ and K. The set of all such allocations will be denoted $\mathcal{A}^{IF}(\Phi, K)$.

I will assume that the planner places a Pareto weight $\alpha(T) = e^{-\rho_S T}$ on the utility of an agent born at date $T \geq 0$. This assumption ensures that the planner values an agent's utility at any given

date the same, regardless of the agent's date of birth. It may be viewed as a kind of generalized utilitarianism across generations and is equivalent to assuming the following social welfare function:

$$U^P = \int_0^\infty \left(e^{-\rho t} \underline{U}_t + \int_0^t e^{-\rho[t-T]} e^{-\rho_S T} U_t^T dT \right) dt$$

where $\underline{U}_t$ and U_t^T refers to aggregate flow utility experienced by the initial and Tth generations at date $t \geq 0$.[9] I may now specify the planning problem.

Definition 3.4. Given an initial distribution Φ and capital stock K, the problem of the planner is

$$V(\Phi, K) = \max_{A \in \mathcal{A}^{IF}(\Phi, K)} U^P(A).$$

The problem in Definition 3.4 is intractable for an arbitrary initial distribution, so I will restrict attention to solutions in which cross-sectional distributions are constant over time. I will characterize these solutions using the method employed in Farhi and Werning (2007) and consider, in succession, *relaxed* and *generational* planner's problems. The relaxed planner's problem has the same objective and state variable as the above planner's problem but allows for intertemporal trade at rate ρ_S.

Definition 3.5. Given an initial capital stock K and distribution Φ over promised utility, the relaxed planner's problem is

$$V^R(\Phi, K) = \max_{A \in \mathcal{A}^{IC}(\Phi)} U^P(A)$$

$$\int_0^\infty e^{-\rho_S t}[C_t(A) + \dot{K}_t(A)]dt \leq \int_0^\infty e^{-\rho_S t} Y_t(A)dt$$

$$\int_0^\infty e^{-\rho_S t} L_t(A)dt \leq \int_0^\infty e^{-\rho_S t} L dt$$

$$K_0 = K.$$

Note that if an allocation solves the relaxed planner's problem and the associated distribution of promised utility is constant over time, then this allocation also solves the original planner's problem beginning at that distribution. Further, it is easy to see that the subjective rate of discount ρ_S is the only intertemporal price for which such stationarity may arise, for all other prices would induce an increasing or decreasing trend in utility across generations. Therefore, in order to characterize stationary solutions to the original planner's problem, it suffices to consider problems of the form in Definition 3.5, and find the Φ and K such that stationarity arises.

Although the relaxed planner's problem continues to take an entire distribution as an argument, it is much simpler than the original planner's problem as there are only two resource constraints, rather than two for each instant in time. Lagrange's theorem then implies that there exists a pair of multipliers $\lambda := (\lambda_R, \lambda_L)$ such that the allocation A that solves the relaxed planner's problem maximizes the Lagrangian

$$V_\lambda(\Phi) = \max_{A \in \mathcal{A}^{IC}} \int_0^\infty e^{-\rho t} \left(\underline{\mathcal{L}}_t dt + \int_0^t e^{-\rho_S T} e^{-\rho(t-T)} \mathcal{L}_t^T dT \right) dt$$

---

[9]Again formal definitions are given in Appendix C.1.

13

where $\underline{\mathcal{L}}_t(\lambda)$ and $\mathcal{L}_t^T(\lambda)$ are the contributions of the initial and Tth generations to the Lagrangian

$$\begin{aligned}
\underline{\mathcal{L}}_t &= \underline{U}_t + \lambda_R[\underline{Y}_t - \rho_S\underline{K}_t - \underline{C}_t + \lambda_L[L - \underline{L}_t]] \\
\mathcal{L}_t^T &= U_t^T + \lambda_R[Y_t^T - \rho_S\underline{K}_t - C_t^T + \lambda_L[L - L_t^T]]
\end{aligned}$$

(13)

and $(\underline{C}_t, \underline{Y}_t, \underline{K}_t, \underline{L}_t)$ and $(C_t^T, Y_t^T, \underline{K}_t^T, L_t^T)$ refer to consumption, output, and labor assignments of initial and Tth generations, respectively, at date $t \geq 0$.[10] Although the state variable remains an entire distribution, the objective may be maximized pointwise as all interdependence across agents is captured by the multipliers. One may then treat each generation in isolation and vary the multipliers until the resource constraints hold in the implied stationary distribution. I will refer to the problem of dealing with a single generation of newborns as the *generational planner's problem*.

Definition 3.6. Given multipliers $\lambda := (\lambda_R, \lambda_L)$ the generational planner's problem is

$$V_\lambda^G = \max_{A \in \mathcal{A}_G^{IC}} \int_0^\infty e^{-\rho t}(\underline{U}_t + \lambda_R[\underline{Y}_t - \rho_S\underline{K}_t - \underline{C}_t + \lambda_L[L - \underline{L}_t]])dt.$$

The choice of assigning labor to entrepreneurs is purely static and depends solely upon λ_L. Conditional on assigning a newborn to be an entrepreneur, the problem of the generational planner is equivalent to the principal-agent problem analyzed in Section 2 with the marginal product of capital now a function of λ_L.

Lemma 3.1. *Given multipliers $\lambda := (\lambda_R, \lambda_L)$, the generational planner's problem may be written*

$$V_\lambda^G = \rho^{-1}\lambda_R\lambda_L + \max_{W \in \mathbb{R}} W + \lambda_R V(W, \Pi(\lambda_L))$$

where $\Pi(\lambda_L) := \max_{l \geq 0} Al^{1-\alpha} - \lambda_L l - \delta = \lambda_L^{1-1/\alpha}\alpha A[(1-\alpha)A]^{1/\alpha-1} - \delta$ is the marginal product of capital and $V(W, \Pi(\lambda_L))$ denotes the value function of a principal given in Definition 2.3 with the marginal product of capital Π who discounts at rate ρ and faces investment tax $\tau_I = -\rho_D$.

Each λ_R corresponds to a level of utility promised to newborns and each λ_L corresponds to a level of the marginal product of capital. One may then solve for the stationary distributions associated with each pair of multipliers by using the policy functions found in Section 2. Proposition 3.2 is the first main result of this paper. It shows that characterizing stationary efficient allocations reduces to finding an appropriate level of the marginal product of capital.

Proposition 3.2. *The stationary level of $\overline{\Pi}$ is a solution to the equation*

$$\underline{c}(\Pi) = \frac{1}{\alpha}(\Pi + (1-\alpha)\delta)\underline{k}(\Pi)$$

where $\underline{c}$ and $\underline{k}$ are given in Proposition 2.2 with $r = \rho$ and $\tau_I = -\rho_D$, provided the no-absconding constraint holds as a strict inequality for this $\overline{\Pi}$.

---

[10]Detailed definitions are found in Appendix C.1.

Proof. We wish to find the λ_L and λ_R such that the resource constraints are satisfied in the stationary distributions implied by the solution to the generational planner's problem. First, note that by Proposition 2.2 the process $(u_t)_{t\geq 0}$ has zero drift when the no-absconding constraint holds as a strict inequality, and so its mean is simply u_0. Aggregating over all agents, aggregate capital is

$$K = L\underline{k}(\Pi(\lambda_L))u_0. \tag{14}$$

Using the function $\Pi(\cdot)$ defined in Lemma 3.1, given the multiplier λ_L, the flow production from the firm of a u-type net of depreciation is

$$AK(u)^\alpha L(u)^{1-\alpha} - \delta K(u) = A(\underline{k}(\Pi(\lambda_L))u)^\alpha(\underline{k}(\Pi(\lambda_L))l(\lambda_L)u)^{1-\alpha} - \delta\underline{k}(\Pi(\lambda_L))u$$
$$= (\Pi(\lambda_L) + \lambda_L l(\lambda_L))\underline{k}(\Pi(\lambda_L))u.$$

Aggregating over all such agents, the labor resource constraint implies $1 = l(\lambda_L)\underline{k}(\Pi(\lambda_L))u_0$. Using $l(\lambda_L) = [A(1-\alpha)]^{1/\alpha}\lambda_L^{-1/\alpha}$, the goods resource constraint becomes

$$\underline{c}(\Pi(\lambda_L))u_0 = (\Pi(\lambda_L) + \lambda_L l(\lambda_L))\underline{k}(\Pi(\lambda_L))u_0 = (\Pi(\lambda_L) + [A(1-\alpha)]^{1/\alpha}\lambda_L^{1-1/\alpha})\underline{k}(\Pi(\lambda_L))u_0$$

Simplifying the above resource constraints and using $l(\lambda_L) = [A(1-\alpha)]^{1/\alpha}\lambda_L^{-1/\alpha}$ gives the desired equation for Π. The expressions for the multiplier and level of the capital stock then follow by combining (14) with the static labor assignment function. $\square$

The simplicity of the characterization given in Proposition 3.2 is due partly to the homotheticity of preferences and partly to the welfare criterion adopted in this paper that weights the flow utility of an agent the same independently of her birth date. Several papers in the literature on dynamic contracting with private information, such as Atkeson and Lucas (1992) or Phelan (1994), consider component planner's problems similar to the above generational planner's problems, but either adopt a welfare criterion with zero discounting or they place weight solely upon the first generation. Such an approach necessitates solving a component planning problem for an arbitrary interest rate that is then varied until resources are balanced. In contrast, with the welfare criterion of this paper, it is immediate that the only price for which stationarity may arise is the subjective discount rate of the agents, as all other prices induce a trend in utility across generations. Together with the assumption of Logarithmic utility and the implication that consumption then follows a martingale, this implies that changes in technology have no effect on the trend in consumption and they affect the efficient allocation only insofar as they alter the marginal product of capital.

The linearity of the policy functions for both capital and consumption also allows for a simple characterization of the stationary distribution of consumption. In general, the stationary distribution of a killed geometric Brownian motion will depend on the mean and volatility of the growth rate and the hazard rate of death. Since death is exogenous and mean consumption growth is zero by the inverse Euler equation, the stationary distribution is solely determined by the risk borne by entrepreneurs. When combined with the defining equation for Π given in Proposition 3.2 this in turn

allows us to determine how changes in technological parameters affect efficient long-run inequality. By (14) and the expression for capital in Proposition 3.2 the initial level of the consumption of entrepreneurs is given by

$$\underline{c} = \frac{\underline{c}(\Pi)K}{\underline{k}(\Pi)} = \frac{\underline{c}(\Pi)}{\underline{k}(\Pi)}\left(\frac{\alpha A}{\Pi + \delta}\right)^{\frac{1}{1-\alpha}}.$$

Combining the above with standard results from the theory of diffusion processes gives the following characterization of the stationary distribution.

Corollary 3.3. *The stationary distribution of the consumption of entrepreneurs associated with the constrained-efficient allocation has density given by*

$$f(C) = \begin{cases} D_1 C^{\beta_+ - 1} & \text{if } C \leq \underline{c} \\ D_2 C^{\beta_- - 1} & \text{if } C \geq \underline{c} \end{cases}$$

with the exponents $\beta_\pm$ given by

$$\beta_\pm = -\frac{1}{2} \pm \sqrt{\frac{1}{4} + \frac{2\rho_D}{\rho x(\Pi)^2}},$$

where $\Pi = \Pi(\lambda_L)$ is the profit level given in Proposition 3.2, and D_1 and D_2 are determined by the requirement that the density integrate to unity and be continuous.

Before turning to the decentralization it is useful to summarize the main points of the above characterization. The efficient allocation is completely described by the following requirements: all newborns attain W units of utility, agents have zero drift in consumption, and volatility is equal to that implied by the marginal product of capital given in Proposition 3.2. The task of the next section is to characterize the taxes that ensure these properties arise in a stationary competitive equilibrium.

4   Decentralization

The preceding sections characterized efficient allocations for consumption and investment, with all co-ordination implicitly conducted by a benevolent social planner executing the direct mechanism. In order to allow this analysis to address the motivating questions stated in the introduction, I now turn to the question of how such allocations may be implemented with trade in decentralized markets. Such an analysis is necessarily contingent upon the contracts agent are assumed capable of writing. For this reason I will consider two separate decentralizations that embody opposite extreme assumptions on the extent to which the private sector may share risk. In the first, contracts are restricted only by informational asymmetries, whereas in the second, agents are assumed able to only trade a risk-free bond and are subject to taxes on savings and reported profits. In both cases all taxes are linear and independent of history.

4.1 Unrestricted contracts

In this section I will suppose there exists a competitive sector of financial intermediaries with whom agents may contract. Intermediaries are risk-neutral, infinitely lived and have the ability to commit to any contract. They compete with one another to provide insurance contracts to newborn agents, in return for rights to future labor income and government transfers. As in previous sections, all agents are endowed with the ability to operate a risky production technology subject to private depreciation shocks, and at any time may abscond with a fraction of the capital stock delegated to them. The intermediaries are subject to the same informational asymmetries as the planner of previous sections and so must structure their payments to the agent to be incentive-compatible.

I will suppose that the government may levy taxes on the output of intermediaries, issue government debt, and transfer lump-sum payments to agents at birth. Although this is a restricted set of instruments, Proposition 4.1 shows that they suffice to decentralize the efficient allocation.

Definition 4.1 (Intermediaries' problem). Given sequences of wages $w = (w_t)_{t \geq 0}$, interest rates $r = (r_t)_{t \geq 0}$, and taxes $(\tau_{Yt})_{t \geq 0}$ on output, the problem of a financial intermediary facing a newborn at time $t \geq 0$ with outside option u is defined to be

$$W(u, t; w, r, \tau_Y) = \max_{(c, K, L) \in \mathcal{A}^{IC}(u)} \int_t^\infty e^{-\int_t^{t'}(r_s + \rho_D)ds} \mathbb{E}[(1 - \tau_{Yt'})(\Pi(w_{t'}) - r_{t'})K_{t'} - c_{t'}]dt'$$

where $\mathcal{A}^{IC}(u)$ denotes the set of all triples for consumption, capital, and labor allocations that are incentive-compatible and given the agent's normalized utility u.

Notice that although intermediaries are infinitely lived and care only about the present discount value of profits, the death rate ρ_D appears in their objective, as it determines the probability that a given agent will be alive to receive the promised consumption and produce the planned output. The relevant state variable for the intermediary is the utility associated with the agent's outside option as this is the level that must be delivered to induce the agent to sign the contract.

The notion of competitive equilibrium in this environment is standard: given the behavior of the government and the sequences of wages and interest rates, consumers and intermediaries maximize utility and profits, respectively, and markets clear.

Definition 4.2 (Competitive equilibrium). Given a choice of transfers to newborns $(T_t)_{t \geq 0}$ and taxes $(\tau_{Yt})_{t \geq 0}$ imposed on the output of intermediaries, a competitive equilibrium consists of a collection of contracts for the initial and future generations such that all agents choose the best contract available to them, intermediaries maximize profits, and markets clear.

The first decentralization result shows that when agents are unrestricted in the contracts they may sign, the role of the government is relegated to distributing wealth equally across generations.

Proposition 4.1. *Under the assumptions of Proposition 3.2, the stationary constrained efficient allocation coincides with the competitive equilibrium in which $\tau_{Yt} = 0$ for all $t \geq 0$ and the government owns the capital stock and gives all agents a flow transfer of $\rho_S K$ every instant. The associated interest rates and wages are constant at $r_t = \rho_S$ and $w = (1-\alpha)AK^\alpha L^{1-\alpha}$.*

Proof. When $r_t = \rho_S$ and $\tau_{Yt} = 0$ for all $t \geq 0$ the problem of the intermediary given in Definition 4.1 is identical to that of the generational planner, and for this interest rate, the government budget constraint is satisfied at every instant, since $\rho_S K$ is the interest earned on the aggregate capital stock (the negative of the government debt).

It remains to show that the intermediaries make zero profits when the marginal product of capital coincides with the efficient level and must give agents u_0 units of lifetime utility. Proposition 3.2 implies that this efficient Π solves $\underline{c}(\Pi) = (\Pi/\alpha + (1/\alpha - 1)\delta)\underline{k}(\Pi)$ where $\underline{c}(\Pi)$ and $\underline{k}(\Pi)$ are the policy function (per unit of normalized utility) associated with the intermediaries' problem when $r = \rho_S$. If u_0 is normalized utility given to newborns, the fact that consumption has zero drift implies $K = \overline{k}(\Pi)u_0$ and so the profit function of the principal is $v(u) = \overline{v}u$ where $\rho\overline{v}u = ([\Pi - \rho_S]\overline{k} - \overline{c})u$. Rearranging and using the equation in Proposition 3.2 again gives

$$\overline{v} = \frac{1}{\rho}\left(\Pi - \rho_S - \frac{1}{\alpha}(\Pi + (1-\alpha)\delta)\right)\underline{k}(\Pi).$$

The discounted value of wages and transfers is $\rho^{-1}[wL + \rho_S K] = \rho^{-1}[(1-\alpha)AK^\alpha L^{1-\alpha} + \rho_S K]$ while $\Pi = \alpha A K^{\alpha-1} L^{1-\alpha} - \delta$. It follows that the discounted value of profits of the intermediary is

$$\frac{1}{\rho}([\Pi + \delta](1 - 1/\alpha) - \rho_S)K + \frac{1}{\rho}[(1-\alpha)AK^\alpha L^{1-\alpha} + \rho_S K] = 0$$

as claimed. □

Theorem 4.1 is reminiscent of the seminal results of Prescott and Townsend (1984) and Atkeson and Lucas (1992), who show that competitive equilibria in endowment economies are efficient if all asymmetric information arises after the signing of contracts. The logic of these papers continues to go through here except for the fact that the initial wealth of agents depends on government policy.

4.2 Restricted contracts

The decentralization given in Section 4.1 is a natural benchmark as it proceeds from the assumption that the government is no more capable of overcoming informational asymmetries than the private sector. However, the assumption of a perfectly competitive sector of financial intermediaries capable of providing such contracts may be unreasonably strong, as such insurance opportunities may simply be absent in the real world. Therefore, a benevolent government may wish to provide additional social insurance using some combination of taxes and transfers. To complement the foregoing analysis, this section implements the efficient allocation with a market structure that captures the opposing extreme assumption of no risk-sharing in the private sector. Specifically, I consider an

environment with idiosyncratic production risk in which agents may trade only a risk-free bond in zero net supply, and show that the constrained-efficient allocation of the previous section may be implemented using linear taxes on (reported) profits and risk-free savings.

4.3 Market structure and equilibrium characterization

All agents are endowed by the government with wealth at birth, may save in a risk-free bond, and contract with a life insurance company to insure against length-of-life risks. Agents in turn rent capital on behalf of their business at the risk-free rate, are unable to issue shares in the profits of their business, and pay taxes on output net of depreciation and interest payments (profits). In order to respect the informational asymmetry between the agents and the government, it must be the case that the taxes depend only upon observable quantities. Therefore, the profits tax must be viewed as a tax on *reported* profits. At any instant, the labor-hiring decision is purely static and so may be solved independently of all savings and investment decisions and so the problems of agents facing constant linear prices are given as follows.

Definition 4.3. Given taxes τ_s and τ_Π on risk-free savings and profits, wage w, and the risk-free rate r, the problem of the agent with a units of assets is

$$V(a) = \max_{(c_t, k_t)_{t \geq 0}} \rho \int_0^\infty e^{-\rho t} \mathbb{E}[\ln c_t] dt$$

$$da_t = [(1 - \tau_s)(r + \rho_D)a_t - c_t + w]dt + (1 - \tau_\Pi)k_t dR_t$$

where $dR_t = [\Pi(w) - r]dt + \sigma dB_t$.

Notice that since dR_t may be negative, the above problem embodies the assumption that the agent receives a tax refund if her firm sustains losses. To aid the reader in understanding the market structure implicit in the above, a discrete-time formulation is contained in Appendix A.2.

Definition 4.3 makes no mention of incentive-compatibility and writes the law of motion of wealth under the assumption of truthful reporting. Clearly not every choice of profits tax will be incentive-compatible. I will show that one may choose taxes such that the resulting competitive equilibrium coincides with the allocation in Section 3, from which incentive-compatibility will be immediate. Finally, I will also assume that newborn agents inherit ηK, where K is the aggregate capital stock and η a parameter chosen by the government, and that these transfers are funded by the interest from government ownership of a fraction ν of the capital stock.

Definition 4.4 (Tax distorted stationary equilibrium)**.** Given taxes τ_s and τ_Π on risk-free savings and profits of entrepreneurs, respectively, a stationary competitive equilibrium consists of an aggregate capital stock K, an inheritance level η, a fraction ν of the capital stock owned by the government, a wage rate w, and a risk-free rate r such that agents maximize; markets for labor, capital, and goods clear; and the government budget constraint is satisfied.

As with the agency problem considered earlier, the homotheticity of preferences and homogeneity of the law of motion of wealth ensure that individual problems admit homogeneous solutions for any choice of taxes.

Lemma 4.2. *The value function of the agents is given by*

$$V(a) = \ln \rho + \ln (a + h) + \rho^{-1} \left(\mu_a - \sigma_a^2/2 \right)$$

where μ_a and σ_a denote the drift and diffusion in the evolution of assets

$$
\begin{aligned}
\mu_a &:= \mu_a(w, r) = (1 - \tau_s)(r + \rho_D) - \rho + \frac{(\Pi(w) - r)^2}{\sigma^2} \\
\sigma_a &:= \sigma_a(w, r) = \frac{\Pi(w) - r}{\sigma}
\end{aligned}
\tag{15}
$$

and $h := h(w, \tau_s, r) = w/[(1 - \tau_s)(r + \rho_D)]$ is the human wealth of the agent. The policy functions for consumption and capital are

$$c(a) = \rho(a + h) \qquad\qquad k(a) = \frac{[\Pi(w) - r]}{\sigma^2(1 - \tau_\Pi)}(a + h). \tag{16}$$

Proof. See Appendix D.1. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Notice that the law of motion of wealth is entirely independent of the profits tax. While this may seem counterintuitive, it simply follows from the linearity of the production technology and the symmetric treatment of profits and losses, illustrating an effect first highlighted by Domar and Musgrave (1944).

Prior to the complete decentralization of the allocation, it is instructive to relate the incentive-compatibility constraints to the value function of the agent. When considering whether to divert the depreciation shock to private consumption, the agent compares the marginal utility $\rho\phi\sigma k u'(c)$ of diverting with the marginal utility $\sigma(1 - \tau_\Pi)k V'(a)$ of faithfully reporting and paying the tax $\sigma\tau_\Pi k$. Since $\rho u'(c) = V'(a)$ by the envelope theorem, the incentive-compatibility of the allocation in Definition 4.3 is equivalent to $\tau_\Pi \leq 1 - \phi$. Since we know that the incentive-constraint does indeed bind after every history (the elasticity of consumption with respect to output is never zero), this suggests setting the profits tax to $\tau_\Pi = 1 - \phi$.

To formalize this claim, recall that the constrained-efficient allocation is characterized by the following conditions: agents have zero drift in consumption, the volatility of consumption growth is given by $\sqrt{\rho}x(\Pi)$, and the marginal product of capital in the competitive equilibrium coincides with that given in Proposition 3.2. The following is the second main result of this paper after Proposition 3.2. It characterizes the taxes that decentralize the efficient allocation found in Section 3.

Proposition 4.3. *The stationary constrained-efficient allocation coincides with the stationary competitive equilibrium allocation in which the taxes on risk-free savings and profits are given by*

$$1 - \tau_s = \frac{\rho[1 - x(\Pi)^2]}{\Pi - \sqrt{\rho}\sigma x(\Pi) + \rho_D} \qquad\qquad 1 - \tau_\Pi = \phi,$$

respectively, and in which each agent is endowed with a fraction

$$\eta = \frac{\Pi - \sqrt{\rho}\phi\sigma x(\Pi)}{\rho(1 - x(\Pi)^2)}$$

of the aggregate capital stock at birth.

Proof. The three market-clearing conditions for labor, capital, and goods are

$$L = \phi_l(w)K$$
$$K = \overline{k}(w,r)\left(\eta K + \frac{wL}{(1-\tau_s)(r+\rho_D)}\right) \tag{17}$$
$$AK^\alpha L^{1-\alpha} - \delta K = \rho\left(\eta K + \frac{wL}{(1-\tau_s)(r+\rho_D)}\right).$$

We wish to find the η, τ_s and τ_Π such that the solution to the above triple of equations implies that the capital stock and the drift and volatility in consumption coincide with their efficient counterparts. First, note that for this to occur, the ratio of capital to consumption in the equilibrium must coincide with the corresponding ratio in the constrained-efficient allocation, and so

$$\frac{\underline{k}(\Pi)}{\underline{c}(\Pi)} = \frac{\overline{k}(w,r)}{\overline{c}(w,r)}. \tag{18}$$

Equating the volatility of consumption in the competitive equilibrium and efficient allocation gives

$$\frac{\rho\phi\sigma\underline{k}(\Pi)}{\underline{c}(\Pi)} = (1-\tau_\Pi)\sigma\overline{k}(w,r). \tag{19}$$

Since $\overline{c}(w,r) = \rho$, combining (18) and (19) implies $1 - \tau_\Pi = \phi$. Next note that in order for the volatility of the agent's wealth to coincide with the level in the efficient allocation, we require

$$(1-\tau_\Pi)\sigma\overline{k} = \frac{\Pi(w) - r}{\sigma} = \sqrt{\rho}x(\Pi)$$

and so the risk-free rate must satisfy $r = \alpha AK^{\alpha-1}L^{1-\alpha} - \delta - \sqrt{\rho}\sigma x(\Pi) = \Pi - \sqrt{\rho}\sigma x(\Pi)$, as claimed. The requirement that the drift in consumption vanishes becomes $0 = (1-\tau_s)(r+\rho_D) - \rho + \rho x(\Pi)^2$, which rearranges to the desired expression for τ_s upon substitution of the above interest rate. Combining the second and third market-clearing conditions and substituting the above expressions gives $AK^{\alpha-1}L^{1-\alpha} - \delta = \rho(\eta + (wL/K)/[(1-\tau_s)(r+\rho_D)])$ or

$$\frac{1}{\alpha}(\Pi + (1-\alpha)\delta) = \rho\eta + \frac{(1/\alpha - 1)[\Pi + \delta]}{1 - x(\Pi)^2}$$

which rearranges to the claimed expression for η. Finally, the government budget constraint is automatically satisfied by Walras' law. $\square$

It is instructive to verify that the government budget constraint is automatically satisfied by using the explicit expressions given above. The revenue raised by the profits tax is

$$\tau_\Pi[\Pi - r]K = (1-\phi)\sqrt{\rho}\sigma x(\Pi)K$$

while the revenue raised from the savings tax is

$$\tau_s(r + \rho_D)\eta K = \big(\Pi - \sqrt{\rho}\sigma x(\Pi) + \rho_D - \rho[1 - x(\Pi)^2]\big)\eta K.$$

The interest revenue from the asset holdings of the government is $r(1 - \eta)K = (\Pi - \sqrt{\rho}\sigma x(\Pi))(1 - \eta)K$, while the expenditures of the government consist only of the transfers to newborn agents. The flow of revenue minus expenditures of the government is therefore

$$\big[\Pi - \sqrt{\rho}\sigma x(\Pi) - \rho[1 - x(\Pi)^2]\eta + (1 - \phi)\sqrt{\rho}\sigma x(\Pi)\big]K = [\Pi - \sqrt{\rho}\phi\sigma x(\Pi) - (\Pi - \sqrt{\rho}\phi\sigma x(\Pi))]K$$

which vanishes, as expected. For another plausibility check, consider the behavior of taxes as the agency problem vanishes. Inspection of the equation for Π in Proposition 3.2 shows that $\Pi \approx \rho_S$ and $x(\Pi) \approx 0$. The above taxes are then approximately $\tau_\Pi = 1 - \phi \approx 1$ and $\tau_s \approx 0$, and the risk-free rate is approximately $r \approx \rho_S$. The associated fraction of the capital stock owned by the government and inherited by agents is then $\eta = \rho_S/\rho$, and $\nu = \rho_D/\rho$, respectively. Therefore, as agency frictions vanish, the interest rate approaches its complete-markets value, and taxes on savings and labor income vanish, as expected. The tax on profits approaches 100 percent, but net revenue collected is negligible because net business profits $\Pi - r$ also approach zero.

Now recall that Lemma 2.3 in the principal-agent (partial equilibrium) setting of Section 2 established that the wedge on the risk-free asset always exceeds that on the risky asset. It is important to note that these statements regarding relative magnitudes of *wedges* do not immediately imply analogous statements for the relative magnitudes of *taxes*, since the interest rate in the incomplete markets setting does not coincide with the interest rate given in the relaxed planner's problem. Indeed, inspection of the expressions found in Proposition 4.3 shows that neither the savings tax nor the profits tax everywhere exceeds the other. As noted in Section 3, the only intertemporal rate for the relaxed planner consistent with stationarity is the subjective rate of discount, as all other rates induce a trend in consumption across different cohorts. However, this introduction of an intertemporal price is only an expositional device to characterize the efficient allocation and does not represent the return on an asset available to any agent. To illustrate this point, we can use the expression for the interest rate found in Proposition 4.3 to show the following.

Corollary 4.4. *The interest rate in the stationary competitive equilibrium that decentralizes the efficient allocation is always lower than the subjective rate of discount.*

Proof. From the expression found in Proposition 4.3 we have $r \le \rho_S$ if and only if $\Pi - \sqrt{\rho}\sigma x(\Pi) \le \rho_S$. From the definition of x this will be assured as long as $[\Pi - \rho_S]/(\sqrt{\rho}\phi\sigma) \le x(\Pi)$ or

$$R(\Pi) \le \frac{1}{2R(\Pi)}\Big[1 - \sqrt{1 - 4R(\Pi)^2}\Big]$$

which is always true so long as $R(\Pi) \in (0, 1)$. $\qquad\square$

Lemma 2.3 shows that the principal always wishes to distort the agent's return on savings so that it is below the risk-free rate. However, Corollary 4.4 shows that the latter is in turn below the

rate available to the generational planner, so the sign of the savings tax is in general ambiguous. The following shows that the tax on savings may indeed assume either sign, reinforcing the point that in this general equilibrium environment, wedges are not identical to taxes.

Lemma 4.5. *The tax on savings is negative for sufficiently small (but positive) ϕ.*

Proof. For this proof write $\Pi(\phi)$ for the efficient marginal product of capital as a function of ϕ. From Proposition 4.3 the savings tax is negative if and only if

$$\sqrt{\rho}\sigma x(\Pi) > \Pi - \rho_S + \rho x(\Pi)^2. \tag{20}$$

The defining equation for Π implies $[\Pi/\alpha + (1/\alpha - 1)\delta]x(\Pi) = \sqrt{\rho}\phi\sigma$ and so simplification gives

$$\big(\rho\phi[\Pi/\alpha + (1/\alpha - 1)\delta] - \rho^2\phi^2\big)\sigma^2 \geq [\Pi - \rho_S][\Pi/\alpha + (1/\alpha - 1)\delta]^2.$$

We know $\Pi(\phi)$ solves $\left[1 - \sqrt{1 - 4R(\Pi)^2}\right][\Pi/\alpha + \delta(1/\alpha - 1)] = 2R(\Pi)\sqrt{\rho}\phi\sigma$, which reduces to

$$0 = \rho\phi^2\sigma^2[\Pi/\alpha + \delta(1/\alpha - 1)] - \rho\phi^2\sigma^2[\Pi - \rho_S] - [\Pi - \rho_S][\Pi/\alpha + \delta(1/\alpha - 1)]^2.$$

Using this and dividing through by ϕ, the inequality (20) becomes

$$\phi[(1 - \phi)[\Pi(\phi)/\alpha + (1/\alpha - 1)\delta] + \phi[\Pi(\phi) - \rho_S - \rho]] > 0$$

which is satisfied for all sufficiently small (positive) ϕ by the continuity of $\Pi(\phi)$ in ϕ. $\qquad\square$

In light of the simplicity of the expressions for taxes in Proposition 4.3, it is natural to wonder about the extent to which the result is sensitive to changes in the assumptions and modeling devices adopted in this paper. Regardless of the tax structure, the envelope theorem implies that the marginal utility of consumption coincides with the marginal utility of wealth, so $1 - \phi$ is the only value of the profits tax for which the incentive-constraint binds. This argument is simple and appears quite general: Proposition D.3 in the appendix shows that Proposition 4.3 extends to preferences with constant relative risk aversion. However, this immediately leads one to question why the results presented here differ from those of Albanesi (2007) and Shourideh (2012), who conduct similar optimal taxation analyses in environments with dynamic agency. The three differ models differ in a number of ways, and so it is difficult to conduct a point-by-point comparison of assumptions and conclusions. Nonetheless, one can isolate some salient differences and their role in affecting the findings.

In Albanesi (2007), the agency problem arises because entrepreneurs incur an additive cost of exerting effort, independently of the amount of capital or consumption delegated to them. This differs from the formulation in this paper and implies the above envelope argument is inapplicable, as the marginal benefit from deviating from the recommended action is no longer proportional to the marginal utility of consumption. The contrast with Shourideh (2012) is more subtle. Both Shourideh (2012) and the current paper model the agency problem as arising from the ability of

the agent to divert capital to private consumption, but differ in the timing of the uncertainty (continuous versus discrete-time), the demographic structure of the agents (perpetual youth versus two-period overlapping generations), and the distribution of uncertainty (Gaussian versus gamma-distributed shocks). Each of these differences appears to play a role. To see why the demographic structure is important, recall the perturbation argument from Section 2 that leads to the derivation of the inverse Euler equation. The hypothesized perturbation scaled up *current* consumption and delegated capital and scaled down *future* consumption and delegated capital in a way that left the agent indifferent. The calculated change in the principal's payoff then assumed that delegated capital was just as productive in each period. This is true in the current paper but not in the two-period overlapping generations model of Shourideh (2012), where the old agents are unable to produce.[11] The choice of modeling the uncertainty via a diffusion process is also important, as it allows us to restrict attention to infinitesimal diversions of the capital stock to private consumption. The above envelope argument would not hold if agents must divert capital in discrete amounts, as incentive-compatibility would then depend upon non-local properties of the utility and value functions.

One major simplifying assumption adopted in this paper thus far is the absence of any (ex-ante) heterogeneity in productivity. This decision is made primarily for tractability, as the incorpora-tion of unobservable ex-ante and ex-post heterogeneity leads to an extremely difficult problem. However, it is worth noting that the analysis of this paper extends easily to environments with a particular kind of ex-ante heterogeneity. Suppose for instance that agents are no longer capable of overseeing production and working simultaneously, and that there are some agents, workers, who are only capable of the latter. In this case the planner need not worry about "double deviations" (in which agents misreport type and subsequent deviation), since workers cannot pretend to be entrepreneurs and entrepreneurs who pretend to be workers are thereafter privy to no private infor-mation. Incentive-compatibility then simply requires consumption, and the capital delegated follows the prescriptions of the preceding analysis and that entrepreneurs be given enough utility to reveal their type. Virtually all of the preceding analysis carries over and so we obtain the following. The proof is omitted because it follows from logic identical to that given in Proposition 2.2, Proposition 3.2, and Proposition 4.3.

Proposition 4.6. *Suppose that a mass L_E of agents are entrepreneurs and that a mass $L_W :=$ $L - L_E$ are workers. The marginal product of capital in the stationary constrained efficient allocation is the solution to the equation*

$$c(\Pi) + L_W/L_E = \frac{1}{\alpha}[\Pi + (1 - \alpha)\delta]k(\Pi) \tag{21}$$

where $c(\Pi)$ and $k(\Pi)$ are given in Proposition 2.2, provided the no-absconding constraint is a strict inequality for this Π. The efficient allocation can be decentralized with occupation-specific linear

---

[11]Appendix A.1 elaborates upon this point by recasting the current model in discrete-time and allowing the life-cycle structure of productivity to be arbitrary.

taxes on savings, and taxes on the reported profits of entrepreneurs. The expressions for taxes on entrepreneurs are identical to those given in Proposition 4.3 (although Π differs), while the taxes on worker's savings are given by

$$1 - \tau_{sW} = \frac{\rho}{\Pi - \sqrt{\rho}x(\Pi) + \rho_D}. \tag{22}$$

Similar results carry over to the case of arbitrary *observable* ex-ante heterogeneity. Reasoning similar to that in Proposition 4.6 then implies that taxes on profits need not be progressive, as all agents are subject to the flat rate $\tau_\Pi = 1 - \phi$. More productive agents will devote a greater fraction of their wealth to the risky asset and will therefore pay more in profit taxes. However, the envelope argument preceding Proposition 4.3 remains applicable and so the profits tax is independent of type. Consequently, in order for the drift in consumption to be zero (and hence the inverse Euler equation to be satisfied), the savings tax of the more productive agent must be higher. Although all agents face type-dependent linear taxes, in some sense the model of this paper implies progressivity of savings taxes, as richer agents will (on average) face higher taxes.

Figure 2 plots savings taxes for workers and entrepreneurs as a function of the agency friction for the example parameters $(\delta, L_E/L_W, \alpha, \sigma^2, \rho, \rho_D) = (0.058, 0.88, 0.33, 0.3, 0.145, 0.022)$. The exponent α, discount factor ρ, depreciation δ and fraction of entrepreneurs $L_E/(L_E + L_W)$ are taken from Cagetti and De Nardi (2006), while ρ_D is chosen such that the average life-span is 45 years (interpreted as the working life-span). The volatility term σ^2 is the average of the two values considered by Angeletos (2007). As implied by the expressions in Proposition 4.6, the tax on entrepreneurs' savings is everywhere above that of the tax on workers' savings and all agents face a subsidy on their savings when agency frictions are small.

Finally, note that the expression for tail inequality in Corollary 3.3 continues to hold in the presence of both workers and entrepreneurs, with the sole effect of the presence of workers being to alter the efficient level of Π. Figure 3 depicts the Pareto exponent for the same parameters used in Figure 2. Further, using the expression for the efficient marginal product of capital given in Proposition 4.6 may derive some simple comparative statics illustrating the effect of changes in technology and ex-ante inequality (in type) on the efficient degree of consumption inequality and optimal taxes.

Corollary 4.7. *The marginal product of capital Π and the Pareto exponent β_- of the upper tail in the stationary distribution are both increasing in the degree of the agency problem $\phi\sigma$, the number of workers per entrepreneur L_W/L_E, and the factor share α.*

The analysis in this paper has been of the theoretical determinants of capital taxation in the presence of dynamic agency. The model in its current form is too stylized to immediately inform policy. Therefore, I have emphasized the novel *qualitative* features of the analysis, such as the invariance of profits taxes to technology and demographics and the fact that savings taxes are increasing in productivity. However, the back-of-the-envelope calculations in this section do suggest
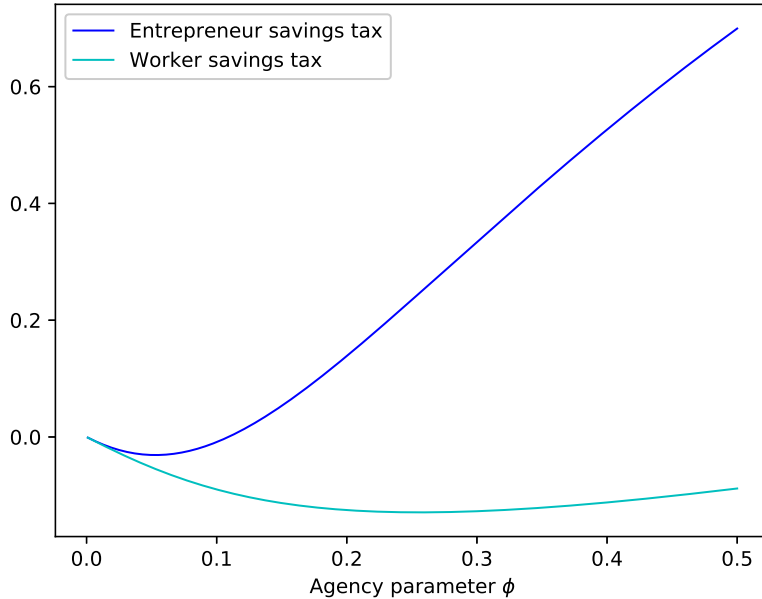
Figure 2: Taxes as function of agency friction

agency frictions have a non-negligible impact on both the efficient degree of wealth inequality and the optimal choice of taxes. Indeed, the savings taxes on entrepreneurs in Figure 2 are well above their empirical counterparts for most agency frictions, and the range of Pareto exponents depicted in Figure 3 encompasses the value $\beta_- \approx -1.54$ of the US wealth distribution estimated by Gabaix et al. (2016) from the 2010 wave of the Survey of Consumer Finances.

5 Conclusion

The United States tax code currently levies different taxes on different forms of capital income. This paper has provided a model in which the desirability of this differential asset taxation emerges naturally from the presence of agency frictions. I consider a dynamic economy with physical capital, entrepreneurs, workers, and endogenous firm formation, in which the returns on capital are privately observed and may be diverted to consumption. In this setting, I assume a generalized utilitarian objective across generations, and allow the government to choose any allocation satisfying the incentive constraints arising from this asymmetry of information. In spite of the great deal of freedom this grants the government, I obtain a simple and sharp characterization of the efficient allocations and the taxes necessary for their implementation.

Specifically, when agents are assumed able to trade a risk-free bond is zero net supply, the efficient allocations may be implemented with linear, time-independent taxes on (reported) profits
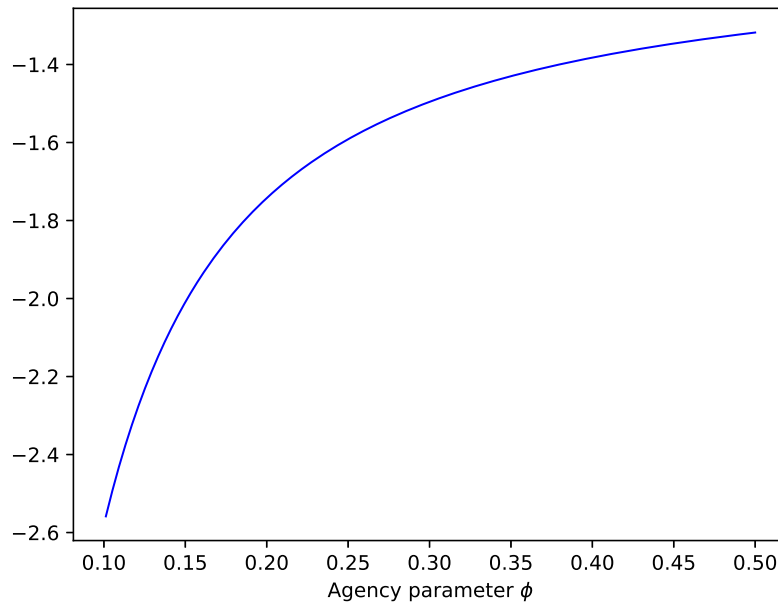
Figure 3: Pareto parameter as function of agency friction

and savings. The profits tax is the highest value consistent with incentive compatibility, and depends solely on the severity of the agency friction. In particular, it is independent of ex-ante ability and is common to all agents. In contrast, the tax on savings is always higher for more productive agents, but in general may assume either sign. To isolate the role played by agency frictions I have assumed that ability is permanent and observable. An interesting direction for future work would be to explore the extent to which the findings of this paper extend to settings in which ability evolves over time and in which there is a non-trivial margin for entry.

References

Albanesi, Stefania (2007). "Optimal taxation of entrepreneurial capital with private information." Discussion Paper 0607-11, Columbia University Department of Economics. doi:10.3386/w12419.

Angeletos, George-Marios (2007). "Uninsured idiosyncratic investment risk and aggregate saving." Review of Economic Dynamics, 10(1), pp. 1–30. doi:10.1016/j.red.2006.11.001.

Atkeson, Andrew and Robert E. Lucas (1992). "On efficient distribution with private information." The Review of Economic Studies, 59(3), pp. 427–453. doi:10.2307/2297858.

Auerbach, Alan J. (2002). "Chapter 19 - taxation and corporate financial policy." In Alan J. Auerbach and Martin Feldstein, editors, Handbook of Public Economics, volume 3, pp. 1251–1292. Elsevier. doi:10.1016/S1573-4420(02)80023-4.

Benhabib, Jess and Alberto Bisin (2018). "Skewed wealth distributions: Theory and empirics." Journal of Economic Literature, 56(4), pp. 1261–1291. doi:10.1257/jel.20161390.

Cagetti, Marco and Mariacristina De Nardi (2006). "Entrepreneurship, frictions, and wealth." Journal of Political Economy, 114(5), pp. 835–870. doi:10.1086/508032.

Chari, V.V. and Patrick J. Kehoe (1999). "Chapter 26 - optimal fiscal and monetary policy." In Handbook of Macroeconomics, pp. 1671–1745. Elsevier. doi:10.1016/s1574-0048(99)10039-9.

Di Tella, Sebastian and Yuliy Sannikov (2016). "Optimal asset management with hidden savings." Research Papers 3429, Stanford University. URL https://ideas.repec.org/p/ecl/stabus/3429.html.

Domar, Evsey D. and Richard A. Musgrave (1944). "Proportional income taxation and risk-taking." The Quarterly Journal of Economics, 58(3), pp. 388–422. doi:10.2307/1882847.

Evans, David (). "Optimal taxation with persistent idiosyncratic investment risk." mimeo. URL http://economics.uoregon.edu/wp-content/uploads/sites/4/2015/07/Evans_David.pdf.

Farhi, Emmanuel and Iván Werning (2007). "Inequality and social discounting." Journal of Political Economy, 115(3), pp. 365–402. doi:10.1086/518741.

Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll (2016). "The dynamics of inequality." Econometrica, 84(6), pp. 2071–2111. doi:10.3982/ecta13569.

Golosov, Mikhail, Narayana Kocherlakota, and Aleh Tsyvinski (2003). "Optimal indirect and capital taxation." Review of Economic Studies, 70(3), pp. 569–587. doi:10.1111/1467-937x.00256.

Gordon, Roger H. and Joel Slemrod (1988). "Do we collect any revenue from taxing capital income?" Tax Policy and the Economy, 2, pp. 89–130. doi:10.1086/tpe.2.20061774.

Mirrlees, J. A. (1971). "An exploration in the theory of optimum income taxation." The Review of Economic Studies, 38(2), pp. 175–208. doi:10.2307/2296779.

Panousi, Vasia and Catarina Reis (2017). "A unified framework for optimal taxation with undiversifiable risk." Working Paper 64, ADEMU. URL http://hdl.handle.net/10230/32771.

Phelan, Christopher (1994). "Incentives and aggregate shocks." The Review of Economic Studies, 61(4), pp. 681–700. doi:10.2307/2297914.

Phelan, Thomas (2019). "Efficient wealth inequality and differential asset taxation with dynamic agency." 2019 Meeting Papers 1350, Society for Economic Dynamics. URL https://ideas.repec.org/p/red/sed019/1350.html.

Prescott, Edward C. and Robert M. Townsend (1984). "Pareto optima and competitive equilibria with adverse selection and moral hazard." Econometrica, 52(1), pp. 21–45. doi:10.2307/1911459.

Rogerson, William P. (1985). "Repeated moral hazard." Econometrica, 53(1), pp. 69–76. doi:10.2307/1911724.

Sannikov, Yuliy (2008). "A continuous-time version of the principal–agent problem." Review of Economic Studies, 75(3), pp. 957–984. doi:10.1111/j.1467-937x.2008.00486.x.

Shourideh, Ali (2012). "Optimal taxation of wealthy individuals." Working paper, Wharton School, University of Pennsylvania. URL https://faculty.wharton.upenn.edu/wp-content/uploads/2013/01/CaptaxEnt.pdf.

A Discrete-time formulation

The purpose of this section is to outline a discrete-time environment that approximates the continuous-time models given in the main text. It is intended to aid the reader and also allow clearer comparison with existing discrete-time environments with private information.

Time is indefinite and discrete, assuming values in the set $\{\Delta, 2\Delta, 3\Delta, \dots\}$. The economy consists of a single risk-averse agent and a risk-neutral principal, both of whom live forever. The preferences of the agent over stochastic sequences of consumption $c := (\Delta c_n)_{n=0}^{\infty}$ are represented by the function

$$U^A\left((\Delta c_n)_{n=0}^{\infty}; \Delta\right) = U^A(c; \Delta) := \left(1 - e^{-\Delta\rho}\right) \sum_{n=0}^{\infty} e^{-n\Delta\rho} \mathbb{E}[\ln c_n]. \tag{23}$$

The appearance of Δc_n rather than c_n in (23) is simply a normalization.[12] The principal possesses a constant-returns-to-scale technology that only the agent has the ability to operate. At each time $n\Delta$ the principal chooses how much physical capital will be installed in the technology for the interval $[n\Delta, (n+1)\Delta]$. If K_n is the amount of capital installed at time $n\Delta$ then the amount $\left[\Delta\Pi + \sqrt{\Delta}x_n\right]K_n$ is produced at time $(n+1)\Delta$, where $(x_n)_{n=0}^{\infty}$ is an i.i.d. sequence of mean zero random variables assuming $\pm\overline{x}$ with probability $1/2$. The capital stock is subject to depreciation, with the fraction of the stock K_n remaining at time $(n+1)\Delta$ equal to $e^{-\Delta\delta}$. At time $(n+1)\Delta$ the principal also chooses additional investment I_{n+1} and so the total amount of installed capital at time $(n+1)\Delta$ is

$$K_{n+1} = I_{n+1} + e^{-\Delta\delta}K_n.$$

The present discounted value of output minus investment is then

$$\sum_{n=0}^{\infty} e^{-(n+1)\Delta r}\mathbb{E}\left[\left[\Delta\Pi + \sqrt{\Delta}x_n\right]K_n - I_{n+1}\right] = \sum_{n=0}^{\infty} e^{-(n+1)\Delta r}\mathbb{E}\left[\left[\Delta\Pi + \sqrt{\Delta}x_n\right]K_n + e^{-\Delta\delta}K_n - K_{n+1}\right]$$

$$= \sum_{n=1}^{\infty} e^{-(n+1)\Delta r}\mathbb{E}\left[\left(\Delta\Pi + \sqrt{\Delta}x_n + e^{-\Delta\delta} - e^{\Delta r}\right)K_n\right]$$

where I used $K_0 = 0$. Notice that the nth term in the above summand satisfies

$$\left(\Delta(\Pi + x_n) + e^{-\Delta\delta} - e^{\Delta r}\right)K_n \approx \Delta(\Pi - \delta - r)K_n + \sqrt{\Delta}x_n K_n. \tag{24}$$

---

[12] The cost of consuming Δc every period when the discount rate is $e^{-\Delta r}$ is $\sum_{n=0}^{\infty} e^{-n\Delta r}\Delta c = \Delta c/[1 - e^{-\Delta r}]$, which tends to c/r as $\Delta \to 0$. The utility from this consumption plan is $\left(1 - e^{-\Delta\rho}\right) \sum_{n=0}^{\infty} e^{-n\Delta\rho} \ln c = \ln c$, so the normalization (23) simply ensures that utility is bounded as $\Delta \to 0$ whenever the present discounted value of consumption is bounded.

The principal wishes to maximize the expected present discounted value of output minus consumption given to the agent. Their preferences over stochastic sequences of capital delegation and consumption $(K_n, c_n)_{n=0}^{\infty}$ are therefore represented by the function

$$U^P(K, c; \Delta) := \sum_{n=1}^{\infty} e^{-(n+1)\Delta r} \mathbb{E}\big[\big(\Delta(\Pi + x_n) + e^{-\Delta\delta} - e^{\Delta r}\big)K_n - \Delta C_n\big].$$

Using (24) the objective of the planner is approximately

$$U^P(K, c; \Delta) \approx \sum_{n=1}^{\infty} \Delta e^{-(n+1)\Delta r} \mathbb{E}[(\Pi - r)K_n - C_n] \approx \int_0^{\infty} e^{-rt}[(\Pi - r)K_t - C_t]dt$$

for the principal's objective, and

$$U^A(c; \Delta) = \big(1 - e^{-\Delta\rho}\big) \sum_{n=0}^{\infty} e^{-n\Delta\rho} \mathbb{E}[\ln c_n] \approx \rho \int_0^{\infty} e^{-\rho t} \mathbb{E}[\ln c]dt$$

for the agent's objective. It follows that the objectives of both the planner and the agent in the main text may be interpreted as limits of their corresponding objectives in this environment.

I will assume that delegated capital is publicly observable but that both output and consumption are privately observable by the agent. Consumption and delegated capital must therefore be functions only of the reported output shocks $(x_n)_{n=0}^{\infty}$. For each $n \geq 0$, write $x^n := (x_0, \ldots, x_n)$ for the history of realizations of the output shocks up to and including date n and denote by $\mathcal{X}_n$ the set of all such histories. I will restrict attention to allocations in which the principal recommends that the agent not divert any delegated capital to consumption. This is obviously without loss when characterizing efficient allocations.

Definition A.1 (Allocations and strategies). An allocation consists of a stochastic sequence of consumption and capital delegation $(K, c) = (K_n, c_n)_{n=0}^{\infty}$ where for each $n \geq 0$ we have $K_{n+1}, c_{n+1} : \mathcal{X}_n \to \mathbb{R}_+$. A strategy of the agent is a sequence of reports $X = (X_n)_{n=1}^{\infty}$ where for each $n \geq 0$ we have $X_{n+1} : \mathcal{X}_n \to \mathbb{R}_+$.

The utility of an agent confronted with an allocation (K, c) when adhering to a strategy X is given by

$$U^A(c, K; X) := \big(1 - e^{-\Delta\rho}\big) \sum_{n=0}^{\infty} e^{-n\Delta\rho} \mathbb{E}\Big[\ln\Big(c_n + \sqrt{\Delta}(X_n - x_n)K_n\Big)\Big].$$

Further, associated with each allocation (K, c) define continuation utility $W := (W_n)_{n=0}^{\infty}$ by

$$W_n(c) := \big(1 - e^{-\Delta\rho}\big) \sum_{N=n}^{\infty} e^{-n\Delta\rho} \mathbb{E}[\ln c_N].$$

As in the model of the main text, I will assume that the agent may abscond with a fixed fraction of assets, which implies that $K_n \leq \omega \exp W_n(c)$ for all n almost surely for some exogenous ω.

Definition A.2. An allocation (K, c) is incentive-compatible if $U^A(c, K; 0) \geq U^A(c, K; X)$ for all agent strategies X and if $K_n \leq \omega \exp W_n(c)$ almost surely for all $n \geq 1$. The set of all incentive-compatible allocations will be denoted $\mathcal{A}^{IC}$.

Since the output shocks assume only two values it is without loss to suppose that the reporting strategies assume only two values (either report the truth or report the other possible shock), since all other deviations will be detected immediately. As is well-known it suffices to impose temporary incentive-compatibility

constraints that dissuade one-shot deviations. The principal then announces two possible future values $W^{\pm}$ for promised utility. For any allocation $C = (C_n)_{n=0}^{\infty}$ the temporary incentive-compatibility constraints are then

$$\left(1 - e^{-\Delta\rho}\right)\ln C + e^{-\Delta\rho}W^{+} \geq \left(1 - e^{-\Delta\rho}\right)\ln\left(C + 2\sqrt{\Delta\bar{x}}K\right) + e^{-\Delta\rho}W^{-}$$
$$\left(1 - e^{-\Delta\rho}\right)\ln C + e^{-\Delta\rho}W^{-} \geq \left(1 - e^{-\Delta\rho}\right)\ln\left(C + 2\sqrt{\Delta\bar{x}}K\right) + e^{-\Delta\rho}W^{+}. \tag{25}$$

The first constraint in (25) is truth-telling for the high-shock, while the second is truth-telling for the low shock. Rearrangement of (25) then gives

$$e^{-\Delta\rho}\left[W^{+} - W^{-}\right] \geq \left(1 - e^{-\Delta\rho}\right)\left[\ln\left(C + 2\sqrt{\Delta\bar{x}}K\right) - \ln C\right]$$
$$\left(1 - e^{-\Delta\rho}\right)\left[\ln C - \ln\left(C - 2\sqrt{\Delta\bar{x}}K\right)\right] \geq e^{-\Delta\rho}\left[W^{+} - W^{-}\right]$$

It is easy to see that the first constraint must bind and that by the concavity of the natural logariprop the second is therefore redundant. Promise-keeping and incentive-compatibility then reduce to the following pair of equations

$$W = \left(1 - e^{-\Delta\rho}\right)\ln C + e^{-\Delta\rho}\left[W^{-} + W^{+}\right]/2$$
$$W^{+} = W^{-} + \left(e^{\Delta\rho} - 1\right)\left[\ln\left(C + 2\sqrt{\Delta\bar{x}}K\right) - \ln C\right]. \tag{26}$$

Also note that simplification of (26) gives

$$W^{\pm} \approx e^{\Delta\rho}W - \left(e^{\Delta\rho} - 1\right)\ln C \pm \frac{1}{2}\left(e^{\Delta\rho} - 1\right)\left[\ln\left(C + 2\sqrt{\Delta\bar{x}}K\right) - \ln C\right]. \tag{27}$$

Using the fact that $\ln\left(C + 2\sqrt{\Delta\bar{x}}K\right) - \ln C \sim 2\sqrt{\Delta\bar{x}}K/C$ as $\Delta \to 0$, the expressions in (27) may be written

$$\frac{W^{\pm} - W}{\Delta} \approx \left(\frac{e^{\Delta\rho} - 1}{\Delta}\right)\left(W - \ln C \pm \sqrt{\Delta\bar{x}}K/C\right).$$

which in turn may be written as

$$dW_t \approx \rho(W - \ln C)\Delta + \rho\sqrt{\Delta\bar{x}}[K/C]X_t$$

where X_t has mean zero, is independent over time and assumes the values ± 1, which approximates the law of motion of promised utility in the continuous-time model.

A.1 Relation with other agency models and the inverse Euler equation

In order to relate the model of the main text with others in the literature, this section will outline a slightly more general model in which productivity and discount factors may be time-dependent. To this end suppose that the preferences of the consumer and the principal over sequences $(\Delta c_n)_{n=0}^{\infty}$ are now given by

$$U^A((\Delta c_n)_{n=0}^{\infty}; \Delta) = (1 - \bar{\beta})\sum_{n=0}^{\infty}\beta_n\mathbb{E}[\ln c_n]$$
$$U^P(K, c; \Delta) = \Delta\sum_{n=1}^{\infty}e^{-(n+1)\Delta r}\mathbb{E}[(\Pi_n - \delta - r)K_n - C_n] \tag{28}$$

for some sequences $(\beta_n)_{n=0}^{\infty}$ and $(\Pi_n)_{n=0}^{\infty}$, where I have abbreviated $\bar{\beta} = 1 - \left(\sum_{n=0}^{\infty}\beta_n\right)^{-1}$. For convenience I will normalize $\beta_0 = 1$. For arbitrary $n \geq 0$ define the continuation utility from period n onwards by

$$W_n := (1 - \bar{\beta})\sum_{N=n}^{\infty}(\beta_N/\beta_n)\mathbb{E}[\ln c_N]. \tag{29}$$

Note that under the specification (29) we have a recursive relation involving continuation utilities

$$W_n = (1 - \overline{\beta}) \sum_{N=n}^{\infty} (\beta_N/\beta_n) \mathbb{E}[\ln c_N] = (1 - \overline{\beta}) \mathbb{E}[\ln c_n] + (1 - \overline{\beta}) \sum_{N=n+1}^{\infty} (\beta_N/\beta_n) \mathbb{E}[\ln c_N]$$

$$= (1 - \overline{\beta}) \mathbb{E}[\ln c_n] + (\beta_{n+1}/\beta_n)(1 - \overline{\beta}) \sum_{N=n+1}^{\infty} (\beta_N/\beta_{n+1}) \mathbb{E}[\ln c_N]$$

$$= (1 - \overline{\beta}) \mathbb{E}[\ln c_n] + (\beta_{n+1}/\beta_n) \mathbb{E}[W_{n+1}].$$

Following standard arguments in the dynamic contracting literature, it suffices to consider temporary incentive-constraints. For any allocation $C = (C_n)_{n=0}^{\infty}$ the temporary incentive-constraints are then

$$\begin{aligned}
(1 - \overline{\beta}) \ln C_n + (\beta_{n+1}/\beta_n) W_{n+1}^+(C) &\geq (1 - \overline{\beta}) \ln(C_n + 2\overline{x} K_n) + (\beta_{n+1}/\beta_n) W_{n+1}^-(C) \\
(1 - \overline{\beta}) \ln C_n + (\beta_{n+1}/\beta_n) W_{n+1}^-(C) &\geq (1 - \overline{\beta}) \ln(C_n - 2\overline{x} K_n) + (\beta_{n+1}/\beta_n) W_{n+1}^+(C).
\end{aligned} \tag{30}$$

It is easy to see that only the first constraint in (30) will bind and so the problem of the principal is then the following.

Definition A.3. The principal's problem is given by

$$V(W) = \Delta \max_{C,K} \sum_{n=0}^{\infty} e^{-\Delta n r} \mathbb{E}[(\Pi_n - \delta - r) K_n - C_n]$$

$$W = (1 - \overline{\beta}) \sum_{n=0}^{\infty} \beta_n \mathbb{E}[\ln C_n]$$

$$(\beta_{n+1}/\beta_n) W_{n+1}^+(C) \geq (1 - \overline{\beta}) \ln(1 + 2\overline{x} K_n/C_n) + (\beta_{n+1}/\beta_n) W_{n+1}^-(C)$$

where $W_{n+1}(C)$ denotes the continuation utility associated with the sequence $C = (C_n)_{n=0}^{\infty}$.

The promise-keeping and incentive-compatibility constraints are given by

$$(1 - \overline{\beta}) \ln C_n + (\beta_{n+1}/\beta_n) W_{n+1}^+ = (1 - \overline{\beta}) \ln(C_n + 2\overline{x} K_n) + (\beta_{n+1}/\beta_n) W_{n+1}^-$$

$$W_n = (1 - \overline{\beta}) \ln C + \frac{1}{2} (\beta_{n+1}/\beta_n) [W_{n+1}^- + W_{n+1}^+]$$

Simplification gives

$$W_{n+1}^+ = W_{n+1}^- + (\beta_{n+1}/\beta_n)^{-1} (1 - \overline{\beta}) \ln(1 + 2\overline{x} K_n/C_n)$$

$$W_n = (1 - \overline{\beta}) \ln C_n + \frac{1}{2} (\beta_{n+1}/\beta_n) [W_{n+1}^- + W_{n+1}^+]$$

$$= (1 - \overline{\beta}) \ln C_n + (\beta_{n+1}/\beta_n) W_{n+1}^- + \frac{1}{2} (1 - \overline{\beta}) \ln(1 + 2\overline{x} K_n/C_n)$$

$$= (1 - \overline{\beta}) \ln C_n + (\beta_{n+1}/\beta_n) W_{n+1}^+ - \frac{1}{2} (1 - \overline{\beta}) \ln(1 + 2\overline{x} K_n/C_n).$$

Solving for $W_{n+1}^{\pm}$ then gives

$$W_{n+1}^{\pm} = (\beta_n/\beta_{n+1}) W_n + (\beta_n/\beta_{n+1})(1 - \overline{\beta}) \left(-\ln C_n \pm \frac{1}{2} \ln(1 + 2\overline{x} K_n/C_n) \right). \tag{31}$$

Note that if $C_n = \underline{c}_n \exp W_n$ and $K_n = \underline{k}_n \exp W_n$ then (31) implies

$$W_{n+1}^{\pm} = (\beta_n/\beta_{n+1}) \left[\overline{\beta} W_n + (1 - \overline{\beta}) \left(-\ln \underline{c}_n \pm \frac{1}{2} \ln(1 + 2\overline{x} \underline{k}_n/\underline{c}_n) \right) \right]. \tag{32}$$

32

Scaling $\underline{c}_n$ by $\exp u$ changes W_{n+1} by $-u(\beta_n/\beta_{n+1})(1-\overline{\beta})$, which implies

$$\Delta W_{n+2} = (\beta_{n+1}/\beta_{n+2})\overline{\beta}\Delta W_{n+1} = -(\beta_n/\beta_{n+2})\overline{\beta}(1-\overline{\beta})u$$

$$\Delta W_{n+3} = (\beta_{n+2}/\beta_{n+3})\overline{\beta}\Delta W_{n+2} = -(\beta_n/\beta_{n+3})\overline{\beta}^2(1-\overline{\beta})u$$

$$\vdots$$

$$\Delta W_{n+k} = -(\beta_n/\beta_{n+k})\overline{\beta}^{k-1}(1-\overline{\beta})u.$$

This discussion is summarized in the following.

Lemma A.1 (Homogeneity of the principal's problem). *The value function of the principal is of the form* $V(W) = -\Omega \exp W$ *for some* $\Omega > 0$ *and the policy functions for consumption and capital are of the form*

$$K_n = \underline{k}_n \exp W_n \quad C_n = \underline{c}_n \exp W_n$$

for some sequence $(\underline{k}, \underline{c}) := (\underline{k}_n, \underline{c}_n)_{n=0}^{\infty}$, *while the policy functions for promised utility are*

$$W_{n+1}^{\pm} = \frac{\overline{\beta}W_n}{\beta_{n+1}/\beta_n} - \left(\frac{1-\overline{\beta}}{\beta_{n+1}/\beta_n}\right)\ln \underline{c}_n \pm \frac{1}{2}\left(\frac{1-\overline{\beta}}{\beta_{n+1}/\beta_n}\right)\ln(1 + 2\Delta\overline{x}\underline{k}_n/\underline{c}_n). \tag{33}$$

Now consider two successive periods, n and $n+1$, and define the following perturbation: scale $\underline{c}_n$ and $\underline{k}_n$ by $\exp u$ and $\underline{c}_{n+1}$ and $\underline{k}_{n+1}$ by $\exp(-u\overline{\beta}\beta_n/\beta_{n+1})$, for some arbitrary u. To motivate this perturbation, note that by (33) in Lemma A.1, if we scale $(\underline{c}_n, \underline{k}_n)$ by $\exp u$ and $(\underline{c}_{n+1}, \underline{k}_{n+1})$ by $\exp \overline{u}$ then the change in W_{n+1} will be $\Delta W_{n+1} = -(\beta_n/\beta_{n+1})(1-\overline{\beta})u$ and so the change in W_{n+2} will be $\Delta W_{n+2} = (\beta_{n+1}/\beta_{n+2})[\overline{\beta}\Delta W_{n+1} - (1-\overline{\beta})\overline{u}]$. It follows that $\Delta W_{n+2} = 0$ if and only if $\overline{u} = \overline{\beta}\Delta W_{n+1}/(1-\overline{\beta}) = -(\overline{\beta}\beta_n/\beta_{n+1})u$. This implies that the above perturbation only affects quantities in periods n and $n+1$, with all other periods unaffected. The associated change in the utility from consumption at date $t+1$ is then

$$-\left(\frac{1-\overline{\beta}}{\beta_{n+1}/\beta_n}\right)u + \overline{u} = -\left(\frac{1-\overline{\beta}}{\beta_{n+1}/\beta_n}\right)u - \left(\frac{\overline{\beta}}{\beta_{n+1}/\beta_n}\right)u = -(\beta_n/\beta_{n+1})u.$$

The payoff to the principal from periods n and $n+1$ from this perturbation is

$$F(u) := ([\Pi_n - \delta - r]\underline{k}_n - \underline{c}_n)\exp(W + u) + e^{-\Delta r}([\Pi_{n+1} - \delta - r]\underline{k}_{n+1} - \underline{c}_{n+1})\mathbb{E}[\exp(W' - u\beta_n/\beta_{n+1})].$$

The necessary condition $F'(0) = 0$ then becomes

$$(\beta_{n+1}/\beta_n)e^{\Delta r}\underline{c}_n \exp W = \left(\frac{[\Pi_{n+1} - \delta - r]\underline{k}_{n+1} - \underline{c}_{n+1}}{[\Pi_n - \delta - r]\underline{k}_n - \underline{c}_n}\right)\underline{c}_n\mathbb{E}[\exp W']. \tag{34}$$

If $u(x) := \ln x$ then $1/u'(x) = x$ and so the inverse Euler equation in this case is

$$(\beta_{n+1}/\beta_n)e^{\Delta r}\underline{c}_n \exp W = \underline{c}_{n+1}\mathbb{E}[\exp W']. \tag{35}$$

Expression (36) clarifies the difference between the results of this paper and that in Shourideh (2012). The timing in this agency problem may be summarized as follows:

1. The agent begins period n with utility (or outside option) W_n.

2. The principal assigns K_n units of capital and C_n units of consumption to the agent.

3. Output produced within period is $\Delta\Pi K_n$.

4. Fraction $1 - e^{-\Delta\delta} + \sqrt{\Delta}x_n K_n$ of capital depreciates during the period.

5. Agent reports x_n and consumes C_n plus any diverted capital.

6. Principal assigns utility W_{n+1} for next period depending upon reported level of output.

In Shourideh (2012) agents live for two periods and the timing is as follows:

1. Principal assigns K_n units of capital and C_n units of consumption to the agent.

2. Agent consumes C_n plus any capital diverted.

3. Output tomorrow is publicly observed and equal to $(\Delta\Pi + \sqrt{\Delta}x_n)\underline{k}_n$ where $\underline{k}_n$ is amount of capital actually invested and x_n is random and exogenous.

4. Principal assigns consumption in second period. Agent eats and the world ends.

The above agency problems are obviously similar and so it is instructive to outline why the associated intertemporal distortions differ. Since Shourideh (2012) adopts a different specification of shocks, it is difficult to directly compare the two models. However, if one adopts the two-period life-cycle structure of Shourideh (2012) (and the above timing) but assumes that shocks take only two values with equal probability, then the resulting model coincides with that given in this section with discount rates and productivities given by $\beta_0 = e^{-\Delta\rho}, \beta_n = 0, \Pi_0 = \Pi$ and $\Pi_n = 0$ for all $n \geq 1$. In contrast, the model of this paper corresponds to that given in the previous section with $\beta_n = e^{-\Delta n\rho}$ and $\Pi_n = \Pi$ for all $n \geq 1$. Combining (34) and (35) shows that the inverse Euler equation holds if and only if

$$\frac{\underline{c}_n}{\underline{c}_{n+1}} = \frac{[\Pi_n - \delta - r]\underline{k}_n - \underline{c}_n}{[\Pi_{n+1} - \delta - r]\underline{k}_{n+1} - \underline{c}_{n+1}}. \tag{36}$$

I hope that equation (36) clarifies things for the reader. In my infinite-horizon setting, we have $\beta_n = \beta^n$ and $\Pi_n \equiv \Pi$ and hence $\underline{k}_{n+1} = \underline{k}_n$ and $\underline{c}_{n+1} = \underline{c}_n$ for all $n \geq 0$. The equality (36) then obviously holds. In Shourideh (2012), the agent lives for two periods (say, $t = 0, 1$) and so $\underline{k}_1 = 0 \neq \underline{k}_0$. The right-hand side of (36) for $n = 0$ then becomes

$$\frac{[\Pi_0 - \delta - r]\underline{k}_n - \underline{c}_n}{[\Pi_{n+1} - \delta - r]\underline{k}_{n+1} - \underline{c}_{n+1}} = \frac{\underline{c}_0 - [\Pi_0 - \delta - r]\underline{k}_0}{\underline{c}_1}$$

which is strictly less than the left-hand side of (36).

A.2   Decentralized economy

In this section I outline a discrete-time environment that approximates the incomplete-markets model of Section 4.2. Suppose again that time assumes the values $\Delta, 2\Delta, \ldots$ for some $\Delta > 0$. At time $n\Delta$ the agent receives interest and principal from savings, receives after-tax profits from his private business, and observes the amount of the capital stock that depreciated since the time $(n-1)\Delta$. He then discovers whether he will live or die. If he lives he receives an annuity from the life insurance company, and if he dies, all his wealth is transferred to a life insurance company.

Morning of the nth period: wake up with a_n potatoes, eat Δc_n and deposit remainder in a bank that promises to pay back $(1 + \Delta r)(a_n - \Delta c_n)$ tomorrow. While at the bank stop by the local life insurance company and write the following contract: tomorrow if I am still alive you will transfer $\Delta\rho_D(a_n - \Delta c_n)$ potatoes to me; otherwise you may take possession of all my wealth. Afternoon: I arrive at work, rent k units of potatoes from the bank invest them in my backyard and hire workers at wage w, finding time during the day to supply labor and so earning a wage myself. Evening: the amount of output produced by my

workers and net of labor payments is $\Delta\Pi k$, and the amount of capital that depreciates is $(\delta + \sqrt{\Delta}\sigma X_n)k$, where X_n is i.i.d. and stochastic. Next period wealth is therefore

$$a_{n+1} = (1 + \Delta(r + \rho_D))(a_n - \Delta c_n) + \Delta w + \Delta[\Pi - r]k_n + \sqrt{\Delta}\sigma X_n k_n$$

and so the change in wealth satisfies $a_{n+1} - a_n = -\Delta^2(r + \rho_D)c_n + [(r + \rho_D)a_n - c_n + w + [\Pi - r]k_n]\Delta + \sqrt{\Delta}\sigma X_n k_n$. As $\Delta \to 0$ this becomes equivalent to the law given in Definition 4.3. Note that the insurance company pays $\Delta\rho_D(a_n - \Delta c_n)$ with probability $1 - \Delta\rho_D$ and receives a_{n+1} with probability $\Delta\rho_D$. Expected profits are therefore

$$
\begin{aligned}
-\Delta\rho_D(a_n - \Delta c_n)(1 - \Delta\rho_D) - \Delta\rho_D\mathbb{E}[a_{n+1}] &= \Delta\rho_D[(a_n - \Delta c_n)(1 - \Delta\rho_D) - a_n \\
&\quad - [-\Delta^2(r + \rho_D)c_n + [(r + \rho_D)a_n - c_n + w + [\Pi - r]k_n]\Delta]] \\
&= \Delta\rho_D\Delta[-c_n(1 - \Delta\rho_D) - \rho_D a_n + \Delta(r + \rho_D)c_n - (r + \rho_D)a_n + c_n - w - [\Pi - r]k_n] \\
&= \Delta\rho_D\Delta[c_n\Delta\rho_D - \rho_D a_n + \Delta(r + \rho_D)c_n - (r + \rho_D)a_n - w - [\Pi - r]k_n]
\end{aligned}
$$

which is $o(\Delta)$ as $\Delta \to 0$. Therefore, the profit over any fixed interval (which behaves like the above multiplied by $1/\Delta$) also tends to zero with Δ. We have the standard labor market-clearing and goods market-clearing. The bank makes zero profits, so the return promised to depositors is equal to the return demanded from entrepreneurs. The total amount of potatoes the agents deposit at the bank in the morning must equal the total amount rented by entrepreneurs. Therefore, market-clearing in the discrete-time environment is

$$\int_0^\infty (a - \Delta c(a))g(da) = \int_0^\infty k(a)g(da).$$

Since the $\Delta c(a)$ term is negligible as $\Delta \to 0$, this approximates the market-clearing condition given in the proof of Proposition 4.3.

B   Agency problem

This section contains proofs of all statements pertaining to agency problems in partial equilibrium. Recall that Proposition 2.2 in the main text covers the logarithmic case for an arbitrary interest rate. Proposition B.3 deals with the case of general constant relative risk aversion.

B.1   Logarithmic utility

Proof of Proposition 2.2. Recall that the Hamilton-Jacobi-Bellman equation is given by

$$rv(u) = \max_{\substack{c,k \geq 0 \\ k \leq \omega}} [\Pi - r - \tau_I]k - c + \left(-\rho \ln c/u + \frac{1}{2}(\rho\phi\sigma k/c)^2\right)uv'(u) + \frac{1}{2}(\rho\phi\sigma k/c)^2 u^2 v''(u).$$

The proof amounts to verifying that the solution to the Hamilton-Jacobi-Bellman equation and associated policy functions assumes the form $v(u) = vu, c(u) = cu$ and $k(u) = ku$ for some constants $\bar{v}, k$ and c. To this end, it is convenient to first change the choice variables from (k, c) to (x, c), where

$$x := \sqrt{\rho}\phi\sigma k c^{-1} \tag{37}$$

is the volatility of utility, and to also write

$$R(\Pi) := \frac{\Pi - r - \tau_I}{\sqrt{\rho}\phi\sigma} \tag{38}$$

so that we seek a solution to the fixed-point equation

$$rv = \max_{\substack{c,x \geq 0 \\ xc \leq \overline{w}}} (R(\Pi)x - 1)c + \rho\left(-\ln c + \frac{x^2}{2}\right)v. \tag{39}$$

where $\overline{w} = \omega\sqrt{\rho}\phi\sigma$. I will characterize the solution to (39) by first fixing x arbitrarily and solving the fixed-point equation associated with a principal who is unable to vary x. The associated restricted value function $v(x)$ solves

$$rv(x) = -\frac{\rho x^2}{2}[-v(x)] + \max_{\substack{c \geq 0 \\ xc \leq \overline{w}}} (R(\Pi)x - 1)c + \rho[-v(x)]\ln c =: -\frac{\rho x^2}{2}[-v(x)] + T[v](x). \tag{40}$$

For arbitrary v and x, the first-order condition for the maximization problem for T becomes

$$c = \frac{\rho[-v]}{1 - [R(\Pi) - \lambda(v,x)]x} \tag{41}$$

where $\lambda(x)$ is the Lagrange multiplier on the constraint $xc \leq \overline{\omega}$. Substitution then gives

$$(R(\Pi)x - 1)c + \rho[-v]\ln c = \frac{\rho[-v](R(\Pi)x - 1)}{1 - [R(\Pi) - \lambda(v,x)]x} + \rho[-v]\ln(\rho[-v]) - \rho[-v]\ln(1 - [R(\Pi) - \lambda(v,x)]x)$$

where $\lambda(v,x)$ is determined by the complementary slackness condition that requires is to be either zero or the value such that the no-absconding constraint binds at the consumption (41), which rearranges to

$$\lambda(v,x) = \max\{R(\Pi) - 1/x + \rho[-v]/\overline{\omega}, 0\}.$$

Substitution implies $1 - [R(\Pi) - \lambda(v,x)]x = 1 - R(\Pi)x + \max\{R(\Pi)x - 1 + \rho[-v]x/\overline{\omega}, 0\} = \max\{\rho[-v]x/\overline{\omega}, 1 - R(\Pi)x\}$, and so T given by the right-hand side of (40) simplifies to

$$T[v(x)] = \frac{\rho[-v](R(\Pi)x - 1)}{1 - [R(\Pi) - \lambda(v,x)]x} + \rho[-v]\ln(\rho[-v]) - \rho[-v]\ln(1 - [R(\Pi) - \lambda(v,x)]x)$$

$$= \frac{\rho[-v](R(\Pi)x - 1)}{\max\{\rho[-v]x/\overline{\omega}, 1 - R(\Pi)x\}} + \rho[-v]\ln(\rho[-v]) - \rho[-v]\ln(\max\{\rho[-v]x/\overline{\omega}, 1 - R(\Pi)x\}).$$

For any given volatility x, the payoff to the principal is then a fixed-point of the equation

$$\frac{x^2}{2} - \frac{r}{\rho} = \frac{R(\Pi)x - 1}{\max\{\rho[-v]x/\overline{\omega}, 1 - R(\Pi)x\}} + \ln(\rho[-v]) - \ln(\max\{\rho[-v]x/\overline{\omega}, 1 - R(\Pi)x\}). \tag{42}$$

Abbreviating $J(\Pi, x) = \rho x\overline{\omega}^{-1}/[1 - R(\Pi)x]$, this gives

$$\ln(1 - R(\Pi)x) + \frac{x^2}{2} - \frac{r}{\rho} - \ln\rho = -\frac{[-v]^{-1}}{\max\{J(\Pi, x), [-v]^{-1}\}} - \ln\max\{J(\Pi, x), [-v]^{-1}\}.$$

which may be written more succinctly as

$$\ln(1 - R(\Pi)x) + \frac{x^2}{2} - \frac{r}{\rho} - \ln\rho = \begin{cases} -1 + \ln[-v] & \text{if } J(\Pi, x)[-v] < 1 \\ -(J(\Pi, x)[-v])^{-1} - \ln J(\Pi, x) & \text{if } J(\Pi, x)[-v] > 1. \end{cases} \tag{43}$$

The right-hand side of (43) is increasing in $[-v]$ and diverges to negative infinity as $v \to 0$. A solution therefore exists if and only if $\ln(1 - R(\Pi)x) + x^2/2 - r/\rho - \ln\rho < -\ln J(\Pi, x)$ or

$$x\overline{\omega}^{-1}\exp\left(x^2/2 - r/\rho\right) < 1. \tag{44}$$

Notice that $\lambda(v,x) = 0$ if and only if $J(\Pi,x)[-v] < 1$. To this end, denote $v^*(x) = -1/J(\Pi,x)$ and note the no-absconding constraint will therefore be strict if and only if the right-hand side of (43) exceeds the left-hand side at $v = v^*(x)$, or $\rho x^2/2 + \rho - r + \rho \ln(x/\varpi) < 0$ in which case the value function becomes

$$v(x) = \frac{1}{\rho}(R(\Pi)x - 1)\exp\left(\frac{x^2}{2} + \frac{\rho - r}{\rho}\right). \tag{45}$$

If (45) is violated and (44) is satisfied, then the value function is

$$\ln(1 - R(\Pi)x) + \frac{x^2}{2} - \frac{r}{\rho} - \ln\rho = -\frac{1}{J(\Pi,x)[-v]} - \ln J(\Pi,x)$$

$$-\frac{\rho x^2}{2} + r - \rho\ln x\varpi^{-1} = \frac{1 - R(\Pi)x}{[-v]x\varpi^{-1}}.$$

In this case we have

$$v = \frac{(R(\Pi)x - 1)\varpi/x}{r - \rho x^2/2 - \rho\ln(x/\varpi)}$$

and the principal's problem may be written

$$\max_{x \geq 0} \; v(x,\Pi)$$
$$x\exp(x^2/2 - r/\rho) \leq \varpi \tag{46}$$

where for any given x we have

$$v(x,\Pi) = \frac{1}{\rho}(R(\Pi)x - 1)\begin{cases} \exp\left(1 - r/\rho + x^2/2\right) & \text{if } 1 - r/\rho + x^2/2 + \ln(x/\varpi) < 0 \\ (\varpi/x)[r/\rho - x^2/2 - \ln(x/\varpi)]^{-1} & \text{if } 0 < 1 - r/\rho + x^2/2 + \ln(x/\varpi) < 1. \end{cases} \tag{47}$$

This problem will be well-defined if and only if $R(\Pi)x - 1 < 0$ for all x satisfying $x^2/2 + \ln(x/\varpi) < r/\rho$, a condition equivalent to the inequality

$$1 > \varpi R(\Pi)\exp\left(r/\rho - \frac{1}{2R(\Pi)^2}\right). \tag{48}$$

Translated into original variables (48) is $[\Pi - r]\varpi\exp\left(r/\rho - \rho\phi^2\sigma^2/(2[\Pi - r]^2)\right) < 1$, which is obviously satisfied for all sufficiently small Π. Finally, if the no-absconding inequality is strict then the multiplier λ must then vanish. By (41) consumption is given by

$$c = \frac{\rho[-v]}{1 - R(\Pi)x} = \exp\left(1 - r/\rho + x^2/2\right). \tag{49}$$

In the case of logarithmic utility, the inverse Euler equation is equivalent to the drift in consumption equalling $r - \rho$. By (49) the drift in consumption is given by

$$\mu_c = \rho(-\ln c + x^2/2) = r - \rho - \rho x^2/2 + \rho x^2/2$$

as claimed. In the event that the no-absconding constraint is strict the optimal choice of x in (47) may be found by solving the first-order condition $0 = (d/dx)(R(\Pi)x - 1)\exp\left(x^2/2\right) = [R(\Pi)x^2 - x + R(\Pi)]\exp\left(x^2/2\right)$, which has solution

$$x(\Pi) = \frac{1}{2R(\Pi)}\left[1 - \sqrt{1 - 4R(\Pi)^2}\right].$$

This implies that consumption and capital policy functions are given by

$$c(\Pi) = \exp\left(1 - r/\rho + x(\Pi)^2/2\right) \qquad\qquad k(\Pi) = \frac{c(\Pi)x(\Pi)}{\sqrt{\rho}\phi\sigma}.$$

Substituting into (47) gives the claimed value function. $\qquad\qquad\square$

The following gives sufficient conditions for the no-absconding constraint to not bind in the $r = \rho$ case (the only case relevant to the examples in the main text).

Lemma B.1. *The no-absconding constraint will hold as a strict inequality if* (48) *and*

$$x(\Pi)^2/2 + \ln(x(\Pi)/[\sqrt{\rho}\phi\sigma]) < \ln\omega \qquad (50)$$

hold strictly.

Lemma B.1 shows that the no-absconding constraint holds for an open set of values of ω. All the examples in the paper hold when $\omega = 1/\rho$ (which corresponds to the ability to abscond with the whole capital stock).

B.2   General constant relative risk aversion

I now consider the case of general constant relative risk aversion with parameter $\gamma > 1$. The main point of this section is to establish that the results regarding optimal taxation and intertemporal distortions do not depend in a significant way on the choice of utility function. For simplicity I will only state this for $r = \rho$, as this is the only case relevant to the stationary efficient allocations considered in this paper. Furthermore, to avoid the main thrust of the paper becoming obscured by complicated algebra, I will not provide the most general treatment of this case and instead provide sufficient conditions that may be checked ex-post.

Lemma B.2. *When the agent absconds with K units of capital, the utility from having access to a bond market with return rate r is given by $W = (\rho + [\rho - r](1/\gamma - 1))^{-\gamma}K^{1-\gamma}/(1 - \gamma)$.*

Set $u := [(1 - \gamma)W]^{\frac{1}{1-\gamma}}$ so that if $V(W) = v(u)$ then we have $u'(W) = [(1 - \gamma)U]^{\frac{\gamma}{1-\gamma}} = u^\gamma, V'(W) = u^\gamma v'(u)$ and $V''(W) = u^{2\gamma-1}[\gamma v'(u) + u v''(u)]$. Standard arguments imply that the Hamilton-Jacobi-Bellman equation for the principal simplifies to

$$rv(u) = \max_{\substack{\bar{c},\bar{k} \geq 0 \\ \bar{k} \leq \omega u}} \left([\Pi - r]\bar{k} - \bar{c}\right)u + \left(\frac{\rho(1 - \bar{c}^{1-\gamma})}{1 - \gamma} + \frac{\gamma(\rho\phi\sigma)^2}{2}(\bar{k}\bar{c}^{-\gamma})^2\right)uv'(u) + \frac{(\rho\phi\sigma)^2}{2}(\bar{k}\bar{c}^{-\gamma})^2 u^2 v''(u).$$

It is easy to see that as in the Logarithmic case, the above Bellman equation will admit a linear solution with linear policy functions if a solution exists at all. However, the characterization of the value function with general relative risk aversion is more delicate here and as such it is convenient to define a number of auxiliary quantities prior to the characterization. Once again change variables from (k, c) to (x, c) where $x := \sqrt{\rho}\phi\sigma k/c^\gamma$ is proportional to the marginal utility of diverting a unit of capital to consumption. Using this and the shorthand $R(\Pi) := [\Pi - \rho_S]/(\sqrt{\rho}\phi\sigma)$, the Hamilton-Jacobi-Bellman equation becomes

$$\rho v = \max_{\substack{cx \geq 0 \\ xc^\gamma \leq \bar{\omega}}} R(\Pi)xc^\gamma - c + \rho\left(\frac{1 - c^{1-\gamma}}{1 - \gamma} + \frac{\gamma x^2}{2}\right)v \qquad (51)$$

where $\bar{\omega} := \omega/[\sqrt{\rho}\phi\sigma]$. Finally, define $\underline{v}$ to be the value function of a principal subject to both the technological and incentive-constraints given above, but who in addition must satisfy the no-absconding constraint with equality. This function must satisfy the equation

$$\rho v = \max_{\substack{cx \geq 0 \\ xc^\gamma = \bar{\omega}}} R(\Pi)xc^\gamma - c + \rho\left(\frac{1 - c^{1-\gamma}}{1 - \gamma} + \frac{\gamma x^2}{2}\right)v. \qquad (52)$$

Obviously then $\underline{v} \leq v$ as the constraint set of the latter is strictly larger than for the former. The following assumption collects together the technical conditions necessary for the extension of the results in the main text.

Assumption B.1. Suppose that the cubic

$$0 = (\gamma - 1)(\gamma - 1/2)x^3 + (3\gamma - 1)\frac{x^2 R(\Pi)}{2\gamma} - x + \frac{R(\Pi)}{\gamma} \tag{53}$$

has a positive solution, and that

$$\rho \underline{v} < \left[(3\gamma - 1)x^2/2 + 1\right]^{\frac{\gamma}{1-\gamma}} \left(\frac{R(\Pi)}{x\gamma}\right)^{\frac{1}{1-\gamma}} \tag{54}$$

$$x[1 - (\gamma - 1)(\gamma - 1/2)x^2]^{\frac{\gamma}{1-\gamma}} < \overline{\omega}$$

where x denotes the lower of the two positive solutions to (53).

I can now state the following analogue of Proposition 2.2.

Proposition B.3. *If the conditions in Assumption* (B.1) *are satisfied, then the inverse Euler equation holds and consumption and capital coefficients are given by*

$$c = \left[1 - (\gamma - 1/2)(\gamma - 1)x^2\right]^{\frac{1}{1-\gamma}}$$

$$k = \frac{x}{\sqrt{\rho\phi\sigma}}\left[1 - (\gamma - 1/2)(\gamma - 1)x^2\right]^{\frac{\gamma}{1-\gamma}} \tag{55}$$

where again x denotes the lower of the two positive solutions to (53). In this case consumption evolves according to a diffusion process of the form $dc_t = \mu_C c_t dt + \sigma_C c_t dZ_t$, where μ_C and σ_C are given by

$$\mu_C = \frac{\rho}{2}(1 - \gamma)x^2 \qquad\qquad \sigma_C = \sqrt{\rho}x.$$

Proof. Denoting the Lagrange multiplier for the right-hand side of (51) (for an arbitrary v) by $\lambda(v)$, the Hamilton-Jacobi-Bellman equation reduces to the system of equations

$$\frac{\rho\gamma}{1-\gamma}[-v] = [R(\Pi) - \lambda(v)]xc^\gamma - c + \frac{\rho c^{1-\gamma}}{1-\gamma}[-v] - \frac{\rho\gamma x^2}{2}[-v] + \lambda(v)\overline{\omega}$$

$$x = \frac{[R(\Pi) - \lambda(v)]c^\gamma}{\rho\gamma[-v]} \tag{56}$$

$$c = \frac{[R(\Pi) - \lambda(v)]^2 c^{2\gamma}}{\rho[-v]} + \rho c^{1-\gamma}[-v].$$

We search for a solution to the above system with $\lambda(v)$ and check ex-post that this is valid. Eliminating x from the system gives a system of two equations in two unknowns:

$$\frac{\rho\gamma}{1-\gamma}[-v] = \frac{R(\Pi)^2 c^{2\gamma}}{2\rho\gamma[-v]} - c + \frac{\rho c^{1-\gamma}}{1-\gamma}[-v]$$

$$c = \frac{R(\Pi)^2 c^{2\gamma}}{\rho[-v]} + \rho c^{1-\gamma}[-v]. \tag{57}$$

Substituting the second into the first gives

$$\frac{\rho\gamma}{1-\gamma}[-v] = \frac{1}{2\gamma}(c - \rho c^{1-\gamma}[-v]) - c + \frac{\rho c^{1-\gamma}}{1-\gamma}[-v]$$

$$\rho\gamma^2[-v] = (\gamma - 1)(\gamma - 1/2)c + (3\gamma/2 - 1/2)\rho c^{1-\gamma}[-v].$$

Eliminating the c term from the first equation in (57) then gives

$$-\rho\gamma[-v](\gamma - 1/2) = \frac{R(\Pi)^2 c^{2\gamma}}{2\rho\gamma[-v]}(\gamma - 1)(\gamma - 1/2) - \rho c^{1-\gamma}[-v](\gamma - 1/2) - (\gamma - 1)(\gamma - 1/2)c$$

$$-\rho\gamma(\gamma - 1/2) = \frac{R(\Pi)^2 c^{2\gamma}}{2\rho\gamma[-v]^2}(\gamma - 1)(\gamma - 1/2) - \rho c^{1-\gamma}(\gamma - 1/2) + \rho[-\gamma^2 + (3\gamma/2 - 1/2)c^{1-\gamma}]$$

$$\rho\gamma/2 = \frac{R(\Pi)^2 c^{2\gamma}}{2\rho\gamma[-v]^2}(\gamma - 1)(\gamma - 1/2) + \rho(3\gamma/2 - 1/2 - \gamma + 1/2)c^{1-\gamma}$$

and so the system (57) may be written

$$1 - c^{1-\gamma} = (\gamma - 1)(\gamma - 1/2)\frac{R(\Pi)^2 c^{2\gamma}}{\rho^2\gamma^2[-v]^2}$$

$$\frac{c}{[-v]} = \frac{R(\Pi)^2 c^{2\gamma}}{\rho[-v]^2} + \rho c^{1-\gamma}. \tag{58}$$

The first equation in (58) implies

$$[-v] = \frac{[(\gamma - 1)(\gamma - 1/2)]^{1/2}}{\rho\gamma(1 - c^{1-\gamma})^{1/2}}R(\Pi)c^\gamma. \tag{59}$$

Substituting into the second equation in (58) then gives the following equivalent equations

$$\frac{\rho\gamma c^{1-\gamma}(1 - c^{1-\gamma})^{1/2}}{[(\gamma - 1)(\gamma - 1/2)]^{1/2}R(\Pi)} = \frac{R(\Pi)^2\rho^2\gamma^2(1 - c^{1-\gamma})c^{2\gamma}}{\rho(\gamma - 1)(\gamma - 1/2)R(\Pi)^2 c^{2\gamma}} + \rho c^{1-\gamma}$$

$$\frac{\gamma c^{1-\gamma}(1 - c^{1-\gamma})^{1/2}}{[(\gamma - 1)(\gamma - 1/2)]^{1/2}} = \left(\frac{\gamma^2(1 - c^{1-\gamma})}{(\gamma - 1)(\gamma - 1/2)} + c^{1-\gamma}\right)R(\Pi)$$

$$\gamma[1 - (\gamma - 1)(\gamma - 1/2)x^2]x = \left(\gamma^2 x^2 + 1 - (\gamma - 1)(\gamma - 1/2)x^2\right)R(\Pi)$$

which simplifies to the cubic (53). Finally, using $(\gamma - 1)(\gamma - 1/2)x^2 = 1 - c^{1-\gamma}$ and equation (59), the coefficient of the value function satisfies

$$[-v] = \frac{R(\Pi)}{\rho x\gamma}\left[1 - (\gamma - 1)(\gamma - 1/2)x^2\right]^{\frac{\gamma}{1-\gamma}} = \rho^{-1}\left[(3\gamma - 1)x^2/2 + 1\right]^{\frac{\gamma}{1-\gamma}}\left(\frac{R(\Pi)}{x\gamma}\right)^{\frac{1}{1-\gamma}}. \tag{60}$$

It follows that the conditions (54) are equivalent to the conditions $v > \underline{v}$ and $xc^\gamma < \overline{\omega}$, respectively, which shows that these requirements imply the policy functions in the statement of the proposition.

Finally, note that the inverse Euler equation in this setting is equivalent to $0 = \mu_c + (\gamma - 1)\sigma_c^2/2$ where μ_c and σ_c denote the mean and volatility of the growth of consumption. In terms of the above choice of variables, we have

$$\mu_c = \frac{\rho(1 - \overline{c}^{1-\gamma})}{1 - \gamma} + \frac{\gamma(\rho\phi\sigma)^2}{2}(\overline{k}\overline{c}^{-\gamma})^2 \qquad\qquad \sigma_c = \rho\phi\sigma\overline{k}\overline{c}^{-\gamma}.$$

Substituting the expressions for $\overline{c}$ and $\overline{k}$ gives

$$\mu_c = \frac{\rho}{1 - \gamma}(\gamma - 1/2)(\gamma - 1)x^2 + \frac{\rho\gamma x^2}{2} \qquad\qquad \sigma_c = \sqrt{\rho}x$$

which gives the claimed expressions and also implies

$$\mu_c + (\gamma - 1)\frac{\sigma_c^2}{2} = -\rho(\gamma - 1/2)x^2 + \frac{\rho\gamma x^2}{2} + (\gamma - 1)\frac{\rho x^2}{2} = \rho(-(\gamma - 1/2) + \gamma/2 + (\gamma - 1)/2)x^2 = 0$$

as desired. $\qquad\square$

B.3 Wedges

Recall that for a consumption process $(c_t)_{t\geq 0}$ and asset A with return process $(R_t^A)_{t\geq 0}$ the associated wedge ν^A is defined implicitly by

$$u'(c_0) = \exp(-\rho t)\mathbb{E}\big[\exp(-\nu^R t)R_t^A u'(c_t)\big]. \tag{61}$$

Note that the associated return processes R^K and R^B are given by

$$R_t^K = \exp\left(\left[\Pi - \sigma^2/2\right]t + \sigma B_t\right) \qquad\qquad R_t^B = \exp\left(\rho t\right)$$

for all $t \geq 0$ and denote by ν^K and ν^B the associated wedges. Proposition B.3 implies that consumption evolves according to a diffusion process of the form $dc_t = \mu_C c_t dt + \sigma_C c_t dZ_t$, where μ_C and σ_C are given by

$$\mu_C = \frac{\rho}{2}(1-\gamma)x^2 \qquad\qquad \sigma_C = \sqrt{\rho}x.$$

Consumption then has the explicit representation

$$c_t = c_0 \exp\left(-\rho\gamma x(\Pi)^2 t/2 + \sqrt{\rho}x(\Pi)B_t\right). \tag{62}$$

This closed-form expression for consumption allows for a sharp characterization of the intertemporal wedges, leading to the following generalization of Lemma 2.3.

Lemma B.4. *The wedges on risky and risk-free capital are given by*

$$\nu^K = \Pi - \rho_S + \rho\gamma^2 x(\Pi)^2 - \sqrt{\rho}\gamma\sigma x(\Pi) \qquad\qquad \nu^B = \rho\gamma^2 x(\Pi)^2.$$

Furthermore, $\nu^B \geq \nu^K$.

Proof. Substituting R^K into (61) and rearranging gives

$$c_0^{-\gamma} = \exp(-[\rho+\nu^K]t)\mathbb{E}\big[\exp\left(\left[\Pi-\sigma^2/2\right]t + \sigma B_t\right)c_t^{-\gamma}\big]$$

$$1 = \exp(-[\rho+\nu^K]t)\mathbb{E}\big[\exp\left(\left[\Pi-\sigma^2/2\right]t + \sigma B_t\right)\exp\left(\rho\gamma^2 x(\Pi)^2 t/2 - \sqrt{\rho}\gamma x(\Pi)B_t\right)\big]$$

$$\exp(\nu^K t) = \exp\left(\left[\Pi-\rho-\sigma^2/2 + \rho\gamma^2 x(\Pi)^2/2\right]t\right)\mathbb{E}[\exp\left(\left[\sigma - \gamma\sqrt{\rho}x(\Pi)\right]B_t\right)].$$

Using $\mathbb{E}[\exp(zB_t)] = \exp(z^2 t/2)$ and taking logs of both sides gives

$$\nu^K = \Pi - \rho + \rho\gamma^2 x(\Pi)^2 - \sqrt{\rho}\gamma\sigma x(\Pi).$$

Similarly, substituting R^B into (61) and rearranging gives

$$c_0^{-\gamma} = \exp(-[\rho+\nu^B]t)\mathbb{E}\big[\exp\left(\rho t\right)c_t^{-\gamma}\big]$$

$$\exp(\nu^B t) = \mathbb{E}\big[\exp\left(\rho\gamma^2 x(\Pi)^2 t/2 - \sqrt{\rho}\gamma x(\Pi)B_t\right)\big] = \exp\left(\rho\gamma^2 x(\Pi)^2 t\right)$$

and so $\nu^B = \rho\gamma^2 x(\Pi)^2$. It remains to show $\nu^B \geq \nu^K$, or equivalently, $\Pi - \rho \leq \sqrt{\rho}\gamma\sigma x(\Pi)$. In the earlier notation this requires $R(\Pi) \leq x(\Pi)\gamma/\phi$ where $x(\Pi)$ is defined as the lower of the two positive roots of the polynomial

$$P(x) = (\gamma-1)(\gamma-1/2)x^3 + (3\gamma-1)\frac{x^2 R(\Pi)}{2\gamma} - x + \frac{R(\Pi)}{\gamma}$$

under the maintained assumption that such roots exist. Obviously it will suffice to show this in the case $\phi = 1$. Substituting $x = R(\Pi)/\gamma$ into the above gives

$$P(R(\Pi)/\gamma) = (\gamma-1)(\gamma-1/2)[R(\Pi)/\gamma]^3 + (3\gamma-1)\frac{[R(\Pi)/\gamma]^2 R(\Pi)}{2\gamma} - [R(\Pi)/\gamma] + R(\Pi)/\gamma$$

$$= (\gamma-1)(\gamma-1/2)[R(\Pi)/\gamma]^3 + (3\gamma-1)\frac{[R(\Pi)/\gamma]^3}{2} = ((\gamma-1)(\gamma-1/2) + 3\gamma/2 - 1/2)[R(\Pi)/\gamma]^3$$

$$= \left(\gamma^2 - 3\gamma/2 + 1/2 + 3\gamma/2 - 1/2\right)[R(\Pi)/\gamma]^3 = \frac{R(\Pi)^3}{\gamma} > 0.$$

It will suffice to show that $R(\Pi)/\gamma \leq z$ where z is the larger of the two roots to $0 = P'(x) = 3(\gamma - 1)(\gamma - 1/2)x^2 + (3\gamma - 1)xR(\Pi)/\gamma - 1$, or

$$z = \frac{1}{6(\gamma - 1)(\gamma - 1/2)}\left[-(3\gamma - 1)R(\Pi)/\gamma + \sqrt{(3\gamma - 1)^2 R(\Pi)^2/\gamma^2 + 12(\gamma - 1)(\gamma - 1/2)}\right].$$

This requires

$$R(\Pi) \leq \frac{\gamma}{6(\gamma - 1)(\gamma - 1/2)}\left[-(3\gamma - 1)R(\Pi)/\gamma + \sqrt{(3\gamma - 1)^2 R(\Pi)^2/\gamma^2 + 12(\gamma - 1)(\gamma - 1/2)}\right]$$

$$1 \leq \frac{1}{6(\gamma - 1)(\gamma - 1/2)}\left[-(3\gamma - 1) + \sqrt{(3\gamma - 1)^2 + 12\gamma^2(\gamma - 1)(\gamma - 1/2)/R(\Pi)^2}\right]$$

which is equivalent to

$$6(\gamma - 1)(\gamma - 1/2) + (3\gamma - 1) \leq \sqrt{(3\gamma - 1)^2 + 12\gamma^2(\gamma - 1)(\gamma - 1/2)/R(\Pi)^2}$$
$$[6(\gamma - 1)(\gamma - 1/2)]^2 + 12(\gamma - 1)(\gamma - 1/2)(3\gamma - 1) \leq 12\gamma^2(\gamma - 1)(\gamma - 1/2)/R(\Pi)^2$$
$$3(\gamma - 1)(\gamma - 1/2) + (3\gamma - 1) \leq \gamma^2/R(\Pi)^2$$

and hence

$$R(\Pi)^2 \leq \frac{2\gamma^2}{6\gamma^2 - 3\gamma + 1}. \tag{63}$$

I want to show that this follows from the assumption $P(z) \leq 0$. To this end note that the defining equality for z gives

$$P'(z) = 0 = 3(\gamma - 1)(\gamma - 1/2)z^2 + (3 - 1/\gamma)R(\Pi)z - 1$$
$$3(\gamma - 1)(\gamma - 1/2)z^2 = 1 - (3 - 1/\gamma)R(\Pi)z.$$

Using this repeatedly implies

$$3P(z) = 3(\gamma - 1)(\gamma - 1/2)z^3 + \frac{3}{2}(3 - 1/\gamma)R(\Pi)z^2 - 3z + \frac{3R(\Pi)}{\gamma}$$
$$= -(3 - 1/\gamma)R(\Pi)z^2 + \frac{3}{2}(3 - 1/\gamma)R(\Pi)z^2 - 2z + \frac{3R(\Pi)}{\gamma}$$
$$= \frac{1}{2}(3 - 1/\gamma)R(\Pi)\frac{[1 - (3 - 1/\gamma)R(\Pi)z]}{3(\gamma - 1)(\gamma - 1/2)} - 2z + \frac{3R(\Pi)}{\gamma}.$$

The requirement $P(z) \leq 0$ is then equivalent to

$$\frac{(3\gamma - 1)}{3(\gamma - 1)(\gamma - 1/2)} + 6 \leq \left(\frac{4\gamma}{R} + \frac{(3\gamma - 1)(3 - 1/\gamma)R}{3(\gamma - 1)(\gamma - 1/2)}\right)z. \tag{64}$$

It will suffice to show that (64) implies (63). To this end I will show that the strictly weaker inequality

$$\frac{(3\gamma - 1)}{3(\gamma - 1)(\gamma - 1/2)} + 6 \leq \frac{4\gamma z}{R} + \frac{(3\gamma - 1)(3 - 1/\gamma)R}{3(\gamma - 1)(\gamma - 1/2)}z' \tag{65}$$

also implies (63), where $(3\gamma - 1)Rz' = \gamma$. Inequality (65) rearranges to

$$\frac{3R}{2} \leq \gamma z = \frac{\gamma}{6(\gamma - 1)(\gamma - 1/2)}\left[-(3 - 1/\gamma)R + \sqrt{(3 - 1/\gamma)^2 R^2 + 12(\gamma - 1)(\gamma - 1/2)}\right]$$
$$9(\gamma - 1)(\gamma - 1/2) + (3\gamma - 1) \leq \sqrt{(3\gamma - 1)^2 + 12\gamma^2(\gamma - 1)(\gamma - 1/2)/R^2}$$
$$R^2 \leq \frac{\gamma^2}{27(\gamma - 1)(\gamma - 1/2)/4 + 3(3\gamma - 1)/2}.$$

42

This will imply (63) if $\gamma^2/[27(\gamma-1)(\gamma-1/2)/4 + 3(3\gamma-1)/2] \leq \gamma^2/(3\gamma^2 - 3\gamma/2 + 1/2)$ or

$$3\gamma^2 - 3\gamma/2 + 1/2 \leq 27(\gamma-1)(\gamma-1/2)/4 + 3(3\gamma-1)/2$$
$$24\gamma^2 - 12\gamma + 4 \leq 54\gamma^2 - 81\gamma + 27 + 12(3\gamma - 1)$$

which rearranges to $0 \leq 30\gamma^2 - 33\gamma + 11$, which is always true. $\qquad\square$

C   Stationary allocations

C.1   Aggregate resource constraints

Aggregate consumption, labor, and output at any date are comprised of contributions from the initial generation and subsequent generations. I will write them in this fashion for clarity.

C.1.1   Single types

Aggregate consumption, capital assigned and output at any date $t \geq 0$ are

$$\underline{C}_t := \int_{\mathbb{R}} \mathbb{E}[c_t^v]\Phi(dv), \quad C_t^T := L\mathbb{E}[c_t^T]$$

$$C_t := e^{-\rho_D t}\underline{C}_t + \int_0^t e^{-\rho_D[t-T]}C_t^T dT$$

$$K_t := e^{-\rho_D t}\int_{\mathbb{R}} \mathbb{E}[K_t^v]\Phi(dv) + \int_0^t e^{-\rho_D[t-T]}\mathbb{E}[K_t^T]dT$$

$$\underline{Y}_t := \int_{\mathbb{R}} \mathbb{E}[F(K_t^v, L_t^v) - \delta K_t^v]\Phi(dv), \quad Y_t^T := \mathbb{E}[F(K_t^T, L_t^T) - \delta K_t^T]$$

$$Y_t := e^{-\rho_D t}\underline{Y}_t + \int_0^t e^{-\rho_D[t-T]}Y_t^T dT$$

where I have used the notation $F(K, L) := AK^\alpha L^{1-\alpha}$. Aggregate labor assigned to entrepreneurs is

$$\underline{L}_t := \int_X \mathbb{E}[L_t^v]\Phi(dv), \quad L_t^T := e^{-\rho_D[t-T]}\mathbb{E}[L_t^T]$$

$$L_t := e^{-\rho_D t}\underline{L}_t + \int_0^t e^{-\rho_D[t-T]}L_t^T dT.$$

I will also use the following notation for the Pareto-weighted flow utility experienced by each generation

$$\underline{U}_t = \int_{\mathbb{R}} \mathbb{E}[u(c_t^v)]\Phi(dv) \qquad\qquad U_t^T = L\mathbb{E}[u(c_t^T)].$$

Now let us go through the details concerning the relationship between the planner's problem and the principal-agent problem. Given an initial distribution Φ over promised utility and types, when the planner discounts at rate ρ_S, the relaxed problem is defined to be

$$V^R(\Phi) = \max_{A \in \mathcal{A}^{IC}(\Phi)} \int_0^\infty \left(e^{-\rho t}\underline{U}_t + \int_0^t e^{-\rho[t-T]}e^{-\rho_S T}U_t^T dT \right)dt$$

$$\int_0^\infty e^{-\rho_S t}[C_t(A) - Y_t(A)]dt \leq 0.$$

$$\int_0^\infty e^{-\rho_S t}[L_t(A) - L]dt \leq 0.$$

43

Using the pair of multipliers $\lambda := (\lambda_R, \lambda_L)$ we form the Lagrangian by combining the above objective with the weights on the resource constraints

$$\mathcal{L} = \int_0^\infty \left(e^{-\rho t}\underline{U}_t + \int_0^t e^{-\rho[t-T]}e^{-\rho_S T}U_t^T dT \right)dt + \lambda_R \int_0^\infty e^{-\rho_S t}[Y_t(A) - C_t(A)]dt + \lambda_R\lambda_L \int_0^\infty e^{-\rho_S t}[L - L_t(A)]dt$$

$$= \int_0^\infty \left(e^{-\rho t}\underline{U}_t + \int_0^t e^{-\rho[t-T]}e^{-\rho_S T}U_t^T dT + \lambda_R e^{-\rho_S t}[Y_t(A) - C_t(A)] + \lambda_R\lambda_L e^{-\rho_S t}[L - L_t(A)] \right)dt.$$

Now we substitute in the expressions for aggregate consumption, output, and labor assigned. This gives

$$\mathcal{L} = \int_0^\infty \left(e^{-\rho t}\underline{U}_t + \int_0^t e^{-\rho[t-T]}e^{-\rho_S T}U_t^T dT \right)dt + \lambda_R \int_0^\infty e^{-\rho_S t}[Y_t(A) - C_t(A)]dt + \lambda_R\lambda_L \int_0^\infty e^{-\rho_S t}[L - L_t(A)]dt$$

$$= \int_0^\infty \left(e^{-\rho t}\underline{U}_t + \int_0^t e^{-\rho[t-T]}e^{-\rho_S T}U_t^T dT + \lambda_R e^{-\rho_S t}\left[-e^{-\rho_D t}\underline{C}_t + \int_0^t e^{-\rho_D[t-T]}C_t^T dT + e^{-\rho_D t}\underline{Y}_t + \int_0^t e^{-\rho_D[t-T]}Y_t^T dT \right] \right.$$

$$\left. + \lambda_R\lambda_L e^{-\rho_S t}\left[-e^{-\rho_D t}\underline{L}_t - \int_0^t e^{-\rho_D[t-T]}L_t^T dT + L \right] \right)dt$$

$$= \int_0^\infty \left(e^{-\rho t}[\underline{U}_t + \lambda_R(\underline{C}_t - \underline{Y}_t + \lambda_L\underline{L}_t)] + \int_0^t \left[e^{-\rho[t-T]}e^{-\rho_S T}U_t^T + \lambda_R e^{-\rho_S t}e^{-\rho_D[t-T]}\left(Y_t^T - C_t^T - \lambda_L L_t^T \right) \right]dT \right)dt + \frac{\lambda_R\lambda_L}{\rho_S}$$

Using the fact that $e^{-\rho[t-T]}e^{-\rho_S T} = e^{-\rho_S t}e^{-\rho_D[t-T]}$ for all $t \geq T \geq 0$, and switching the order of integration,

$$\mathcal{L} = \max_{A \in \mathcal{A}^{IC}} \int_0^\infty e^{-\rho t}\underline{\mathcal{L}}_t dt + \int_0^\infty \int_T^\infty e^{-\rho_S T}e^{-\rho(t-T)}\mathcal{L}_t^T dt dT$$

where $\underline{\mathcal{L}}_t(\lambda)$ and $\mathcal{L}_t^T(\lambda)$ are the contributions of the initial and Tth generations to the Lagrangian

$$\underline{\mathcal{L}}_t = \underline{U}_t + \lambda_R[\underline{Y}_t - \underline{C}_t + \lambda_L[L - \underline{L}_t]]$$
$$\mathcal{L}_t^T = U_t^T + \lambda_R\left[Y_t^T - C_t^T + \lambda_L\left[L - L_t^T\right]\right]$$

(66)

and $(\underline{C}_t, \underline{Y}_t, \underline{L}_t)$ and (C_t^T, Y_t^T, L_t^T) refer to consumption, output, and labor assignments of the initial and Tth generations, respectively, at date $t \geq 0$. The above decomposition illustrates that for any given choice of the multipliers, for each $T > 0$ the planner solves

$$V_T = \max_{A \in \mathcal{A}^{IC}} \int_T^\infty e^{-\rho(t-T)}\mathcal{L}_t^T dt.$$

(67)

In other words, the planner behaves exactly as per the principal of Section 2 who discounts at the rate $r = \rho$.

C.2 Stationary efficient distributions

D Decentralization

Section D.1 states the value functions associated with individual agents and Section D.2 characterizes the taxes that decentralize the efficient allocation.

D.1 Agent's problems

Recall that Lemma 4.2 solved the agent's problem when facing linear taxes on savings and profits.

Proof of Lemma 4.2. Given taxes on profits τ_Π and risk-free savings τ_s, the Hamilton-Jacobi-Bellman equation for the agent's value function is

$$\rho V(a) = \max_{c,k \geq 0} \rho \ln c + ((1 - \tau_s)(r + \rho_D)a - c + (1 - \tau_\Pi)(\Pi - r)k + w)V'(a) + \frac{\sigma^2}{2}(1 - \tau_\Pi)^2 k^2 V''(a).$$

Substitution of the assumed form $V(a) = \ln(a + h) + D$ for some constant D into the right-hand side gives

$$\max_{c,k \geq 0} \rho \ln c + (1 - \tau_s)(r + \rho_D) - c/(a + h) + (1 - \tau_\Pi)(\Pi - r)k/(a + h) - \frac{[\sigma(1 - \tau_\Pi)]^2}{2}(k/(a + h))^2.$$

Optimal consumption is then $c(a) = \rho a$, and optimal capital is

$$k(a) := \underline{k}(a + h) = \frac{(\Pi - r)(a + h)}{\sigma^2(1 - \tau_\Pi)}.$$

The constant D then satisfies

$$\rho D = \rho \ln \rho + (1 - \tau_s)r - \rho + \frac{(\Pi - r)^2}{\sigma^2} - \frac{[\sigma(1 - \tau_\Pi)]^2}{2} \frac{(\Pi - r)^2}{\sigma^4(1 - \tau_\Pi)^2}$$

which reduces to the desired expression for V. $\qquad\square$

Notice that Lemma 2.1 is simply a special case of Lemma 4.2 with $\tau_s = w = 0$ and $\Pi = r$. Now, for general CRRA parameters, the Hamilton-Jacobi-Bellman equation is given by

$$\rho V(a) = \max_{c,k} \frac{c^{1-\gamma}}{1 - \gamma} + [(1 - \tau_s)(r + \rho_D)a - c + (1 - \tau_\Pi)(\Pi - r)k + w]V'(a) + (1 - \tau_\Pi)^2 \frac{\sigma^2 k^2}{2} V''(a).$$

Lemma D.1. *The value function of the agent is $V(a) = \overline{V} a^{1-\gamma}/(1 - \gamma)$ for some $\overline{V}$, with policy functions*

$$c(a) = \overline{c}(a + h) = \left(\frac{1}{\gamma}[\rho - (1 - \gamma)(1 - \tau_s)(r + \rho_D)] - \frac{(\Pi - r)^2}{2\gamma^2\sigma^2}(1 - \gamma) \right)(a + h)$$

$$k(a) = \overline{k}(a + h) = \frac{(\Pi - r)(a + h)}{\gamma\sigma^2(1 - \tau_\Pi)}$$

where $h = w/[(1 - \tau_s)(r + \rho_D)]$. The associated law of motion of wealth is $da = \mu_a(a_t + h)dt + \sigma_a(a_t + h)dZ_t$, where

$$\mu_a = \frac{(1 - \tau_s)(r + \rho_D) - \rho}{\gamma} + \frac{(\Pi - r)^2}{2\gamma^2\sigma^2}(1 + \gamma) \qquad\qquad \sigma_a = \frac{\Pi - r}{\gamma\sigma}.$$

Proof. Upon substituting the assumed form, the Hamilton-Jacobi-Bellman equation becomes

$$\frac{\rho\overline{V}}{1 - \gamma} = \max_{\overline{c},\overline{k}} \frac{\overline{c}^{1-\gamma}}{1 - \gamma} + \overline{V}\left[(1 - \tau_s)(r + \rho_D) - \overline{c} + (1 - \tau_\Pi)(\Pi - r)\overline{k}\right] - \gamma(1 - \tau_\Pi)^2 \overline{V}\frac{\sigma^2\overline{k}^2}{2}.$$

First-order conditions for capital and consumption give

$$\overline{k} = \frac{\Pi - r}{\gamma(1 - \tau_\Pi)\sigma^2} \qquad\qquad \overline{c} = \overline{V}^{-1/\gamma}.$$

Substituting into the Hamilton-Jacobi-Bellman equation gives

$$\frac{\rho\overline{V}}{1 - \gamma} = \frac{\overline{V}^{1-1/\gamma}}{1 - \gamma} + \overline{V}[(1 - \tau_s)(r + \rho_D) - \overline{c}] + \overline{V}\left[(1 - \tau_\Pi)(\Pi - r)\overline{k} - \frac{\gamma}{2}[(1 - \tau_\Pi)\sigma]^2\overline{k}^2\right]$$

$$= \frac{\gamma\overline{V}^{1-1/\gamma}}{1 - \gamma} + \overline{V}(1 - \tau_s)(r + \rho_D) + \overline{V}\left[\frac{(\Pi - r)^2}{\gamma\sigma^2} - \frac{\gamma}{2}[(1 - \tau_\Pi)\sigma]^2\left(\frac{\Pi - r}{\gamma(1 - \tau_\Pi)\sigma^2}\right)^2\right]$$

$$\frac{\rho}{1 - \gamma} = \frac{\gamma\overline{V}^{-1/\gamma}}{1 - \gamma} + (1 - \tau_s)(r + \rho_D) + \frac{(\Pi - r)^2}{2\gamma\sigma^2}$$

which rearranges to

$$\overline{c} = \overline{V}^{-1/\gamma} = \frac{1}{\gamma}[\rho - (1-\gamma)(1-\tau_s)(r+\rho_D)] - \frac{(\Pi-r)^2}{2\gamma^2\sigma^2}(1-\gamma)$$

as claimed. The law of motion of wealth is then

$$
\begin{aligned}
da_t &= [(1-\tau_s)(r+\rho_D)a_t + w - c_t + (1-\tau_\Pi)(\Pi-r)k_t]dt + (1-\tau_\Pi)\sigma k_t dZ_t \\
&= \left[(1-\tau_s)(r+\rho_D) - \overline{c} + (1-\tau_\Pi)(\Pi-r)\overline{k}\right](a_t+h)dt + (1-\tau_\Pi)\sigma\overline{k}(a_t+h)dZ_t \\
&= \left[(1-\tau_s)(r+\rho_D) - \frac{1}{\gamma}[\rho-(1-\gamma)(1-\tau_s)(r+\rho_D)] + \frac{(\Pi-r)^2}{2\gamma^2\sigma^2}(1-\gamma) + \frac{(\Pi-r)^2}{\gamma\sigma^2}\right](a_t+h)dt \\
&\quad + \frac{(\Pi-r)}{\gamma\sigma}(a_t+h)dZ_t.
\end{aligned}
$$

This clearly implies $\sigma_a = (\Pi-r)/[\gamma\sigma]$, while μ_a simplifies to

$$\mu_a = -\frac{\rho}{\gamma} + \frac{1}{\gamma}(1-\tau_s)(r+\rho_D) + \frac{(\Pi-r)^2}{\gamma\sigma^2}\left(\frac{1}{2}(1/\gamma - 1) + 1\right)$$

as claimed. $\qquad\square$

The value function associated with absconding with the capital stock corresponds to the $\Pi = r$ case with no taxes.

Lemma D.2 (Utility from absconding). *The value function associated with absconding with the stock of capital is given by $V(K) = (\rho + [\rho - r](1/\gamma - 1))^{-\gamma}K^{1-\gamma}/(1-\gamma)$.*

D.2 Decentralization with taxes

Recall that the drift and diffusion for wealth are given by

$$\mu_a = \frac{(1-\tau_s)(r+\rho_D) - \rho}{\gamma} + \frac{(\Pi-r)^2}{2\gamma^2\sigma^2}(1+\gamma) \qquad\qquad \sigma_a = \frac{\Pi-r}{\gamma\sigma}.$$

Recall that the efficient allocation is again characterized by three properties: the marginal product of capital coincides with the solution to the stationary form of the goods resource constraint, and the mean and volatility of the growth in consumption coincide with those in the efficient allocation.

Proposition D.3. *The marginal product of capital that obtains in the efficient stationary distribution is the solution to the equation*

$$c(\Pi) = \frac{1}{\alpha}[\Pi + (1-\alpha)\delta]k(\Pi) \tag{68}$$

provided the no-absconding constraint holds as a strict inequality for this $\overline{\Pi}$. In this case the efficient allocations may be decentralized with linear taxes on savings and profits

$$1 - \tau_s = \frac{\rho[1 - x(\overline{\Pi})^2\gamma^2]}{\overline{\Pi} - \sqrt{\rho}\gamma\sigma x(\overline{\Pi}) + \rho_D} \qquad\qquad \tau_\Pi = 1 - \phi$$

and in which each agent is endowed with the fraction

$$\eta = (1 - (\rho/\rho_D)(1-\gamma)x(\Pi)^2/2)\left(\frac{\Pi/\alpha + (1/\alpha - 1)\delta}{\rho(1 - (1-\gamma)(1/2 - \gamma)x(\Pi)^2)} - \frac{(\Pi+\delta)(1/\alpha - 1)}{\rho[1 - \gamma^2 x(\Pi)^2]}\right)$$

of the aggregate capital stock at birth.

46

Proof of Proposition D.3. Denote by ψ the fraction of the aggregate capital stock that is owned by private agents. Note that ψ is related to the inheritance share η via the Gordon growth relation

$$\psi = \frac{\rho_D \eta}{\rho_D - \mu_c} \tag{69}$$

where μ_c is the (yet-to-be-characterized) growth rate of consumption. The stationary form of the three market-clearing conditions for labor, capital and goods may be written

$$L = \phi_l(w)K$$
$$K = \overline{k}(w,r)(\psi K + h(w,r)L)$$
$$(A\phi_l(w)^{1-\alpha} - \delta)\overline{k}(w,r) = \overline{c}(w,r).$$

The competitive equilibrium is characterized by this system of three equations and three unknowns, with the inheritance rate being determined residually from the requirement that net wealth be constant. The labor resource constraint implies $\phi_l(w) = [A(1-\alpha)/w]^{1/\alpha}$, so the wage is simply the marginal product of labor $w = A(1-\alpha)(K/L)^\alpha$, and so the marginal product of capital is then $\Pi = \alpha A[K/L]^{\alpha-1} - \delta$. The total level of human wealth satisfies

$$h(w,r)L = \frac{wL}{(1-\tau_s)(r+\rho_D)} = \frac{(\Pi+\delta)(1/\alpha-1)K}{(1-\tau_s)(r+\rho_D)}. \tag{70}$$

The bond and goods market-clearing conditions then become

$$1 = \left[\psi + \frac{(\Pi+\delta)(1/\alpha-1)}{(1-\tau_s)(r+\rho_D)}\right]\overline{k}(w,r)$$
$$\frac{1}{\alpha}[\Pi + (1-\alpha)\delta]\overline{k}(w,r) = \overline{c}(w,r). \tag{71}$$

In order for the efficient allocation to be implemented, three conditions must be met: the marginal product of capital must coincide with the solution to (68), and the mean and volatility of consumption growth must coincide with those values attained in the efficient allocation. Combining Proposition B.3 with Lemma D.1, this implies

$$\frac{(1-\tau_s)(r+\rho_D)-\rho}{\gamma} + \frac{[\Pi-r]^2}{2\gamma^2\sigma^2}(1+\gamma) = \frac{\rho}{2}(1-\gamma)x(\Pi)^2$$
$$\sigma(1-\tau_\Pi)\overline{k}(w,r) = \sqrt{\rho}x(\Pi). \tag{72}$$

Using Lemma D.1 and the second condition of (72) we find

$$r = \Pi - \sqrt{\rho}\gamma\sigma x(\Pi).$$

Substituting this interest rate into the first condition of (72) then gives

$$1 - \tau_s = \frac{\rho[1-\gamma^2 x(\Pi)^2]}{\Pi - \sqrt{\rho}\gamma\sigma x(\Pi) + \rho_D}$$

as claimed. Now, using the explicit forms for the policy functions gives

$$\overline{c} = \left(\frac{\rho}{\gamma}(1-(1-\gamma)[1-\gamma^2 x(\Pi)^2]) - \frac{\rho x(\Pi)^2}{2}(1-\gamma)\right) = \rho\big(1-(1-\gamma)(1/2-\gamma)x(\Pi)^2\big)$$
$$\overline{k} = \frac{\Pi-r}{\gamma\sigma^2(1-\tau_\Pi)} = \frac{\sqrt{\rho}x(\Pi)}{\sigma(1-\tau_\Pi)}.$$

47

It follows that

$$\frac{1}{\alpha}[\Pi + (1-\alpha)\delta] = \bar{c}(w,r)/\bar{k}(w,r) = \frac{\sqrt{\rho}\sigma(1-\tau_\Pi)}{x(\Pi)}\left(1 - (\gamma-1)(\gamma-1/2)x(\Pi)^2\right).$$

Using Proposition B.3 we have

$$c(\Pi)/k(\Pi) = \frac{\sqrt{\rho}\phi\sigma}{x(\Pi)}\left[1 - (\gamma-1/2)(\gamma-1)x(\Pi)^2\right].$$

Equating $c(\Pi)/k(\Pi) = \bar{c}(w,r)/\bar{k}(w,r)$ (which must hold if the goods resource constraint and goods market-clearing hold for the same Π) implies $\tau_\Pi = 1 - \phi$. Finally, using the goods market-clearing condition and the above expressions for τ_s, r and $\bar{c}$ we have

$$\frac{1}{\alpha}[\Pi + (1-\alpha)\delta] = \bar{c}(w,r)\left(\psi + \frac{(\Pi+\delta)(1/\alpha-1)}{(1-\tau_s)(r+\rho_D)}\right)$$

$$\psi = \frac{\Pi/\alpha + (1/\alpha-1)\delta}{\rho(1-(1-\gamma)(1/2-\gamma)x(\Pi)^2)} - \frac{(\Pi+\delta)(1/\alpha-1)}{\rho[1-\gamma^2 x(\Pi)^2]}.$$

Since $\mu_c = \rho(1-\gamma)x(\Pi)^2/2$ it follows from (69) that

$$\eta = (\rho_D - \mu_c)\psi = (1 - (\rho/\rho_D)(1-\gamma)x(\Pi)^2/2)\left(\frac{\Pi/\alpha + (1/\alpha-1)\delta}{\rho(1-(1-\gamma)(1/2-\gamma)x(\Pi)^2)} - \frac{(\Pi+\delta)(1/\alpha-1)}{\rho[1-\gamma^2 x(\Pi)^2]}\right)$$

as claimed. □

Corollary D.4 relates the interest rate that obtains in the incomplete markets model to the subjective rate of discount and is the analogue of Corollary 4.4 in the general relative risk aversion case.

Corollary D.4. *The interest rate in the stationary competitive equilibrium that decentralizes the efficient allocation is always lower than the subjective rate of discount.*

Proof. Notice that by Proposition D.3, $r \leq \rho_S$ is equivalent to

$$\bar{\Pi} - \sqrt{\rho}\gamma\sigma x(\Pi) \leq \rho_S \tag{73}$$

which follows from Lemma B.4. □