Efficient wealth inequality and differential asset taxation with dynamic agency

Thomas Phelan

Abstract

This paper characterizes a class of stationary constrained-efficient allocations and optimal taxes in an economy with endogenous firm formation and dynamic moral hazard. 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, high realizations of output must be accompanied by high future consumption, leading to ex-post inequality in the efficient allocation. I show that the distributions of consumption and wealth associated with the stationary efficient allocation exhibit thick right (Pareto) tails, with the degree of inequality monotonically increasing in the number of workers per entrepreneur. This constrained-efficient allocation is then implemented in a general equilibrium model using linear taxes on labour income, risk-free savings and 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 demographic parameters.

Keywords: Dynamic moral hazard, optimal contracting, entrepreneurship, optimal taxation.

1 Introduction

Business owners are disproportionately represented amongst the very wealthy and are exposed to substantial idiosyncratic risk.[1] Given this, how should a utilitarian government balance a concern for redistribution between business owners and workers with the need to provide incentives for both business creation and investment? In particular, should the government tax business profits, the (risk-free) savings of business owners, or some combination of both? To address these and related questions, this paper studies optimal differential asset taxation and efficient levels of wealth inequality in a dynamic economy in which the risk borne by business owners arises endogenously from agency frictions.

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). Only some agents have the ability to run a business and this is private information. In addition, entrepreneurship is subject to two agency frictions: physical capital is subject to (privately-observed) output shocks and may be diverted to (privately-observed) consumption; and entrepreneurs may abscond with a fixed fraction of delegated capital. An allocation in this environment must specify the occupation of every agent and the amount of capital and labour delegated to each business. The ability

[1] See Smith, Yagan, Zidar and Zwick [18] and Phelan [12] and the references therein for evidence on this point.

of entrepreneurs to divert capital to private consumption will imply that consumption must depend upon business performance, which leads to ex-post inequality in the efficient allocation. The ability of entrepreneurs to abscond with a fraction of their assets will limit the amount of capital that may be delegated to them.

I characterize a class of constrained-efficient allocations in which aggregate quantities are constant over time. The associated stationary distributions of consumption and firm size may be characterized in closed-form and exhibit thick right (Pareto) tails. I then show that these stationary efficient allocations may be implemented in a general equilibrium model with linear taxes on the labour income of workers, the risk-free savings of workers and entrepreneurs and the profits of entrepreneurs' businesses. The savings taxes on entrepreneurs may assume either sign, while the tax on profits depends solely upon the extent of the agency problem (specifically, the fraction of diverted capital that may be converted into consumption).

An extensive literature, surveyed in Chari and Kehoe [5], has analyzed optimal capital and labour 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 and consumption. Recent contributions to this tradition, such as Panousi and Reis [11] and Evans [6], extend this analysis to consider optimal linear taxation in economies with a continuum of agents subject to idiosyncratic risk with exogenously incomplete markets. In contrast, the analysis of this paper builds upon the seminal contribution of Mirrlees [10] and the New Dynamic Public Finance literature beginning with Golosov, Kocherlakota and Tsyvinski [9] to characterize efficient allocations when the only restrictions placed upon government policy are those implied by informational asymmetries. However, the majority of this literature has focused on the implications of private-information in labour productivity on the structure of optimal tax schedules and has not explicitly accounted for entrepreneurial activity. Three notable exceptions are Albanesi [1], Shourideh [17] and my companion paper Phelan [12], each of which explicitly model entrepreneurial activity in the presence of agency frictions. It is therefore instructive to outline how the modeling assumptions and findings of these papers differ from those of this paper.

Albanesi [1] considers a two-period model in which there are no workers, initial wealth is exogenous and common across entrepreneurs, and the returns to entrepreneurial activity depend upon unobserved effort. She finds that in general the efficient intertemporal wedge differs from the case with unobserved labour productivity and may assume either sign. Further, decentralization of the efficient allocation with an exogenous market structure may require double taxation at both the firm and individual level. I recover this latter result in my decentralization, but extend the model to an infinite-horizon setting in which wealth is endogenous and so am able to illustrate how private information affects the efficient level of long-run inequality.

More closely related with the current paper is Shourideh [17], 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 lifecycle structure that leads to simpler characterizations of both efficient allocations and their decentralization.[2] The modeling of the agency problem also qualitatively changes the nature of efficient intertemporal distortions. In contrast to the findings of both Albanesi [1] and Shourideh [17], I show that the inverse Euler equation of Rogerson [14] and Golosov, Kocherlakota and Tsyvinski [9] continues to hold in the presence of production risk for a wide range of parameter values. Further, the increased tractability allows me to extend the

[2] 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.

literature in several ways.

First, in addition to characterizing a stationary efficient allocation, I show how it may be implemented in a general equilibrium model with exogenously incomplete markets. The decentralization requires only linear, time-independent taxes and so optimal policy may be completely specified by five scalars, each of which admits a closed-form representation in terms of the solution to a single non-linear equation. Further, the finding that the optimal profit tax is independent of all technological parameters relies on general equilibrium effects and appears to have no antecedent in the literature on optimal taxation with endogenously incomplete markets.

Second, the model of this paper contains both workers and entrepreneurs, which has important implications for both the efficient allocation and its decentralization. I show that when the number of workers per entrepreneur increases, capital per worker falls and inequality increases in the associated stationary distribution. The intuition for this result is as follows: the fewer entrepreneurs there are, the higher is their marginal product, and hence the more capital and labour is delegated to each individual entrepreneur. In order to preserve incentives to not divert assets to private consumption, each entrepreneur must bear more risk and so inequality increases. However, the inclusion of workers has asymmetric effects on taxes, with changes in the number of workers affecting only the tax on entrepreneurs' savings, leaving the tax on business profits unchanged.

Third, the model also allows for a sharper characterization of the long-run distributions of utility and consumption and their determinants. Indeed, in this paper the stationary distribution of consumption admits a closed-form density of the 'double-Pareto' form[3] with the degree of inequality in the upper tail determined by the amount of risk borne by each entrepreneur. In addition to the aforementioned role played by the number of workers per entrepreneur, this allows me to show how inequality and taxes depend upon the severity of agency frictions, the amount of exogenous uncertainty, and the returns-to-scale of capital in production.

My companion paper Phelan [12] considers an environment in which the productivity (or human capital) of entrepreneurs grows randomly over time and depends (partly) on unobserved effort. The focus of Phelan [12] is the characterization of efficient allocations in an environment with a novel agency problem involving human capital rather than physical capital, with no discussion of how such an allocation may be implemented with exogenously incomplete markets. Although the current paper and Phelan [12] 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 with exogenously incomplete markets.

From a methodological point of view, this paper draws upon the continuous-time contracting literature, and in particular the martingale techniques pioneered in Sannikov [15], for the recursive analysis of the principal-agent problem. The method by which the principal-agent analysis is embedded within a macroeconomic setting follows that outlined in my companion paper Phelan [12], which in turn is an extension of the techniques of Farhi and Werning [7].

The outline of this paper is as follows: Section 2 analyzes a principal-agent model in which both the productivity

[3] Piecewise polynomial on the positive halfline.

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 labour resource constraints, and derives a number of comparative statics results; and Section 4 decentralizes these stationary efficient allocations in a general equilibrium model with exogenously incomplete markets and linear taxes on labour income of workers and the (risk-free) savings and profits of entrepreneurs and plots a numerical example. Technical proofs and a discrete-time version of the environment that relates the findings of the main text with 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 and a risk-neutral principal in an environment where the agent may operate a risky production technology, their consumption is private-information and they may abscond with a fraction of the physical assets under their control. The environment is a slight variation on that considered in Sannikov and Di Tella [16].[4] This principal-agent problem will later be embedded into a macroeconomic model in which flow payoffs to the principal are endogenously determined by aggregate resource constraints. Most of the technical details will be relegated to the appendix.

Formal setup

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 consumer 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 productivity. The principal may delegate capital to the agent so that it may be invested in their technology. When 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

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

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 (exogenous) constant Π in (1) may be interpreted as the marginal product of capital and will be made endogenous in Section 3, while δ denotes the rate of capital depreciation and r the rate at which the principal discounts. 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

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

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

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

[4]The problem considered here is simpler in one respect because savings are observable. However, as Sannikov and Di Tella [16] observe, the principal's problem often gives infinite profits in the absence of hidden savings. Rather than allowing for hidden savings I instead assume the agent may abscond with delegated capital and thereafter trade only a risk-free bond.

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 output 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 consumer may at any time take a fraction ι of the capital delegated to him 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 diverts to private consumption, after every history of output. To be more 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 they alter the law of motion of output and so change the measure used to evaluate the random variables $(K_t, c_t, \tilde{a})_{t\geq 0}$. Denote the measure associated with strategy 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

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

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 $W = \ln K + \ln \rho + [r - \rho]/\rho$.*

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

$$k_t \leq [\iota\rho]^{-1} \exp(1 - r/\rho) \exp W_t =: \omega \exp W_t \tag{4}$$

for all $t \geq 0$ almost surely. 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

$$U^A(K, c, \tilde{a}; a) \geq U^A(K, c, \tilde{a}; \tilde{a})$$

for all agent strategies a 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 $\mathcal{A}^{IC}$.

[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 $t \geq 0$.

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. I may now define the principal's problem. Note that it is indexed by the utility associated with the outside option available to the agent.

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

$$V(W) = \max_{(K,c)\in\mathcal{A}^{IC}} \int_0^\infty e^{-rt}\mathbb{E}[(\Pi - \delta - r)K_t - c_t]dt$$
$$W = \int_0^\infty \rho e^{-\rho t}\mathbb{E}[\ln c_t]dt.$$

As is well-known from the theory of optimal contracting, the principal's problem is naturally recursive in the state variable W, interpreted as promised utility. Standard arguments from the continuous-time contracting literature[7] then ensure that promised utility follows a diffusion process with volatility at least as large as the marginal benefit of diverting output to private consumption. Specifically, the requirement that the allocation (K, c) be incentive compatible may be replaced by the explicit law of motion

$$dW_t = \rho(W_t - \ln c_t)dt + \rho\phi\sigma(k_t/c_t)dB_t. \tag{5}$$

Note that the drift term in (5) is the law of motion of W_t that would obtain in the absence of any uncertainty, as can be seen by simply differentiating the expression (3) with respect to time. The sensitivity of utility to output shocks (the coefficient of the increments of the Brownian motion) 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. Standard arguments then ensure that the value function for the principal solves the following Hamilton-Jacobi-Bellman equation

$$rV(W) = \max_{\substack{c,k\geq 0\\ k\leq\omega\exp W}} [\Pi - \delta - r]k - c + \rho(W - \ln c)V'(W) + \frac{(\rho\phi\sigma k/c)^2}{2}V''(W).$$

The following shows that under certain conditions, the value function of the principal admits a simple closed-form solution.

Theorem 2.2. *The problem of the principal is finite-valued for all sufficiently small* Π. *Further, if the no-absconding constraint does not hold with equality and* $\Pi < \delta + \rho + \sqrt{\rho}\phi\sigma/2$, *then the value function of the principal is of the form* $V(W) = -\Omega(\Pi)\exp W$ *where*

$$\Omega(\Pi) = \frac{1}{\rho[h(\Pi)^2 + 1]}\exp\left(\frac{h(\Pi)^2}{2} + \frac{\rho - r}{\rho}\right)$$

and the policy functions are $c(W) = \underline{c}(\Pi)\exp W$ *and* $k(W) = \underline{k}(\Pi)\exp W$, *where*

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

and for brevity of notation I have defined

$$h(\Pi) := \frac{1 - \sqrt{1 - (2[\Pi - \delta - r]/[\sqrt{\rho}\phi\sigma])^2}}{2[\Pi - \delta - r]/[\sqrt{\rho}\phi\sigma]}.$$

[7] See e.g. Sannikov [15] and Di Tella and Sannikov [16].

Proof. See Appendix B. □

Corollary 2.3. *When the no-absconding constraint does not hold with equality, the volatility of the consumption of the agent is given by* $\sqrt{\rho}h(\Pi)$.

Proof. This is immediate from the policy functions given in Theorem 2.2 and the expression (5) for the law of motion of promised utility. □

Corollary 2.3 establishes the intuitive claim 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 of Section 3 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.

As mentioned in the introduction, a large literature has analyzed constrained-efficient allocations in dynamic environments with privately-observed labour productivity shocks. An important result in this literature, established by Rogerson [14] in a principal-agent setting and extended to a dynamic Mirrleesian environment by Golosov, Kocherlakota and Tsyvinski [9], is the inverse Euler equation for intertemporal distortions. The following shows that this result carries over to the model of the current paper as long as the no-absconding constraint does not hold with equality.

Corollary 2.4. *The stochastic process* $\left(e^{(\rho-r)t}c_t\right)_{t\geq 0}$ *is a martingale in any efficient allocation when the no-absconding constraint does not hold with equality. In other words, the inverse Euler equation holds.*

Proof. From Theorem 2.2 the law of motion of W_t is

$$dW_t = \rho(W_t - \ln c_t)dt + \rho\phi\sigma(K_t/c_t)dB_t = -\rho \ln \underline{c}\; dt + \rho\phi\sigma(\underline{k}/\underline{c})dB_t,$$

Ito's lemma implies that the law of motion of $c_t = \underline{c}\exp W_t$ is then

$$dc_t = \left(-\rho\ln\underline{c} + \frac{1}{2}(\rho\phi\sigma\underline{k}/\underline{c})^2\right)c_t dt + \rho\phi\sigma(\underline{k}/\underline{c})c_t dB_t.$$

Using the expressions for $\underline{c}$ and $\underline{k}$ found in Theorem 2.2, the drift μ_c of c_t becomes

$$\mu_c = -\rho\ln\underline{c} + \frac{1}{2}(\rho\phi\sigma\underline{k}/\underline{c})^2 = -\frac{\rho h(\Pi)^2}{2} + r - \rho + \frac{1}{2}\left(\frac{\rho\phi\sigma h(\Pi)}{\sqrt{\rho}\phi\sigma}\right)^2 = r - \rho$$

as desired. □

The proof of Corollary 2.4 clearly relies on knowledge of the explicit form of the policy functions derived in Theorem 2.2. To gain an intuitive understanding for the emergence of the inverse Euler equation it is therefore instructive to relate Corollary 2.4 with the more direct arguments of Rogerson [14] and Golosov, Kocherlakota and Tsyvinski [9]. Constant-returns-to-scale and homotheticity of preferences imply policy functions of the form $K(W) = \underline{k}\exp W$ and $C(W) = \underline{c}\exp W$ for some $\underline{k}, \underline{c} > 0$. The Hamilton-Jacobi-Bellman equation then reduces to the single equation for the scalar Ω

$$-r\Omega = \max_{\substack{\underline{c}\geq 0 \\ \underline{k}\leq\omega}} (\Pi - \delta - r)\underline{k} - \underline{c} + \rho\Omega\ln\underline{c} - \frac{[\rho\phi\sigma]^2}{2}(\underline{k}/\underline{c})^2\Omega.$$

By the expression (5) for the evolution of promised utility, each choice of $\underline{c}$ and $\underline{k}$ completely specifies the allocation of consumption and capital for any history of output. Given any (not necessarily optimal) scalars $\underline{c}$ and $\underline{k}$, the associated cost Ω per unit of utility $\exp W$ in consumption terms is given by

$$\Omega(\underline{k}, \underline{c}) = \frac{\underline{c} - (\Pi - \delta - r)\underline{k}}{r + \rho \ln \underline{c} - (\rho\phi\sigma\underline{k}/\underline{c})^2/2}.$$

For the optimal $\underline{c} \equiv \underline{c}(\Pi)$ and $\underline{k} \equiv \underline{k}(\Pi)$, define $\overline{\Omega}(u)$ to be the cost associated with the coefficients

$$(\underline{c}(u), \underline{k}(u)) = (\exp(u)\underline{c}, \exp(u)\underline{k}).$$

Note that the allocations $(\underline{c}(u), \underline{k}(u))$ are analogous to the perturbations considered in Rogerson [14] and Golosov, Kocherlakota and Tsyvinski [9] in the sense that they increase flow utility this instant by u utils whilst preserving both promise-keeping and incentive compatibility. By the assumed optimality of $(\underline{c}, \underline{k})$ and the assumption that the no-absconding constraint does not hold with equality, we must have $\overline{\Omega}'(0) = 0$, where $\overline{\Omega}(u)$ is defined to be

$$\overline{\Omega}(u) := \Omega(\underline{c}(u), \underline{k}(u)) = \frac{\exp(u)[(\Pi - \delta - r)\underline{k} - \underline{c}]}{(\rho\phi\sigma\underline{k}/\underline{c})^2/2 - r - \rho \ln \underline{c} - \rho u}.$$

Now note that

$$\overline{\Omega}'(u) = \frac{\exp(u)[(\Pi - \delta - r)\underline{k} - \underline{c}]}{[(\rho\phi\sigma\underline{k}/\underline{c})^2/2 - r - \rho \ln \underline{c} - \rho u]^2}\big((\rho\phi\sigma\underline{k}/\underline{c})^2/2 - r - \rho \ln \underline{c} - \rho u + \rho\big).$$

It follows that $\overline{\Omega}'(u) = 0$, or

$$-\rho \ln \underline{c} + \frac{1}{2}(\rho\phi\sigma\underline{k}/\underline{c})^2 = r - \rho,$$

which is exactly the continuous-time form of 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 $\overline{\Omega}'(0) \geq 0$ and so we still have the inequality $\mu_c \leq r - \rho$.

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 definition of intertemporal wedges adopted below, first note that if an agent may invest in an asset with (gross) return R over the interval $[t, t + \Delta]$, then intertemporal optimization implies

$$u'(c_t) = \exp(-\rho\Delta)\mathbb{E}[Ru'(c_{t+\Delta})|\mathcal{F}_t]. \tag{6}$$

The intertemporal wedges defined in Definition 2.4 measure the extent to which the relation (6) fails for an arbitrary stochastic return.

Definition 2.4. Given the consumption process $(c_t)_{t\geq 0}$ in the principal-agent problem, for each asset A with return process $(R_t^A)_{t\geq 0}$ define the associated wedge ν^A implicitly by

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

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

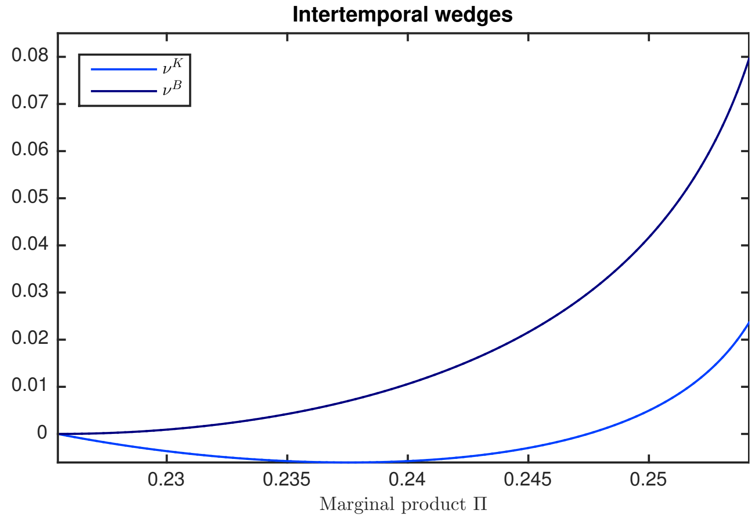

Figure 1: Intertemporal wedges

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

$$R_t^K = \exp\left(\left[\Pi - \delta - \frac{\sigma^2}{2}\right]t + \sigma B_t\right) \qquad\qquad R_t^B = \exp(rt)$$

for all $t \geq 0$. One may also show that the intertemporal wedges are the unique constants ν^K and ν^B such that the solution to the problem

$$\max_{(c_t,k_t)_{t\geq 0}} \rho \int_0^\infty e^{-\rho t}\mathbb{E}[\ln c_t]dt$$
$$da_t = \left[(r - \nu^B)(a_t - k_t) - c_t + (\Pi - \delta - \nu^K)k_t\right]dt + \sigma k_t dB_t$$

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. Combining Corollary 2.4 with Corollary 2.3 shows that consumption may be written explicitly as

$$c_t = c_0 \exp\left(\left[r - \rho - \frac{\rho h(\Pi)^2}{2}\right]t + \sqrt{\rho}h(\Pi)B_t\right).$$

This closed-form expression for consumption allows for a sharp characterization of the intertemporal wedges.

Lemma 2.5. *The intertemporal wedges for risky capital and the risk-free bond are given by*

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

Further, we have the inequalities $\nu^B \geq \nu^K$ and $\nu^B \geq 0$, while the risky wedge ν^K may assume either sign.

Proof. See Appendix B. □

Figure 1 depicts the intertemporal wedges for both risky capital and the risk-free bond as a function of the marginal product Π, given the parameters:

$$\rho = r = 0.145 \qquad \phi = 0.5 \qquad \sigma = 0.3 \qquad \delta = 0.058.$$

As noted in Lemma 2.5, 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.

Throughout this paper I will restrict attention to parameters for which the no-absconding constraint does not hold with equality. The following outlines sufficient conditions for this assumption to be satisfied.

Lemma 2.6. *The no-absconding constraint will hold as a strict inequality whenever* $2(\Pi-\delta-r) \leq \sqrt{\rho}\phi\sigma$, $\underline{k}(\Pi) < \omega$ *and*

$$0 > \max_{c\geq 0} \ (\rho \ln c + r)\Omega(\Pi) - c + (\Pi - \delta - r)\omega - \frac{1}{2}\Omega(\Pi)[\rho\phi\sigma\omega]^2 c^{-2}. \tag{7}$$

Proof. See Appendix B. ☐

Note that the condition (7) is necessarily violated for sufficiently large ω. For instance, if one sets $c = (\Pi - \delta - r)\omega$ then the right-hand side of (7) becomes

$$(\rho \ln \omega + \rho \ln (\Pi - \delta - r) + r)\Omega(\Pi) - \frac{\Omega(\Pi)}{2}\left(\frac{\rho\phi\sigma}{\Pi - \delta - r}\right)^2$$

which diverges to infinity as ω becomes arbitrarily large. It is also very important to note that although there is a range of parameters such that the no-absconding constraint does not hold with equality, such a constraint is *always* necessary to ensure that the problem of the principal is finite-valued.[9] This is easy to see by setting $\underline{k}(\omega) = \omega$ and $\underline{c} = \underline{c}(\omega)$ solving

$$\rho\underline{c}(\omega)^2 \ln \underline{c}(\omega) = \frac{1}{2}(\rho\phi\sigma\underline{k})^2 = \frac{1}{2}(\rho\phi\sigma\omega)^2 \tag{8}$$

and noting that the associated cost coefficient is

$$\Omega(\underline{c}(\omega), \underline{k}(\omega)) = \frac{\underline{c} - (\Pi - \delta - r)\underline{k}}{r + \rho \ln \underline{c} - (\rho\phi\sigma\underline{k}/\underline{c})^2/2} = \underline{c}(\omega)\rho^{-1}\left(1 - \left(\frac{\Pi - \delta - r}{\rho\phi\sigma}\right)\sqrt{2\rho \ln \underline{c}(\omega))}\right)$$

which is negative for sufficiently large ω. If the value function of the principal is convex and increasing in W (as is the case if the cost coefficient is negative) then they may make arbitrarily large expected profits by first offering the agent a lottery over different levels of promised utility, and so their problem fails to be finite-valued.

What lessons to take out of this partial equilibrium analysis before turning to the characterization of stationary efficient allocations? The capital assignment $\underline{k}(\Pi)$, consumption $\underline{c}(\Pi)$ and the capital/consumption ratio $\underline{k}(\Pi)/\underline{c}(\Pi)$ (and hence risk borne by the agent) are increasing in the marginal product of capital Π. In the following section this value of Π will be determined by the physical resource constraints of an economy with a continuum of agents.

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.

[9] This was observed by Di Tella and Sannikov [16], who show that without the no-absconding constraint the principal's problem is well-defined if and only if the elasticity of intertemporal substitution exceeds 2 (and here it is 1).

Formal setup

Time is again indefinite and continuous. At any moment there is a continuum of agents who do not care for their descendants and die at constant rate ρ_D. Agents may engage in one of two activities: one subject to idiosyncratic risk and requiring special (innate) ability, and one subject to no idiosyncratic risk that may be performed by anyone. The first activity is identified with entrepreneurial activity (running a business), and the second with wage labour (working for someone else). A mass L_E of the agents may engage in the entrepreneurial activity and the remaining mass L_W cannot, and this ability is private information at birth. For simplicity I will assume that once an agent is assigned to one of these activities they lose the ability to engage in the other. All agents have preferences over sequences $(c_t)_{t\geq 0}$ represented by

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

Agents who engage in the entrepreneurial activity have access to a risky production technology that produces consumption using physical capital and labour. Specifically, if an entrepreneur assigns capital and labour to their technology according to the processes $(K_t, L_t)_{t\geq 0}$, then the law of motion of physical capital is given by the following

$$dY_t = \big(AK_t^\alpha L_t^{1-\alpha} - \delta K_t\big)dt + \sigma K_t dB_t$$

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 pairs (v,i) of promised utility and type, where $i \in \{E, W\}$ denotes whether or not an agent is born with the ability to be an entrepreneur or worker. The formal definition is then the following.

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

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

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

Incentive compatibility for an allocation now imposes the additional requirement that promises made to the initial generation be satisfied and that entrepreneurs be given an incentive to reveal their private information at birth.

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

Feasibility is defined in terms of aggregate quantities C, Y and L associated with the allocations. Denote by $C_t(A), Y_t(A)$ and $L_t(A)$ the aggregate amount of consumption, output and labour assigned at date t given the allocation A. Formal definitions are given in Appendix B.2.

Definition 3.3. An allocation A is resource feasible if $C_t(A) \leq Y_t(A)$ and $L_t(A) \leq L_W$ for all $t \geq 0$. The set of such allocations will be denoted $\mathcal{A}^{RF}$. An allocation is incentive feasible given Φ if it is both resource feasible and incentive compatible given Φ. The set of all such allocations will be denoted $\mathcal{A}^{IF}(\Phi)$.

I will assume the planner places a Pareto weight $\alpha(T,i) = \Gamma_i e^{-\rho T}$ on the agents' utility where $T \geq 0$ denotes date of birth and $i \in \{E, W\}$ denotes whether or not the agent is a worker or entrepreneur. This specification ensures that the planner values the utility experienced by an agent at any given date the same regardless of their date of birth. This welfare criterion may be viewed as a kind of generalized utilitarianism across generations and is equivalent to assuming the following social welfare function:

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

where $\underline{U}_t$ and U_t^T refer to Pareto-weighted aggregate flow utility experienced by the initial and Tth generations at date $t \geq 0$.[10] I may now specify the planning problem.

Definition 3.4. Given an initial distribution Φ, the problem of the planner is given by

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

The problem defined in Definition 3.4 is intractable for an arbitrary initial distribution so I will restrict attention to solutions in which aggregate consumption and output are constant over time. I will characterize such distributions using the method outlined in Farhi and Werning [7] and employed in my companion paper Phelan [12], and consider, in succession, *relaxed* and *generational* planner problems. The relaxed planner problem has the same objective and state variable as in the above planner problem, but allows for intertemporal trade in goods and labour at the subjective rate of discount ρ.

Definition 3.5 (Relaxed planner's problem)**.** Given an initial distribution Φ over promised utility and types, the relaxed planner's problem is defined to be

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

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

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

Note that if an allocation solves the relaxed planner problem and the associated implied distributions of promised utility and types are constant over time at Φ, then this allocation also solves the original planner problem given the distribution Φ. Further, it is easy to see that the subjective rate of discount ρ 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. Lagrange's Theorem 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

$$V_\lambda(\Phi) = \max_{A \in \mathcal{A}^{IC}} \int_0^\infty e^{-[\rho+\rho_D]t} \left(\underline{U}_t + \lambda_R \left[\underline{Y}_t - \underline{C}_t + \lambda_L \left[L_W - \underline{L}_t^E \right] \right] \right) dt$$
$$+ \int_0^\infty \int_0^t e^{-\rho T} e^{-[\rho+\rho_D][t-T]} \Big(U_t^T + \lambda_R \Big[Y_t^T - C_t^T + \lambda_L \Big[L_W - L_t^{E,T} \Big] \Big] \Big) dT dt$$

where the triples $(\underline{C}_t, \underline{Y}_t, \underline{L}_t^E)$ and $(C_t^T, Y_t^T, L_t^{E,T})$ refer to consumption, output and labour assignments of initial and Tth generations, respectively, at date $t \geq 0$.[11] Although the state variable is still an entire distribution, the objective $V_\lambda(\Phi)$ may be maximized pointwise because all interdependence across agents is captured by the multipliers. One may then treat each generation in isolation and vary the multipliers λ_L and λ_R 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 problem*.

[10] Again formal definitions are given in Appendix B.2.
[11] Detailed definition are found in Appendix B.2.

Definition 3.6. The problem of a generational planner given multipliers $\lambda := (\lambda_R, \lambda_L)$ is defined to be

$$V^G_\lambda = \max_{A \in \mathcal{A}^{IC}_G} \int_0^\infty e^{-[\rho+\rho_D]t} \big(\underline{U}_t + \lambda_R \big[\underline{Y}_t - \underline{C}_t + \lambda_L \big[L_W - \underline{L}^E_t \big] \big] \big) dt.$$

The choice of assigning labour to entrepreneurs is purely static and depends solely upon the multiplier 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 Section 2 with the marginal product of capital now a function of the multiplier on the labour resource constraint.

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

$$V^G_\lambda = \max_{\substack{W_W, W_E \\ W_E \geq W_W}} \Gamma_E L_E W_E + \Gamma_W L_W W_W + \lambda_R \big(L_E V(W_E, \Pi(\lambda_L)) + L_W \rho^{-1} [\lambda_L - \exp W_W] \big)$$

where $\Pi(\lambda_L) := \max_{L \geq 0} A l^{1-\alpha} - \lambda_L l = \alpha(1-\alpha)^{-1} [A(1-\alpha)]^{1/\alpha} \lambda_L^{1-1/\alpha}$.

I will assume that the weight on the utility of workers is sufficiently large that the type-revelation constraint $W_E \geq W_W$ holds with equality. Each choice of λ_R then corresponds to a level of utility promised to newborns and each choice of λ_L corresponds to a level of the marginal product of capital $\Pi \equiv \Pi(\lambda_L)$ in the principal-agent problem. One may then solve for the stationary distributions associated with each pair of multipliers by using the policy functions found in Section 2. Theorem 3.2 is the first main result of this paper. It shows that determining the multipliers for which stationarity obtains reduces to finding an appropriate level of the marginal product of capital Π.

Theorem 3.2. *The stationary level of Π is the solution to the equation*

$$\underline{c}(\Pi) + L_W / L_E = (\Pi/\alpha - \delta) \underline{k}(\Pi),$$

provided the no-absconding constraint does not hold with equality for this Π. The associated multiplier λ_L, level of the capital stock K and initial level of utility W are then

$$\lambda_L = (1-\alpha) A^{\frac{1}{1-\alpha}} (\Pi/\alpha)^{-\frac{\alpha}{1-\alpha}} \qquad K = (\alpha A / \Pi)^{\frac{1}{1-\alpha}} L_W \qquad W = \ln \left(\frac{K/L_E}{\underline{k}(\Pi)} \right).$$

Proof. We wish to characterize the values of λ_L and λ_R such that the labour and goods resource constraints are satisfied in the stationary distributions implied by the solutions to the generational planner problem. First note that by Corollary 2.4 the process $(\exp W_t)_{t \geq 0}$ has no drift, and so its mean is simply $\exp W$, where W is the initial level of utility. It follows that aggregate capital may be written

$$K = L_E \underline{k}(\Pi(\lambda_L)) \exp W. \tag{9}$$

Combining (9) with the static labour assignment function $l(\lambda_L) = [A(1-\alpha)]^{1/\alpha} \lambda_L^{-1/\alpha}$ the labour and goods resource constraints become

$$L_W = L_E [A(1-\alpha)]^{1/\alpha} \lambda_L^{-1/\alpha} \underline{k}(\Pi(\lambda_L)) \exp W$$
$$(L_W + L_E \underline{c}(\Pi(\lambda_L))) \exp W = L_W \lambda_L + L_E \left(\frac{\alpha}{1-\alpha} [A(1-\alpha)]^{1/\alpha} \lambda_L^{1-1/\alpha} - \delta \right) \underline{k}(\Pi(\lambda_L)) \exp W.$$

Simplifying the above resource constraints gives the desired equation for Π. The expressions for the multiplier and the stationary level of the capital stock then follow by combining (9) with the static labour assignment function. ☐

How does the above depend upon the elasticity between labour and capital? My intuition is that nothing more comes from this beyond the claim that anything that makes the marginal product of capital higher will increase inequality in the stationary distribution. I suspect the following: the degree of risk borne always increases with the number of workers, but is lower when the elasticity is higher. Perhaps the better intuition is the following: anything that increases the capital stock (or capital per entrepreneur) in the efficient allocation will increase the amount of risk borne. Careful: risk borne depends on consumption too.

Note that the simplicity of the characterization given in Theorem 3.2 is due partly to the assumption of logarithmic preferences and partly to the welfare criterion adopted in this paper that weights flow utility of an agent the same independently of their birth date. Other papers in the literature on dynamic economies with private information, such as Atkeson and Lucas [3] or Phelan [13], consider component planner problems similar to the above generational planner problems, but adopt a welfare criterion with either zero discounting or place weight solely upon the first generation. Such an approach necessitates solving a component planning problem for an arbitrary interest rate which is then varied until resources are balanced. In contrast, with the welfare criterion of this paper is it immediate that the only intertemporal rate for which stationarity may arise is the subjective discount rate ρ of the agents, for all other intertemporal rates would induce a trend in utility. Together with the assumption of logarithmic utility and the implication derived in Corollary 2.4 that consumption then follows a martingale, this implies changes in technology have no effect on the trend in consumption and affect the constrained-efficient allocation only insofar as they alter the marginal product of capital and hence risk borne by entrepreneurs.

Now, the homogeneity of the planner's policy functions for both capital and consumption allows for a simple characterization of the stationary distribution of consumption associated with the above constrained-efficient allocation. In general, the stationary distribution of a killed geometric Brownian motion will depend on the average growth rate, the volatility and the hazard rate of death. Since the rate of death is exogenous, and Corollary 2.4 shows the growth rate of consumption to be zero independent of all technological parameters, it follows that the stationary distribution is solely determined by the risk $\sqrt{\rho}h(\Pi)$ borne by entrepreneurs. When combined with the defining equation for Π given in Theorem 3.2 this in turn allows us to determine how changes in technological parameters affect efficient long-run inequality. First note that by (9) and the expression for capital in Theorem 3.2 the initial level of the consumption of entrepreneurs is given by

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

Combining the above observations 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

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

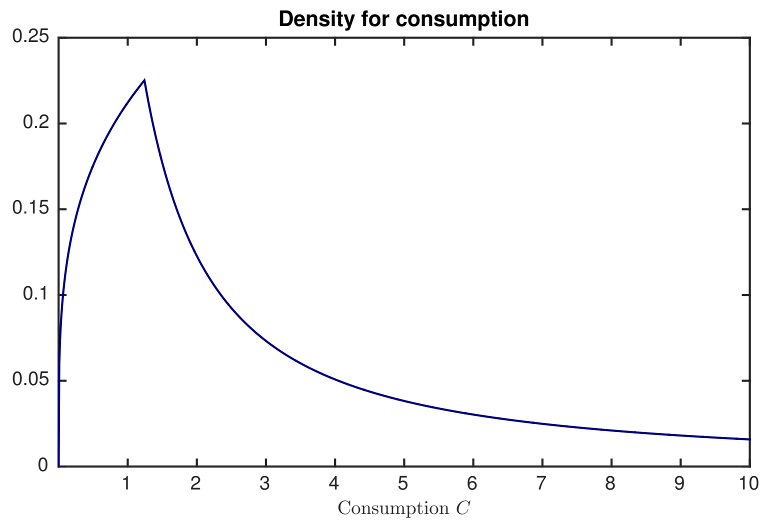

Figure 2: Density of consumption

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

Figure 2 depicts the stationary distribution of consumption for the following parameters:

$$\phi = 0.5 \qquad \delta = 0.058 \qquad L_E = 0.115 \qquad L_W = 0.885$$
$$\alpha = 0.33 \qquad \sigma^2 = 0.3 \qquad \rho = 0.145 \qquad \rho_D = 0.022.$$

In this example, 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 [4], while the death rate ρ_D is chosen such that the average lifespan is 45 years (interpreted as working lifespan). The volatility term σ^2 is the average of the two values considered by Angeletos [2]. Note that the upper tail of the stationary distribution in this example has Pareto exponent $\beta_- \approx -1.47$, which is close to the exponent $\beta_- \approx -1.54$ of the wealth distribution US estimated by Gabaix, Moll, Lasry and Lions [8] from the 2010 wave of the Survey of Consumer Finances.

Analysis of the defining equation for Π given in Theorem 3.2 allows us determine how both aggregate capital and the degree of inequality vary in response to changes in technological parameters.

Corollary 3.4. *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 ratio L_W/L_E, and the factor share α.*

Proof. Using the change of variables

$$x(\Pi) := \frac{2[\Pi - \delta - \rho]}{\sqrt{\rho}\phi\sigma},$$

the defining equation for Π may be rewritten in terms of x,

$$L_E + L_W \exp\left(\frac{1}{2} - \left(1 - \sqrt{1-x^2}\right)x^{-2}\right) = \frac{1}{\alpha}\left(1 - \sqrt{1-x^2}\right)\left(\frac{1}{2} + \frac{(1-\alpha)\delta + \rho}{\sqrt{\rho}\phi\sigma x}\right)L_E. \tag{10}$$

It is easy to check that the left- and right-hand sides of (10) are decreasing and increasing in x, respectively. The claim that Π increases with $\phi\sigma$ then follows from the fact that the right-hand side of (10) is decreasing in $\phi\sigma$.

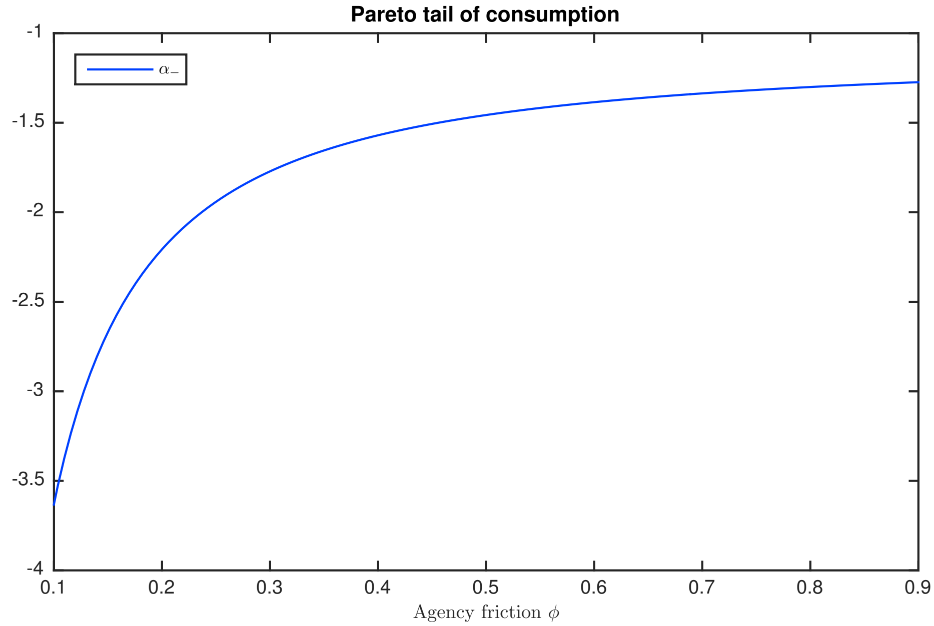

Figure 3: Pareto parameter as function of agency problem

Further, since the volatility $\sqrt{\rho}h(\Pi)$ may be written as an increasing function of $x(\Pi)$, this also shows that the Pareto exponent β_- is increasing in $\phi\sigma$.

To establish the assertions regarding the population ratio L_W/L_E and factor share α, first note that we can write the defining equation for Π as follows

$$0 = \frac{\underline{c}(\Pi) + L_W/L_E}{\underline{k}(\Pi)} - \frac{\Pi}{\alpha} + \delta. \tag{11}$$

The right-hand side of (11) is decreasing in Π and increasing in both L_W/L_E and α, which shows that the associated roots are also increasing in these variables. The remaining claims then follow from the fact that the volatility $\sqrt{\rho}h(\Pi)$ is increasing in Π. □

Notice that the efficient allocation characterized in this section depends on the severity of the agency problem ϕ and the level of exogenous uncertainty σ only through the product $\phi\sigma$. Increasing ϕ is therefore equivalent to increasing σ, and vice versa. However, this is not the case in the decentralization of the following section, in which ϕ and σ play qualitatively distinct roles in the determination of taxes.

To illustrate the importance of agency frictions, Figure 3 plots the Pareto parameter β_- as a function of the parameter ϕ, with all other parameters given by the example distribution plotted in Figure 2. The domain has been restricted to [0.1 0.9] because the decay in consumption diverges to $-\infty$ as $\phi \to 0$ and the no-absconding constraint begins to bind around 0.9. Similarly, Figure 4 plots the Pareto parameter of the upper tail as a function of the ratio L_W/L_E indicating the number of workers per entrepreneur. As with Figure 3, all other parameters are fixed at those given in the example depicted in Figure 2.

Before turning to the question of 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

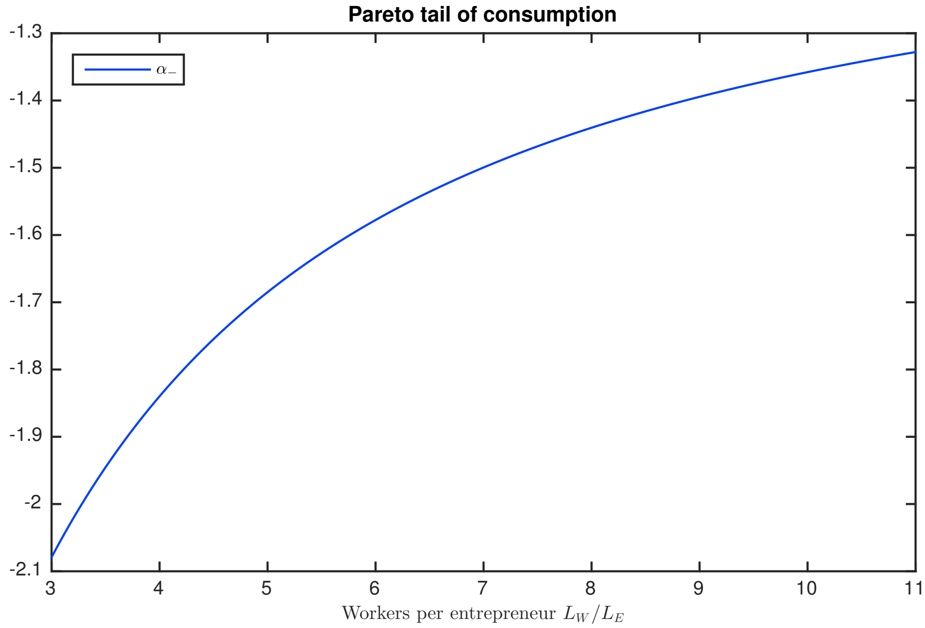

Figure 4: Pareto parameter as function of worker per entrepreneur

W units of utility, workers and entrepreneurs have zero drift in consumption, and consumption of entrepreneurs has volatility equal to $\sqrt{\rho}h(\Pi)$, where Π is given in Theorem 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

This section lays out a general equilibrium model with (irreversible) occupational choice and idiosyncratic production risk with exogenously incomplete markets. I characterize stationary competitive equilibria and show how the constrained-efficient allocation of the previous section may be implemented using linear taxes on profits, risk-free savings and labour income. To be consistent with the environment described in previous sections I will again suppose that at birth all agents make an irreversible (once-and-for-all) decision to either run their own business or work for another person's business.

Market structure and equilibrium characterization

I will assume that all agents are endowed with occupation-specific levels of wealth at birth and may save in physical capital that earns a risk-free return. Entrepreneurs in turn rent capital on behalf of their business at the same rate, are unable to issue shares in the profits of their business and pay taxes on profits net of both depreciation and interest payments.

At any instant the labour-hiring decision is purely static and so may be solved independently of all savings and investment decisions. To this end define $\Pi \equiv \Pi(w) := \max_{l\geq 0}\ Al^{1-\alpha} - wl$ and note that from the point of view of the entrepreneur production is linear in capital with marginal product Π. The problems of agents facing constant linear prices are given as follows.

Definition 4.1. Given taxes $\tau_{LW}, \tau_{sW}, \tau_{sE}$ and τ_Π on labour income, risk-free savings and profits, the wage w, and risk-free rate r, the problems of an entrepreneur and worker endowed with a units of assets are, respectively,

$$\begin{aligned} V_E(a) &= \max_{(c_t,k_t)_{t\geq 0}} \rho \int_0^\infty e^{-\rho t}\mathbb{E}[\ln c_t]dt \\ da_t &= [(1-\tau_{sE})ra_t - c_t]dt + (1-\tau_\Pi)k_t([\Pi - \delta - r]dt + \sigma dB_t) \\ V_W(a) &= \max_{(c_t)_{t\geq 0}} \rho \int_0^\infty e^{-\rho t}\mathbb{E}[\ln c_t]dt \\ da_t &= [(1-\tau_{sW})ra_t - c_t]dt + (1-\tau_L)wdt. \end{aligned}$$

To understand the market structure and timing assumptions implicit in the above laws of motion of wealth it is instructive to consider a discrete-time analogue of this environment. Suppose that at the beginning of the nth period an entrepreneur is equipped with a_n units of (physical) wealth. They rent this wealth to businesses who agree to pay back $(1+\Delta r)a_n$ units at the end of the day, at which time they must pay a savings tax on the interest payments Δr. At the same time, the entrepreneur, on behalf of his private firm, rents k_n units of capital for the day and hires labour at wage Δw. Production takes place during the day. At the end of the day a fraction $\Delta\delta$ of the capital has depreciated and the entrepreneur pays back $(1+\Delta r)k_n$ to the bank. All agents eat at the end of the day. Some agents die during the night, and those remaining repeat the above scenario.

The pre-tax profits net of depreciation and borrowing costs when the wage is w, the interest rate is r and there is no uncertainty are then

$$\Delta\left(\max_{l\geq 0} Ak_n^\alpha l^{1-\alpha} - wl\right) + (1-\delta\Delta)k_n - (1+\Delta r)k_n = \Delta[\Pi - \delta - r]k_n. \tag{12}$$

The law of motion in Definition 4.1 is then obtained by supposing that in addition to the above, production is also subject to proportional i.i.d. shocks, with the profits tax being levied on output net of rental costs and depreciation.

Finally, I will assume that newborn entrepreneurs and workers inherit $\eta_E K$ and $\eta_W K$ at birth, respectively, where K is the aggregate capital stock and η_E and η_W are chosen by the government. At death, all of an agent's wealth is returned to the government. The following notion of competitive equilibrium is standard.

Definition 4.2. Given taxes $(\tau_{LW}, \tau_{sW}, \tau_{sE}, \tau_\Pi)$ on labour income, risk-free savings and profits of entrepreneurs and workers, respectively, a stationary competitive equilibrium consists of inheritance levels η_E and η_W for entrepreneurs and workers, an aggregate capital stock K, wage rate w and risk-free rate r such that agents maximize, the markets for labour, capital and goods clear, and the government budget constraint is satisfied.

As with the agency problem considered earlier, the homotheticity of preferences and the log-linearity of the evolution of wealth ensures that both the worker and entrepreneur problems admit homogeneous solutions for any choice of (linear) taxes.

Lemma 4.1. *Given the wage w, risk-free rate r and depreciation δ, and taxes $\tau_{LW}, \tau_{sW}, \tau_{sE}$ and τ_Π, the value functions for entrepreneurs and workers are given by*

$$\begin{aligned} V_E(a) &= \ln\rho + \ln a + \rho^{-1}\left(\mu_E(w,r) - \frac{\sigma_E(w,r)^2}{2}\right) \\ V_W(a) &= \ln\rho + \ln\left(a + \frac{(1-\tau_{LW})w}{(1-\tau_{sW})r}\right) + \rho^{-1}\mu_W(r) \end{aligned}$$

where μ_E and σ_E denote the drift and diffusion in the wealth of entrepreneurs

$$\mu_E(w,r) = (1-\tau_{sE})r - \rho + \left(\frac{\Pi(w)-\delta-r}{\sigma}\right)^2 \qquad \sigma_E(w,r) = \frac{\Pi(w)-\delta-r}{\sigma} \tag{13}$$

and $\mu_W(r) = (1-\tau_{sW})r - \rho$ denotes the drift in worker wealth. The policy functions for consumption are

$$c_E(a) = \rho a \qquad c_W(a) = \rho\left(a + \frac{(1-\tau_{LW})w}{(1-\tau_{sW})r}\right)$$

and the policy function for capital is

$$\frac{k_E(a)}{a} = \frac{\Pi(w)-\delta-r}{\sigma^2(1-\tau_\Pi)} =: \overline{k}(w,r). \tag{14}$$

Proof. See Appendix C.1. □

Lemma 4.1 implies that the laws of motion of wealth for entrepreneurs and workers are given by

$$\begin{aligned} da_t^E &= \mu_E(w,r)a_t^E dt + \sigma_E(w,r)a_t^E dB_t \\ da_t^W &= \mu_W(r)\left(a_t^W + \frac{(1-\tau_{LW})w}{(1-\tau_{sW})r}\right)dt. \end{aligned} \tag{15}$$

Theorem 4.2 characterizes stationary equilibrium prices and aggregate quantities given arbitrary levels of taxes and inheritance of entrepreneurs.

Theorem 4.2. *When the inheritance levels for workers and entrepreneurs are given by η_W and η_E, respectively, and taxes are $(\tau_\Pi, \tau_{sE}, \tau_{sW}, \tau_{LW})$, the wage w and risk-free rate r that prevail in the associated stationary equilibrium solve the pair of equations*

$$\begin{aligned} 1 &= \frac{\rho_D \eta_E L_E \overline{k}(w,r)}{\rho_D - \mu(w,r)} \\ \frac{(1-\tau_{sW})r}{1-\tau_{LW}} &= \frac{\rho(1-\alpha)\Pi(w)}{\Pi(w)-\alpha(\delta+\rho)}. \end{aligned} \tag{16}$$

The associated capital stock is given by $K = [(1-\alpha)A/w]^{-1/\alpha}$.

Proof. There are three markets that must clear in the stationary competitive equilibrium: labour, capital and goods. It is convenient to first define κ to be the fraction of the aggregate capital stock owned by entrepreneurs, as this quantity enters each of the market-clearing equations. If each entrepreneur inherits a multiple η_E of the aggregate capital stock K then the fraction of wealth entrepreneurs own in the stationary distribution is

$$\kappa = \frac{\rho_D \eta_E L_E}{\rho_D - \mu_E(w,r)}$$

where μ_E is the drift in entrepreneurial wealth given by (13). The stationary form of the three market clearing conditions for labour, capital and goods may then be written

$$\begin{aligned} L_W &= \phi_l(w)\overline{k}(w,r)\kappa K \\ 1 &= \kappa \overline{k}(w,r) \\ (A\phi_l(w)^{1-\alpha} - \delta)\overline{k}(w,r)\kappa K &= \rho\left(K/L_W + \frac{(1-\tau_{LW})w}{(1-\tau_{sW})r}\right)L_W, \end{aligned}$$

respectively. Using the above expression for κ, the capital market-clearing equation reduces to the first equation in (16). Using this, the goods market clearing condition simplifies to

$$(A\phi_l(w)^{1-\alpha} - \delta - \rho)[K/L_W] = \frac{\rho(1-\tau_{LW})w}{(1-\tau_{sW})r}$$

and hence

$$\rho(1-\tau_{LW})(1-\alpha)A[K/L_W]^{\alpha-1} = (1-\tau_{sW})r(A[K/L_W]^{\alpha-1}-\delta-\rho).$$

which reduces to the second equality in (16). Since $\phi_l(w) = [(1-\alpha)A/w]^{1/\alpha}$, the labour market clearing condition becomes $[(1-\alpha)A/w]^{1/\alpha}K = L_W$ and so the wage is simply the marginal product of labour, $w = (1-\alpha)A[K/L_W]^{\alpha}$, which gives the desired expression for the aggregate capital stock. The government budget constraint is automatically satisfied by Walras' law. □

Decentralization

We now turn to the characterization of those taxes and inheritance share that decentralize the efficient allocation. First recall the constrained-efficient allocation is characterized by six conditions: workers have zero drift in consumption; entrepreneurs have zero drift in consumption; the volatility of the consumption of entrepreneurs is equal to $\sqrt{\rho}h(\Pi)$; workers and entrepreneurs both have lifetime utility given by W in Theorem 3.2; and the level of capital (equivalently, marginal product of capital) in the competitive equilibrium coincides with that given in the stationary efficient allocation.

The constrained-efficient allocation was characterized by the solution to a single equation in the marginal product of capital, Π, and so it will be convenient to express all aggregate quantities in the incomplete markets model in terms of this variable. Since the wage is simply the marginal product of labour, rearrangement gives

$$w = (1-\alpha)A^{\frac{1}{1-\alpha}}[\alpha/\Pi]^{\frac{\alpha}{1-\alpha}}, \qquad K/L_W = [A\alpha/\Pi]^{\frac{1}{1-\alpha}}. \tag{17}$$

The following is the second main result of this paper, after the characterization of the stationary efficient allocation given in Theorem 3.2. It characterizes the inheritance level and taxes that decentralize the stationary constrained-efficient allocation found in the previous section. Note the remarkable fact that the tax on profits is independent of all technological and demographic parameters and depends solely upon the parameter ϕ.

Theorem 4.3. *The stationary constrained-efficient allocation coincides with the stationary competitive equilibrium allocation in which the inheritance levels η_E and η_W are given by*

$$\eta_E = \frac{\phi\sigma}{L_E\sqrt{\rho}h(\Pi)}, \qquad \eta_W = \frac{1}{L_W}\left(1-\frac{\phi\sigma}{\sqrt{\rho}h(\Pi)}\right)$$

where Π denotes the marginal product of capital that obtains in the stationary efficient allocation and characterized in Theorem 3.2. The profits tax is given by $\tau_\Pi = 1-\phi$, the tax on the labour income of workers is given by

$$\tau_{LW} = \frac{\alpha(\delta+\rho-\Pi)}{(1-\alpha)\Pi},$$

and the taxes on the savings of workers and entrepreneurs satisfy

$$\tau_{sW} = 1-\frac{\rho}{\Pi-\delta-\sqrt{\rho}\sigma h(\Pi)} \qquad \tau_{sE} = 1-\frac{\rho[1-h(\Pi)^2]}{\Pi-\delta-\sqrt{\rho}\sigma h(\Pi)}.$$

Finally, the associated wage and interest rate are

$$w = (1-\alpha)A^{\frac{1}{1-\alpha}}[\alpha/\Pi]^{\frac{\alpha}{1-\alpha}} \qquad r = \Pi-\delta-\sqrt{\rho}\sigma h(\Pi).$$

Proof. Decentralizing the efficient allocation with taxes and inheritance level ultimately amounts to solving a system of eight equations in eight unknowns. The eight equations are the six conditions characterizing the efficient allocation and the two market-clearing conditions from Theorem 4.2, while the eight unknowns are the four taxes, two inheritance levels, the interest rate r and wage rate w.

First note that the ratio of capital to consumption in the equilibrium must coincide with the corresponding ratio in the constrained-efficient allocation, so

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

Further, in order for the volatility of consumption to be equal across the competitive equilibrium and constrained-efficient allocation we need

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

Since $\overline{c}(w,r) = \rho$, combining (18) and (19) implies $1-\tau_\Pi = \phi$. Next, the marginal product of capital must coincide with the level Π in the constrained-efficient allocation, which immediately gives the desired expression for the wage w. From Lemma 4.1 we know that the volatility of entrepreneurs' wealth is independent of taxes and given by

$$\sigma_E(r) = \frac{\Pi - \delta - r}{\sigma}.$$

The desired expression for the interest rate then follows by equating $\sigma_E(r)$ with that given in the constrained-efficient allocation, $\sqrt{\rho}h(\Pi)$. The taxes on the risk-free savings of workers and entrepreneurs are determined by combining this expression for the risk-free rate with the requirement that both the respective drifts in consumption vanish. For workers this implies

$$1-\tau_{sW} = \frac{\rho}{r} = \frac{\rho}{\Pi - \delta - \sqrt{\rho}\sigma h(\Pi)}$$

as claimed, and for entrepreneurs this implies

$$\rho - (1-\tau_{sE})r = \rho - (1-\tau_{sE})(\Pi - \delta - \sqrt{\rho}\sigma h(\Pi)) = \rho h(\Pi)^2,$$

which reduces to the claimed expression for τ_{sE}. Now recall that by Theorem 3.2 the utility attained in the constrained-efficient allocation is given by

$$W = \ln\left(\frac{K/L_E}{\underline{k}(\Pi)}\right) = \ln\rho + \ln\left(\frac{[K/L_E]\phi\sigma}{\sqrt{\rho}h(\Pi)}\right) - \frac{h(\Pi)^2}{2}. \tag{20}$$

From Lemma 4.1 if $\mu_E(w,r) = 0$ and $\sigma_E(w,r) = \sqrt{\rho}h(\Pi)$ then the utility entrepreneurs experience from $\eta_E K$ units of wealth is

$$W_{CE} = \ln\rho + \ln(\eta_E K) + \rho^{-1}\left(\mu_E(w,r) - \frac{\sigma_E(w,r)^2}{2}\right) = \ln\rho + \ln(\eta_E K) - \frac{h(\Pi)^2}{2} \tag{21}$$

which coincides with (20) if and only if η_E is given by the above expression. By again using the fact that $(1-\tau_{sW})r = \rho$, the goods market-clearing condition from Theorem 4.2 becomes

$$\frac{\rho}{1-\tau_{LW}} = \frac{\rho(1-\alpha)\Pi}{\Pi - \alpha(\delta+\rho)},$$

which gives the claimed expression for the labour tax. Finally, we show that for the above taxes entrepreneurs and workers attain the same level of promised utility. First note that by (17), the present discounted value of (after-tax) wages simplifies to

$$\frac{(1-\tau_{LW})w}{(1-\tau_{sW})r} = \rho^{-1}(1-\tau_{LW})(1-\alpha)A^{\frac{1}{1-\alpha}}[\alpha/\Pi]^{\frac{\alpha}{1-\alpha}}. \tag{22}$$

Combined with the earlier expression for τ_{LW} this gives

$$\frac{(1-\tau_{LW})w}{(1-\tau_{sW})r} = \rho^{-1}A^{\frac{1}{1-\alpha}}[\alpha/\Pi]^{\frac{\alpha}{1-\alpha}}\left(\frac{\Pi-\alpha(\delta+\rho)}{\Pi}\right) = \rho^{-1}[A\alpha/\Pi]^{\frac{1}{1-\alpha}}(\Pi/\alpha-\delta-\rho).$$

and so the value function for the worker simplifies to

$$V_W(a) = \ln\rho + \ln\left(a + \rho^{-1}[A\alpha/\Pi]^{\frac{1}{1-\alpha}}(\Pi/\alpha-\delta-\rho)\right).$$

Using the expression for aggregate capital given in (17), individual workers are endowed at birth with $(1-\eta_E L_E)K/L_W = (1-\eta_E L_E)[A\alpha/\Pi]^{\frac{1}{1-\alpha}}$, and so their inherited wealth plus the present discounted value of labour income is

$$\begin{aligned}
\big(1-\eta_E L_E + \rho^{-1}(\Pi/\alpha-\delta-\rho)\big)K/L_W &= \big(-\eta_E L_E + \rho^{-1}(\Pi/\alpha-\delta)\big)K/L_W \\
&= \left(-\frac{\phi\sigma}{\sqrt{\rho}h(\Pi)} + \rho^{-1}(\Pi/\alpha-\delta)\right)K/L_W \\
&= \left(-1 + \frac{h(\Pi)}{\sqrt{\rho}\phi\sigma}(\Pi/\alpha-\delta)\right)\left(\frac{\phi\sigma K/L_W}{\sqrt{\rho}h(\Pi)}\right).
\end{aligned}$$

Using the defining equation for Π given Theorem 3.2 this simplifies to

$$\big(-1+1+\exp(-h(\Pi)^2/2)L_W/L_E\big)\left(\frac{\phi\sigma K/L_W}{\sqrt{\rho}h(\Pi)}\right) = \frac{\phi\sigma K/L_E}{\sqrt{\rho}h(\Pi)\exp(h(\Pi)^2/2)}.$$

Substituting this expression for total wealth into the value function of the worker gives

$$V_W((1-\eta_E L_E)K/L_W) = \ln\left(\frac{\sqrt{\rho}\phi\sigma K/L_E}{h(\Pi)\exp(h(\Pi)^2/2)}\right)$$

which coincides with the expression (20) for $V_E(\eta_E K)$ given above. □

It is instructive to check that the government budget constraint is automatically satisfied in the equilibrium characterized in Theorem 4.3. Summing the tax revenue collected from entrepreneurs and workers gives

$$T = \tau_{sW}r(1-\eta_E L_E)K + L_W(\delta+\rho-\Pi)[A\alpha/\Pi]^{\frac{1}{1-\alpha}} + \tau_{sE}r\eta_E L_E K + (1-\phi)\sqrt{\rho}\sigma h(\Pi)K.$$

Using the expressions for the savings taxes given in Theorem 4.3 we have

$$\tau_{sW}r+\rho = \Pi-\delta-\sqrt{\rho}\sigma h(\Pi) \qquad\qquad \tau_{sE}-\tau_{sW} = \frac{\rho h(\Pi)^2}{\Pi-\delta-\sqrt{\rho}\sigma h(\Pi)}$$

$$(\tau_{sE}-\tau_{sW})r = \rho h(\Pi)^2 \qquad\qquad \eta_E L_E = \frac{\sigma\phi}{\sqrt{\rho}h(\Pi)},$$

and so the expression for aggregate taxes becomes

$$\begin{aligned}
T/K &= \tau_{sW}r + \rho + (\tau_{sE}-\tau_{sW})r\eta_E L_E + \delta - \Pi + (1-\phi)\sqrt{\rho}\sigma h(\Pi) \\
&= \Pi-\delta-\sqrt{\rho}\sigma h(\Pi) + \sqrt{\rho}\sigma\phi h(\Pi) + \delta - \Pi + (1-\phi)\sqrt{\rho}\sigma h(\Pi) = 0
\end{aligned}$$

as desired. As another plausibility check, consider the behaviour of taxes as the agency problem vanishes $\phi \to 0$. In this case, inspection of the defining equation for Π given in Theorem 3.2 shows that $\Pi \approx \delta+\rho$ and $h(\Pi) \approx 0$. The above taxes are then approximately

$$\tau_\Pi = 1-\phi \approx 1 \qquad\qquad \tau_{LW} \approx \tau_{sW} \approx \tau_{sE} \approx 0.$$

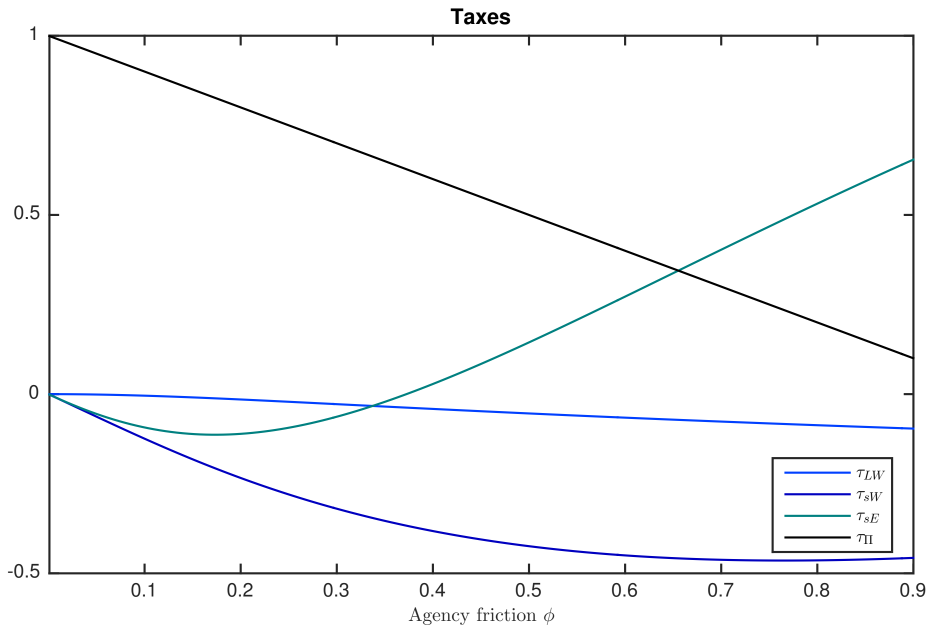

Figure 5: Taxes as function of agency problem

and the risk-free rate is approximately $r \approx \rho$. Therefore, as agency frictions vanish, the interest rate approaches its complete-markets value, and taxes on savings, taxes on labour income, as expected. The tax profits approaches one hundred percent, but net revenue collected is negligible because net business profits $\Pi - \delta - r$ also approach zero.

It is important to note that the interest rate calculated in the exogenously incomplete markets setting does not coincide with the interest rate given in the relaxed planner problem. As noted in the discussion following Theorem 3.2, the only intertemporal rate for the relaxed planner consistent with stationarity is the subjective rate of discount ρ, as all other rates induce a trend in consumption across different cohorts. The introduction of an intertemporal price in Section 3 is simply a mathematical device to characterize the efficient allocation and does not represent the return on an asset available to any agent. Indeed, using the expression for the interest rate found in Theorem 4.3, we can 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. See Appendix C. □

One consequence of Corollary 4.4 is that although the intertemporal wedge in the principal-agent is always positive, it does not necessarily follow that the entrepreneurs must face a positive tax on savings in the decentralization. Figure 5 plots the taxes on savings, labour income and profits as a function of the parameter ϕ affecting the degree of the agency problem, for the parameters employed in Figure 2 (apart from ϕ of course). Notice that for low levels of ϕ entrepreneurs actually face a subsidy on their risk-free savings. Also note that the taxes on labour income and worker savings are negative for all values in the domain.

In contrast, Figure 6 plots the taxes on labour income and savings as a function of the level of exogenous uncertainty σ for the same set of parameters as in Figure 2. Once again the taxes on workers are everywhere

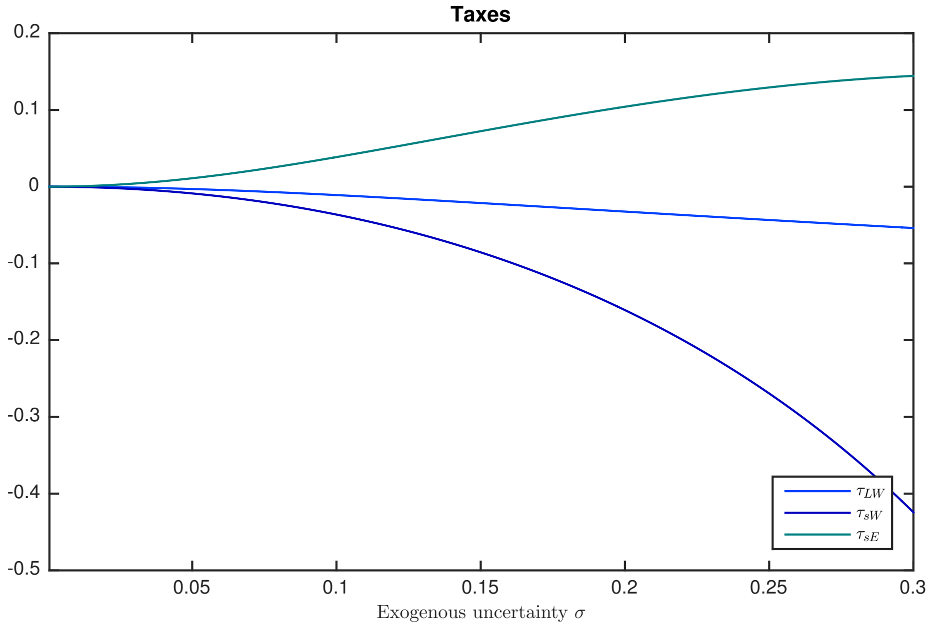

Figure 6: Taxes as function of exogenous uncertainty

negative. Comparison of Figure 6 and Figure 5 shows that although changes in ϕ and σ have identical effects on the efficient allocation, they do not have identical effects on taxes.

5 Conclusion

This paper characterizes a stationary constrained-efficient allocation in a dynamic economy with physical capital, entrepreneurs, workers and repeated moral hazard. I characterize the inheritance levels and taxes necessary for decentralization in a general equilibrium model. The decentralization is noteworthy both for its simplicity and the invariance of the profits tax to technological changes. All other taxes are linear, time-independent, and admit closed-form representations in terms of the root of a single non-linear equation. For simplicity, I have restricted attention to parameters for which the no-absconding constraint holds as a strict inequality. An interesting direction for future work would be to explore the extent to which the findings of this paper extend to more general preferences and parameters and to incorporate a non-trivial margin for entry.

References

[1] Albanesi, Stefania. "Optimal Taxation of Entrepreneurial Capital with Private Information." (2007).

[2] Angeletos, George-Marios. "Uninsured idiosyncratic investment risk and aggregate saving." Review of Economic dynamics 10.1 (2007): 1-30.

[3] Atkeson, Andrew, and Robert E. Lucas Jr. "On efficient distribution with private information." The Review of Economic Studies 59.3 (1992): 427-453.

[4] Cagetti, Marco, and Mariacristina De Nardi. "Entrepreneurship, frictions, and wealth." Journal of political Economy 114.5 (2006): 835-870.

[5] Chari, Varadarajan V., and Patrick J. Kehoe. "Optimal fiscal and monetary policy." Handbook of macroeconomics 1 (1999): 1671-1745.

[6] Evans, David. Optimal taxation with persistent idiosyncratic investment risk. mimeo, 2014.

[7] Farhi, Emmanuel, and Ivan Werning. "Inequality and social discounting." Journal of Political Economy 115.3 (2007): 365-402.

[8] Gabaix, Xavier, et al. "The dynamics of inequality." Econometrica 84.6 (2016): 2071-2111.

[9] Golosov, Mikhail, Narayana Kocherlakota, and Aleh Tsyvinski. "Optimal indirect and capital taxation." The Review of Economic Studies 70.3 (2003): 569-587.

[10] Mirrlees, James A. "An exploration in the theory of optimum income taxation." The review of economic studies 38.2 (1971): 175-208.

[11] Panousi, Vasia, and Catarina Reis. "A unified framework for optimal taxation with undiversifiable risk." (2017).

[12] Phelan, Thomas, On the efficient level of inequality in business income, 2018.

[13] Phelan, Christopher. "Incentives and aggregate shocks." The Review of Economic Studies 61.4 (1994): 681-700.

[14] Rogerson, William P. "Repeated moral hazard." Econometrica: Journal of the Econometric Society (1985): 69-76.

[15] Sannikov, Yuliy. "A continuous-time version of the principal-agent problem." The Review of Economic Studies 75.3 (2008): 957-984.

[16] Di Tella, Sebastian, and Yuliy Sannikov. Optimal asset management with hidden savings. Working Paper, Princeton University, 2016.

[17] Shourideh, Ali. "Optimal taxation of wealthy individuals." Work. pap. U. of Pennsylvania (2012).

[18] Smith, M., Yagan, D., Zidar, O., & Zwick, E. (2017). Capitalists in the Twenty-first Century. UC Berkeley and University of Chicago Working Paper.

A Discrete-time formulation

The purpose of this section is to outline a discrete-time environment that approximates the continuous-time model 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 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

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

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

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

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

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

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 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 not divert any delegated capital to consumption. This is obviously without loss when characterizing efficient allocations.

[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 $\big(1 - e^{-\Delta\rho}\big)\sum_{n=0}^{\infty} e^{-n\Delta\rho}\ln c = \ln c$, so the normalization (23) simply ensures utility is bounded as $\Delta \to 0$ whenever the present discounted value of consumption is bounded.

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

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

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

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

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 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

$$\begin{aligned}\left(1 - e^{-\Delta\rho}\right)\ln C + e^{-\Delta\rho}W^+ &\geq \left(1 - e^{-\Delta\rho}\right)\ln\Big(C + 2\sqrt{\Delta}\overline{x}K\Big) + e^{-\Delta\rho}W^- \\ \left(1 - e^{-\Delta\rho}\right)\ln C + e^{-\Delta\rho}W^- &\geq \left(1 - e^{-\Delta\rho}\right)\ln\Big(C + 2\sqrt{\Delta}\overline{x}K\Big) + e^{-\Delta\rho}W^+.\end{aligned} \tag{25}$$

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

$$\begin{aligned} e^{-\Delta\rho}\left[W^+ - W^-\right] &\geq \left(1 - e^{-\Delta\rho}\right)\Big[\ln\Big(C + 2\sqrt{\Delta}\overline{x}K\Big) - \ln C\Big] \\ \left(1 - e^{-\Delta\rho}\right)\Big[\ln C - \ln\Big(C - 2\sqrt{\Delta}\overline{x}K\Big)\Big] &\geq e^{-\Delta\rho}\left[W^+ - W^-\right]\end{aligned}$$

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

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

Also note that simplification of (26) gives

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

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

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

which in turn may be written as

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

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. So suppose that the preferences of the consumer and the principal over sequences $(\Delta c_n)_{n=0}^{\infty}$ are now given by

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

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

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

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

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

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

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

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

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

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

$$
\begin{aligned}
(1-\overline{\beta})\ln C_n + (\beta_{n+1}/\beta_n)W_{n+1}^{+} &= (1-\overline{\beta})\ln(C_n + 2\overline{x}K_n) + (\beta_{n+1}/\beta_n)W_{n+1}^{-} \\
W_n &= (1-\overline{\beta})\ln C + \frac{1}{2}(\beta_{n+1}/\beta_n)\left[W_{n+1}^{-} + W_{n+1}^{+}\right]
\end{aligned}
$$

Simplification gives

$$
\begin{aligned}
W_{n+1}^{+} &= W_{n+1}^{-} + (\beta_{n+1}/\beta_n)^{-1}(1-\overline{\beta})\ln(1+2\overline{x}K_n/C_n) \\
W_n &= (1-\overline{\beta})\ln C_n + \frac{1}{2}(\beta_{n+1}/\beta_n)\left[W_{n+1}^{-} + W_{n+1}^{+}\right] \\
&= (1-\overline{\beta})\ln C_n + (\beta_{n+1}/\beta_n)W_{n+1}^{-} + \frac{1}{2}(1-\overline{\beta})\ln(1+2\overline{x}K_n/C_n) \\
&= (1-\overline{\beta})\ln C_n + (\beta_{n+1}/\beta_n)W_{n+1}^{+} - \frac{1}{2}(1-\overline{\beta})\ln(1+2\overline{x}K_n/C_n).
\end{aligned}
$$

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

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

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

$$
\begin{aligned}
\Delta W_{n+2} &= (\beta_{n+1}/\beta_{n+2})\overline{\beta}\Delta W_{n+1} = -(\beta_n/\beta_{n+2})\overline{\beta}(1-\overline{\beta})u \\
\Delta W_{n+3} &= (\beta_{n+2}/\beta_{n+3})\overline{\beta}\Delta W_{n+2} = -(\beta_n/\beta_{n+3})\overline{\beta}^2(1-\overline{\beta})u \\
&\vdots \\
\Delta W_{n+k} &= -(\beta_n/\beta_{n+k})\overline{\beta}^{k-1}(1-\overline{\beta})u.
\end{aligned}
$$

This discussion is summarized in the following.

Lemma A.1 (Homogeneity of 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*

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

I will now contrast the analysis of this paper with that given in Shourideh [17]. 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 [17] 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 [17] adopts a different specification of shocks it is difficult to directly compare the two models. However, if one adopts the two-period lifecycle structure of Shourideh [17] (and the above timing) but assume 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 $n \geq 1$.

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 $\Delta W_{n+2} = (\beta_{n+1}/\beta_{n+2})\big[\overline{\beta}\Delta W_{n+1} - (1-\overline{\beta})\overline{u}\big]$. It follows that $\Delta W_{n+2} = 0$ if and only if $\overline{u} = \overline{\beta}\Delta W_{n+1}/(1-\overline{\beta}) = -\big(\overline{\beta}\beta_n/\beta_{n+1}\big)u$. 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

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

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

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

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 expression (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 [17], 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).

B Agency problem

This section contains proofs of Lemma 2.2, Lemma 2.5 and Lemma 2.6.

Proof of Lemma 2.2. The Hamilton-Jacobi-Bellman equation for the principal's profit function is

$$rV(W) = \max_{\substack{c \geq 0 \\ k \leq \omega \exp W}} (\Pi - \delta - r)k - c + \rho(W - \ln c)V'(W) + \frac{[\rho\phi\sigma]^2}{2}(k/c)^2 V''(W).$$

First consider the case where $\phi\sigma = 0$. This will serve as an upper bound on the true function and therefore ensure that the problem of the principal is finite-valued for sufficiently small Π. The Hamilton-Jacobi-Bellman equation is then

$$rV(W) = \max_{\substack{c \geq 0 \\ k \leq \omega \exp W}} (\Pi - \delta - r)k - c + \rho(W - \ln c)V'(W).$$

Assuming a solution of the form $-\Omega \exp W$ for some $\Omega > 0$, the Hamilton-Jacobi-Bellman equation becomes

$$-r\Omega \exp W = (\Pi - \delta - r)\omega \exp W + \max_{c \geq 0} -c - \rho\Omega(W - \ln c) \exp W.$$

The first-order condition implies $c = \rho\Omega \exp W$ and so the Hamilton-Jacobi-Bellman equation reduces to

$$\left(\frac{\rho - r}{\rho}\right)\rho\Omega = (\Pi - \delta - r)\omega + \rho\Omega \ln(\rho\Omega)$$

The minimum of $x \mapsto (1 - r/\rho)x + x \ln x$ occurs when $1 - r/\rho + \ln x + 1 = 0$ or $x = \exp(-2 + r/\rho)$, where the minimum is $(1 - r/\rho)\exp(-2 + r/\rho) + (-2 + r/\rho)\exp(-2 + r/\rho) = -\exp(-2 + r/\rho)$. The solution is therefore well-defined when $\phi\sigma = 0$ if $(\Pi - \delta - r)\omega \leq \exp(-2 + r/\rho)$. It is easy to check that for such Π the above equation has two solutions and that the lower of the two violates the transversality condition.[13]

Now return to the general case where $\phi\sigma > 0$ and again assume a solution of the form $V(W) = -\Omega \exp W$. Then the Hamilton-Jacobi-Bellman equation reduces to

$$0 = r\Omega + \max_{\substack{\underline{c} \geq 0 \\ \underline{k} \leq \omega}} (\Pi - \delta - r)\underline{k} - \frac{\underline{k}^2}{2}\left(\rho\phi\sigma/\underline{c}\right)^2 \Omega - \underline{c} + \rho\Omega \ln \underline{c} =: T(\Omega). \tag{37}$$

The above maximand is concave in $\underline{k}$ for any positive Ω so the optimal level of capital is

$$\underline{k} = \min\left\{\omega, \frac{(\Pi - \delta - r)}{[\sqrt{\rho}\phi\sigma]^2}\frac{\underline{c}^2}{\rho\Omega}\right\}. \tag{38}$$

[13] In this case the Inverse Euler equation does not hold. The perturbation argument assumes one may scale up the level of capital but this will not be incentive compatible if the no-absconding constraint holds with equality.

Now define a pair of functions T_1 and T_2:

$$T_1(\Omega) = \max_{\underline{c}} \ (\rho \ln \underline{c} + r)\Omega - \underline{c} + (\Pi - \delta - r)\omega - \frac{\omega^2}{2}\frac{[\rho\phi\sigma]^2\Omega}{\underline{c}^2}$$
$$\underline{c}^2 \geq \frac{\omega[\rho\phi\sigma]^2\Omega}{\Pi - \delta - r}$$
$$T_2(\Omega) = \max_{\underline{c}} \ (\rho \ln \underline{c} + r)\Omega - \underline{c} + \frac{(\Pi - \delta - r)^2\underline{c}^2}{2\Omega[\rho\phi\sigma]^2}$$
$$0 \leq \underline{c}^2 \leq \frac{\omega[\rho\phi\sigma]^2\Omega}{\Pi - \delta - r}$$

and note that equation (37) may be written $0 = \max(T_1(\Omega), T_2(\Omega)) = T(\Omega)$.[14] Changing variables to $y = \underline{c}/\Omega$, $T_2(\Omega)/\Omega$ simplifies to

$$\begin{aligned}
\frac{T_2(\Omega)}{\Omega} &= \max_{\underline{c}^2 \leq \frac{\Omega\omega[\rho\phi\sigma]^2}{\Pi-\delta-r}} \rho \ln \underline{c} + r - \frac{\underline{c}}{\Omega} + \frac{1}{2}\left(\frac{\underline{c}}{\Omega}\right)^2 \left(\frac{\Pi - \delta - r}{\rho\phi\sigma}\right)^2 \\
&= \max_{y^2 \leq \frac{[\rho\phi\sigma]^2\omega/\Omega}{\Pi-\delta-r}} \rho \ln(y\Omega) + r - y + \frac{y^2}{2}\left(\frac{\Pi - \delta - r}{\rho\phi\sigma}\right)^2 \\
&= \rho \ln \Omega + r + \max_{y^2 \leq \frac{[\rho\phi\sigma]^2\omega/\Omega}{\Pi-\delta-r}} G(\Pi, y).
\end{aligned}$$

for the maximand $G(\Pi, y)$ given by

$$G(\Pi, y) = \rho \ln y - y + \frac{y^2}{2}\left(\frac{\Pi - \delta - r}{\rho\phi\sigma}\right)^2.$$

If the no-absconding constraint does not hold with equality then the maximum of $G(\Pi, \cdot)$ will be attained at an interior point. The critical points of $G(\Pi, y)$ solve $\rho/y = 1 - ([\Pi - \delta - r]/[\rho\phi\sigma])^2 y$ and hence are given by

$$0 = y(\Pi)^2 \left(\frac{\Pi - \delta - r}{\rho\phi\sigma}\right)^2 - y(\Pi) + \rho. \tag{39}$$

The solutions to (39) are

$$y(\Pi) = \frac{1 \pm \sqrt{1 - \rho(2(\Pi - \delta - r)/[\rho\phi\sigma])^2}}{2((\Pi - \delta - r)/[\rho\phi\sigma])^2} = 2\rho\left(\frac{1 \pm \sqrt{1 - (2(\Pi - \delta - r)/[\sqrt{\rho}\phi\sigma])^2}}{(2(\Pi - \delta - r)/[\sqrt{\rho}\phi\sigma])^2}\right). \tag{40}$$

Note that only the lower value of y corresponds to a local maxima because $G(\Pi, y) \to \infty$ as $y \to \infty$. Recalling the definition of the function h,

$$h(\Pi) := \frac{1 - \sqrt{1 - (2[\Pi - \delta - r]/[\sqrt{\rho}\phi\sigma])^2}}{2[\Pi - \delta - r]/[\sqrt{\rho}\phi\sigma]},$$

the function y simplifies to

$$y(\Pi) = \frac{2\rho}{2(\Pi - \delta - r)/[\sqrt{\rho}\phi\sigma]}\left(\frac{1 - \sqrt{1 - (2(\Pi - \delta - r)/[\sqrt{\rho}\phi\sigma])^2}}{2(\Pi - \delta - r)/[\sqrt{\rho}\phi\sigma]}\right) = \frac{\rho h(\Pi)}{(\Pi - \delta - r)/[\sqrt{\rho}\phi\sigma]}.$$

[14] One may think of T_1 and T_2 as being the maxima on the right-hand side of (37) under the additional restrictions $\omega \leq \frac{(\Pi-\delta-r)\underline{c}^2}{\Omega[\rho\phi\sigma]^2}$ and $\omega \geq \frac{(\Pi-\delta-r)\underline{c}^2}{\Omega[\rho\phi\sigma]^2}$ respectively.

Further simplification gives

$$h(\Pi)^2+1=\frac{2-2\sqrt{1-(2(\Pi-\delta-r)/[\sqrt{\rho}\phi\sigma])^2}}{(2(\Pi-\delta-r)/[\sqrt{\rho}\phi\sigma])^2}=\frac{h(\Pi)}{(\Pi-\delta-r)/[\sqrt{\rho}\phi\sigma]}$$
$$y(\Pi)=\frac{\rho h(\Pi)}{(\Pi-\delta-r)/[\sqrt{\rho}\phi\sigma]}=\rho(h(\Pi)^2+1). \tag{41}$$

Using (39) and (41), we have

$$G(\Pi,y(\Pi))=\rho\ln y(\Pi)-y(\Pi)+\frac{1}{2}\left(\frac{\Pi-\delta-r}{\rho\phi\sigma}\right)^2 y(\Pi)^2=\rho\ln y(\Pi)-y(\Pi)+\frac{y(\Pi)-\rho}{2}$$
$$=\rho\ln(\rho(h(\Pi)^2+1))-\frac{\rho(h(\Pi)^2+2)}{2}=\rho[\ln\rho-1]+\rho\ln(h(\Pi)^2+1)-\frac{\rho h(\Pi)^2}{2}.$$

It follows that whenever the maximum of $G(\Pi,\cdot)$ occurs in the interior of the constraint set, $T_2(\Omega)/\Omega$ may be written

$$\frac{T_2(\Omega)}{\Omega}=\rho\ln\Omega+r-\rho+\rho\ln\rho+\rho\ln(h(\Pi)^2+1)-\frac{\rho h(\Pi)^2}{2}$$

and so if the no-absconding constraint does not bind then $\Omega(\Pi)$ solves

$$\Omega(\Pi)=\frac{1}{\rho(h(\Pi)^2+1)}\exp\left(\frac{h(\Pi)^2}{2}+\frac{\rho-r}{\rho}\right).$$

Recalling (38), the policy functions are given by

$$\underline{c}(\Pi)=y(\Pi)\Omega(\Pi)=\exp\left(\frac{h(\Pi)^2}{2}+\frac{\rho-r}{\rho}\right)$$
$$\underline{k}(\Pi)=\frac{(\Pi-\delta-r)\underline{c}^2}{[\sqrt{\rho}\phi\sigma]^2\rho\Omega}=\frac{(\Pi-\delta-r)}{[\sqrt{\rho}\phi\sigma]^2}[h(\Pi)^2+1]\exp\left(\frac{h(\Pi)^2}{2}+\frac{\rho-r}{\rho}\right)$$

which simplifies to the claimed expressions. □

Now consider the case where the no-absconding constraint binds. Then the value function is

$$rV(W)=\max_{c\geq 0}\,(\Pi-\delta-r)\omega\exp W-c+\rho(W-\ln c)V'(W)+\frac{[\rho\phi\sigma]^2}{2}(c^{-1}\omega\exp W)^2V''(W).$$

Assume solution of the form $-\Omega\exp W$ and policy functions of the form $c(W)=\underline{c}\exp W$. Then substitution gives

$$-r\Omega=\max_{c\geq 0}\,(\Pi-\delta-r)\omega-\underline{c}+\rho\Omega\ln\underline{c}-\frac{\Omega[\rho\phi\omega\sigma]^2}{2\underline{c}^2}.$$

First-order condition and subsequent rearrangement gives

$$\underline{c}=\rho\Omega+\Omega[\rho\phi\omega\sigma]^2\underline{c}^{-2}$$
$$-r\Omega=(\Pi-\delta-r)\omega-\underline{c}+\rho\Omega\ln\underline{c}-\frac{\Omega}{2}[\rho\phi\omega\sigma]^2\underline{c}^{-2}.$$

This is a pair of equations in two unknowns. The first allows us to write Ω in terms of $\underline{c}$, which then may be substituted into the second to obtain a single nonlinear equation.

$$\Omega=\frac{\underline{c}^3}{\rho\underline{c}^2+[\rho\phi\omega\sigma]^2}$$
$$-r\Omega=(\Pi-\delta-r)\omega-\underline{c}+\rho\Omega\ln\underline{c}-\frac{1}{2}(\underline{c}-\rho\Omega)$$

$$
\begin{aligned}
0 &= (\Pi - \delta - r)\omega - \frac{3}{2}\underline{c} + + \Big(r + \frac{\rho}{2} + \rho \ln \underline{c}\Big)\Omega \\
&= (\Pi - \delta - r)\omega - \frac{3}{2}\underline{c} + \Big(r + \frac{\rho}{2} + \rho \ln \underline{c}\Big)\frac{\underline{c}^3}{\rho\underline{c}^2 + [\rho\phi\omega\sigma]^2}.
\end{aligned}
$$

What happens when $\Pi = \delta + r$? Or $\phi = 0$? In the latter case we should have

$$
\begin{aligned}
\frac{3}{2}\underline{c} &= (\Pi - \delta - r)\omega + \Big(r + \frac{\rho}{2} + \rho \ln \underline{c}\Big)\frac{\underline{c}}{\rho} \\
\Big(\frac{\rho - r}{\rho}\Big)\underline{c} &= (\Pi - \delta - r)\omega + \underline{c}\ln \underline{c}.
\end{aligned}
$$

This has no explicit solution but that's fine. In the former situation we have

$$
\begin{aligned}
\frac{3}{2} &= \Big(r + \frac{\rho}{2} + \rho \ln \underline{c}\Big)\frac{\underline{c}^2}{\rho\underline{c}^2 + [\rho\phi\omega\sigma]^2}. \\
\Big(\frac{\rho - r}{\rho}\Big)\underline{c}^2 + \frac{3}{2}\rho[\phi\omega\sigma]^2 &= \underline{c}^2 \ln \underline{c}.
\end{aligned}
$$

This last one doesn't really make any sense, because one would never be delegated capital in this case.

Proof of Lemma 2.5. For each $t \geq 0$ and return process R I want to determine the law of motion of the process

$$
Y_{t,\Delta} := \mathbb{E}\Big[\exp(-\rho\Delta)R_{t,t+\Delta}\frac{u'(c_{t+\Delta})}{u'(c_t)}\Big|\mathcal{F}_t\Big]. \tag{42}
$$

To this end note that the processes for consumption and marginal utility imply

$$
\begin{aligned}
c_{t+\Delta} &= c_t \exp\Big(\Big[r - \rho - \frac{\rho h(\Pi)^2}{2}\Big]\Delta + \sqrt{\rho}h(\Pi)[B_{t+\Delta} - B_t]\Big) \\
u'(c_{t+\Delta}) &= u'(c_t)\exp\Big(\Big[\rho - r + \frac{\rho h(\Pi)^2}{2}\Big]\Delta - \sqrt{\rho}h(\Pi)[B_{t+\Delta} - B_t]\Big).
\end{aligned}
$$

It follows that for risky capital the process Y satisfies

$$
\begin{aligned}
\exp(-\rho\Delta)R_{t,t+\Delta}\frac{u'(c_{t+\Delta})}{u'(c_t)} &= \exp\Big(\Big[\Pi - \delta - r + \frac{\rho h(\Pi)^2 - \sigma^2}{2}\Big]\Delta\Big)\mathbb{E}[\exp((\sigma - \sqrt{\rho}h(\Pi))B_\Delta)] \\
&= \exp\Big(\Big[\Pi - \delta - r + \frac{\rho h(\Pi)^2}{2} - \frac{\sigma^2}{2} + \frac{(\sigma - \sqrt{\rho}h(\Pi))^2}{2}\Big]\Delta\Big) \\
&= \exp\big(\big[\Pi - \delta - r + \rho h(\Pi)^2 - \sqrt{\rho}\sigma h(\Pi)\big]\Delta\big)
\end{aligned}
$$

which gives ν^K. For the risk-free bond we have

$$
\exp((r-\rho)\Delta)\frac{u'(c_{t+\Delta})}{u'(c_t)} = \exp\Big(\frac{\rho h(\Pi)^2\Delta}{2} - \sqrt{\rho}h(\Pi)[B_{t+\Delta} - B_t]\Big) = \exp\big(\rho h(\Pi)^2\Delta\big)
$$

which gives ν^B. The inequality $\nu^B \geq \nu^K$ follows from the algebra

$$
\begin{aligned}
\nu^K &= \sqrt{\rho}\sigma\left(\phi[\Pi - \delta - r]/[\sqrt{\rho}\phi\sigma] + \frac{\sqrt{1 - (2[\Pi - \delta - r]/[\sqrt{\rho}\phi\sigma])^2} - 1}{2[\Pi - \delta - r]/[\sqrt{\rho}\phi\sigma]}\right) + \nu^B \\
\nu^K - \nu^B &\leq -\frac{\sqrt{\rho}\sigma(1 - \sqrt{1 - (2[\Pi - \delta - r]/[\sqrt{\rho}\phi\sigma])^2})^2}{2[\Pi - \delta - r]/[\sqrt{\rho}\phi\sigma]} < 0
\end{aligned}
$$

while the inequality $\nu^B \geq 0$ is obvious. □

The solution to the principal-agent problem in Theorem 2.2 is valid provided $T_1(\Omega(\Pi)) < 0$, the local maximum $y(\Pi)$ given by (40) is well-defined, lies in the constraint set of T_2 and is in fact a global maximum. Lemma 2.6 shows how these requirements may be simplified.

Proof of Lemma 2.6. First note that the requirement that $T_1(\Omega(\Pi)) < 0$ is weaker than the inequality (7) in the statement of the lemma. Next observe that the maximum value of y in the constraint set of the function T_2 satisfies

$$\begin{aligned}
\overline{y}^2 &= \frac{1}{\Omega(\Pi)} \frac{\rho\omega[\sqrt{\rho}\phi\sigma]^2}{(\Pi - \delta - r)} = \rho(h(\Pi)^2 + 1) \exp\left(-\frac{h(\Pi)^2}{2} + \frac{r - \rho}{\rho}\right) \frac{\rho\omega[\sqrt{\rho}\phi\sigma]^2}{(\Pi - \delta - r)} \\
&= \frac{\rho h(\Pi)}{(\Pi - \delta - r)/[\sqrt{\rho}\phi\sigma]} \exp\left(-\frac{h(\Pi)^2}{2} + \frac{r - \rho}{\rho}\right) \frac{\rho\omega[\sqrt{\rho}\phi\sigma]^2}{(\Pi - \delta - r)} \\
&= \sqrt{\rho} h(\Pi) \frac{\omega[\rho\phi\sigma]^3}{(\Pi - \delta - r)^2} \exp\left(-\frac{h(\Pi)^2}{2} - \frac{(\rho - r)}{\rho}\right).
\end{aligned} \tag{43}$$

Since $y(\Pi) = \rho\phi\sigma\sqrt{\rho}h(\Pi)/(\Pi - \delta - r)$, the inequality $y(\Pi) < \overline{y}$ is equivalent to

$$\begin{aligned}
[\rho\phi\sigma]^2 \left(\frac{\sqrt{\rho}h(\Pi)}{\Pi - \delta - r}\right)^2 &< \frac{\omega\sqrt{\rho}h(\Pi)[\rho\phi\sigma]^3}{(\Pi - \delta - r)^2} \exp\left(-\frac{h(\Pi)^2}{2} - \frac{(\rho - r)}{\rho}\right) \\
\sqrt{\rho}h(\Pi) &< \omega\rho\phi\sigma \exp\left(-\frac{h(\Pi)^2}{2} - \frac{(\rho - r)}{\rho}\right) \\
(h(\Pi)^2 + 1)\frac{(\Pi - \delta - r)}{\sqrt{\rho}\phi\sigma} \exp\left(\frac{h(\Pi)^2}{2} + \frac{\rho - r}{\rho}\right) &< \omega\sqrt{\rho}\phi\sigma
\end{aligned}$$

which reduces to $\underline{k}(\Pi) < \omega$. The next observation is that the requirement $G(\Pi, y(\Pi)) > G(\Pi, \overline{y})$ is implied by the inequality $T_1(\Omega(\Pi)) < 0$. To see this, note that using the change of variables $y = c/\Omega$, T_1 and T_2 may be written

$$\begin{aligned}
T_1(\Omega) &= \max_{y \geq 0} \; (\rho \ln y + \rho \ln \Omega + r)\Omega - y\Omega + (\Pi - \delta - r)\omega - \frac{\omega^2}{2} \frac{[\rho\phi\sigma]^2/\Omega}{y^2} \\
&\quad y^2 \geq \frac{[\rho\phi\sigma]^2\omega/\Omega}{\Pi - \delta - r} \\
T_2(\Omega) &= (\rho \ln \Omega + r)\Omega + \Omega G(\Pi, y).
\end{aligned}$$

By the construction of $\Omega(\Pi)$ we have $0 = (\rho \ln \Omega(\Pi) + r)\Omega(\Pi) + \Omega(\Pi)G(\Pi, y(\Pi))$. Evaluating the maximand in T_1 at the boundary point of its constraint set gives

$$\begin{aligned}
T_1(\Omega(\Pi)) &\geq (\rho \ln \Omega(\Pi) + r)\Omega(\Pi) + \Omega(\Pi)(\rho \ln \overline{y} - \overline{y}) + (\Pi - \delta - r)\omega - \frac{\omega^2}{2} \frac{[\rho\phi\sigma]^2/\Omega(\Pi)}{\overline{y}^2} \\
&= (\rho \ln \Omega(\Pi) + r)\Omega(\Pi) + \Omega(\Pi)(\rho \ln \overline{y} - \overline{y}) + (\Pi - \delta - r)\omega/2.
\end{aligned}$$

So if $T_1(\Omega(\Pi)) < 0$ then we have

$$\begin{aligned}
\Omega(\Pi)[\rho \ln \Omega(\Pi) + r + G(\Pi, y(\Pi))] &= 0 \\
&> \Omega(\Pi)[\rho \ln \Omega(\Pi) + r + \rho \ln \overline{y} - \overline{y}] + (\Pi - \delta - r)\omega/2 \\
G(\Pi, y(\Pi)) &> \rho \ln \overline{y} - \overline{y} + \frac{1}{2}(\Pi - \delta - r)\omega/\Omega(\Pi) = G(\Pi, \overline{y}).
\end{aligned}$$

It follows that in addition to $\underline{k}(\Pi) < \omega$, we need only impose the requirement $T_1(\Omega(\Pi)) < 0$, which is strictly weaker than the requirement (7). $\square$

B.1 General CRRA utility

Now suppose that the preferences of the consumer are represented by the utility function

$$U(c) := \int_0^\infty \rho e^{-\rho t} \frac{\mathbb{E}[c_t^{1-\gamma}]}{1-\gamma} dt \tag{44}$$

for some $\gamma \neq 1$. Standard arguments ensure that the value function solves the Hamilton-Jacobi-Bellman equation

$$rV(W) = \max_{\substack{c,k \geq 0 \\ k \leq \omega \exp W}} [\Pi - \delta - r]k - c + \rho\left(W - \frac{c^{1-\gamma}}{1-\gamma}\right)V'(W) + \frac{(\rho\phi\sigma k c^{-\gamma})^2}{2} V''(W).$$

Theorem B.1. *The value function of the principal is finite-valued for all sufficiently small* Π*, and in this case admits a homogeneous solution of the form* $V(W) = -\Omega[(1-\gamma)W]^{\frac{1}{1-\gamma}}$ *for some* $\Omega > 0$*. Further, if the no-absconding constraint holds as a strict inequality then the inverse Euler equation holds.*

Proof. Suppose that the value function and policy functions are of the form $V(W) = -\Omega[(1-\gamma)W]^{\frac{1}{1-\gamma}}, c(W) = \underline{c}[(1-\gamma)W]^{\frac{1}{1-\gamma}}$ and $k(W) = \underline{k}[(1-\gamma)W]^{\frac{1}{1-\gamma}}$ for some $\Omega, \underline{c}$ and $\underline{k}$. Then the Hamilton-Jacobi-Bellman equation reduces to a single nonlinear equation for the (scalar) coefficient Ω

$$-r\Omega = \max_{\substack{\underline{k}, \underline{c} \geq 0 \\ \underline{k} \leq \omega}} (\Pi - \delta - r)\underline{k} - \underline{c} - \rho\left(\frac{1 - \underline{c}^{1-\gamma}}{1-\gamma}\right)\Omega - \frac{\gamma\Omega(\rho\phi\sigma)^2 \underline{k}^2}{2\underline{c}^{2\gamma}}.$$

The right-hand side of the above is quadratic and concave in $\underline{k}$ and so the optimal assignment of capital is simply

$$\underline{k} = \min\left\{\omega, \frac{(\Pi - \delta - r)\underline{c}^{2\gamma}}{\gamma(\rho\phi\sigma)^2\Omega}\right\}. \tag{45}$$

If the no-absconding constraint does not bind then the Hamilton-Jacobi-Bellman equation becomes

$$\begin{aligned}
-r\Omega &= \max_{\underline{c} \geq 0} \frac{(\Pi - \delta - r)^2 \underline{c}^{2\gamma}}{\gamma(\rho\phi\sigma)^2\Omega} - \underline{c} - \rho\left(\frac{1 - \underline{c}^{1-\gamma}}{1-\gamma}\right)\Omega - \frac{\gamma\Omega(\rho\phi\sigma)^2}{2\underline{c}^{2\gamma}} \frac{\Pi^2 \underline{c}^{4\gamma}}{\gamma^2(\rho\phi\sigma)^4\Omega^2} \\
&= \max_{\underline{c} \geq 0} \frac{1}{2} \frac{(\Pi - \delta - r)^2 \underline{c}^{2\gamma}}{\gamma(\rho\phi\sigma)^2\Omega} - \underline{c} - \rho\left(\frac{1 - \underline{c}^{1-\gamma}}{1-\gamma}\right)\Omega.
\end{aligned}$$

The first-order condition implies

$$0 = \frac{(\Pi - \delta - r)^2 \underline{c}^{2\gamma - 1}}{(\rho\phi\sigma)^2\Omega} - 1 + \rho\underline{c}^{-\gamma}\Omega.$$

Substituting this into the Bellman equation gives us the pair of equations

$$\begin{aligned}
\frac{(\Pi - \delta - r)^2 \underline{c}^{2\gamma}}{\gamma(\rho\phi\sigma)^2\Omega} &= \frac{\underline{c}}{\gamma} - \frac{\rho}{\gamma}\underline{c}^{1-\gamma}\Omega \\
\left(\frac{\rho}{1-\gamma} - r\right)\Omega &= \left(\frac{1}{2\gamma} - 1\right)\underline{c} + \left(-\frac{1}{2\gamma} + \frac{1}{1-\gamma}\right)\rho\underline{c}^{1-\gamma}\Omega.
\end{aligned} \tag{46}$$

It remains to show that the solution to the above pair (46) satisfies the inverse-Euler equation, which is equivalent to the drift of $(c_t^\gamma)_{t \geq 0}$ vanishing. From the law of motion for utility, $dW_t = \rho(1 - \underline{c}^{1-\gamma})W_t dt + \rho\phi\sigma\underline{kc}^{-\gamma}(1-\gamma)W_t dB_t$ and applying Ito's lemma to the function $f(W) = [(1-\gamma)W]^{\frac{\gamma}{1-\gamma}}$, we may obtain the law of motion of $c_t := f(W_t)$. Differentiation gives $f'(W) = \gamma[(1-\gamma)W]^{\frac{\gamma}{1-\gamma}-1}$ and $f''(W) = \gamma(2\gamma - 1)[(1-\gamma)W]^{\frac{\gamma}{1-\gamma}-2}$, so

$$df_t = \gamma\left(\rho\left(\frac{1 - \underline{c}^{1-\gamma}}{1-\gamma}\right) + \frac{1}{2}(2\gamma - 1)(\rho\phi\sigma\underline{kc}^{-\gamma})^2\right)[(1-\gamma)W_t]^{\frac{\gamma}{1-\gamma}} dt + \rho\phi\sigma\underline{kc}^{-\gamma}\gamma(2\gamma-1)[(1-\gamma)W_t]^{\frac{\gamma}{1-\gamma}} dB_t.$$

The reciprocal of marginal utility solves

$$dc_t^\gamma = \gamma\left(\rho\frac{\left(1-\underline{c}^{1-\gamma}\right)}{1-\gamma} + \frac{1}{2}(2\gamma-1)(\rho\phi\sigma\underline{kc}^{-\gamma})^2\right)c_t^\gamma dt + \rho\phi\sigma\underline{kc}^{-\gamma}c_t^\gamma dB_t. \tag{47}$$

The drift $\mu(\gamma)$ in this variable satisfies

$$\frac{\mu(\gamma)}{\gamma} = \frac{\rho(1-\underline{c}^{1-\gamma})}{1-\gamma} + \left(\gamma - \frac{1}{2}\right)(\rho\phi\sigma)^2(\underline{kc}^{-\gamma})^2 = \frac{\rho(1-\underline{c}^{1-\gamma})}{1-\gamma} + \left(1-\frac{1}{2\gamma}\right)\frac{(\Pi-\delta-r)^2\underline{c}^{2\gamma}}{\gamma(\rho\phi\sigma)^2\Omega^2}.$$

where for the second equality I used (45). The first-order equation implies (multiplying through by $\underline{c}[\gamma\Omega]^{-1}$)

$$\frac{(\Pi-\delta-r)^2\underline{c}^{2\gamma}}{\gamma(\rho\phi\sigma)^2\Omega^2} = \frac{\underline{c}}{\gamma\Omega} - \frac{\rho}{\gamma}\underline{c}^{1-\gamma}$$

Substituting this into the drift gives

$$\frac{\mu(\gamma)}{\gamma} = \frac{\rho(1-\underline{c}^{1-\gamma})}{1-\gamma} + \left(1-\frac{1}{2\gamma}\right)\left(\frac{\underline{c}}{\gamma\Omega} - \frac{\rho}{\gamma}\underline{c}^{1-\gamma}\right) = \frac{\rho}{1-\gamma} + \left(1-\frac{1}{2\gamma}\right)\frac{\underline{c}}{\gamma\Omega} - \left(\frac{\rho}{1-\gamma} + \left(1-\frac{1}{2\gamma}\right)\frac{\rho}{\gamma}\right)\underline{c}^{1-\gamma} \tag{48}$$

Using the Bellman equation once more we have

$$\frac{\rho}{1-\gamma} - \rho = \left(\frac{1}{2\gamma} - 1\right)\frac{\underline{c}}{\Omega} + \left(-\frac{1}{2\gamma} + \frac{1}{1-\gamma}\right)\rho\underline{c}^{1-\gamma}.$$
$$\left(1-\frac{1}{2\gamma}\right)\frac{\underline{c}}{\gamma\Omega} = -\frac{\rho}{1-\gamma} + \frac{1}{\gamma}\left(-\frac{1}{2\gamma} + \frac{1}{1-\gamma}\right)\rho\underline{c}^{1-\gamma}.$$

which upon substitution into (48) gives

$$\begin{aligned}\mu(\gamma) &= \frac{\rho}{1-\gamma} - \left(\frac{1}{1-\gamma} + \left(1-\frac{1}{2\gamma}\right)\frac{1}{\gamma}\right)\rho\underline{c}^{1-\gamma} - \frac{\rho}{1-\gamma} + \frac{1}{\gamma}\left(-\frac{1}{2\gamma} + \frac{1}{1-\gamma}\right)\rho\underline{c}^{1-\gamma} \\ &= \left(-\frac{\gamma}{1-\gamma} - \left(1-\frac{1}{2\gamma}\right) - \frac{1}{2\gamma} + \frac{1}{1-\gamma}\right)\frac{\rho\underline{c}^{1-\gamma}}{\gamma} = 0\end{aligned}$$

as claimed. □

Theorem B.1 shows that the solution to the principal's problem in the general CRRA case is finite-valued for sufficiently small marginal product of capital and admits homogeneous capital and consumption policy functions. To characterize the policy function in the case where the no-absconding inequality is strict, we simplify the Hamilton-Jacobi-Bellman equation by using the fact that the inverse Euler equation holds with equality. Together with the first-order condition this gives the pair of equations

$$0 = \frac{(\Pi-\delta-r)^2\underline{c}^{2\gamma-1}}{(\rho\phi\sigma)^2\Omega} - 1 + \rho\underline{c}^{-\gamma}\Omega$$
$$0 = \frac{\rho(1-\underline{c}^{1-\gamma})}{1-\gamma} + \left(1-\frac{1}{2\gamma}\right)\frac{(\Pi-\delta-r)^2\underline{c}^{2\gamma}}{\gamma(\rho\phi\sigma)^2\Omega^2}.$$

One may think of the second equation as defining $\Omega = \Omega(c)$. Rearranging gives

$$\Omega(c)^2 = \frac{\underline{c}^{2\gamma}}{\rho\gamma}\left(1-\frac{1}{2\gamma}\right)\left(\frac{\Pi-\delta-r}{\rho\phi\sigma}\right)^2\left(\frac{1-\gamma}{\underline{c}^{1-\gamma}-1}\right).$$

Substituting this into the first-order condition then gives a single equation for $\underline{c}$:

$$\Omega(c) = \frac{(\Pi-\delta-r)^2\underline{c}^{2\gamma-1}}{(\rho\phi\sigma)^2} + \rho\underline{c}^{-\gamma}\Omega(c)^2 = \left(\frac{\Pi-\delta-r}{\rho\phi\sigma}\right)^2\underline{c}^{2\gamma-1} + \frac{\underline{c}^{2\gamma-1}}{\gamma}\left(1-\frac{1}{2\gamma}\right)\left(\frac{\Pi-\delta-r}{\rho\phi\sigma}\right)^2\left(\frac{1-\gamma}{1-\underline{c}^{\gamma-1}}\right). \tag{49}$$

Simplification then gives the following.

Theorem B.2. *The optimal consumption and capital coefficients are given by*

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

where x is the lower of the positive solutions to the cubic

$$0 = \gamma(\gamma - 1)(\gamma - 1/2)x^3 + \left(\frac{3\gamma - 1}{2}\right)\frac{(\Pi - \delta - r)}{\sqrt{\rho}\phi\sigma}x^2 - \gamma x + \frac{\Pi - \delta - r}{\sqrt{\rho}\phi\sigma}$$

if they exist.

Proof. Proceeding from (49) and simplifying,

$$\frac{\sqrt{\rho}\phi\sigma}{(\Pi - \delta - r)}\left(\frac{1}{\gamma} - \frac{1}{2\gamma^2}\right)^{1/2}\left(\frac{1 - \gamma}{\underline{c}^{1-\gamma} - 1}\right)^{1/2} = \underline{c}^{\gamma - 1} + \left(\frac{1}{\gamma} - \frac{1}{2\gamma^2}\right)\left(\frac{1 - \gamma}{\underline{c}^{1-\gamma} - 1}\right). \tag{50}$$

Now define

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

and note that $\underline{c}^{1-\gamma} = 1 - (1/2 - \gamma)(1 - \gamma)x^2$ and so (50) may be written

$$\begin{aligned}
\frac{\sqrt{\rho}\phi\sigma}{(\Pi - \delta - r)}\frac{1}{\gamma x} &= \frac{1}{1 - (1/2 - \gamma)(1 - \gamma)x^2} + \frac{1}{\gamma^2 x^2} \\
\frac{\sqrt{\rho}\phi\sigma}{(\Pi - \delta - r)}\gamma x[1 - (1/2 - \gamma)(1 - \gamma)x^2] &= \gamma^2 x^2 + 1 - (1/2 - \gamma)(1 - \gamma)x^2 \\
\gamma x - \gamma(1/2 - \gamma)(1 - \gamma)x^3 &= \frac{\Pi - \delta - r}{\sqrt{\rho}\phi\sigma} + \left(\frac{3\gamma - 1}{2}\right)\frac{(\Pi - \delta - r)}{\sqrt{\rho}\phi\sigma}x^2
\end{aligned}$$

which simplifies to the desired cubic. □

The following shows that the drift and diffusion of consumption are simple transformations of the solution of the cubic given in Theorem B.2.

Lemma B.3. *Consumption evolves according to a diffusion process of the form $dc_t = \mu_C c_t dt + \sigma_C c_t dZ_t$, where the coefficients μ_C and σ_C are given by*

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

Proof. It suffices to determine the law of motion of $[(1 - \gamma)W_t]^{\frac{1}{1-\gamma}}$. Write $f(W) = [(1 - \gamma)W]^{\frac{1}{1-\gamma}}$ so that

$$f'(W) = [(1 - \gamma)W]^{\frac{\gamma}{1-\gamma}} \qquad f''(W) = \gamma[(1 - \gamma)W]^{\frac{\gamma}{1-\gamma} - 1}.$$

If $dW_t = \mu_W(W_t)dt + \sigma_W(W_t)dZ_t$ then $f(W_t)$ evolves according to

$$df_t = \left(\mu_W(W)f'(W_t) + \frac{\sigma_W^2(W_t)}{2}f''(W_t)\right)dt + \sigma_W(W_t)f'(W_t)dZ_t.$$

In our case we have $c(W) = \underline{c}[(1 - \gamma)W]^{\frac{1}{1-\gamma}}$ and $k(W) = \underline{k}[(1 - \gamma)W]^{\frac{1}{1-\gamma}}$, so

$$\begin{aligned}
\mu_W(W_t) &= \rho\left(W_t - \frac{c(W_t)^{1-\gamma}}{1 - \gamma}\right) = \rho\left(\frac{1 - \underline{c}^{1-\gamma}}{1 - \gamma}\right)[(1 - \gamma)W_t] \\
\sigma_W(W_t) &= \rho\phi\sigma k(W_t)c(W_t)^{-\gamma} = \rho\phi\sigma\underline{k}\underline{c}^{-\gamma}[(1 - \gamma)W_t].
\end{aligned}$$

The law of motion of consumption in the optimal contract is then

$$df_t(W_t) = \left(\rho\left(\frac{1-\underline{c}^{1-\gamma}}{1-\gamma}\right) + \frac{\gamma(\rho\phi\sigma\underline{k}\underline{c}^{-\gamma})^2}{2}\right) f_t(W_t)dt + \rho\phi\sigma\underline{k}\underline{c}^{-\gamma} f(W_t)dZ_t. \tag{51}$$

The drift in consumption is then

$$\mu_C = \rho\left(\frac{1-\underline{c}^{1-\gamma}}{1-\gamma}\right) + \frac{\gamma(\rho\phi\sigma\underline{k}\underline{c}^{-\gamma})^2}{2}. \tag{52}$$

Now recall

$$\Omega(c) = \frac{\underline{c}^{\gamma}}{\rho}\left(\frac{1}{\gamma} - \frac{1}{2\gamma^2}\right)^{1/2}\left(\frac{\Pi-\delta-r}{\sqrt{\rho}\phi\sigma}\right)\left(\frac{1-\gamma}{\underline{c}^{1-\gamma}-1}\right)^{1/2}$$

$$\underline{k} = \frac{1}{\rho\gamma}\left(\frac{\Pi-\delta-r}{\sqrt{\rho}\phi\sigma}\right)\frac{1}{\sqrt{\rho}\phi\sigma}\frac{\underline{c}^{2\gamma}}{\Omega(\underline{c})}$$

Therefore we have

$$\sigma_C = \rho\phi\sigma\underline{k}\underline{c}^{-\gamma} = \frac{1}{\sqrt{\rho}\gamma}\left(\frac{\Pi-\delta-r}{\sqrt{\rho}\phi\sigma}\right)\frac{\underline{c}^{\gamma}}{\Omega(\underline{c})} = \sqrt{\rho}\left(\gamma-\frac{1}{2}\right)^{-1/2}\left(\frac{\underline{c}^{1-\gamma}-1}{1-\gamma}\right)^{1/2} = \sqrt{\rho}x.$$

Substituting into (52) then gives

$$\mu_C = -\rho(\gamma-1/2)x^2 + \gamma\rho x^2/2 = \rho(1-\gamma)\frac{x^2}{2} \tag{53}$$

as claimed. □

Now consider the case in which the no-absconding constraint holds as an equality. In this case one may use the first-order condition to write Ω as a function of c and then substitute into the Hamilton-Jacobi-Bellman equation. The first-order condition $1 = \rho\underline{c}^{-\gamma}\Omega + (\gamma\rho\phi\sigma\omega)^2\underline{c}^{-2\gamma-1}\Omega$ may be rearranged to give

$$\Omega(\underline{c}) = \left(\rho\underline{c}^{-\gamma} + (\gamma\rho\phi\sigma\omega)^2\underline{c}^{-2\gamma-1}\right)^{-1}. \tag{54}$$

Using (54) the Hamilton-Jacobi-Bellman equation becomes

$$\left(\frac{\rho}{1-\gamma} - r\right)\Omega(\underline{c}) = (\Pi-\delta-r)\omega - \underline{c} + \frac{\rho\underline{c}^{1-\gamma}\Omega(\underline{c})}{1-\gamma} - \frac{\gamma}{2}(\rho\phi\sigma\omega)^2\underline{c}^{-2\gamma}\Omega(\underline{c})$$

$$(\underline{c} - (\Pi-\delta-r)\omega)\Omega(\underline{c})^{-1} = r - \frac{\rho}{1-\gamma} + \frac{\rho\underline{c}^{1-\gamma}}{1-\gamma} - \frac{\gamma}{2}(\rho\phi\sigma\omega)^2\underline{c}^{-2\gamma}.$$

This simplifies as follows.

Lemma B.4. *When the no-absconding constraint holds as an equality the coefficient of consumption is a solution to the following equation*

$$(\underline{c} - (\Pi-\delta-r)\omega)\left(\rho\underline{c}^{-\gamma} + (\gamma\rho\phi\sigma\omega)^2\underline{c}^{-2\gamma-1}\right) = r - \frac{\rho}{1-\gamma} + \frac{\rho\underline{c}^{1-\gamma}}{1-\gamma} - \frac{\gamma}{2}(\rho\phi\sigma\omega)^2\underline{c}^{-2\gamma}. \tag{55}$$

Theorem B.5. *The profits tax that decentralizes the efficient allocation is simply $\tau_\Pi = 1-\phi$.*

Proof. We must have $\overline{k}/\overline{c} = \underline{k}/\underline{c}$ and $\rho\phi\sigma\underline{k}\underline{c}^{-\gamma} = (1-\tau_\Pi)\sigma\overline{k}$. Combining the above with gives

$$\rho\phi[1-(1/2-\gamma)(1-\gamma)x^2] = \rho\phi\underline{c}^{1-\gamma} = (1-\tau_\Pi)\overline{c}. \tag{56}$$

The volatility of the wealth (and hence consumption) of entrepreneurs is

$$\sqrt{\rho}x = \frac{\Pi-\delta-r}{\gamma\sigma}$$

Drift and volatility of consumption are

$$\mu_c = \frac{1}{\gamma}\left[(1-\tau_s)r - \rho + \frac{1}{2}\left(\frac{\Pi-\delta-r}{\sigma}\right)^2(1/\gamma+1)\right] \qquad\qquad \sigma_c = \frac{\Pi-\delta-r}{\gamma\sigma}$$

Now note that when $dc_t = \mu_c c_t dt + \sigma_c c_t dZ_t$, the law of motion of $\overline{c}^\gamma$ is given by $dc_t^\gamma = \gamma\left[\mu_c + (\gamma-1)\frac{\sigma_c^2}{2}\right]t + \gamma\sigma_c c_t^\gamma dZ_t$. The inverse-Euler equation then implies $0 = \mu_c + (\gamma-1)\sigma_c^2/2$, or

$$\begin{aligned}
0 &= \gamma\left[(1-\tau_s)r - \rho + \frac{1}{2}\left(\frac{\Pi-\delta-r}{\sigma}\right)^2(1/\gamma+1)\right] + \frac{(\gamma-1)}{2}\left(\frac{\Pi-\delta-r}{\sigma}\right)^2 \\
0 &= \gamma[(1-\tau_s)r-\rho] + \frac{1}{2}[(1+\gamma)+(\gamma-1)]\left(\frac{\Pi-\delta-r}{\sigma}\right)^2 \\
\rho - (1-\tau_s)r &= \left(\frac{\Pi-\delta-r}{\sigma}\right)^2.
\end{aligned}$$

Recall that in general this is

$$\begin{aligned}
\overline{c} &= \rho + \left(\rho - (1-\tau_s)r - \frac{\gamma^{-1}}{2}\left(\frac{\Pi-\delta-r}{\sigma}\right)^2\right)(1/\gamma-1) = \rho - (\gamma-1/2)\left(\frac{\Pi-\delta-r}{\gamma\sigma}\right)^2(\gamma-1) \\
&= \rho[1-(\gamma-1/2)(\gamma-1)x^2].
\end{aligned}$$

Substituting this last expression into (56) then gives the results. □

B.2 Aggregate resource constraints

Aggregate consumption, labour 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. Defining $X := \mathbb{R}\times\{E,W\}$, where the indices $i\in\{E,W\}$ indicate whether or not an agent may be an entrepreneur or a worker, aggregate consumption and output at any date $t\geq 0$ are then,

$$\begin{aligned}
\underline{C}_t &:= \int_X \mathbb{E}\left[c_t^{v,i}\right]\Phi(dv,di), \qquad C_t^T := L_E\mathbb{E}\left[c_t^{T,E}\right] + L_W c_t^{T,W} \\
C_t &:= e^{-\rho_D t}\underline{C}_t + \int_0^t e^{-\rho_D[t-T]}C_t^T dT \\
\underline{Y}_t &:= \int_X \mathbb{E}\left[F\left(K_t^{v,i}, L_t^{v,i}\right) - \delta K_t^{v,i}\right]\Phi(dv,di), \qquad Y_t^T := \mathbb{E}\left[F\left(K_t^{T,E}, L_t^{T,E}\right) - \delta K_t^{T,E}\right] \\
Y_t &:= e^{-\rho_D t}\underline{Y}_t + \int_0^t e^{-\rho_D[t-T]}Y_t^T dT
\end{aligned}$$

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

$$\begin{aligned}
\underline{L}_t^E &:= \int_X \mathbb{E}\left[L_t^{v,i}\right]\Phi(dv,di), \qquad L_t^{T,E} := e^{-\rho_D[t-T]}\mathbb{E}\left[L_t^{T,E}\right] \\
L_t^E &:= e^{-\rho_D t}\underline{L}_t^E + \int_0^t e^{-\rho_D[t-T]}L_t^{E,T}dT.
\end{aligned}$$

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

$$\underline{U}_t = \int_X \Gamma_i\mathbb{E}\left[u\left(c_t^{v,i}\right)\right]\Phi(dv\times di) \qquad\qquad U_t^T = \Gamma_E L_E\mathbb{E}\left[u(c_t^{T,E})\right] + \Gamma_W L_W\mathbb{E}\left[u(c_t^{T,W})\right].$$

C Decentralization

C.1 Agent problems

Proof of Lemma 4.1. Given taxes on profits τ_Π and risk-free savings τ_{sE}, the Hamilton-Jacobi-Bellman equation for the entrepreneur's value function is

$$\rho V_E(a) = \max_{c,k\geq 0} \rho \ln c + ((1-\tau_{sE})ra - c + (1-\tau_\Pi)(\Pi - \delta - r)k)V_E'(a) + \frac{[\sigma(1-\tau_\Pi)]^2}{2} k^2 V_E''(a).$$

Substitution of the assumed form $V_E(a) = \ln a + D_E(w,r)$ into the right-hand side gives

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

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

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

The constant $D_E(w,r)$ then satisfies

$$\rho D_E(w,r) = \rho \ln \rho + (1-\tau_{sE})r - \rho + \frac{(\Pi-\delta-r)^2}{\sigma^2} - \frac{[\sigma(1-\tau_\Pi)]^2}{2}\frac{(\Pi-\delta-r)^2}{\sigma^4(1-\tau_\Pi)^2}$$

which reduces to the desired expression for V_E. The Hamilton-Jacobi-Bellman equation for the worker value function is given by

$$\rho V_W(a) = \max_{c\geq 0} \rho \ln c + ((1-\tau_{sW})ra - c + (1-\tau_{LW})w)V_W'(a).$$

Assuming a solution of the form

$$V_W(a) = \ln\left(a + \frac{(1-\tau_{LW})w}{(1-\tau_{sW})r}\right) + D_W(w,r)$$

substitution in the Hamilton-Jacobi-Bellman equation gives

$$D_W(w,r) = \ln \rho + \frac{(1-\tau_{sW})r - \rho}{\rho}$$

as claimed. □

Proof of Corollary 4.4. From the expression found in Theorem 4.3 we have $r \leq \rho$ if and only if $\Pi - \delta - \rho \leq \sqrt{\rho}\sigma h(\Pi)$. From the definition of h this will be assured as long as

$$1 - \sqrt{1 - (2[\Pi-\delta-\rho]/[\sqrt{\rho}\phi\sigma])^2} \geq \frac{\phi}{2}(2[\Pi-\delta-\rho]/[\sqrt{\rho}\phi\sigma])^2$$

which is always true because $\phi \in (0,1]$ and the inequality $1 - \sqrt{1-x} \geq x/2$ holds for all $x \in [0,1]$, as can be seen by rewriting it as $1 - x/2 \geq \sqrt{1-x}$ and squaring both sides. □