On the Optimality of Differential Asset Taxation

Thomas Phelan*
Federal Reserve Bank of Cleveland

January 30, 2025

Abstract

In this paper I study the optimality of differential asset taxation in an environment with entrepreneurs and workers in which output is stochastic and entrepreneurs can misreport profits and abscond with capital. I show that a stationary efficient allocation may be implemented as an equilibrium with endogenous collateral constraints, transfers to newborns, and linear taxes on profits, investment, and interest. Further, these taxes differ from one another and serve distinct purposes. The profits tax shares risk and depends solely on the severity of the misreporting friction, while the remaining instruments determine the efficient mean and variance of entrepreneurs' consumption growth.

Keywords: Optimal taxation, moral hazard, optimal contracting.

JEL Codes: D61, D63, E62.

*The views expressed in this paper are those of the author and do not necessarily reflect the position of the Federal Reserve Bank of Cleveland or the Federal Reserve System. This paper is based on the second chapter of my PhD dissertation, completed at the University of Minnesota under the guidance of Chris Phelan (no relation). I have benefited from discussions with Dan Carroll, V.V. Chari, Sebastian Di Tella, Matthew Knowles, Narayana Kocherlakota, R. Vijay Krishna, Dirk Krueger, Guillaume Sublet, Pierre Yared, Gustavo Ventura and especially Sergio Ocampo and Vasia Panousi. All remaining errors are of course my own. Replication files may be found at https://github.com/tphelanECON/diff_cap_tax.

1 Introduction

The optimal taxation of capital income has long been a contentious issue in both policy debates and the academic literature. The majority of studies of optimal taxation divide income into labor or capital and assume that a common tax is levied on all capital income.[1] However, capital income can assume a variety of distinct forms, including interest and business profits.[2] Further, a recent literature has documented the rising importance of private business income and has shown that it differs from interest income in several ways. For instance, DeBacker et al. (2023) show that business income is risky, with a variance more than 60 times the variance of labor income, while Smith et al. (2019) show that it is concentrated among top income groups and often falls significantly upon the owner's death, suggesting that it is not solely the passive return on savings.[3] Motivated by these facts, this paper characterizes the optimal lump-sum transfers to newborns and taxes on profits, investment, and interest in an environment in which business income is risky and owners cannot diversify. I find that the optimal taxes are constant, linear, and serve distinct purposes: the tax on profits shares risk, while the remaining instruments jointly determine the interest rate and the efficient mean and variance of entrepreneurs' consumption growth.

I consider a perpetual youth environment in which individuals may either run their own business or work for someone else. Only some individuals are born with the ability to run businesses and this is not observable by the government; so tax policy must provide sufficient incentives for firm formation. Firm output exhibits constant-returns-to-scale in capital and labor and is subject to two agency frictions. First, business profits are subject to idiosyncratic risk and are not observable by the government, and owners may choose to misreport profits and divert a fraction to their consumption. Second, at any time, entrepreneurs may abscond with a fraction of the capital invested in their firm, and thereafter trade only a risk-free bond with exogenous return. These two frictions are motivated by the undiversified nature of business ownership together with the observed presence of collateral constraints.[4]

[1] For recent surveys on optimal taxation see Golosov and Tsyvinski (2015) and Stantcheva (2020).

[2] See Bastani and Waldenström (2020) for cross-country evidence on various forms of capital taxation.

[3] For instance, on page 1678 of Smith et al. (2019) the authors note that in their dataset business profits fall by an average of 82 percent upon the unexpected death of an owner.

[4] For evidence of a lack of diversification, Table I on page 1694 of Smith et al. (2019) shows that the median number of owners of pass-through firms with an owner between the top 1 percent and top 0.1 percent is 2.0. For the empirical relevance of collateral constraints, see, e.g., Cagetti and De Nardi (2006).

I first characterize a particular constrained-efficient allocation in which aggregate capital and the distributions of consumption and firm size are constant over time and the government discounts the welfare of future generations at a constant rate. In doing so I do not restrict attention to a fixed set of instruments but instead allow the government to choose any allocation that respects the constraints imposed by the above agency frictions and private information. The ability of entrepreneurs to misreport profits limits risk-sharing because their consumption must depend on the risky output of their firm in order to induce truthful reporting. Similarly, the ability of entrepreneurs to abscond with capital leads to a "no-absconding constraint" that limits the amount of capital that may be delegated to them. In a partial equilibrium setting with a single principal and a single entrepreneur, these two agency frictions imply constant wedges on both the risk-free and risky assets, and the no-absconding constraint either holds with equality at every date or never holds with equality. In the perpetual youth model with a continuum of agents, the constancy of wedges then implies that the stationary efficient allocation is completely characterized by the aggregate capital stock, the initial consumption of all agents, and the constant mean and variance of entrepreneurs' consumption growth.

I then implement this allocation as a competitive equilibrium in an economy with incomplete markets in which the market structure and taxation policy are chosen to respect the above agency frictions and private information. All agents trade a risk-free bond in zero net supply and are subject to endogenous collateral constraints, which are the most relaxed constraints consistent with the ability of entrepreneurs to abscond with capital. The government chooses lump-sum transfers for newborns, issues government debt, and levies linear taxes/subsidies on interest income, investment, and reported profits, where the latter is firm revenue net of wages, interest on debt, investment taxes, depreciation, and any amount that the owner misreports.[5] The government may choose any policy respecting incentive compatibility and so the transfers and taxes can differ by occupation as long as the entrepreneurs have the incentive to reveal their type and start a firm.

The simple characterization of the efficient allocation is mirrored by an equally simple set of implementations, with the optimal taxes all constant, linear, and admitting closed-form expressions.[6] Because equilibrium allocations are incentive compatible, an optimal

[5]Note that the linearity of the tax on profits implies that the entrepreneur's overall tax liability falls when their firm incurs losses (i.e. earn negative profits).

[6]This paper does not attempt to characterize all implementations of the stationary efficient allocation, but instead focuses on a set of implementations within a relatively standard market structure.

policy is any policy that ensures that the equilibrium consumption processes coincide with their efficient counterparts. The optimal tax on profits depends solely on the severity of the agency frictions and is simply the highest level consistent with incentive compatibility because this maximizes risk-sharing. Further, because the leverage of the entrepreneur depends endogenously on tax policy, the tax on profits does not necessarily reduce the after-tax income of the entrepreneur or primarily serve a redistributive role. In contrast, the taxes on interest and investment and the lump-sum transfers to newborns are chosen to ensure that all agents obtain the efficient level of lifetime utility and that the equilibrium mean and variance of consumption growth coincide with their efficient counterparts. For workers, the latter pair of requirements amounts to ensuring that their consumption is constant, while for entrepreneurs, two separate conditions must be met.

First, when the no-absconding constraint does not hold with equality, a variation of the perturbation argument of Rogerson (1985) implies that the planner wishes to distort the after-tax return on savings below the discount rate in order to reduce the future cost of providing utility. In this case, the model does, in a qualified sense, imply progressive taxes on savings, because entrepreneurs (who are typically richer) face a lower after-tax safe return than workers. Second, to implement the efficient level of investment, the cost of borrowing faced by an entrepreneur's business must fall below the subjective discount rate when the no-absconding constraint does not hold with equality. Because the government can affect the equilibrium interest rate using a tax or subsidy on investment, there are multiple ways in which the above two conditions can be met, and therefore multiple implementations of the efficient allocation. To illustrate this indeterminacy, I then describe two specific implementations in detail. In the first, workers face no taxes and the investment tax is chosen to ensure that the interest rate equals the subjective discount rate. In the second, workers' savings are subsidized and the investment tax is chosen such that the firms' cost of borrowing coincides with the interest rate.

For the benchmark environment described above I suppose that there is no private risk-sharing and that all entrepreneurs are equally productive on average. I subsequently discuss how the results change when one relaxes these assumptions. First, when firm owners can write short-term state-contingent contracts with competitive investors, the optimal tax on profits is zero because these contracts serve the same risk-sharing role. In this case the presence of private risk-sharing alters the equilibrium interest rate, and so relative to the above implementations taxes must adjust so that the after-tax returns and risk borne by entrepreneurs are unchanged. Second, when entrepreneurs differ in their expected returns and the no-absconding constraint holds with equality for no entrepreneur, similar logic to the above continues to apply provided that we allow taxes to depend on productivity. In this case, the efficient cost of borrowing faced by entrepreneurs' firms is decreasing in productivity, and so the model generates a qualified kind of regressivity with respect to business income.

Finally, I conclude the paper with a series of numerical examples to illustrate how the optimal wedges and tax revenue raised vary with the severity of the agency frictions and the number of workers per entrepreneur. The main point of these examples is that although for standard parameters the model can generate large optimal taxes on interest, these often overstate the total tax revenue raised. The primary reason for this is that the tax on interest is only levied on the risk-free component of capital income, and that the allocation of wealth between bonds and capital is chosen by the entrepreneur. Indeed, as agency frictions rise, the fall in the after-tax return on savings is partially offset by an increase in the excess return on capital, and the wedge on the risk-free asset reaches its highest value when agency frictions are at their most severe and profits are not taxed.

Related literature. A vast literature, often referred to as the "Ramsey" approach in honor of Ramsey (1927), has studied optimal taxation in environments in which the government has access to an exogenous set of taxes on capital and labor. As first shown by Chamley (1986) and Judd (1985), in environments with a representative agent it is typically the case that the optimal linear tax on capital income is zero in the long run.[7] The Ramsey framework has been extended to include uninsurable labor income risk by Aiyagari (1995), Conesa et al. (2009) and Dyrda and Pedroni (2023) and to include uninsurable capital income risk by Panousi and Reis (2012), Evans (2014) and Panousi and Reis (2021). In contrast to the current paper, these papers assume that a common tax is levied on all capital income.

A separate literature, beginning with Golosov et al. (2003) and sometimes referred to as the *New Dynamic Public Finance*, considers dynamic extensions of Mirrlees (1971) and considers all allocations that satisfy incentive constraints arising from informational asymmetries.[8] However, the majority of this literature has focused on environments in which the primary source of risk is labor productivity and capital income represents the risk-free return on saving. I follow the approach of considering all allocations that satisfy

[7] However, see Straub and Werning (2020) for some important qualifications of this result.

[8] For a review of this literature see Golosov and Tsyvinski (2015).

incentive constraints, but I allow for multiple assets and heterogeneous returns to capital.[9] A small but growing number of papers share this approach. Albanesi (2006) considers a two-period model with risky returns on capital and unobservable effort, while Shourideh (2013) considers a discrete-time model in which agents may divert capital to consumption prior to investment. I extend Shourideh (2013) by allowing entrepreneurs to abscond with capital and by analyzing several different implementations with incomplete markets. Gerritsen et al. (2024) and Boadway and Spiritus (2025) consider two-period models with exogenous and heterogeneous returns on capital, while Phelan (2023) studies an environment in which output depends on the history of unobserved effort. This paper is also related to recent papers that study the benefits of taxing capital income and wealth, such as Boar and Knowles (2024) and Guvenen et al. (2023), who characterize the optimal linear taxes on capital income and wealth in environments with entrepreneurs and collateral constraints. In contrast to these two papers, I do not impose linearity of taxes as a restriction, but instead show that linear taxes and lump-sum transfers implement a constrained-efficient allocation in which the instruments available to the government are microfounded by agency frictions.

The agency problem I consider is similar to that in Di Tella and Sannikov (2021) except that I omit the possibility of hidden savings and allow entrepreneurs to abscond with a fraction of the capital stock, after which they may trade only a risk-free bond. This latter friction is reminiscent of the literature on limited commitment (see, e.g., Kocherlakota (1996)), because it imposes the restriction that the agent never have an incentive to permanently leave the relationship with the principal.[10] Using arguments adapted from Farhi and Werning (2007), I show that the problem of a planner facing a continuum of agents and a perpetual-youth demographic structure (as in Blanchard (1985)) is isomorphic to the principal's problem for a given pair of multipliers on resource constraints, which are then varied until the resource constraints are satisfied in the stationary allocation. In the decentralization, I assume that entrepreneurs face a continuous-time portfolio problem as in Panousi (2010) and Angeletos and Panousi (2009) that is augmented to incorporate taxes on profits, investment and savings and to include collateral constraints. Finally, the equilibrium notion is reminiscent of Alvarez and Jermann (2000), in the sense that the collateral constraints are the most relaxed constraints consistent with the underlying friction

[9]This notion of constrained efficiency contrasts with Davila et al. (2012), whose paper is similar in spirit to Geanakoplos and Polemarchakis (1985) and assumes that markets are exogenously incomplete.

[10]However, it differs from typical models of limited commitment because the benefit from leaving depends on the delegated capital chosen by the principal, and not on an exogenous stream of endowments.

that the entrepreneurs may abscond with capital. These collateral constraints are therefore endogenous to tax policy, because taxes affect the utility from not absconding.

The outline of this paper is as follows: Section 2 analyzes a principal-agent model with an exogenous interest rate and productivity of capital; Section 3 characterizes stationary efficient allocations in an environment with a continuum of agents; Section 4 implements this allocation in a general equilibrium model with incomplete markets; Section 5 provides intuition for the main results and discusses various extensions; Section 6 computes a series of numerical examples; and Section 7 concludes.

2 Principal-agent model

This section characterizes the optimal risk-sharing arrangement between a risk-averse entrepreneur (she) and a risk-neutral principal (he) in an environment where the entrepreneur may operate a risky production technology, she may divert output to private consumption, and she may abscond with a fraction of the capital under her control. Labor is absent from production, and both the marginal product of capital and the interest rate are exogenous. This problem will later be embedded into a macroeconomic model with a continuum of entrepreneurs subject to idiosyncratic risk and workers in which the marginal product of capital depends on the aggregate resource constraints for both labor and capital.

Preferences and technology. Time is continuous and extends indefinitely. Both the principal and the entrepreneur live forever and discount at the common rate $\rho > 0$. The preferences of the entrepreneur over positive consumption processes $c = (c_t)_{t \geq 0}$ are represented by the function

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

The entrepreneur may operate a linear technology that takes capital as the sole input and is subject to stochastic depreciation shocks. Only the entrepreneur may operate the production function and so the principal must delegate capital to the entrepreneur in order for production to take place. In addition, the entrepreneur may divert output to consumption at a rate of $s_t \in [0, \overline{s}]$ per unit of capital for some $\overline{s} > 0$, and so if the capital delegated follows the process $k = (k_t)_{t \geq 0}$ then the net output received by the principal $Y := (Y_t)_{t \geq 0}$ evolves according to

$$dY_t = (\Pi - \rho - \tau_k - s_t)k_t dt + \sigma k_t dB_t \tag{1}$$

where $(B_t)_{t\geq 0}$ is a standard Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}, P)$ satisfying the usual conditions.[11] In the law of motion (1), the constant Π is the marginal product of capital, ρ is the cost of borrowing, τ_k is a tax on capital, and $\sigma > 0$ is the volatility of the shocks. Both Π and τ_k are fixed exogenously in this section in order to first understand the optimal allocation in partial equilibrium. In the perpetual youth economy with a continuum of agents in Section 3, the marginal product of capital Π will be determined by an aggregate production function and the number of workers in the economy and the tax τ_k will capture the extent to which capital affects the welfare of future generations.[12]

Agency frictions. The delegated capital k_t is observable to both the principal and the agent, but only the agent observes consumption c_t and the diverted output $s_t k_t$. However, the entrepreneur may only consume a fraction $\phi \in (0, 1)$ of the diverted flow $s_t k_t$ per unit of time dt, so that $1-\phi$ may be interpreted as the deadweight loss from diversion. Further, the entrepreneur has no access to a savings technology, and so any diverted output must be immediately consumed. I also assume that the entrepreneur may, at any time, take a fraction $\iota \in (0, 1)$ of the capital delegated to her and abscond, and after doing so trade only the same risk-free bond available to the principal.

Allocations and strategies. An allocation must specify the consumption c of the entrepreneur, the capital k delegated by the principal, and the amount of output $\tilde{s}k$ that the principal recommends the entrepreneur divert to consumption, all as functions of the observed history of output. However, because $\phi < 1$, some output is destroyed whenever it is diverted by the entrepreneur, and so to characterize efficient allocations it is without loss of generality to assume that the principal always recommends $\tilde{s} = 0$. For brevity of notation in what follows I therefore omit reference to $\tilde{s}$.[13]

Definition 2.1. *An allocation is a pair of $\mathcal{F}$-adapted processes (k, c) satisfying $k_t \geq 0$ and $c_t > 0$ for all $t \geq 0$. An allocation (k, c) is admissible if $\mathbb{E}[\int_0^\infty e^{-\rho t} \ln c_t dt], \mathbb{E}[\int_0^\infty e^{-\rho t} c_t dt]$ and $\mathbb{E}[\int_0^\infty e^{-\rho t} k_t dt]$ are well-defined and finite.*

An allocation may be interpreted as a choice of the principal indicating delegated capital

[11] In this section all stochastic processes are assumed to be adapted to the filtration generated by the Brownian motion B.

[12] I emphasize that τ_k is introduced here in order to later relate the problem of the planner in Section 3 to a principal-agent problem and is distinct from the taxes imposed on agents in Section 4. For the welfare notion of this paper, τ_k will turn out to be negative (i.e., a subsidy).

[13] Technically, the principal is also free to recommend that the entrepreneur abscond with capital. However, such allocations cannot be efficient (because $\iota < 1$) and so without loss of generality I restrict attention to allocations for which the principal recommends no absconding in addition to $\tilde{s} = 0$.

and recommended consumption. Given an allocation chosen by the principal, a strategy s of the entrepreneur is then the choice of how much output to divert to consumption. When the entrepreneur varies her strategy s, she alters the law of motion of observed output and changes the measure used to evaluate output paths. This leads to the following definition.

Definition 2.2. *A strategy is an $\mathcal{F}$-adapted process s assuming values in $[0, \overline{s}]$. Denoting the corresponding expectation operator by $\mathbb{E}^s$, the utility from adhering to a strategy s is*

$$U^A(k, c, s) := \mathbb{E}^s\left[\rho \int_0^\infty e^{-\rho t} \ln(c_t + \phi s_t k_t) dt\right]. \tag{2}$$

Given an admissible allocation (k, c), a strategy s is feasible if it vanishes beyond some fixed time T and the utility given in (2) is finite, in which case we define an associated utility process $W^s = (W_t^s)_{t \geq 0}$ by

$$W_t^s := \mathbb{E}^s\left[\rho \int_t^\infty e^{-\rho(t'-t)} \ln(c_{t'} + \phi s_{t'} k_{t'}) dt' \,\middle|\, \mathcal{F}_t\right].$$

Incentive compatibility. Because the entrepreneur's consumption is unobservable, the allocation must be incentive compatible, in the sense that the entrepreneur must not wish to either divert output to consumption or abscond with the delegated capital. An entrepreneur equipped with k units of capital and access only to a bond market with interest rate ρ experiences lifetime utility $\ln(\rho k)$. When utility follows $(W_t)_{t \geq 0}$, the entrepreneur will therefore not abscond provided that capital $(k_t)_{t \geq 0}$ satisfies the "no-absconding constraint"

$$k_t \leq \omega e^{W_t}$$

for all $t \geq 0$, where $\omega := (\rho\iota)^{-1}$. This leads to the following definition.

Definition 2.3. *An admissible allocation (k, c) is incentive compatible if $U^A(k, c, 0) \geq U^A(k, c, s)$ and $k_t \leq \omega e^{W_t^s}$ for all feasible strategies s and $t \geq 0$ almost surely. The set of incentive compatible allocations that give utility W to the entrepreneur is denoted $\mathcal{A}^{IC}(W)$.*

The principal is risk-neutral and so his preferences over incentive compatible allocations are represented by the objective function

$$U^P(k, c) := \mathbb{E}\left[\int_0^\infty e^{-\rho t}[(\Pi - \rho - \tau_k)k_t - c_t]dt\right] \tag{3}$$

which is well-defined and finite whenever the allocation is admissible. The problem of the principal is then defined formally as follows.

Definition 2.4. *Given initial utility W, the problem of the principal is defined to be*

$$V(W) = \sup_{(k,c)\in\mathcal{A}^{IC}(W)} U^P(k,c) \tag{4}$$

and a pair (k,c) attaining the supremum in (4) is termed an optimal (or efficient) allocation.

It will be convenient to write utility in consumption-equivalent units, $u_t := e^{W_t}$, and to denote the principal's associated value function and set of incentive compatible allocations that deliver consumption-equivalent utility u to the entrepreneur by $v(u)$ and $\tilde{\mathcal{A}}^{IC}(u)$, respectively. Before stating the formal characterization of the optimal allocation, I first partially characterize the optimal allocation using homogeneity and perturbation arguments.

Informal characterization. First, it follows directly from Definition 2.3 that for any scalars $\lambda, u > 0$, $(k,c) \in \tilde{\mathcal{A}}^{IC}(u)$ if and only if $(\lambda k, \lambda c) \in \tilde{\mathcal{A}}^{IC}(\lambda u)$. Because the principal's objective in (3) is homogeneous of degree one in (k,c), this implies that the value and policy functions of the principal are linear in u, at least if the value function is finite. In this case the problem of the principal reduces to choosing two scalars, $\overline{k} := k_t/u_t$ and $\overline{c} := c_t/u_t$, denoting capital and consumption per unit of consumption-equivalent utility, and so the capital-to-consumption ratio k_t/c_t is constant over time and independent of history.

Second, a perturbation argument may be employed to derive the optimal intertemporal distortions. An important observation in the dynamic contracting literature, established by Rogerson (1985) in a principal-agent setting and by Golosov et al. (2003) in a dynamic Mirrleesian setting, is that intertemporal distortions often satisfy an inverse Euler equation. This result rests on the insight that if an allocation is efficient, it cannot be possible to perturb it in such a way that it remains incentive compatible, delivers the same utility to the entrepreneur, and increases the payoff to the principal. A similar argument is applicable here: if (k,c) is the efficient allocation then for any scalars z, t_0, t_1 and dt with $t_0 + dt < t_1$ and $dt > 0$, we define the following process for an arbitrary history after time t_0,

$$\eta_t^z = \begin{cases} e^z & \text{if } t \in [t_0, t_0 + dt] \\ e^{-ze^{\rho(t_1 - t_0)}} & \text{if } t \in [t_1, t_1 + dt]. \end{cases}$$

Because preferences are logarithmic and (k,c) is incentive compatible, for any scalar z the allocation $(\eta^z k, \eta^z c)$ satisfies $U^A(\eta^z k, \eta^z c, 0) \geq U^A(\eta^z k, \eta^z c, s)$ for all feasible strategies s and delivers the same utility to the entrepreneur as (k,c). The fact that (k,c) is efficient then implies that expected profits must be maximized at $z = 0$, and so differentiating with respect to z, evaluating at zero, and applying the above argument to any history after time

t_0 gives $k_{t_0} = \mathbb{E}_{t_0}[k_{t_1}]$. Because the capital-to-consumption ratio is independent of history by the above homogeneity argument, this implies $c_{t_0} = \mathbb{E}_{t_0}[c_{t_1}]$, which is the inverse Euler equation under logarithmic utility.

However, there are two potential problems with the above arguments. First, the principal's problem may fail to be finite-valued. For example, if $\phi = 0$ and $(\Pi - \rho - \tau_k)\omega > 1$, then the payoff associated with $(\overline{k}, \overline{c}) = (\omega, 1)$ is increasing and linear in u, and therefore convex and unbounded in W, and so the principal could obtain arbitrarily high profits by offering the entrepreneur an initial lottery over allocations of the above form.[14] Second, even if the value function were finite, if the no-absconding constraint holds with equality, then the above perturbed allocation is only incentive compatible if $z \leq 0$. The main content in Proposition 2.1 below is that when the excess return on capital is sufficiently small, both of these technical problems do not arise and the above claims can be formally justified.

Formal characterization. I now state the key properties of the optimal allocation, leaving formal proofs to Appendix A. In any admissible allocation, promised utility evolves according to $dW_t = \rho(W_t - \ln c_t)dt + \tilde{\sigma}_{W,t} dB_t$ for some process $\tilde{\sigma}_{W,t}$, and the entrepreneur will choose not to divert output if and only if

$$0 \in \arg\max_{s \geq 0} \rho \ln(c_t + \phi k_t s) - s\tilde{\sigma}_{W,t}/\sigma. \tag{5}$$

The minimal value of the diffusion term necessary to dissuade diversion is $\tilde{\sigma}_{W,t} = \rho\phi\sigma k_t/c_t$, and so $\tilde{\sigma}_{W,t} dB_t$ may be viewed as the product of three terms: the shocks to output, $\sigma k_t dB_t$, the marginal utility of consumption, ρ/c_t, and the fraction ϕ of each unit of diverted output that the entrepreneur may actually consume. The term $\tilde{\sigma}_{W,t}$ may be interpreted as the "skin-in-the-game" necessary to align the incentives of the principal and entrepreneur.

As noted above, wherever they are well-defined, the optimal policies of the principal are of the form $c_t = \overline{c}u_t$ and $k_t = \overline{k}u_t$ for some constants $\overline{c}, \overline{k} > 0$. An application of Ito's lemma then shows that consumption satisfies

$$dc_t = \mu_c c_t dt + \sigma_c c_t dB_t \tag{6}$$

for $\mu_c = \mu_c(\overline{c}, x) := \rho\big(-\ln \overline{c} + x^2/2\big)$ and $\sigma_c := \sqrt{\rho}x$, where I changed variables to

$$x := \sqrt{\rho}\phi\sigma\overline{k}/\overline{c}. \tag{7}$$

[14] As noted by Di Tella and Sannikov (2021), the value function of the principal is not finite-valued for any $\Pi > \rho + \tau_k$ when utility is logarithmic and there is no hidden savings and the entrepreneur cannot abscond with capital, and so the no-absconding constraint is essential for the problem considered here.

One natural guess for the value function of the principal is that it is found by maximizing over the constants $\overline{c}$ and x that satisfy the no-absconding constraint and ensure a finite cost of consumption. If I define the two parameters

$$S := \frac{\Pi - \rho - \tau_k}{\sqrt{\rho}\phi\sigma} \qquad\qquad \overline{\omega} := \frac{\sqrt{\rho}\phi\sigma}{\rho\iota} \tag{8}$$

then the expected flow output per unit of u is $(\Pi - \rho - \tau_k)\overline{k} = Sx\overline{c}$, and so by the Gordon growth formula, one candidate for the value function of the principal is $v(u) \equiv \overline{v}u$, where

$$\overline{v} := \sup_{\substack{\overline{c}>0, x\geq 0 \\ x\overline{c}\leq\overline{\omega}, \mu_c(\overline{c},x)<\rho}} \frac{(Sx-1)\overline{c}}{\rho - \mu_c(\overline{c},x)}. \tag{9}$$

Note that the variable S is the ratio of the excess return on capital, $\Pi - \rho - \tau_k$, to a measure of the severity of the agency friction, $\sqrt{\rho}\phi\sigma$. In what follows I will write (9) and the associated optimal choices as $\overline{v}(S,\overline{\omega}), \overline{c}(S,\overline{\omega})$ and $x(S,\overline{\omega})$, respectively, whenever there is a need to emphasize the dependence on S and $\overline{\omega}$.[15]

The maximization in (9) is subtle both because the maximand is not concave in the choice variables $\overline{c}$ and x and because the constraint set is unbounded. Indeed, the existence of values attaining the supremum (or even the finiteness of the supremum) is not assured without further assumptions. Proposition 2.1 shows that the principal's value function is finite-valued (and in fact, negative-valued) and coincides with the conjectured value $\overline{v}u$ provided that the parameter S is sufficiently small.

Proposition 2.1. *For any $\overline{\omega} > 0$, there exist $\overline{S}(\overline{\omega}), \tilde{S}(\overline{\omega}) > 0$ satisfying $\tilde{S}(\overline{\omega}) \leq \overline{S}(\overline{\omega})$ such that the following holds:*

1. *The principal's value function is everywhere negative and given by $v(u) = \overline{v}u$ for all $u > 0$ if and only if $S \leq \overline{S}(\overline{\omega})$; and*

2. *The no-absconding constraint is strict if $S < \tilde{S}(\overline{\omega})$.*

Further, $x(S,\overline{\omega})$ is increasing in S wherever it is well-defined.

Proof. See Appendix A.2. □

[15] In the event that there are multiple solutions to the maximization problem in (9), I assume that the principal chooses the one for which the corresponding value of x is the lowest. Multiple solutions do not appear to arise in any of the numerical examples computed in this paper.

The principal's problem reduces to making just two choices, $\overline{c}$ and x, whenever the optimal policy is well-defined. This implies that the no-absconding constraint $\overline{c}x \leq \overline{\omega}$ either never holds with equality after any history or holds with equality after every history. When it holds as a strict inequality, the above perturbation argument is applicable and consumption is a martingale. In contrast, when the no-absconding constraint holds with equality, the principal wishes to backload utility in order to relax the future no-absconding constraints, and therefore introduces an upward drift in consumption.

Wedges. 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 and the assumed market structure. To isolate the role of agency frictions independently of a particular implementation, I will first analyze optimal wedges in partial equilibrium. If $u(c) \equiv \ln c$ and the return from continually investing in an asset over the interval $[t, t+\Delta]$ is $R = R_{t,t+\Delta}$, then intertemporal optimization implies $u'(c_t) = e^{-\rho\Delta}\mathbb{E}[Ru'(c_{t+\Delta})|\mathcal{F}_t]$. The following notion measures the extent to which this relationship fails for an arbitrary return.

Definition 2.5. *Given a consumption process $(c_t)_{t\geq 0}$ and asset A with return $(R_t^A)_{t\geq 0}$ the associated wedge $\nu_{t,t'}^A$ between two dates t and $t' > t$ is defined by*

$$u'(c_t) = e^{-\rho(t'-t)}\mathbb{E}_t\Big[e^{-\nu_{t,t'}^A(t'-t)}R_{t,t'}^A u'(c_{t'})\Big]. \tag{10}$$

Denote by ν^K and ν^B the wedges associated with capital and the bond, respectively, and note that the associated log returns are $\ln R_t^K = \big(\Pi - \tau_k - \sigma^2/2\big)t + \sigma B_t$ and $\ln R_t^B = \rho t$. These wedges 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 entrepreneur. In principle, for an arbitrary consumption process the wedges defined in Definition 2.5 could depend on both time and the length of the interval in equation (10). However, the fact that both efficient log consumption and log returns possess constant drift and diffusion terms implies that these wedges are independent of both time and history. Further, when the no-absconding constraint is satisfied as a strict inequality, we have the following comparative statics with respect to the marginal product of capital.

Proposition 2.2. *For the set of Π such that the no-absconding constraint holds as a strict inequality, the wedge on the bond ν^B and the difference in wedges $\nu^B - \nu^K$ are both non-negative and increasing in Π.*

Proof. See Appendix A.2. □

Proposition 2.2 shows that when the no-absconding constraint is strict, it is efficient to distort the return on the entrepreneur's savings below the risk-free rate, and that the magnitude of this distortion is an increasing function of the marginal product of capital. In contrast, the sign of the wedge on the risky asset is in general ambiguous. However, although the wedge of the risky asset cannot be signed, when the no-absconding constraint is strict the risky wedge is always lower than the wedge on the bond and the difference is increasing in the marginal product of capital.

The history-independence of the above wedges anticipates the results in Section 4, where it is shown that in a general equilibrium environment with a continuum of agents, the optimal taxes are linear and independent of both age and wealth. The key point is that when faced with a portfolio problem with linear taxes on all forms of income, an entrepreneur will devote a constant fraction of her (total) wealth to the risky asset (her business), which implies that her consumption evolves according to geometric Brownian motion, just as in the above agency problem. The optimal taxes are then chosen to ensure that the equilibrium process for consumption coincides exactly with the efficient process given above.

Before turning to the environment with a continuum of agents, I now summarize the key insights that emerge from this partial equilibrium setting. First, whenever the principal's problem is finite-valued, the optimal allocation takes a simple form and the entrepreneur's consumption evolves according to geometric Brownian motion. Second, the no-absconding constraint may or may not hold with equality in the optimal allocation, and it either holds with equality after every history or never holds with equality after any history. Third, the risk borne by the entrepreneur is an increasing function of the ratio S of the excess return on capital to the severity of the agency frictions. The task of the next section is to explain how these agency frictions determine the marginal product of capital (and hence S) when some agents (workers) work for others (entrepreneurs) and output is a constant-returns-to-scale function of both capital and labor. Intuitively, when agency frictions are small, the principal wishes to delegate more capital to the entrepreneur, which tends to increase the capital stock and therefore reduce the marginal product of capital. The net effect of a fall in ϕ on the key ratio S is therefore not obvious (because it reduces the numerator and the denominator), and requires a general equilibrium analysis, to which I now turn.

3 Stationary efficient allocations

Section 2 characterized the efficient allocation in an environment with a risk-averse entrepreneur and a risk-neutral principal given an exogenous interest rate and productivity of capital. This section uses the above to characterize a particular stationary efficient allocation in a production economy with a continuum of entrepreneurs and workers subject to idiosyncratic risk.

Physical environment. Time is again continuous and extends indefinitely. At any moment there is a unit mass of agents who discount at rate $\rho_S > 0$, die at rate $\rho_D > 0$ and are endowed with L units of labor. To fix the population at unity, new agents are born at rate ρ_D, and the agents born at a particular date $t \geq 0$ will be referred to as a *generation*. All agents have preferences over consumption represented by

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

where $\rho := \rho_S + \rho_D$. Agents may either run a firm or work for someone else. However, only a fraction $1 - \psi \in [0, 1]$ of each generation, termed entrepreneurs, is capable of running a firm, with the remaining fraction, termed workers, only able to work for someone else. I follow Angeletos (2007) and assume that these activities are not mutually exclusive and that entrepreneurs may perform both simultaneously. Whether an agent is an entrepreneur or a worker will be private information and will be referred to as their *type* and indexed by $i \in \{E, W\}$. Entrepreneurs have access to a production technology that produces consumption using physical capital and labor and is exposed to idiosyncratic risk, and production is subject to the same agency frictions as in Section 2. Specifically, an entrepreneur may abscond with a fraction $\iota \in (0, 1)$ of the capital in her business and after doing so trade only a bond with return ρ.[16] She may also divert a flow of capital, with each unit diverted yielding $\phi \in (0, 1)$ units of consumption. If capital and labor are assigned to an entrepreneur according to the processes $(k_t, l_t)_{t\geq 0}$ and the entrepreneur adheres to the diversion strategy $(s_t)_{t\geq 0}$ then output satisfies

$$dY_t = \big(Ak_t^\alpha l_t^{1-\alpha} - \delta k_t - s_t k_t\big)dt + \sigma k_t dB_t,$$

where $B := (B_t)_{t\geq 0}$ is a standard Brownian motion and $A, \delta > 0$ and $\alpha \in (0, 1)$ are exogenous constants. The Brownian motion B represents depreciation shocks that are idiosyncratic to the entrepreneur, and is assumed to be independent across entrepreneurs. Because

[16]Note that because this return is exogenous (and set to ρ for simplicity), I am interpreting "absconding" as leaving the jurisdiction of the planner.

the production function is constant-returns-to-scale in labor and capital and the shocks are idiosyncratic, for the welfare notion given below and any given wage (or, technically, multiplier on the labor resource constraint) and interest rate the problem of the planner facing an individual entrepreneur will be isomorphic to the principal-agent problem of Section 2.

However, in contrast to the model of Section 2, an allocation is now indexed by an initial distribution Φ over promised utility and types, and must specify the consumption, capital and labor delegated to an entrepreneur as a function of initial utility or date of birth, type, and history of her output. Because agents supply labor inelastically, I will omit labor supply from the definition of an allocation and I will also assume without loss of generality that the planner never recommends that an entrepreneur divert a positive amount of output or abscond with capital. In the following, $(c^v_{it}, k^v_{it}, l^v_{it})$ and $(c^T_{it}, k^T_{it}, l^T_{it})$ refer to the consumption, capital and labor assigned to a given type i at date t, where, for agents alive at the initial date, the superscript indicates promised utility, and for all remaining agents, the superscript indicates their birth date.

Definition 3.1. *Given a distribution Φ over utility and types, an allocation consists of sequences $(c^v_{it}, k^v_{it}, l^v_{it})_{t\geq 0}, (v, i) \in \operatorname{supp}(\Phi)$ for the initial generation and $(c^T_{it}, k^T_{it}, l^T_{it})_{t\geq T\geq 0}$, $i = E, W$, for subsequent generations. An allocation satisfies promise-keeping if $U(c^v_i) = v$ for all $(v, i) \in \operatorname{supp}(\Phi)$, and is incentive compatible if, in addition, $U(c^T_E) \geq U(c^T_W)$ for all $T \geq 0$ and the allocations to entrepreneurs satisfy Definition 2.3 in Section 2.*

The planner need not worry about double deviations, in which the entrepreneur misreports type and then diverts output, because workers cannot pretend to be entrepreneurs and entrepreneurs who pretend to be workers are not entrusted with any capital and so thereafter have no private information. In what follows I denote by $C_t(A), K_t(A), Y_t(A)$ and $L_t(A)$ aggregate consumption, capital, output, and labor assigned at t in allocation A, which are restricted to be differentiable functions of time.[17]

Definition 3.2. *An allocation A is resource feasible given capital stock K if $K_0(A) = K$, $C_t(A) + \dot{K}_t(A) \leq Y_t(A)$ and $L_t(A) \leq L$ for all $t \geq 0$, and is incentive feasible if it is both resource feasible and incentive compatible. The set of all incentive feasible allocations given Φ and K will be denoted $\mathcal{A}^{IF}(\Phi, K)$.*

Welfare notion. I will assume that the planner cares only about workers and values the flow utility of a worker at any date the same regardless of their date of birth, which

[17]The formal expressions are not necessary for the discussion here and so are relegated to Appendix B.1.

amounts to placing weight $e^{-\rho_S T}$ on the Tth generation. This is equivalent to the objective

$$U^P(A) = \int_0^\infty \left(e^{-\rho t} \underline{U}_{Wt}(A) + \rho_D \int_0^t e^{-\rho_S T} e^{-\rho(t-T)} U^T_{Wt}(A) dT \right) dt \tag{11}$$

where $\underline{U}_{Wt}(A)$ and $U^T_{Wt}(A)$ are the flow utility of workers in the initial and Tth generations at time $t \geq 0$, respectively, conditional on being alive, in the allocation A. The fact that the planner only values the utility of workers means that the planner just gives the entrepreneurs the lowest utility necessary to reveal their type.[18] Given (Φ, K), the planner's problem is then

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

In this paper I restrict attention to efficient allocations in which aggregate capital and the cross-sectional distributions of consumption, capital and utility are constant over time. The method by which this is achieved is similar to that followed in Farhi and Werning (2007) and so details are relegated to Appendix B.2. Essentially, one first relaxes the problem of the planner by considering the allocation he or she would choose if he or she could trade intertemporally at the subjective rate of discount. In this way, both the law of motion of capital and the labor resource constraint are replaced with constraints on the present value of resources, and the problem decomposes into many problems all identical in form to the principal-agent problem given in Definition 2.4. If, for some initial utility distribution and capital stock, the implied distributions of consumption and capital are constant and the planner does not wish to trade, then this present value constraint implies that the resource constraint is satisfied every period and we have a found a particular stationary efficient allocation.

Characterization of efficient allocation. Relative to the principal-agent setting in Section 2, the resource constraints affect the analysis in two ways. The production technology and the stock of labor jointly determine the marginal product of capital, while the presence of the Pareto weights on future generations leads the planner to behave as if he or she faced a subsidy on capital. The problem of the relaxed planner facing a newborn entrepreneur is identical to the problem of the principal in Section 2 in which $\tau_k = -\rho_D$ and the marginal product of capital is $\Pi = \alpha A(K/L)^{\alpha-1} - \delta$. The efficient marginal product of capital is then the value such that the associated ratio $S = (\Pi - \rho_S)/(\sqrt{\rho}\phi\sigma)$, the

[18]Relaxing the assumption of unobservable types and then placing positive weights on entrepreneurs would change the initial utility of each type and the resource constraint but not the qualitative features of the allocation (at least if the conditions in Proposition 2.1 ensuring a finite value function are satisfied).

key variable in Section 2, implies that the goods resource constraint holds. One technical subtlety in this process is that such a value for S might fail to exist, because aggregate consumption and capital will diverge if the growth in consumption exceeds the rate of death. In this paper I rule out this possibility by making the following assumption, where I recall that $\mu_c(S,\overline{\varpi})$ and $\sigma_c(S,\overline{\varpi})$ were defined in Section 2.[19]

Assumption 3.1. *There exists a solution $\hat{S}$ to the equation*

$$(1-\psi)\overline{C}(S)+\psi = ((S\sqrt{\rho}\phi\sigma+\rho_S)/\alpha+(1/\alpha-1)\delta)(1-\psi)\overline{K}(S) \tag{12}$$

such that the principal's value function in Section 2 with $\tau_k=-\rho_D$ is finite-valued and $\mu_c(\hat{S},\overline{\varpi})<\rho_D$, where $\overline{C}(S)$ and $\overline{K}(S)$ are the stationary amount of consumption and capital delegated to entrepreneurs per unit of initial utility for a given S.

The explicit expressions for the consumption and capital aggregates $\overline{C}(S)$ and $\overline{K}(S)$ are not important for what follows and so are relegated to Appendix B.3. The following characterizes the stationary efficient allocation for the above welfare notion and parameters and essentially amounts to rearranging the goods resource constraint.

Proposition 3.1. *A solution to equation (12) is unique whenever Assumption 3.1 is satisfied. In this case, an efficient stationary allocation exists in which the capital stock is $\hat{K}=(\alpha A/(\hat{S}\sqrt{\rho}\phi\sigma+\rho_S+\delta))^{\frac{1}{1-\alpha}}L$, the workers' consumption is constant and the entrepreneurs' consumption satisfies*

$$dc_t=\mu_c(\hat{S},\overline{\varpi})c_t dt+\sigma_c(\hat{S},\overline{\varpi})c_t dB_t. \tag{13}$$

Proof. See Appendix B.3. □

Note that when the no-absconding constraint holds as a strict inequality, the resource constraint simplifies because the drift in the consumption of entrepreneurs vanishes. Proposition 3.2 below shows that this will occur when agency frictions are sufficiently small in the following sense. Throughout this paper, when the agency frictions are varied, the parameters governing the diversion and absconding constraints will be assumed to be in fixed proportions to one another, so that $\iota\equiv\phi\bar{\iota}$ for some $\bar{\iota}\in(0,1]$ and all $\phi\in(0,1)$.[20]

[19] Note that in what follows, I distinguish efficient quantities with hat notation, so that, e.g., the efficient value of S is denoted by $\hat{S}$.

[20] Note that when $\bar{\iota}$ is fixed in this manner, the allocation with no agency frictions arises as $\phi\to 0$.

Proposition 3.2. *Assumption 3.1 is satisfied for all sufficiently small agency frictions. The solution $\hat{S}$ is increasing in ϕ wherever it is well-defined and tends to zero as $\phi \to 0$, and so the no-absconding constraint holds as a strict inequality for all sufficiently small ϕ.*

Proof. See Appendix B.3. $\square$

Proposition 3.2 is noteworthy because the comparative statics in the environment with a continuum of agents are the opposite of those that obtain in the partial equilibrium environment of Section 2. When the marginal product of capital is fixed as in Section 2, the parameter S governing the consumption risk borne by the entrepreneur mechanically increases as ϕ falls. However, in the infinite-horizon setting with an aggregate production technology, a reduction in agency frictions increases the incentive to delegate capital to the entrepreneur, which tends to increase the capital stock and therefore reduces the marginal product of capital. Proposition 3.2 shows that the latter force always overwhelms the partial equilibrium effect, so that the risk borne by entrepreneurs in the above stationary efficient allocations is increasing in agency frictions.

The above stationary efficient allocation is completely described by the requirements that all newborns attain the same level of utility, workers' consumption is constant, entrepreneurs' consumption evolves according to (13) and the capital stock is given by Proposition 3.1. The next section characterizes the taxes and transfers that ensure that these properties arise in a stationary competitive equilibrium with collateral constraints.

4 Decentralization

Section 3 characterized a particular stationary efficient allocation, with the distribution of resources implicitly conducted by a benevolent social planner. In this section I show how this allocation may be implemented with taxes and transfers when agents trade in decentralized markets. In order for this to be a coherent exercise, it is essential that the market structure described below and policy instruments respect the incentive constraints inherent in the environment of Section 3. For this reason, the equilibrium notion introduced below in Definition 4.2 will require that entrepreneurs wish to reveal their type at birth and have no incentive to misreport income or abscond with the capital invested in their firm. Further, I deliberately endow the government with more instruments than necessary to implement the efficient allocation in order to emphasize that the efficient allocation is consistent with a variety of different pre-tax prices and taxes.

Tax instruments and market structure. All agents receive a constant flow wL of labor income while alive, where w is the competitively determined wage, and trade a risk-free bond in zero net supply with (endogenous) return r. Agents also contract with life insurance companies, receiving a return ρ_D on their wealth a_t when alive in exchange for forfeiting their wealth at death.[21] The government issues debt, transfers wealth to newborn agents, and imposes taxes on various forms of income. Because I focus on implementing the stationary efficient allocation from Section 3, all of these policy instruments (debt, transfers, and taxes) will be time-invariant constants.

Government debt is denoted by D, and for each $i \in \{E, W\}$, an agent of type i inherits a multiple η_i of the aggregate capital stock at birth and faces constant linear taxes $\tau_{si}, \tau_{Li}, \tau_\pi$ and τ_I on interest income (or savings), labor income, profits, and investment, respectively. For brevity, in what follows I will write

$$r_{si} = (1 - \tau_{si})(r + \rho_D) \tag{14}$$

for the after-tax safe return available to an agent of type i. Entrepreneurs may fund the capital k_t invested in their firm either by reducing their personal bond holdings b_{et} or by taking out a business loan, so that at any time $t \geq 0$ their wealth satisfies $a_t = b_{et} + b_{bt} + k_t$ for some $b_{bt} \in [-k_t, 0]$.[22] The tax τ_π is levied on reported business profits, defined as output net of wages, interest paid on business loans, depreciation, taxes on investment $\tau_I k_t$, and underreported income $s_t k_t$. The entrepreneur's wealth therefore satisfies

$$\begin{aligned} da_t = {} & [(1 - \tau_{sE})\rho_D a_t + (1 - \tau_{sE}) r b_{et} - c_t + (1 - \tau_{LE}) wL] dt \\ & + (1 - \tau_\pi)\big[(A k_t^\alpha l_t^{1-\alpha} - w l_t + r b_{bt} - (\delta + \tau_I + s_t) k_t) dt + \sigma k_t dB_t\big]. \end{aligned} \tag{15}$$

An entrepreneur will choose $b_{bt} = -k_t$ if $(1 - \tau_\pi) r < (1 - \tau_{sE}) r$ and $b_{bt} = 0$ otherwise, and so when $\tau_\pi < 1$, the law of motion (15) can be written more succinctly by first defining the following variable that represents the effective cost of borrowing faced by their firm,

$$r_b := \min\left\{ r, \left(\frac{1 - \tau_{sE}}{1 - \tau_\pi}\right) r \right\} + \tau_I, \tag{16}$$

which coincides with r in the absence of taxes. The after-tax excess return on capital may therefore be written $(1 - \tau_\pi) dR_t$, where

$$dR_t := (A(l_t/k_t)^{1-\alpha} - w l_t / k_t - \delta - r_b - s_t) dt + \sigma dB_t.$$

[21] Note that the ownership of these life insurance companies is irrelevant because they make zero profits.

[22] The restriction $b_{bt} \leq 0$ ensures that b_{bt} represents a loan, while the restriction $b_{bt} \geq -k_t$ implies that the entrepreneur cannot take out a business loan and place the funds in their personal account.

Note that because dR_t may assume both positive and negative values, the above law of motion of wealth embodies the assumption that the entrepreneur receives a tax offset if her firm sustains losses. In this way, the tax on profits provides risk-sharing between the government and the entrepreneur, because only a fraction $1-\tau_\pi$ of the idiosyncratic shock $\sigma k_t dB_t$ passes through to the entrepreneur's income. Similarly, note that because b_{et} may assume either sign, the linearity of the above tax on interest implies that the borrowing of an indebted entrepreneur is subsidized.[23]

Borrowing and collateral constraints. There are no ad-hoc borrowing constraints, and so the only restriction on the possible values of an agent's wealth is that total wealth, the sum of financial wealth a_t and human wealth, must remain non-negative, where human wealth is the present discounted value of after-tax labor income and equal to

$$h_i := \int_0^\infty e^{-r_{si}t}(1-\tau_{Li})wLdt = (1-\tau_{Li})wL/r_{si}, \tag{17}$$

which is well-defined if $r_{si} > 0$. In terms of this human wealth and the excess return on capital, the law of motion (15) may be written $da_t = [r_{sE}(a_t + h_E) - c_t]dt + (1-\tau_\pi)k_t dR_t$.

In addition, because entrepreneurs may abscond with a fraction of the capital stock, the amount of capital invested in their firm will be subject to a collateral constraint, in which the capital invested cannot exceed a multiple $\overline{\omega}_{\mathrm{d}}$ of total wealth.[24] The equilibrium notion adopted below in Definition 4.2 will then impose the requirement that this collateral constraint is the least restrictive value such that no entrepreneur ever wishes to abscond with the capital invested in their business.

Individual problems. The wage, interest rate, and taxes are only relevant to the entrepreneur insofar as they affect her human wealth, h_E, after-tax safe return, r_{sE}, excess return on capital, dR_t, the constant $\overline{\omega}_{\mathrm{d}}$ in the collateral constraint, and the tax on profits τ_π. I therefore write the individual problems and implementation in terms of these quantities because this will simplify the subsequent analysis of optimal policy.

Definition 4.1. *Given taxes* $\tau \equiv (\{\tau_{si}, \tau_{Li}\}_{i\in\{E,W\}}, \tau_I, \tau_\pi)$*, wage* w*, interest rate* r*, and*

[23] Note that when income is taxed at a common rate $\tau = \tau_{LE} = \tau_{sE} = \tau_\pi$, the equation (15) simplifies to $da_t = (1-\tau)\big[(\rho_D a_t + r(b_{et}+b_{bt}) + wL + Ak_t^\alpha l_t^{1-\alpha} - wl_t - (\delta+\tau_I+s_t)k_t)dt + \sigma dB_t\big] - c_t dt$, and the above amounts to allowing the entrepreneur to deduct interest paid on debt from their taxable income.

[24] Wherever relevant, I use the subscript d for quantities appearing in the decentralized environment in order to avoid confusion with their efficient counterparts considered in Section 3.

collateral parameter $\overline{\omega}_{\mathrm{d}}$, the problem of an entrepreneur with wealth a is given by

$$\begin{aligned} V_E(a) &= \max_{(c_t,k_t,l_t,s_t)_{t\geq 0}} \mathbb{E}\left[\rho\int_0^\infty e^{-\rho t}\ln(c_t+\phi s_t k_t)dt\right] \\ da_t &= [r_{sE}(a_t+h_E)-c_t]dt + (1-\tau_\pi)k_t dR_t \\ k_t &\leq \overline{\omega}_{\mathrm{d}}(a_t+h_E), \quad \forall t\geq 0 \\ 0 &\leq a_t+h_E, \quad \forall t \geq 0, \end{aligned}$$

while the problem of a worker with wealth a is given by

$$\begin{aligned} V_W(a) &= \max_{(c_t)_{t\geq 0}} \mathbb{E}\left[\rho\int_0^\infty e^{-\rho t}\ln c_t dt\right] \\ da_t &= [r_{sW}(a_t+h_W)-c_t]dt \\ 0 &\leq a_t+h_W, \quad \forall t\geq 0. \end{aligned}$$

Before characterizing the solution to the individual problems, I make two preliminary observations. First, for any constant wage the optimal choice of labor per unit of capital is constant across entrepreneurs and solves $\Pi := \max_{z\geq 0} Az^{1-\alpha} - wz - \delta$, so that the optimal excess return on capital may be written $dR_t = (\Pi - r_b - s)dt + \sigma dB_t$. Second, the entrepreneur will choose $s_t = 0$ if and only if $\tau_\pi \leq 1-\phi$. As in Section 3 I restrict attention to this case and write a stationary allocation as $\{(c_{it}, k_{it}, l_{it})_{t\geq 0}\}_{i\in\{E,W\}}$. The problem of the entrepreneur then admits the following simple characterization.

Lemma 4.1. *The entrepreneur will choose not to divert capital if and only if $\tau_\pi \leq 1-\phi$, in which case her value function is $V_E(a) = \ln\rho + \ln(a+h_E) + \rho^{-1}\left(\mu_{c,\mathrm{d}} - \sigma^2_{c,\mathrm{d}}/2\right)$ and her consumption satisfies $dc_t = \mu_{c,\mathrm{d}}c_t dt + \sigma_{c,\mathrm{d}}c_t dB_t$, where*

$$\mu_{c,\mathrm{d}} = r_{sE} - \rho + (1-\tau_\pi)(\Pi - r_b)\overline{k}_{\mathrm{d}} \qquad \sigma_{c,\mathrm{d}} = (1-\tau_\pi)\sigma\overline{k}_{\mathrm{d}} \tag{18}$$

and the constant $\overline{k}_{\mathrm{d}}$ is given by

$$\overline{k}_{\mathrm{d}} = \min\left\{\frac{\Pi - r_b}{\sigma^2(1-\tau_\pi)}, \overline{\omega}_{\mathrm{d}}\right\}.$$

The entrepreneur's policy functions are $c_E(a) = \rho(a+h_E)$ and $k_E(a) = \overline{k}_{\mathrm{d}}(a+h_E)$, and the worker's policy and value functions are $c_W(a) = \rho(a+h_W)$ and $V_W(a) = \ln\rho + \ln(a+h_W) + (r_{sW}/\rho - 1)$, respectively.

Proof. See Appendix C.1. ∎

Equilibrium notion. The equilibrium notion I adopt below will impose the familiar conditions that consumers optimize and markets clear, together with two additional requirements motivated by the environment in Section 3. First, the equilibrium utility of entrepreneurs must weakly exceed that of workers at birth; otherwise, entrepreneurs will not reveal their type. Second, the constant $\overline{\omega}_{\mathrm{d}}$ must equal the largest value such that the entrepreneur will never wish to abscond with her firm's capital. This latter requirement is similar in spirit to that imposed in the equilibrium concept introduced in Alvarez and Jermann (2000), in which their solvency constraints are described as "not too tight." I emphasize that this equilibrium condition is consistent with the environment of Section 3, in which an entrepreneur who absconds with capital has access only to a fraction of the stolen capital and a bond with return ρ, but no labor income. The entrepreneur will never wish to abscond with the capital invested in her firm if and only if $\ln(\rho\iota k) \leq V_E(a)$ for all $k \leq \overline{\omega}_{\mathrm{d}}(a + h_E)$. By Lemma 4.1, the most relaxed collateral constraint consistent with no absconding corresponds to the value

$$\overline{\omega}_{\mathrm{d}} = \iota^{-1} e^{\left(\mu_{c,\mathrm{d}} - \sigma^2_{c,\mathrm{d}}/2\right)/\rho} \tag{19}$$

which is endogenous to tax policy because taxes affect the drift and diffusion of consumption and hence the utility from not absconding. Finally, I will write $\kappa_i K := \eta_i K + h_i$ for the initial total wealth of an agent of type $i \in \{E, W\}$. The following is then the notion of equilibrium adopted in this paper.[25]

Definition 4.2. *Given taxes $\tau \equiv (\{\tau_{si}, \tau_{Li}\}_{i\in\{E,W\}}, \tau_I, \tau_\pi)$ satisfying $\tau_\pi \leq 1 - \phi$ and transfers $\{\eta_i\}_{i\in\{E,W\}}$, a stationary competitive equilibrium with endogenous collateral constraints consists of an allocation $\{(c_{it}, k_{it}, l_{it})_{t\geq 0}\}_{i\in\{E,W\}}$, government debt D, capital stock K, wage w, interest rate r, and collateral parameter $\overline{\omega}_{\mathrm{d}}$, such that:*

1. *For $i \in \{E, W\}$, $(c_{it}, k_{it}, l_{it})_{t\geq 0}$ solves the problem of a type i agent in Definition 4.1 given the wage w, interest rate r, taxes τ, transfers η_i, and collateral constraint $\overline{\omega}_{\mathrm{d}}$.*

2. *The after-tax return for workers and the drift in entrepreneurs' consumption satisfy*

[25] To understand the market-clearing conditions in Definition 4.2, note that when agents die at rate ρ_D and the total wealth of type i agents grows at an average rate of $\mu_{ci} < \rho_D$, the average total wealth of type i agents in the stationary distribution is $\rho_D \kappa_i K/(\rho_D - \mu_{ci})$.

$\rho_D - \rho + r_{sW} > 0$ and $\rho_D - \mu_{c,\mathrm{d}} > 0$ and the markets for labor, bonds, and goods clear:

$$L = [(1-\alpha)A/w]^{1/\alpha} K$$
$$K = \frac{(1-\psi)\rho_D \kappa_E}{\rho_D - \mu_{c,\mathrm{d}}} K \bar{k}_{\mathrm{d}}$$
$$AK^\alpha L^{1-\alpha} - \delta K = \rho\left(\frac{(1-\psi)\rho_D \kappa_E}{\rho_D - \mu_{c,\mathrm{d}}} + \frac{\psi \rho_D \kappa_W}{\rho_D - \rho + r_{sW}}\right) K.$$

3. *The government budget constraint is satisfied: the interest on government debt equals the revenue raised from taxes minus the transfers to newborns every instant.*

4. *The initial utility of entrepreneurs weakly exceeds that of workers.*

5. *The constant $\overline{\omega}_{\mathrm{d}}$ in the collateral constraint satisfies (19).*

Equilibrium characterization. I now turn to the main result of this paper, which shows that whenever the stationary efficient allocation characterized in Section 3 exists and the associated efficient value of $\hat{x}$ satisfies $\hat{x} < 1$, the efficient allocation can be implemented as an equilibrium of the form given in Definition 4.2.[26] Specifically, Proposition 4.2 shows that implementing the efficient allocation amounts to finding an interest rate and taxes that solve the system of equations (20) and then choosing the remaining instruments to ensure that the resource and government budget constraints hold.[27] I will use hat notation to denote efficient quantities, so that $\hat{S}$ is the solution to equation (12), $\hat{\Pi}$ and $\hat{K}$ are the associated marginal product of capital and capital stock, respectively, and $\hat{x}, \hat{\mu}_c, \hat{\nu}^B$, and $\hat{\nu}^K$ are the efficient values of the functions defined in Section 2.[28] Finally, because labor is supplied inelastically and agents can borrow up to the natural limit, I state the following characterization in terms of $\kappa_i = \eta_i + h_i/K$ rather than η_i and τ_{Li} separately, because the latter two objects are not uniquely pinned down.

Proposition 4.2. *When Assumption 3.1 holds and $\hat{x} < 1$, the stationary efficient allocation can be implemented as a equilibrium of the form given in Definition 4.2 in which the interest*

[26]Note that the requirement $\hat{x} < 1$ in Proposition 4.2 ensures that the value of human wealth is finite.

[27]Under the conditions stated in Proposition 4.2, there always exists at least one solution to the system of equations (20), in which $r = \rho_S, \tau_{sW} = 0$, $\tau_{sE} = \hat{\nu}^B/\rho$ and $\tau_I = \hat{\nu}^K - \hat{\nu}^B - \min\{0, \big((1-\hat{\nu}^B/\rho)/\phi - 1\big)\rho_S\}$.

[28]That is, the values obtained when evaluated at $S = \hat{S}$.

rate r and taxes τ_{sE}, τ_{sW} and τ_I are any values such that

$$
\begin{aligned}
r_{sW} &= \rho \\
r_{sE} - \rho &= -\hat{\nu}^B \\
r_b - \rho_S &= -\hat{\nu}^B + \hat{\nu}^K,
\end{aligned}
\tag{20}
$$

and the remaining instruments and equilibrium quantities are given as follows:

1. *The tax on profits is $\tau_\pi = 1 - \phi$.*

2. *The transfers and labor taxes are chosen such that κ_E and κ_W satisfy*

$$
\kappa_E = \frac{\phi\sigma(\rho_D - \hat{\mu}_c)}{\sqrt{\rho}\hat{x}(1-\psi)\rho_D} \qquad\qquad \kappa_W = \kappa_E \max\left\{e^{-\hat{x}^2/2}, \hat{x}/\overline{\omega}\right\}.
$$

3. *Given the taxes on labor, the level of government debt is*

$$
D = \Big((\hat{\Pi}/\alpha + (1/\alpha - 1)\delta)/\rho - 1\Big)\hat{K} - ((1-\psi)h_E + \psi h_W) \tag{21}
$$

where for $i \in \{E, W\}$, h_i denotes the human wealth of type i as defined in (17).

Finally, the wage is $w = (1-\alpha)AL^{-\alpha}\hat{K}^\alpha$ and the constant in the collateral constraint is given by $\overline{\omega}_{\mathrm{d}} = \iota^{-1}\max\left\{\hat{x}/\overline{\omega}, e^{-\hat{x}^2/2}\right\}$.

Proof. See Appendix C.1. □

Although the expressions for total wealth and government debt in Proposition 4.2 may appear complicated, the logic underlying the characterization is simple and can be summarized in the following steps. First, the expressions for r_{si} and r_b in Proposition 4.2 ensure that the drift and diffusion of consumption in Lemma 4.1 coincide with their efficient counterparts given in Proposition 3.1. Second, the tax on profits is the highest level consistent with incentive compatibility because this maximizes risk-sharing between the risk-averse entrepreneurs and the government. Third, the transfers and labor taxes ensure that the bond market clears and that all agents obtain the same level of initial utility. Finally, the expression for government debt follows by equating the sum of private and public wealth (the negative of government debt) with the capital stock.[29]

Note that by Proposition 2.2, Proposition 4.2 implies that when the no-absconding constraint is a strict inequality, in any implementation of the efficient allocation, the risk-free return faced by entrepreneurs must fall below the discount rate ρ and the cost of firm

[29] As expected (and verified in Appendix C.1), the government budget constraint automatically holds.

borrowing must fall below the subjective discount rate ρ_S. There are many ways in which the government may ensure that these conditions hold. In the next section I consider two specific implementations and then discuss the main result in greater detail.

5 Special cases and robustness

The proof of Proposition 4.2 amounts to ensuring that the capital stock, initial consumption and the (constant) drift and diffusion of log consumption in the competitive equilibrium coincide with their counterparts in the efficient allocation. In this section I first consider two specific implementations, provide some intuition for the overall approach, and discuss the robustness of the main result to various extensions.[30]

5.1 Specific implementations

As noted above, the optimal policy in Section 4 is indeterminate both because agents care only about after-tax returns (and not pre-tax prices) and because lump-sum transfers to newborns are equivalent to taxes on labor income because labor supply is inelastic. I now consider two specific implementations of the efficient allocation that warrant special attention, in which the expressions in Proposition 4.2 simplify. In the first, workers face no taxes and the interest rate coincides with the subjective discount rate, and in the second, the tax on investment is chosen such that the tax on profits may be interpreted as a tax on the excess return on capital.

Untaxed workers and common human wealth. In this paper, the agency frictions do not directly affect the labor income of any agent or the capital income of workers. For this reason, one natural implementation corresponds to the situation in which human wealth is independent of type and workers face no taxes, which occurs when entrepreneurs' labor taxes satisfy $1 - \tau_{LE} = r_{sE}/\rho$ and the interest rate is $r = \rho_S$. In this case, the second and third equations in (20) reduce to $\tau_{sE} = \hat{\nu}^B/\rho$ and $\tau_I = \hat{\nu}^K - \hat{\nu}^B - \min\{0, \big((1-\hat{\nu}^B/\rho)/\phi - 1\big)\rho_S\}$. Further, in this case the level of government debt and the revenue raised from entrepreneurs as a fraction of income admit the following simple expressions.

Proposition 5.1. *In the implementation in which $r = \rho_S, \tau_{LW} = 0$, and $1 - \tau_{LE} = r_{sE}/\rho$, government debt is $D = (\hat{\Pi}/\rho - 1)\hat{K}$. Further, when the no-absconding constraint is strict,*

[30] I remind the reader of the convention described prior to Proposition 3.2: in this paper as I vary ϕ, I assume that $\iota \equiv \phi\bar{\iota}$ for some fixed $\bar{\iota} \in (0, 1]$, so that both agency frictions vanish as $\phi \to 0$.

the average revenue raised from entrepreneurs as a fraction of income is $\hat{x}^2/(2\hat{x}^2+1)$.

Proof. See Appendix C.2. □

When combined with Proposition 3.2, Proposition 5.1 shows that when the no-absconding constraint is strict, both $D/\hat{K}$ and total tax revenue raised per unit of income are increasing in the parameter ϕ governing the strength of the agency frictions. This monotonicity of revenue is noteworthy because, in contrast, the revenue raised solely by the tax on profits exhibits no such monotonicity, and in fact vanishes when agency frictions are either very high or very low. Indeed, when $\phi = 1$, the tax on profits vanishes and therefore obviously raises no revenue, and when $\phi = 0$, the profits themselves vanish because the marginal product of capital equals the interest rate.

Normal versus excess returns. Boadway and Spiritus (2025) study the optimal taxation of "normal" and "excess" returns on capital income in a two-period environment, where the normal return on capital income is defined as the risk-free rate of return multiplied by the savings of the entrepreneur and excess returns are any deviations from this quantity. This differs from the decentralization in Section 4, in which taxes were levied on different sources of capital income (interest and profits) instead of being imputed from the total stock of savings.[31] However, the taxes on interest and profits in Section 4 may be interpreted as taxes on the normal and excess returns on capital in the special case in which $r = r_b$, so that the cost of firm borrowing faced by the firm is equal to the interest rate.[32] In this implementation, both the pre-tax safe return $r + \rho_D$ and the post-tax safe return r_{sE} fall below the discount rate ρ, and so the sign of the tax on entrepreneurs' interest is not *a priori* obvious. In fact, as the following shows, in this implementation the tax on entrepreneurs' interest income may assume either sign.

Proposition 5.2. *For all sufficiently small* $\phi > 0$, *there exists an implementation of the efficient allocation in which* $r = r_b$. *In this implementation, the tax on workers' interest income is negative when the no-absconding constraint is strict and the tax on entrepreneurs' interest income may assume either sign and is negative when* ϕ *is sufficiently small.*

Proof. See Appendix C.2. □

[31] Taxing normal returns at rate τ_{sE} is equivalent to imposing $b_{et} = a_t$ and $b_{bt} = -k_t$ in (15), and setting $\tau_I = 0$ (this was the implementation and market structure considered in previous drafts).

[32] By equation (20), this implementation corresponds to choosing taxes on interest income and investment satisfying $(1-\tau_{sE})(\rho - \hat{\nu}^B + \hat{\nu}^K) = \rho - \hat{\nu}^B$ and $\tau_I = \min\{0, (1-\tau_{sE})/\phi - 1\}(\hat{\nu}^B - \hat{\nu}^K - \rho_S)$. Such an implementation will exist if $\rho - \hat{\nu}^B + \hat{\nu}^K \neq 0$, which is true for sufficiently small $\phi > 0$.

5.2 Discussion and robustness

In this section I discuss the methodology adopted in this paper, the role of some of the key assumptions, and the robustness of the main insights to several extensions.

Mechanism design and the primal approach. It is worth emphasizing that the mechanism design approach simplifies the characterization of the optimal taxes. This may appear counterintuitive, because modeling incentive constraints necessitates the analysis of a dynamic agency problem seemingly unrelated to the incomplete markets model of Section 4. However, proceeding in this way eliminates the need to understand exactly how competitive equilibria vary with taxes. The government can never do better than the efficient allocation, and so the task of Proposition 4.2 is simply to show that the efficient allocation can be implemented with the above tax instruments, which amounts to solving a finite system of equations.[33] If taxes and transfers were the objects of choice in the planner's problem, then the analysis would be more complicated because the interest rate, capital stock and the constant in the collateral constraint are only defined in terms of the solutions to market-clearing equations. A change in any instrument will have non-obvious effects on all of these objects, but this is irrelevant to the proof of Proposition 4.2.

This reasoning is reminiscent of the primal approach employed in the literature on optimal linear taxation in representative agent economies.[34] Recall that here one rearranges the first-order conditions of the consumer's problem to eliminate prices from the budget constraint to obtain what is termed an "implementability constraint." One can then reverse this procedure and show that any allocation that is resource feasible and satisfies the implementability constraint can be supported as a competitive equilibrium. In this way, there is no need to understand exactly how competitive equilibria vary with taxes, and the planner's problem becomes a standard programming problem. The analogy with the approach of this paper is far from exact, but in both cases one uses the optimality conditions that obtain in competitive equilibria and then chooses among allocations directly.

The role of the welfare notion and preferences. The simplicity of the characterization of the stationary efficient allocation in Proposition 3.1 is due partly to the preferences being logarithmic and partly to the adoption of a welfare criterion that weights the flow utility of an agent the same independently of her birth date. As emphasized in

[33] Appendix D.1 shows that this methodology also applies to preferences exhibiting constant relative risk aversion with $\gamma > 1$, although the analysis is more complicated in this case and the ensuing expressions and harder to interpret.

[34] See, e.g., Chari and Kehoe (1999).

Section 2, the homotheticity of preferences and the fact that technology exhibits constant-returns-to-scale in all variable factors imply that consumption and capital are linear in consumption-equivalent utility u, which permits aggregation over entrepreneurs and implies a simple form of the goods resource constraint. Further, the above welfare notion also implies that transfers and government debt remain necessary to implement the efficient stationary allocation even when $\phi \to 0$ and agency frictions vanish. In this case, $\hat{S} = \hat{x} = 0$ and the goods resource constraint simplifies to $\kappa_E = \kappa_W = (\rho_S/\alpha + (1/\alpha - 1)\delta)/\rho$. The stationary efficient allocation can therefore be implemented as an equilibrium in which the taxes on all forms of income are set to zero and the transfers to newborns are given by $\eta_i = \kappa_i - (wL/\rho)/\hat{K} = \rho_S/\rho$ for $i \in \{E, W\}$. In this case the interest rate is $r = \rho_S$ and government debt is $D = -\rho_D \hat{K}/\rho$.

The role of the tax on profits. The fact that a change in one instrument will, in general, affect all equilibrium quantities makes it difficult to isolate a single, unique effect of each tax. Indeed, in Proposition 4.2, the lump-sum transfers to newborns and the taxes levied on the interest and investment of entrepreneurs jointly determine both the equilibrium interest rate and the drift and diffusion of wealth, and so their effects cannot be neatly separated from one another. However, the role of the profits tax in the above implementations is simple and unambiguous: it serves to maximize risk-sharing subject to incentive compatibility. In particular, its role is not to discourage diversion or to tax away excess returns, because in the absence of such a tax, the entrepreneur would have no incentive to misreport income, as this would only lose her money (because of the deadweight loss). Further, Lemma 4.1 shows that when the collateral constraint does not bind, the evolution of total wealth in partial equilibrium depends on the profits tax only via the cost of borrowing faced by the firm r_b. In particular, for the range of τ_π such that $r_b = r + \tau_I$, the entrepreneur simply chooses her leverage to leave her return on total wealth unaffected by the tax on profits.[35]

Private risk-sharing. Section 4 assumed that agents could only trade a risk-free bond in zero net supply. One alternative to this is to allow for the existence of private contracting arrangements, so that the government is not the only source of risk-sharing. Indeed, although the work of Smith et al. (2019) cited in the introduction shows that business ownership is highly concentrated, the assumption in the above model that every

[35]In a model with a common tax on all capital income, Panousi (2010) makes a similar observation and shows that the capital tax may therefore increase capital accumulation, and relates this insight to Domar and Musgrave (1944), in which the profits tax essentially makes the government a partner in the business.

firm is owned by a single individual is unrealistic. As noted by Di Tella (2017), in the absence of taxes, when entrepreneurs can write short-term contracts with a competitive sector of risk-neutral intermediaries, their exposure to their firms' shocks is multiplied by the factor ϕ and all other parameters are unchanged.[36] These private contracts therefore play a risk-sharing role similar to that of the tax on profits in the above implementations, which is therefore set to zero. The problem of the entrepreneur is identical to that in Definition 4.1 except that $\tau_\pi = 0$ and σ is replaced by $\phi\sigma$. Reasoning identical to that given after Proposition 4.2 implies that the optimal after-tax safe returns and total wealth do not change (because the efficient allocation does not change), but that the effective cost of firm borrowing necessary to ensure the efficient level of risk must now satisfy

$$r_b - \rho_S = \hat{\Pi} - \rho_S - \sqrt{\rho}\phi\sigma\hat{x} \geq \hat{\nu}^K - \hat{\nu}^B. \tag{22}$$

To see this, note that if $\tau_\pi = 0$ and σ is replaced by $\phi\sigma$, Lemma 4.1 shows that when the collateral constraint does not bind, the risk in the entrepreneurs' consumption is given by the Sharpe ratio,

$$\sigma_{c,\mathrm{d}} = \phi\sigma\overline{k}_{\mathrm{d}} = \frac{\hat{\Pi} - r_b}{\phi\sigma}$$

which equals the efficient value $\hat{\sigma}_c = \sqrt{\rho}\hat{x}$ when r_b satisfies (22). Further, in this implementation, we have the following comparative statics with respect to the agency friction when the no-absconding constraint does not hold with equality.

Lemma 5.3. *For the range of ϕ such that the no-absconding constraint is strict, the efficient value of $r_b - \rho_S$ in the implementations with optimal short-term contracts is negative and decreasing in ϕ.*

Proof. See Appendix C.1. □

Heterogeneous entrepreneurs. In the above analysis all entrepreneurs were assumed to be equally productive on average, in the sense that they all operated with the technology represented by the function $F(k, l) = Ak^\alpha l^{1-\alpha}$ for some common A and α. A general analysis of efficient allocations in an environment in which productivity parameters are private information or heterogeneous across time is beyond the scope of the paper. However, it is worth noting that the characterization of optimal tax policy in Proposition 4.2 generalizes in a simple way if the parameter A differs across entrepreneurs but is observable and

[36] See the discussion of the experts' problem on page 2046 of Di Tella (2017) for further details. Boar and Knowles (2024) derive a similar result in their discrete-time model.

constant over time and an analogue of Assumption 3.1 (ensuring existence of a stationary efficient allocation) is satisfied. In this case, the marginal product of capital differs among entrepreneurs, and so the analogue of the resource constraint in Proposition 3.1 becomes more complicated.[37] However, arguments analogous to those employed in Proposition 4.2 remain applicable if the government can allow taxes to depend on productivity and the technical conditions in Section 2 are satisfied for each type.[38] Specifically, efficiency requires that the after-tax safe returns and cost of firm borrowing satisfy the equations in (20) for each type of entrepreneur *separately*, where the efficient wedges $\hat{\nu}^B$ and $\hat{\nu}^K$ now differ across entrepreneurs. By Proposition 2.2, when the no-absconding constraint binds for no entrepreneur, heterogeneity in productivity provides a force for regressivity with respect to business income, because the efficient cost of firm borrowing must be lower for more productive entrepreneurs, and a force for progressivity with respect to the safe asset whenever the no-absconding constraint is strict. However, the optimal tax on profits remains common to all entrepreneurs at $\tau_\pi = 1 - \phi$.

6 Numerical examples

The goal of this paper has been to characterize the optimal taxes on different forms of capital income in a model in which the relevant economic forces are as simple as possible. To conclude the paper, I now compute some examples. I first depict the after-tax returns and the effective cost of firm borrowing appearing in Proposition 4.2 alongside the drift and diffusion for consumption, and then plot the associated revenue raised and transfers for the implementation characterized in Proposition 5.1. Beyond illustrating the basic mechanisms in the model, the main point of this section is that savings taxes and wedges can be large for standard parameters but often substantially overstate the revenue raised from taxation. Throughout I fix $(\alpha, \sigma, \rho_S, \rho_D, \delta) = (0.33, 0.2, 0.04, 0.02, 0.06)$ and vary ϕ, following the convention described prior to Proposition 3.2 in which $\iota \equiv \phi\bar{\iota}$ for some fixed $\bar{\iota} \in (0, 1]$. The capital share α, subjective discount factor ρ_S, rate of death ρ_D and depreciation rate δ are standard, while σ is toward the lower end of the range of values adopted in the literature.[39]

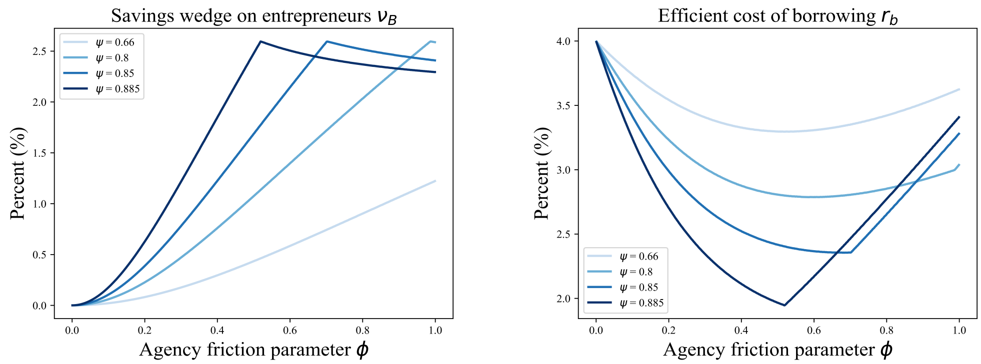
Wedges and after-tax returns. Figure 1 plots the wedge on savings and the efficient cost of borrowing by the firm as a function of ϕ with $\bar{\iota} = 1$, for different values of the fraction

[37] Appendix D.2.1 provides details of the characterization.

[38] The details of the decentralization are given in Appendix D.2.2.

[39] For example, Angeletos (2007) considers $\sigma = 0.2$ and $\sigma = 0.4$ in his quantitative exercises. Choosing a larger σ would only increase the effects of agency frictions.

Figure 1: Wedges and efficient cost of borrowing with tight collateral constraints ($\bar{\iota} = 1.0$)



of agents ψ who are workers.[40] As implied by Proposition 3.2, for small ϕ the no-absconding constraint does not bind and in this region the savings wedge is increasing in ϕ. As shown in Figure 2, the "kinks" in the savings wedge depicted in Figure 1 occur at parameters where the no-absconding constraint begins to bind and the wedge on savings starts to decline. Further, in contrast to the implementation with private risk-sharing considered in Lemma 5.3, this example shows that the efficient cost of borrowing is not in general monotonic in ϕ, even in regions in which the no-absconding constraint does not bind. Figure 3 then complements Figure 1 by depicting the wedges and cost of borrowing for parameters such that the collateral constraints are as relaxed as possible.[41] The wedges on savings coincide with those in Figure 1 for low ϕ but are now increasing everywhere and exhibit no "kinks."

Revenue and transfers. Figure 3 shows that for the above parameters the model generates high optimal wedges on savings when collateral constraints are relaxed and agency frictions are severe. Indeed, in view of Figure 3, the safe return $r_{sE} = \rho - \nu^B$ falls from 6 percent to less than 1 percent as ϕ ranges over the unit interval, a fall of over 80 percent.

However, there are two important points to bear in mind when interpreting the magnitude of the savings wedge and its implications for the tax on interest. The first is that in Section 4, the tax on interest is applied only to the safe return of the entrepreneur and not their total capital income. When their personal bond holdings are negative, the tax on interest can reduce the tax liability of entrepreneurs. Further, the savings wedge is high when

[40] Cagetti and De Nardi (2006) show that roughly 11.5 percent of households in the SCF are active business owners, which corresponds to $\psi = 0.885$. However, this paper abstracts from a corporate sector and so I use this as an upper bound in the numerical examples instead of focusing on a benchmark calibration.

[41] Lemma A.8 in Appendix A.2 shows that this occurs when $\bar{\iota} \approx 0.5$.

Figure 2: Drift and diffusion of consumption with tight collateral constraints ($\bar{\iota} = 1.0$)

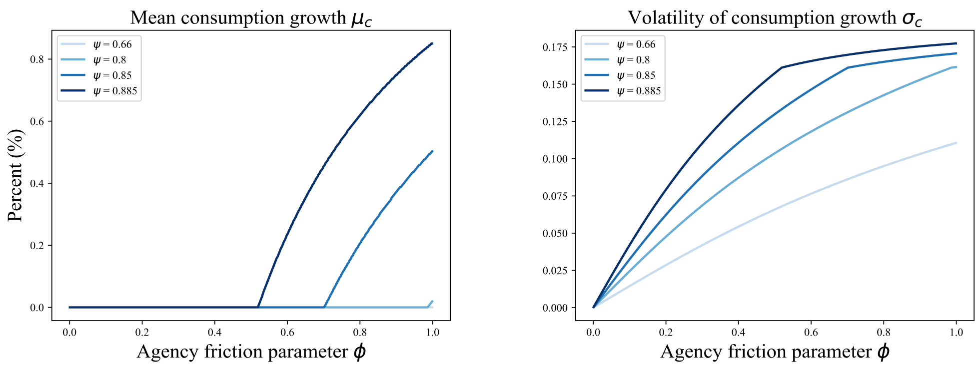


Figure 3: Wedges and efficient cost of borrowing with relaxed collateral constraints ($\bar{\iota} \approx 0.5$)

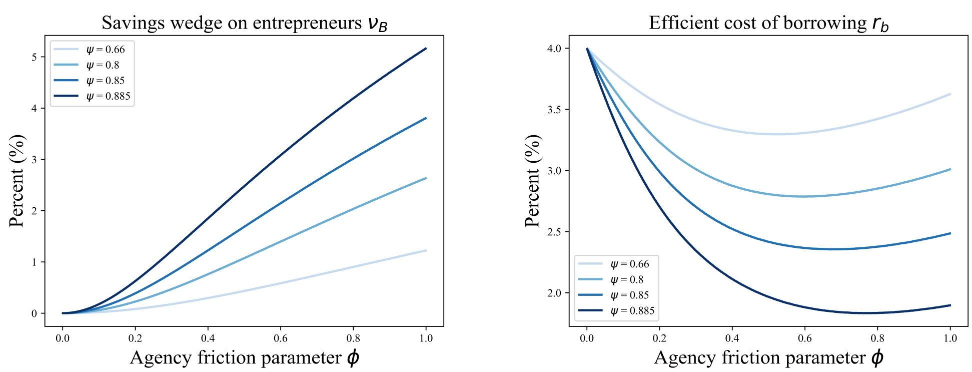
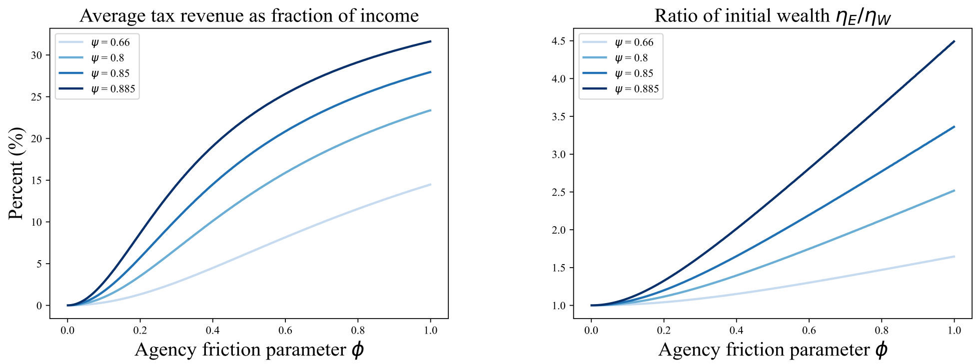


Figure 4: Revenue and inherited wealth with relaxed collateral constraints ($\bar{\iota} \approx 0.5$)



agency frictions are high, which is precisely when the tax on profits is low. The left-hand plot in Figure 4 complements Figure 3 by depicting the average taxes paid by entrepreneurs as a fraction of income for the implementation in Proposition 5.1, and shows that while the tax on interest can exceed 80 percent, the average tax revenue raised never exceeds 32 percent of income. The second point is that in the above implementation, entrepreneurs receive higher transfers than workers at birth in order to compensate them for the risk they bear. The right-hand plot in Figure 4 depicts the relative magnitudes of these transfers by plotting the ratio of initial (financial) wealth η_E/η_W of the two types. In particular, when the savings wedge in Figure 3 reaches its highest value, entrepreneurs begin life with over four times the amount of wealth as workers.

The sensitivity of taxes and transfers to both parameters and the choice of implementation in this environment is why I have computed a range of examples instead of emphasizing one particular calibration. Smith et al. (2019) show that there exists substantial heterogeneity in returns on private businesses, and there is much we do not know regarding the determinants of this income. In particular, although the evidence noted in the introduction suggests that business income appears to reflect owner-specific characteristics, the extent to which these are endogenous to tax policy is not yet clear. They could reflect innate ability (which would be invariant to policy) or the return on past effort and reputation (which would likely be discouraged by high taxes). In this paper I have instead therefore focused primarily on the qualitatively distinct roles played by each instrument in an environment in which their values may be analytically characterized.

7 Conclusion

Capital income can assume many forms, including (but not limited to) interest on personal savings and the profits of private businesses. This paper has provided a model in which the desirability of differential treatment of this income emerges when business owners operate firms exhibiting constant-returns-to-scale, are subject to idiosyncratic risk, may misreport profits, and can abscond with a fraction of borrowed capital. I show that whenever a stationary efficient allocation exists, it may be implemented in a competitive equilibrium with endogenous collateral constraints, lump-sum transfers, and constant, occupation-specific, linear taxes on reported profits, investment, and interest. The main findings regarding these taxes were as follows. First, the profits tax serves only to maximize the level of risk-sharing and is not driven by redistributive concerns. Second, when the no-absconding constraint does not bind, entrepreneurs face lower after-tax returns on the risk-free asset than workers, and so the model generates progressive taxes on savings in a qualified sense. Third, to provide entrepreneurs with efficient incentives for investment, the effective cost of firm borrowing faced by their businesses must fall below the complete markets level when the no-absconding constraint does not hold with equality. Further, there is a degree of freedom in optimal policy, because the (after-tax) return on the bond can be affected either by taxing interest or by altering the equilibrium interest rate with an investment tax. Finally, for an implementation in which workers are not taxed, the optimal taxes on entrepreneurs' interest overstate the overall revenue raised, in part because the distribution of wealth between bonds and capital is endogenously chosen by the entrepreneurs and interest taxes fall on only a subset of capital income.

List of Figures

A Agency problem

This appendix contains formal statements and proofs for all claims pertaining to the principal-agent problem. Appendix A.1 characterizes incentive compatible allocations in terms of diffusion processes for utility and Appendix A.2 characterizes the value function of the principal and the implied wedges on risky and risk-free assets.

A.1 Incentive compatibility

The characterization of incentive compatibility essentially follows from the arguments employed in the online appendix to Di Tella (2019), who considers an environment in which a financial intermediary (analogous to what I have termed an entrepreneur) is subject to a diversion problem as in the current paper but there is no ability to abscond with capital. However, because the situation in the current paper is not a special case of Di Tella (2019) (who considers a class of preferences that does not include CRRA utility), I will spell out some additional details and recall some definitions from the main text in order to aid the reader.

Formally, in this environment an allocation is a triple of processes $(k, c, \tilde{s})$ defined on the filtered probability space $(C[0,\infty], (\mathcal{F}_t)_{t\geq 0}, P)$, where $(\mathcal{F}_t)_{t\geq 0}$ is the filtration generated by the evaluation maps and P is the Wiener measure. Because there is no loss in assuming that the recommended stealing is $\tilde{s}_t = 0$ for all $t \geq 0$, as in the main text I omit $\tilde{s}$ from the definition of an allocation. Further, in this appendix all stochastic processes are adapted to the filtration generated by the underlying Brownian motion. An *allocation* is a pair of $\mathcal{F}$-adapted processes (k, c) satisfying $k_t \geq 0$ and $c_t > 0$ for all $t \geq 0$, while a *strategy* of the agent is an $\mathcal{F}$-adapted process s assuming values in $[0, \overline{s}]$ for all $t \geq 0$. Denoting the corresponding expectation operator by $\mathbb{E}^s$, the utility from adhering to a strategy s is

$$U^A(k, c, s) := \mathbb{E}^s\left[\rho \int_0^\infty e^{-\rho t} \ln(c_t + \phi s_t k_t) dt\right]. \tag{23}$$

Following Di Tella (2019) and Di Tella and Sannikov (2021), I restrict attention to *admissible* allocations and *feasible* strategies, as defined below.[42]

Definition A.1 (Admissible allocations)**.** *An allocation (k, c) is admissible if the associated utility and present discounted value of capital and consumption are all well-defined and finite,*

$$\mathbb{E}\left[\int_0^\infty e^{-\rho t} \ln c_t dt\right] < \infty, \mathbb{E}\left[\int_0^\infty e^{-\rho t} c_t dt\right] < \infty, \mathbb{E}\left[\int_0^\infty e^{-\rho t} k_t dt\right] < \infty.$$

Definition A.2 (Feasible strategies)**.** *Given an admissible allocation (k, c), a strategy s is feasible if it vanishes beyond some fixed time T and the utility in (23) is well-defined and finite.*

[42] Di Tella (2019) calls feasible strategies as defined in Definition A.2 "valid." For this paper I adopt the nomenclature of "feasible" from Di Tella and Sannikov (2021).

Given an elasticity of intertemporal substitution ψ and risk aversion parameter γ, the utility process $(W_t)_{t\geq 0}$ in Di Tella (2019) associated with consumption $(c_t)_{t\geq 0}$ is a solution to

$$W_t := \mathbb{E}_t\left[\int_t^\infty f(c_s, W_s)ds\right] \tag{24}$$

where the Epstein-Zin aggregator is defined by

$$f(c,W) := \frac{\rho}{1-1/\psi}\left(\frac{c^{1-1/\psi}}{[(1-\gamma)W]^{\frac{\gamma-1/\psi}{1-\gamma}}} - (1-\gamma)W\right) \tag{25}$$

when $\psi \neq 1$. CRRA utility corresponds to $\psi = 1/\gamma$ and logarithmic utility arises as $\gamma, \psi \to 1$. Specifically, for the CRRA case the aggregator becomes $f(c,W) := \rho\big(c^{1-\gamma}/(1-\gamma) - W\big)$ and for the logarithmic case the aggregator becomes $f(c,W) := \rho(\ln c - W)$. Now recall the following definitions from the main text.

Definition A.3. *An admissible allocation (k,c) is incentive compatible if $U^A(k,c,0) \geq U^A(k,c,s)$ and $k_t \leq \omega e^{W_t^s}$ for all feasible strategies s and $t \geq 0$ almost surely. The set of incentive compatible allocations that give utility W to the entrepreneur is denoted $\mathcal{A}^{IC}(W)$.*

The principal is risk-neutral and so his preferences over incentive compatible allocations (k,c) are represented by the objective function

$$U^P(k,c) := \mathbb{E}\left[\int_0^\infty e^{-\rho t}[(\Pi - \rho - \tau_k)k_t - c_t]dt\right] \tag{26}$$

and the principal's problem is then the following.

Definition A.4. *Given initial utility W, the problem of the principal is defined to be*

$$V(W) = \sup_{(k,c)\in\mathcal{A}^{IC}(W)} U^P(k,c) \tag{27}$$

and an allocation attaining the supremum in (27) is termed an optimal (or efficient) allocation.

The following is essentially a combination of Lemma 1 and Lemma 2 from the online appendix of Di Tella (2019).

Lemma A.1. *For any admissible allocation (k,c), the promised utility admits the representation $dW_t = \rho(W_t - \ln c_t)dt + \tilde{\sigma}_{W,t}dB_t$ for some process $\tilde{\sigma}_{W,t} \in \mathcal{L}^2$, the set of $\mathcal{F}$-adapted processes x satisfying $\mathbb{E}\Big[\int_0^t x_u^2 du\Big] < \infty$ for all $t \geq 0$. The entrepreneur will choose $s = 0$ if and only if*

$$0 \in \arg\max_{s\geq 0} \rho \ln(c_t + \phi k_t s) - s\tilde{\sigma}_{W,t}/\sigma \tag{28}$$

almost surely for all $t \geq 0$. It follows that when characterizing efficient allocations there is no loss in assuming that utility evolves according to

$$dW_t = \rho(W_t - \ln c_t)dt + \rho\phi\sigma(k_t/c_t)dB_t, \tag{29}$$

which is equivalent to the law of motion of u given by

$$\begin{aligned} du_t &= \rho\big(-\ln(c_t/u_t) + (\sqrt{\rho}\phi\sigma k_t/c_t)^2/2\big)u_t dt + (\rho\phi\sigma k_t/c_t)u_t dB_t \\ &= \rho\big(-\ln \overline{c}_t + x_t^2/2\big)u_t dt + \sqrt{\rho}x_t u_t dB_t \end{aligned} \tag{30}$$

where $\overline{c}_t = c_t/u_t$ and $x_t = \sqrt{\rho}\phi\sigma k_t/c_t$ as in the main text.

Proof. This follows from Lemma 1 and the proof of Lemma 2 in the appendix to Di Tella (2019). Note that although the literal statement of Lemma 2 does not apply to logarithmic utility, the proof does in fact extend to this case, and indeed to all CRRA utility functions, because the only point in the proof at which properties of the aggregator in (25) are relevant is for the existence of a constant κ such that $f(c,y) - f(c,x) \leq \kappa(y-x)$ for any c and all $y \geq x$. For general Epstein-Zin preferences, this requires some restrictions on parameters (such as those imposed in the statement of Lemma 2 of the appendix to Di Tella (2019)), but for the case of CRRA utility considered in this paper, we have $f(c,y) - f(c,x) = \rho(-y+x) \leq 0$ for $y \geq x$, and so this holds automatically. □

A.2 Characterization of value function

To prepare for the following proofs I will first introduce some additional notation and results. I will define $\overline{x} = \overline{x}(\overline{\omega})$ and $\overline{\overline{x}} = \overline{\overline{x}}(\overline{\omega})$ to be the solutions to $\overline{x}e^{\overline{x}^2/2} = \overline{\omega}$ and $\overline{\overline{x}}e^{\overline{\overline{x}}^2/2-1} = \overline{\omega}$, respectively. Note that under the following change of variables adopted in the main text,

$$x := \sqrt{\rho}\phi\sigma\overline{k}/\overline{c} \qquad\qquad S := \frac{\Pi - \rho - \tau_k}{\sqrt{\rho}\phi\sigma} \qquad\qquad \overline{\omega} := \frac{\sqrt{\rho}\phi\sigma}{\rho\iota} \tag{31}$$

the no-absconding constraint $\overline{k} \leq \omega = (\rho\iota)^{-1}$ is equivalent to the inequality $\overline{c}x \leq \overline{\omega}$. Consequently, by (30), $\overline{x}$ is the maximum x for which the no-absconding constraint holds, under the assumption of the inverse Euler equation holding (which for logarithmic utility means zero drift in (30)), and $\overline{\overline{x}}$ is the maximum x for which consumption growth is smaller than the rate of discount when the no-absconding constraint holds with equality. For any $\overline{\omega} > 0$, I will define the set

$$D(\overline{\omega}) := \big\{(\overline{c},x) \in \mathbb{R}^2 \;\big|\; \overline{c} > 0, x \geq 0\ x\overline{c} \leq \overline{\omega},\ -\ln \overline{c} + x^2/2 < 1\big\} \tag{32}$$

and for ease of reference recall the definition

$$\overline{v} \equiv \overline{v}(S,\overline{\omega}) := \sup_{(\overline{c},x)\in D(\overline{\omega})} \frac{(Sx-1)\overline{c}}{\rho(1+\ln\overline{c} - x^2/2)}. \tag{33}$$

The constraint set $D(\overline{\omega})$ is not compact, which complicates the following analysis. For this reason I now note that for any $(\overline{c},x) \in D(\overline{\omega})$ we have $\overline{c} \geq e^{-1}$, and for any $\epsilon > 0$ define

$$D_\epsilon(\overline{\omega}) := \big\{(\overline{c},x) \in \mathbb{R}^2 \;\big|\; \overline{c} \in [e^{-1}, 1/\epsilon], x \geq 0\ x\overline{c} \leq \overline{\omega},\ -\ln \overline{c} + x^2/2 \leq 1 - \epsilon\big\} \tag{34}$$

which is a *compact* subset of $D(\overline{\omega})$. For future reference I isolate the following observation.

Lemma A.2. *For any* $\overline{\omega} > 0$, $x < \overline{\overline{x}}(\overline{\omega})$ *for all* $(\overline{c}, x) \in D(\overline{\omega})$.

Proof. Simply combine the defining inequalities in (32) to note that for any $(\overline{c}, x) \in D(\overline{\omega})$, we either have $x = 0$ or $x > 0$ and $x^2/2 < 1 + \ln \overline{c} \le 1 + \ln(\overline{\omega}/x)$, from which the conclusion follows from the definition of $\overline{\overline{x}}(\overline{\omega})$. □

The proof of Proposition 2.1 will show that when the no-absconding constraint does not hold with equality, consumption is a martingale and satisfies $\overline{c} = e^{x^2/2}$. For this choice of consumption, the maximand in (33) may be written $F(x) := \rho^{-1}(Sx - 1)e^{x^2/2}$, which satisfies

$$\begin{aligned} F'(x) &= \rho^{-1}(Sx^2 - x + S)e^{x^2/2} \\ F''(x) &= \rho^{-1}(Sx^3 - x^2 + 3Sx - 1)e^{x^2/2}. \end{aligned} \tag{35}$$

Consequently, it is easy to check that if $S < 1/2$, the function F has a unique local maximum on the interval $x \in [0, 1/S]$ given by

$$x_{\text{loc}}(S) := \frac{1 - \sqrt{1 - 4S^2}}{2S}. \tag{36}$$

Proposition 2.1 in the main text asserts that the problem of the principal is finite-valued for all sufficiently small S. I now outline the technical conditions defining what "sufficiently small" means, where I write $\overline{v} \equiv \overline{v}(S, \overline{\omega})$ for the quantity defined in (33) to illustrate the dependence on S and $\overline{\omega}$.

Assumption A.1. $S\overline{\overline{x}} \le 1$.

The following shows that Assumption A.1 implies that $\overline{v}$ is negative. This will be a necessary condition for the principal's value function to be finite.

Lemma A.3. *Assumption A.1 holds if and only if* $\overline{v}(S, \overline{\omega}) < 0$, *and if* $S\overline{\overline{x}} < 1$ *then the supremum in* (33) *is attained at some* $(\overline{c}, x)$. *Further, the function* $\overline{v}$ *is continuous in* S *on the interval* $[0, 1/\overline{\overline{x}}]$.

Proof. First note that if $S\overline{\overline{x}} > 1$ then inspection of the definition of $D(\overline{\omega})$ shows that $\overline{v}(S, \overline{\omega}) > 0$. Next, I show that if $S\overline{\overline{x}} < 1$, then for all $\delta > 0$ there exist $\epsilon(\delta) > 0$ such that for all $S \in [0, (1-\delta)/\overline{\overline{x}}]$, the supremum in (33) is unchanged when we restrict the constraint set to $D_\epsilon(\overline{\omega})$,

$$\overline{v}(S, \overline{\omega}) = \sup_{(\overline{c}, x) \in D_\epsilon(\overline{\omega})} \frac{(Sx - 1)\overline{c}}{\rho(1 + \ln \overline{c} - x^2/2)}. \tag{37}$$

To see this, note that because $\overline{v}(S, \overline{\omega}) \ge -1/\rho$, by Lemma A.2 it will suffice to find $\epsilon > 0$ such that

$$\delta \overline{c} > 1 + \ln \overline{c} - x^2/2 \qquad \forall (\overline{c}, x) \in D(\overline{\omega}) \setminus D_\epsilon(\overline{\omega}) \tag{38}$$

and (38) will hold for ϵ satisfying $\delta/\epsilon' > 1 + \ln(1/\epsilon')$ and $\delta e^{-1} > \epsilon'$ for all $\epsilon' \in (0, \epsilon]$. Because $D_\epsilon(\overline{\omega})$ is compact, (37) shows that the supremum in (33) is attained when $S\overline{\overline{x}} < 1$ from which is easily follows that $\overline{v}$ is continuous in S on $[0, 1/\overline{\overline{x}})$. To establish $\overline{v}(1/\overline{\overline{x}}, \overline{\omega}) < 0$, note that if to the contrary there exists $(\overline{c}_n, x_n)_{n=1}^{\infty} \subseteq D(\overline{\omega})$ such that

$$\limsup_{n\to\infty} \frac{(x_n/\overline{\overline{x}} - 1)\overline{c}_n}{\rho(1 + \ln \overline{c}_n - x_n^2/2)} = 0, \tag{39}$$

then we derive a contradiction as follows. The limit (39) cannot hold if $x_n \in [\epsilon, \overline{\overline{x}} - \epsilon]$ for some $\epsilon > 0$ and all sufficiently large $n \geq 1$ because in this case the numerator in (39) is bounded away from zero and the denominator is bounded from above. It will therefore suffice to assume that (39) holds for some sequence satisfying $\lim x_n = 0$ or $\lim x_n = \overline{\overline{x}}$. First, if (39) holds for $\lim x_n = 0$ then we obtain a contradiction because $\overline{c}/(1 + \ln \overline{c}) \geq 1$ for $\overline{c} \in (e^{-1}, \infty)$. Second, if (39) holds for $\lim x_n = \overline{\overline{x}}$ then because $\overline{c} > e^{-1}$ for all $(\overline{c}, x) \in D(\overline{\omega})$, this implies

$$\lim_{n\to\infty} \frac{x_n/\overline{\overline{x}} - 1}{1 + \ln(\overline{\omega}/x_n) - x_n^2/2} = 0 \tag{40}$$

which leads to a contradiction by l'Hopital's rule,

$$\lim_{x\to\overline{\overline{x}}^-} \frac{x/\overline{\overline{x}} - 1}{1 + \ln(\overline{\omega}/x) - x^2/2} = \lim_{x\to\overline{\overline{x}}^-} \frac{1/\overline{\overline{x}}}{-1/x - x} = -\frac{1}{1 + \overline{\overline{x}}^2} < 0.$$

Finally, to note that $\overline{v}(S, \overline{\omega})$ is continuous on $[0, 1/\overline{\overline{x}}]$ (and not just $[0, 1/\overline{\overline{x}})$), note that if $\overline{v}(1/\overline{\overline{x}}, \overline{\omega}) > \lim_{S\to(1/\overline{\overline{x}})^-} \overline{v}(S, \overline{\omega})$, then there exists $\epsilon > 0$ and $(\overline{c}, x) \in D(\overline{\omega})$ such that

$$\frac{(x/\overline{\overline{x}} - 1)\overline{c}}{\rho(1 + \ln(\overline{\omega}/x) - x^2/2)} > \epsilon + \overline{v}(S, \overline{\omega}) \geq \epsilon + \frac{(Sx - 1)\overline{c}}{\rho(1 + \ln(\overline{\omega}/x) - x^2/2)}$$

for all $S < 1/\overline{\overline{x}}$, which contradicts the definition of $\overline{v}(S, \overline{\omega})$ for S sufficiently close to $1/\overline{\overline{x}}$. $\square$

As noted in the main text, $\overline{v}$ may be interpreted as the constant such that $\overline{v}u$ is the value function of a principal who is constrained to choose consumption growth below the subjective rate of discount. The following additional assumption will ensure that $\overline{v}u$ actually solves the principal's problem (with no ad-hoc restrictions on growth).

Assumption A.2. *For all $x \geq \overline{\overline{x}}$ we have*

$$(Sx - 1)\overline{\omega}/x + \rho(-\ln(\overline{\omega}/x) + x^2/2 - 1)\overline{v}(S, \overline{\omega}) \leq 0. \tag{41}$$

Note that it follows from Lemma A.3 that $\overline{v}(S, \overline{\omega})$ is weakly increasing and continuous in S on the interval $[0, 1/\overline{\overline{x}}]$. Further, by definition, $-\ln(\overline{\omega}/x) + x^2/2 - 1 \geq 0$ for $x \geq \overline{\overline{x}}$. Consequently, if Assumption A.2 holds for some $\overline{S}_2(\overline{\omega}) > 0$ then it holds for all $S \in [0, \overline{S}_2(\overline{\omega}))$. Also note that Assumption A.2 holds for all sufficiently small, positive S, because it is implied by Assumption A.1 and the following assumption.

Assumption A.3. $S\overline{\omega} + \rho(1 + S^{-2})\overline{v}(S, \overline{\omega}) \leq 0$.

Lemma A.4. *Assumption A.2 is implied by Assumption A.1 and Assumption A.3.*

Proof. First note that by Lemma A.3, note that the inequality (41) holds automatically when $x \in [\overline{\overline{x}}, 1/S]$ under Assumption A.1 because $\overline{v}(S, \overline{\omega}) \leq 0$, and so it will therefore suffice for the derivative of the left-hand side with respect to x to be negative for all $x \geq 1/S$. This is equivalent to $\overline{\omega}/x + \rho(1+x^2)\overline{v}(S,\overline{\omega}) \leq 0$ for $x \geq 1/S$. Using $\overline{v}(S,\overline{\omega}) \leq 0$ again, this will be true for $x \geq 1/S$ if and only if it is true for $x = 1/S$, which is exactly Assumption A.3. $\square$

I now turn to the characterization of the principal's value function v given in Proposition 2.1. For clarity, I break this into two parts. First, Proposition A.5 shows that the function $v(u) \equiv \overline{v}u$ solves the Hamilton-Jacobi-Bellman (HJB) equation (43) if and only if S is sufficiently small. Second, Theorem A.6 proves a "verification theorem," and shows that a solution to the HJB equation is a solution to the original problem.[43]

By Lemma A.1, the HJB equation for the function v is

$$\rho v(u) = \sup_{\substack{k,c\geq 0 \\ k \leq \omega u}} (\Pi - \rho - \tau_k)k - c + \rho\left(-\ln(c/u) + \frac{(\sqrt{\rho}\phi\sigma k/c)^2}{2}\right)uv'(u) + \frac{(\rho\phi\sigma k/c)^2}{2}u^2 v''(u) \tag{42}$$

which in terms of the variables given in (31) becomes

$$\rho v(u) = \sup_{\substack{\overline{c},x\geq 0 \\ x\overline{c}\leq \overline{\omega}}} (Sx-1)\overline{c}u + \rho\left(-\ln \overline{c} + \frac{x^2}{2}\right)uv'(u) + \frac{\rho}{2}x^2u^2v''(u). \tag{43}$$

Proposition A.5. *Given $\overline{\omega} > 0$, the function $v(u) \equiv \overline{v}u$ solves the HJB equation (43) if and only if S satisfies both Assumption A.1 and Assumption A.2.*

Proof. Substituting $v(u) := \overline{v}u$ into (43) gives

$$0 = \sup_{\substack{\overline{c},x\geq 0 \\ x\overline{c}\leq \overline{\omega}}} (Sx-1)\overline{c} + \rho(-\ln \overline{c} + x^2/2 - 1)\overline{v} =: \sup_{\substack{\overline{c},x\geq 0 \\ x\overline{c}\leq \overline{\omega}}} H(\overline{c}, x, \overline{v})$$

where the second equality defines H. If Assumption A.1 and Assumption A.2 hold, then by the definition of $\overline{v}$, we have $\sup_{(\overline{c},x)\in D(\overline{\omega})} H(\overline{c},x,\overline{v}) = 0$, and so it remains to eliminate the possibility that $H(\overline{c},x,\overline{v}) > 0$ for some $(\overline{c},x)$ satisfying $\overline{c}, x \geq 0$, $x\overline{c} \leq \overline{\omega}$, and $-\ln\overline{c} + x^2/2 - 1 \geq 0$. Because $\overline{v} < 0$ under Assumption A.1 by Lemma A.3, the inequality $H(\overline{c},x,\overline{v}) > 0$ requires $Sx > 1$, and therefore would imply that $(Sx-1)\overline{\omega}/x + \rho(-\ln(\overline{\omega}/x) + x^2/2 - 1)\overline{v} > 0$ for some x satisfying $-\ln(\overline{\omega}/x) + x^2/2 - 1 > 0$ (i.e. $x \geq \overline{\overline{x}}$). It will then suffice to show that

$$(Sx-1)\overline{\omega}/x + \rho(-\ln(\overline{\omega}/x) + x^2/2 - 1)\overline{v} \leq 0 \tag{44}$$

[43] Proposition A.5, which asserts the existence of a solution to (43) for sufficiently small S, is the novel part of the analysis of this paper. The second part of this characterization, Theorem A.6, follows a standard approach. For instance, Theorem A.6 is analogous to Theorem 1 in the online appendix to Di Tella (2019) and Theorem 3 in Di Tella and Sannikov (2021) (although the agency problem here is different).

for all $S \leq \overline{S}(\varpi)$ and $x \geq \overline{\overline{x}}$, which is exactly Assumption A.2. Conversely, if Assumption A.1 fails, then $\overline{v} > 0$ by Lemma A.3, in which case $H(\overline{c}, x, \overline{v})$ diverges to ∞ as $\overline{c} \to 0$, while if Assumption A.2 fails, the inequality (44) fails and the H function again becomes positive. $\square$

Theorem A.6 (Verification theorem). *Suppose that the function $v(u) := \overline{v}u$ solves the HJB equation (42) for some constant $\overline{v} < 0$. Then we have the following:*

1. *For any incentive compatible allocation (k, c) that delivers at least utility $u > 0$ to the entrepreneur, we have $\mathbb{E}\big[\int_0^\infty e^{-\rho t}((\Pi - \rho - \tau_k)k_t - c_t)dt\big] \leq \overline{v}u$.*
2. *Suppose that the optimal policy functions in the HJB are $\overline{k}u$ and $\overline{c}u$ for some $\overline{k}, \overline{c} > 0$ and all $u > 0$, and define (k, c) by $(k_t, c_t) = (\overline{k}u_t, \overline{c}u_t)$ for all $t > 0$, where $(u_t)_{t \geq 0}$ satisfies*

$$du_t = \rho\big(-\ln \overline{c} + (\sqrt{\rho}\phi\sigma\overline{k}/\overline{c})^2/2\big)u_t dt + (\rho\phi\sigma\overline{k}/\overline{c})u_t dB_t \tag{45}$$

and $u_0 = u$. Then (k, c) is an optimal allocation if it is admissible.

Proof. Let (k, c) be an arbitrary incentive compatible allocation (which is then admissible by definition), and note that the associated law of motion of consumption-equivalent utility is given by $du_t = \mu_{ut}u_t dt + \sigma_{ut}u_t dB_t$ where

$$\mu_{ut} = \rho\big(-\ln(c_t/u_t) + (\sqrt{\rho}\phi\sigma k_t/c_t)^2/2\big) \qquad\qquad \sigma_{ut} = \rho\phi\sigma k_t/c_t.$$

If $\hat{v}$ is any C^2 function then applying Ito's lemma to $\tilde{v}_t := e^{-\rho t}\hat{v}(u_t)$ gives

$$d\tilde{v}_t = e^{-\rho t}\big(\big(\mu_{ut}u_t\hat{v}'(u_t) + ((\sigma_{ut}u_t)^2/2)\hat{v}''(u_t)\big) - \rho\hat{v}(u_t)\big)dt + e^{-\rho t}\sigma_{ut}u_t\hat{v}'(u_t)dB_t.$$

Applying this to the function $v(u) \equiv \overline{v}u$, the law of motion of $\tilde{v}_t := e^{-\rho t}\overline{v}u_t$ may be written

$$\begin{aligned} d\tilde{v}_t &= e^{-\rho t}\overline{v}((\mu_{ut} - \rho)u_t dt + \sigma_{ut}u_t dB_t) \\ &= e^{-\rho t}\overline{v}\big(\rho\big(-\ln(c_t/u_t) + (\sqrt{\rho}\phi\sigma k_t/c_t)^2/2 - 1\big)u_t dt + (\rho\phi\sigma k_t/c_t)u_t dB_t\big) \end{aligned} \tag{46}$$

Because the function $v(u) = \overline{v}u$ solves the HJB equation (42), for any positive scalars k, c and u satisfying $k \leq \omega u$, we have

$$-((\Pi - \rho - \tau_k)k - c) \geq \rho\big(-\ln(c/u) + (\sqrt{\rho}\phi\sigma k/c)^2/2 - 1\big)\overline{v}u \tag{47}$$

and so it follows from (47) and (46) and the fact that (k, c) is incentive compatible that

$$d\tilde{v}_t = d[e^{-\rho t}\overline{v}u_t] \leq e^{-\rho t}(-((\Pi - \rho - \tau_k)k_t - c_t)dt + (\rho\phi\sigma k_t/c_t)\overline{v}u_t dB_t) \tag{48}$$

for all $t \geq 0$ almost surely. To prepare for an application of the dominated convergence theorem, for any integer $n \geq 1$, define

$$\tau^n := \inf\left\{T > 0 \;\middle|\; \int_0^T \left|e^{-\rho t}\overline{v}u_t\rho\phi\sigma k_t/c_t\right|^2 dt > n\right\}.$$

Integrating (48) up to τ^n then rearranges to

$$\int_0^{\tau^n} e^{-\rho s}((\Pi-\rho-\tau_k)k_s-c_s)ds \leq \overline{v}u_0 - e^{-\rho\tau^n}\overline{v}u_{\tau^n} + \int_0^{\tau^n} e^{-\rho s}(\rho\phi\sigma k_s/c_s)\overline{v}u_s dB_s. \tag{49}$$

Taking expectations, it follows from (49) that for any $n \geq 1$, the discounted expected payoff to the principal up to time τ_n is bounded by

$$\mathbb{E}\left[\int_0^{\tau^n} e^{-\rho s}((\Pi-\rho-\tau_k)k_s-c_s)ds\right] \leq \overline{v}u_0 + \mathbb{E}\Big[e^{-\rho\tau^n}[-\overline{v}u_{\tau^n}]\Big]. \tag{50}$$

We now note that

$$\begin{aligned}\left|\int_0^{\tau^n} e^{-\rho s}((\Pi-\rho-\tau_k)k_s-c_s)ds\right| &\leq \int_0^{\tau^n} e^{-\rho s}\left|(\Pi-\rho-\tau_k)k_s-c_s\right| ds \\ &\leq \int_0^{\infty} e^{-\rho s}((\Pi-\rho-\tau_k)k_s+c_s)ds\end{aligned}$$

which is integrable by the definition of an admissible allocation. It then follows from the dominated convergence theorem that the limit of the left-hand side of (50) as $n \to \infty$ is

$$\mathbb{E}\left[\int_0^{\infty} e^{-\rho s}((\Pi-\rho-\tau_k)k_s-c_s)ds\right]$$

which is literally the discounted expected payoff to the principal under the allocation (k, c). To see that $\lim_{n\to\infty}\mathbb{E}\big[e^{-\rho\tau^n}u_{\tau^n}\big] = 0$ and conclude the first part of the proof, note that u is the lowest cost of delivering utility u to the entrepreneur without any capital, and so

$$\mathbb{E}\left[e^{-\rho\tau^n}u_{\tau^n} + \int_0^{\tau^n} e^{-\rho t}c_t dt\right] \leq \mathbb{E}\left[\int_0^{\infty} e^{-\rho t}c_t dt\right] < \infty \tag{51}$$

where the last inequality again follows from the definition of an admissible allocation. Applying the monotone convergence theorem to the left-hand side of (51) then gives $\lim_{n\to\infty}\mathbb{E}\big[e^{-\rho\tau^n}u_{\tau^n}\big] = 0$.

The second part of the theorem is then immediate from the fact that if the allocation (k, c) constructed from the solution to the HJB equation is admissible then the payoff to the principal is equal to $\overline{v}u$ by the Gordon growth formula. □

Proof of Proposition 2.1. If $\overline{S}_1(\varpi)$ is the supremum of values for which Assumption A.1 is satisfied and $\overline{S}_2(\varpi)$ is the supremum of values for which Assumption A.2 is satisfied, then Proposition A.5 shows that $\overline{v}u$ solves the HJB equation if and only if $S \in [0, \overline{S}(\varpi)]$, where $\overline{S}(\varpi) = \min\{\overline{S}_1(\varpi), \overline{S}_2(\varpi)\}$. Theorem A.6 then shows that $\overline{v}u$ is actually the value function of the principal, because any other admissible allocation delivering utility u gives lower net profits.

It remains to establish that $x(S, \varpi)$ is increasing in S whenever it is well-defined (i.e. whenever the supremum in the definition of $\overline{v}$ is attained) and that there exists $\tilde{S}(\varpi)$ as in the statement of

Proposition 2.1. To establish that $x(S,\overline{\omega})$ is increasing in S whenever it is well-defined, note that because in this case the maximand in (33) is negative in the constraint set, given x, $\overline{c}$ solves

$$\min_{\substack{\overline{c}\geq 0,\ x\overline{c}\leq\overline{\omega} \\ -\ln\overline{c}+x^2/2<1}} \frac{\overline{c}}{1+\ln\overline{c}-x^2/2}.$$

Changing variables to $C=\ln\overline{c}$ and taking logarithms of this objective (which leaves optimal choices unaffected), the minimization becomes $C-\ln(1+C-x^2/2)$ over the set of real C satisfying $xe^C\leq\overline{\omega}$ and $-C+x^2/2<1$. Since this latter minimand is convex and diverges as $C\to x^2/2-1$ from above, the optimal choice either occurs at the solution $C=x^2/2$ to the first-order condition or the boundary point $C=\ln(\overline{\omega}/x)$. The optimal choice of $\overline{c}$ given $\overline{\omega}$ and x is then

$$\overline{c}=\min\{e^{x^2/2},\overline{\omega}/x\} \tag{52}$$

and the principal's problem may be written

$$\overline{v}=\max_{x\in[0,\overline{\overline{x}})} \frac{(Sx-1)\min\{e^{x^2/2},\overline{\omega}/x\}}{\rho(1+\min\{0,\ln(\overline{\omega}/x)-x^2/2\})} \tag{53}$$

where I remind the reader that $\overline{\overline{x}}$ is the solution to $\overline{\overline{x}}e^{\overline{\overline{x}}^2/2-1}=\overline{\omega}$. The fact that $x(S,\overline{\omega})$ is increasing in S will follow from (53) together with Topkis' theorem, provided that $\partial^2 m/\partial S\partial x\geq 0$, where m is the maximand in equation (53). That is, we wish to show that

$$\frac{\partial m}{\partial S}=\frac{\min\{xe^{x^2/2},\overline{\omega}\}}{\rho(1+\min\{0,\ln(\overline{\omega}/x)-x^2/2\})}$$

is weakly increasing in $x\in[0,\overline{\overline{x}})$. For $x\in[0,\overline{x}]$, this last quantity becomes $xe^{x^2/2}/\rho$, while if $x\in[\overline{x},\overline{\overline{x}})$, it becomes $(\overline{\omega}/\rho)(1-\ln(x/\overline{\omega})-x^2/2)^{-1}$. Both of these expressions are increasing in x and so the assumptions of Topkis' theorem are satisfied.

Finally, note that the no-absconding constraint will hold as a strict inequality if and only if the optimal x in (53) lies in $[0,\overline{x})$, where $\overline{x}$ solves $\overline{x}e^{\overline{x}^2/2}=\overline{\omega}$. Because $x(S,\overline{\omega})$ is increasing in S whenever it is well-defined, it remains to show that $x(S,\overline{\omega})<\overline{x}$ for sufficiently small $S>0$, and then define $\tilde{S}$ to be the supremum of all such points. It will therefore suffice to show that

$$-\frac{1}{\rho}>\max_{x\in[\overline{x},\overline{\overline{x}})} \frac{(Sx-1)\overline{\omega}/x}{\rho(1+\ln(\overline{\omega}/x)-x^2/2)} \tag{54}$$

for sufficiently small $S>0$, because the right-hand side of inequality (54) is the objective of the principal on the region $[\overline{x},\overline{\overline{x}})$ where the no-absconding constraint holds with equality. Now define

$$Z:=\min_{x\in[\overline{x},\overline{\overline{x}})} \frac{\overline{\omega}/x}{1+\ln(\overline{\omega}/x)-x^2/2} \tag{55}$$

and note that the minimum in (55) is attained at some point because the minimand is continuous and diverges to $+\infty$ as $x\to\overline{\overline{x}}$. Further, $Z>1$, because the inequality $\overline{\omega}/x>1+\ln(\overline{\omega}/x)-x^2/2$

holds for all $x \geq \overline{x}$. To verify this last inequality, note that it is equivalent to

$$\overline{\omega} > x(1 + \ln \overline{\omega}) - x \ln x - x^3/2. \tag{56}$$

The definition of $\overline{x}$ implies that $-\overline{x} \ln \overline{x} - \overline{x}^3/2 = -\overline{x} \ln \overline{\omega}$ and so (56) reduces to $\overline{\omega} > \overline{x}$, which is true. Further, the derivative of the right-hand side of (56) is $\ln(\overline{\omega}/x) - 3x^2/2 < \ln(\overline{\omega}/x) - x^2/2$, which is negative for $x > \overline{x}$ and so (56) holds for $x \geq \overline{x}$ and $Z > 1$. It follows that (54) holds for sufficiently small $S > 0$, because as $S \to 0$, the right-hand side tends to $-Z/\rho < -1/\rho$. □

For future reference, I state the following implication of the proof of Proposition 2.1.

Corollary A.7. *Whenever the principal's problem is finite-valued and the optimal policy is well-defined, the policy function for consumption is $c(u) = \overline{c}u$, where $\overline{c} = \min\{e^{x^2/2}, \overline{\omega}/x\}$, and the law of motion of consumption is $dc_t = \mu_c c_t dt + \sigma_c c_t dB_t$, where $\sigma_c = \sqrt{\rho}x$ and $\mu_c = \rho(-\ln \overline{c} + x^2/2)$, or*

$$\mu_c = \rho \max\left\{0, x^2/2 - \ln(\overline{\omega}/x)\right\}.$$

The local maximum in (36) is only well-defined if $S \leq 1/2$, and so the constant $\tilde{S} = \tilde{S}(\overline{\omega})$ in Proposition 2.1 always satisfies $\tilde{S} \leq 1/2$. The following lemma shows that this upper bound is acheived for $\overline{\omega} = e^{1/2}$, which will be useful in ascertaining when collateral constraints are at their most relaxed value in the decentralization.

Lemma A.8. *The function $\tilde{S}$ satisfies $\tilde{S}(e^{1/2}) = 1/2$.*

Proof. Because $\tilde{S} \leq 1/2$ always, it will suffice to show that $1 = x_{\text{loc}}(1/2) = x(1/2, e^{1/2})$. Substituting $S = 1/2$ and $\overline{\omega} = e^{1/2}$ into the principal's problem in (53), the maximum will be attained at $\overline{x} = 1$ if and only if

$$-e^{1/2}/2 \geq \max_{x \in [1, \overline{\overline{x}}]} \frac{(x/2 - 1)e^{1/2}/x}{3/2 - \ln x - x^2/2}$$

which is equivalent to $-(3/2 - \ln x - x^2/2) \geq (x-2)/x$, or $5/2 \leq \ln x + x^2/2 + 2/x$ for $x \in [1, \overline{\overline{x}}]$. This last inequality reduces to $5/2 \leq 5/2$ at $x = 1$, and $(\ln x + x^2/2 + 2/x)' = 1/x + x - 2/x^2 = (x + x^3 - 2)/x^2$, which is non-negative on $x \geq 1$. Finally, note that Assumption A.2 is satisfied for the above parameters, because Assumption A.1 is obviously satisfied (it reduces to $\overline{\overline{x}} < 2$), and we have $\overline{v}(S, \overline{\omega}) = -\rho^{-1}e^{1/2}/2$, which implies that the stronger assumption in Assumption A.3 is satisfied, because it reduces to $e^{1/2}/2 - (1 + (1/2)^{-2})e^{1/2}/2 \leq 0$. □

The following lemma establishes properties of the local maximum given in equation (36), and will be used in the proof of Proposition 2.2.

Lemma A.9. *The function x_{loc} satisfies $\lim_{S \to 0} x_{\text{loc}}(S)/S = 1$ and for all $S \in [0, 1/2]$ we have $S \leq x_{\text{loc}}(S) \leq 2S$ and $x'_{\text{loc}}(S) \geq 1$.*

Proof of Lemma A.9. Using the definition $x_{\text{loc}}(S) := 1/[2S] - \sqrt{1/[4S^2]-1}$, we have

$$x'_{\text{loc}}(S) = -\frac{1}{2S^2} + \frac{1}{4S^3}\frac{1}{\sqrt{1/[4S^2]-1}} = \frac{1}{2S^2}\left(\frac{1}{\sqrt{1-4S^2}} - 1\right).$$

The inequality $x'_{\text{loc}}(S) \geq 1$ is then equivalent to $1 \geq (1+2S^2)\sqrt{1-4S^2}$, and by squaring both sides, this in turn is equivalent to

$$1 \geq (1+4S^2+4S^4)(1-4S^2) = 1 - 12S^4 - 16S^6$$

which is always true. The inequality $S \leq x_{\text{loc}}(S)$ is equivalent to $2S^2 \leq 1 - \sqrt{1-4S^2}$, or, by rearranging and squaring both sides, $1-4S^2 \leq 1-4S^2+4S^4$, while $x_{\text{loc}}(S) \leq 2S$ is equivalent to $1-4S^2 \leq \sqrt{1-4S^2}$, and both of these are true when $S \in [1/2]$. Finally, l'Hopital's rule implies

$$\lim_{S\to 0} x_{\text{loc}}(S)/S = \lim_{S\to 0}\frac{1-\sqrt{1-4S^2}}{2S^2} = \lim_{S\to 0}\frac{4S/\sqrt{1-4S^2}}{4S} = 1$$

as claimed. □

Recall that Proposition 2.2 in the main text claims that the wedge on the bond ν^B, and the difference in wedges, $\nu^B - \nu^K$, are non-negative and increasing in the marginal product of capital when the no-absconding constraint is strict. I first provide general expressions for these wedges before turning to the proof of Proposition 2.2.

Lemma A.10 (Expressions for wedges)**.** *The wedges on the bond and risky capital are given by*

$$\begin{aligned}\nu^B &= \rho x(S,\overline{\omega})^2 - \mu_c(S,\overline{\omega})\\ \nu^K &= \Pi - \rho - \tau_k + \rho x(S,\overline{\omega})^2 - \sqrt{\rho}\sigma x(S,\overline{\omega}) - \mu_c(S,\overline{\omega})\end{aligned}$$

and so the difference in wedges satisfies

$$\nu^K - \nu^B = \Pi - \rho - \tau_k - \sqrt{\rho}\sigma x(S,\overline{\omega}).$$

Proof. First note that the efficient consumption process satisfies

$$\ln(c_t/c_0) = \mu_c(S,\overline{\omega})t - \rho x(S,\overline{\omega})^2 t/2 + \sqrt{\rho}x(S,\overline{\omega})dB_t.$$

For logarithmic utility the defining equation for wedges is

$$c_t^{-1} = e^{-\rho(t'-t)}\mathbb{E}_t\Big[e^{-\nu^A(t'-t)}R^A c_{t'}^{-1}\Big].$$

Substituting the expression for the log return $\ln R_t^K = \big(\Pi - \tau_k - \sigma^2/2\big)t + \sigma B_t$ into Definition 2.5 and taking logarithms gives

$$\nu^K = \Pi - \rho - \tau_k - \sigma^2/2 + \rho x(S,\overline{\omega})^2/2 + \frac{1}{t}\ln\mathbb{E}\Big[e^{(\sigma-\sqrt{\rho}x(S,\overline{\omega}))B_t}\Big] - \mu_c(S,\overline{\omega}).$$

Using $\mathbb{E}[e^{zB_t}] = e^{z^2t/2}$ gives the claimed expression for the wedge on risky capital. Similarly, substitution of R^B gives the expression for the wedge on the bond. □

Proof of Proposition 2.2. When the no-absconding inequality is strict the drift in consumption is zero and so the wedge on the bond is $\nu^B = \rho x_{\text{loc}}(S)^2$, which is obviously non-negative and increasing in S and hence Π. Because $\phi \leq 1$, the difference between the wedges satisfies

$$\nu^B - \nu^K = -\sqrt{\rho}\phi\sigma S + \sqrt{\rho}\sigma x_{\text{loc}}(S) \geq \sqrt{\rho}\phi\sigma(x_{\text{loc}}(S) - S)$$

which is non-negative by Lemma A.9. Finally, for fixed ϕ the derivative of $\nu^B - \nu^K$ with respect to S is

$$\frac{d}{dS}[-\sqrt{\rho}\phi\sigma S + \sqrt{\rho}\sigma x_{\text{loc}}(S)] = -\sqrt{\rho}\phi\sigma + \sqrt{\rho}\sigma x'_{\text{loc}}(S) \geq \sqrt{\rho}\phi\sigma(x'_{\text{loc}}(S) - 1)$$

which is again non-negative by Lemma A.9 and gives the result. □

B Stationary efficient allocations

This appendix contains formal statements and proofs relating to Section 3 in the main text. Appendix B.1 writes out the aggregate resource constraints and objective appearing in the planner's problem, Appendix B.2 shows how the "relaxed planner's problem" (defined below) reduces to a problem identical in form to the principal-agent problem in Section 2, and Appendix B.3 provides proofs of the characterization of efficient allocations.

B.1 Planner's problem

Recall that in Definition 3.1, superscripts indicate date-of-birth (if not alive at the initial date) or promised utility (if alive at the initial date). Aggregate quantities at any date are comprised of contributions from both the initial generation and subsequent generations, and so in what follows I will break things up in this manner for clarity. The aggregate quantities conditional on being alive associated with the initial generation are distinguished by an underline, and the analogous quantities associated with the generation born at date T are distinguished by a T superscript. To understand the following calculations and expressions, note that, e.g. k^T_{Et} is the capital assigned to an entrepreneur at t born at date T (conditional on being alive), and so, by a law of large numbers, the total capital assigned to all such entrepreneurs is $e^{-\rho_D(t-T)}(1-\psi)\mathbb{E}[k^T_{Et}]$.

Conditional on being alive, the consumption, capital, labor assigned to entrepreneurs, and output at any date $t \geq 0$ of each generation is

$$
\begin{aligned}
\underline{C}_t &:= \int_{\mathbb{R}\times\{E,W\}} \mathbb{E}[c^v_{it}]\Phi(dv,i), \quad C^T_t := (1-\psi)\mathbb{E}[c^T_{Et}] + \psi\mathbb{E}[c^T_{Wt}] \\
\underline{K}_t &:= \int_{\mathbb{R}} \mathbb{E}[k^v_{Et}]\Phi(dv,E), \quad K^T_t := (1-\psi)\mathbb{E}[k^T_{Et}] \\
\underline{L}_t &:= \int_{\mathbb{R}} \mathbb{E}[l^v_{Et}]\Phi(dv,E), \quad L^T_t := (1-\psi)\mathbb{E}[l^T_{Et}] \\
\underline{Y}_t &:= \int_{\mathbb{R}} \mathbb{E}[F(k^v_{Et}, l^v_{Et}) - \delta k^v_{Et}]\Phi(dv,E), \quad Y^T_t := (1-\psi)\mathbb{E}[F(k^T_{Et}, l^T_{Et}) - \delta k^T_{Et}]
\end{aligned}
$$

where $F(K,L) := AK^\alpha L^{1-\alpha}$. The corresponding *aggregate* quantities at any date $t \geq 0$ are then found by invoking a law of large numbers together with the fact that all agents die at rate ρ_D,

$$
\begin{aligned}
C_t &:= e^{-\rho_D t}\underline{C}_t + \rho_D \int_0^t e^{-\rho_D(t-T)} C^T_t dT \\
K_t &:= e^{-\rho_D t}\underline{K}_t + \rho_D \int_0^t e^{-\rho_D(t-T)} K^T_t dT \\
L_t &:= e^{-\rho_D t}\underline{L}_t + \rho_D \int_0^t e^{-\rho_D(t-T)} L^T_t dT \\
Y_t &:= e^{-\rho_D t}\underline{Y}_t + \rho_D \int_0^t e^{-\rho_D(t-T)} Y^T_t dT
\end{aligned}
$$

which are restricted to be bounded and smooth functions of time. Note that for any such bounded and smooth function $H(T,t)$, using $e^{-\rho(t-T)}e^{-\rho_S T} = e^{-\rho_S t}e^{-\rho_D(t-T)}$ and interchanging the order of integration gives

$$
\int_0^\infty \int_0^t e^{-\rho_S t}e^{-\rho_D(t-T)} H(T,t)dtdT = \int_0^\infty \int_T^\infty e^{-\rho(t-T)}e^{-\rho_S T} H(T,t)dtdT. \tag{57}
$$

It follows that the present discounted value of consumption when the interest rate is ρ_S is given by

$$
\begin{aligned}
\int_0^\infty e^{-\rho_S t} C_t dt &= \int_0^\infty e^{-\rho_S t}\left(e^{-\rho_D t}\underline{C}_t + \rho_D \int_0^t e^{-\rho_D(t-T)} C^T_t dT\right)dt \\
&= \int_0^\infty e^{-\rho t}\underline{C}_t dt + \rho_D \int_0^\infty e^{-\rho_S T} \int_T^\infty e^{-\rho(t-T)} C^T_t dtdT
\end{aligned} \tag{58}
$$

and similarly for output and labor. Conditional on being alive, the flow utility experienced at time $t \geq 0$ by each type in the initial and Tth generations are

$$
\begin{aligned}
\underline{U}_{Et} &= \int_{\mathbb{R}} \mathbb{E}[\rho \ln(c^v_{Et})]\Phi(dv,E), \quad U^T_{Et} = (1-\psi)\mathbb{E}[\rho\ln(c^T_{Et})]. \\
\underline{U}_{Wt} &= \int_{\mathbb{R}} \mathbb{E}[\rho \ln(c^v_{Wt})]\Phi(dv,W), \quad U^T_{Wt} = \psi\mathbb{E}[\rho\ln(c^T_{Wt})].
\end{aligned}
$$

In the main text I assume that the planner cares only about workers and values their utility at any date the same regardless of their date of birth. This is equivalent to the objective function associated with an allocation A being given by

$$U^P(A) = \int_0^\infty \left(e^{-\rho t}\underline{U}_{Wt} + \rho_D \int_0^t e^{-\rho_S T} e^{-\rho(t-T)} U_{Wt}^T dT \right) dt. \tag{59}$$

It may benefit the reader to note that by (57), an equivalent representation of the objective (59) is

$$U^P(A) = \int_0^\infty e^{-\rho t}\underline{U}_{Wt} dt + \int_0^\infty e^{-\rho_S T} \left(\int_T^\infty e^{-\rho(t-T)} U_{Wt}^T dt \right) \rho_D dT \tag{60}$$

where the term in parentheses represents the lifetime utility (at birth) of workers born at date $T \geq 0$. I will use (59) and (60) interchangeably in what follows. The following is then the planner's problem, where I remind the reader that $\mathcal{A}^{IF}(\Phi, K)$ was defined in Definition 3.2 to be the set of incentive compatible and resource feasible allocations beginning with an initial capital stock K and distribution Φ.

Definition B.1. *Given* (Φ, K)*, the planner's problem is* $V^P(\Phi, K) = \sup_{A \in \mathcal{A}^{IF}(\Phi,K)} U^P(A)$.

In this paper I restrict attention to stationary solutions to the planner's problem. I therefore search for the distribution Φ and capital stock K such that the distributions of utility, consumption and capital implied by the solution to Definition B.1 are constant over time.

B.2 Reduction to principal-agent problem

I will characterize stationary efficient allocations using the ideas outlined in Farhi and Werning (2007) and consider, in succession, *relaxed* and *generational* planner's problems. The relaxed problem differs from the planner's problem by allowing intertemporal trade at rate ρ_S.

Definition B.2. *Given* (Φ, K)*, the relaxed planner's problem is*

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

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

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

$$K_0 = K$$

where $\mathcal{A}^{IC}(\Phi)$ *denotes the set of incentive compatible allocations.*

If an allocation solves the relaxed planner's problem and the distributions of utility and capital are constant over time, then it also solves the planner's problem beginning at that distribution and capital. To characterize stationary solutions to the original planner's problem, it therefore suffices to consider problems of the form in Definition B.2 and find Φ and K such that stationarity arises. The relaxed planner's problem therefore has only two resource constraints instead of two for each instant in time. Further, because K_t remains bounded, integrating by parts implies that

$$\int_0^\infty e^{-\rho_S t}\dot{K}_t(A)dt = -K_0(A) + \rho_S \int_0^\infty e^{-\rho_S t} K_t(A)dt.$$

Given a distribution Φ over utility and types, when the planner discounts at rate ρ_S the relaxed problem in Definition B.2 is then

$$V^R(\Phi) = \sup_{A \in \mathcal{A}^{IC}(\Phi)} \int_0^\infty \left(e^{-\rho t}\underline{U}_{Wt} + \rho_D \int_0^t e^{-\rho_S T} e^{-\rho(t-T)} U^T_{Wt} dT \right) dt.$$

$$\int_0^\infty e^{-\rho_S t}[C_t(A) + \rho_S K_t(A) - Y_t(A)]dt \leq K_0(A)$$

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

Denote by λ_R and $\lambda_R\lambda_L$ the multipliers on the two resource constraints.[44] The Lagrangian for the relaxed problem is

$$\begin{aligned}\mathcal{L} = &\int_{\mathbb{R}} v\Phi(dv, W) + \rho_D \int_0^\infty e^{-\rho_S T}\psi \int_T^\infty e^{-\rho(t-T)}\mathbb{E}[\rho \ln(c^T_{Wt})]dtdT \\ &- \lambda_R \int_0^\infty e^{-\rho_S t}[C_t + \rho_S K_t - Y_t + \lambda_L L_t]dt + \lambda_R K_0 + \lambda_R \lambda_L L.\end{aligned}$$

Using (57) and (58), the terms that do not depend on the initial generation are

$$\begin{aligned}&\int_0^\infty e^{-\rho_S T}\psi \int_T^\infty e^{-\rho(t-T)}\mathbb{E}[\rho \ln(c^T_{Wt})]dtdT \\ &- \lambda_R \int_0^\infty e^{-\rho_S T} \int_T^\infty e^{-\rho(t-T)}[C^T_t + \rho_S K^T_t - Y^T_t + \lambda_L L^T_t]dtdT \\ &= \int_0^\infty e^{-\rho_S T} \int_T^\infty e^{-\rho(t-T)}\big(\psi \mathbb{E}[\rho \ln(c^T_{Wt})] - \lambda_R[C^T_t + \rho_S K^T_t - Y^T_t + \lambda_L L^T_t]\big)dtdT.\end{aligned}$$

The task of maximizing the terms in the above Lagrangian pertaining to a particular generation born at date T will be referred to as the *generational* planner's problem. The above term in parentheses may be written as

$$\psi \mathbb{E}\big[\rho \ln(c^T_{Wt}) - \lambda_R c^T_{Wt}\big] + (1-\psi)\lambda_R \mathbb{E}\big[\big(A(l^T_{Et}/k^T_{Et})^{1-\alpha} - \lambda_L(l^T_{Et}/k^T_{Et}) - \delta - \rho_S\big)k^T_{Et} - c^T_{Et}\big].$$

Because the choice of labor does not affect the agency problem (entrepreneurs are diverting a multiple of the capital stock, which is independent of labor) the optimal choice of labor per unit of

[44] I write the multiplier on the labor constraint as $\lambda_R\lambda_L$ rather than λ_L because this simplifies the subsequent analysis.

capital solves

$$\Pi(\lambda_L) := \max_{z\geq 0} Az^{1-\alpha} - \lambda_L z - \delta = \alpha A^{1/\alpha}[(1-\alpha)/\lambda_L]^{1/\alpha - 1} - \delta \tag{61}$$

whenever $k_{Et}^T > 0$ (and is zero otherwise). Now, writing

$$S(\lambda_L) = \frac{\Pi(\lambda_L) - \rho_S}{\sqrt{\rho}\phi\sigma}, \tag{62}$$

it follows that the problem of the planner facing the Tth generation is

$$\max_{\substack{W_E, W_W \in \mathbb{R} \\ W_E \geq W_W}} \psi W_W - \psi\lambda_R e^{W_W} + \lambda_R(1-\psi)\overline{v}(S(\lambda_L), \overline{\omega})e^{W_E}. \tag{63}$$

Since $\overline{v} < 0$ wherever it is well-defined, it is immediate that in the relaxed planner's problem we have $W_E = W_W$ and so the problem of a planner facing a particular generation reduces to

$$\psi \max_{W\in\mathbb{R}} W - \lambda_R(1 + (1/\psi - 1)[-\overline{v}(S(\lambda_L), \overline{\omega})])e^W. \tag{64}$$

For any λ_L and $\lambda_R > 0$ such that $S(\lambda_L)$ satisfies the conditions in Proposition 2.1 ensuring that the quantity $\overline{v}(S(\lambda_L), \overline{\omega})$ is well-defined and solves the problem of a principal faced with the above prices, the problem in (64) describes the problem of a generational planner. Because $\overline{v}(S(\lambda_L), \overline{\omega}) < 0$ wherever the principal's problem is well-defined, as λ_R varies from 0 to ∞, the associated W varies monotonically from ∞ to $-\infty$.

B.3 Proofs

Given (62), the average consumption and capital per entrepreneur in the stationary distribution per unit of initial utility are

$$\begin{aligned} \overline{C}(S) &= \frac{\rho_D \overline{c}(S, \overline{\omega})}{\rho_D - \mu_c(S, \overline{\omega})} \\ \overline{K}(S) &= \frac{\rho_D \overline{c}(S, \overline{\omega}) x(S, \overline{\omega})}{(\rho_D - \mu_c(S, \overline{\omega}))\sqrt{\rho}\phi\sigma} \end{aligned} \tag{65}$$

where $\overline{c}$ and x are the policy functions in the principal-agent problem from Section 2. Prior to the proof of Proposition 3.1 I record one final observation.[45]

Lemma B.1. *For any fixed $\overline{\omega}$, the function $\overline{C}(S)$ defined in equation (65) is increasing in S wherever it is well-defined.*

Proof. In view of the expression for consumption in equation (52) and the expression for the drift in consumption in Corollary A.7, I want to show that the function

$$x \mapsto \frac{\rho_D \min\{e^{x^2/2}, \overline{\omega}/x\}}{\rho_D - \rho \max\{0, x^2/2 - \ln(\overline{\omega}/x)\}} \tag{66}$$

[45] Note that by Lemma A.3, $\overline{C}(S)$ is well-defined whenever $S\overline{\overline{x}} < 1$ and $\mu_c(S, \overline{\omega}) < \rho_D$.

is increasing in x. This is obviously true on the region in which $xe^{x^2/2} \leq \overline{\omega}$, while for the x such that $xe^{x^2/2} > \overline{\omega}$, by rearranging (66) we see that it suffices to show that the function

$$g(x) := \frac{\rho_D x}{\rho} - \frac{x^3}{2} + x \ln(\overline{\omega}/x)$$

is decreasing. Evaluating the derivative and rearranging gives

$$g'(x) = \frac{\rho_D}{\rho} - \frac{3x^2}{2} + \ln \overline{\omega} - \ln x - 1 = -\frac{\rho_S}{\rho} - x^2 - \left[x^2/2 - \ln(\overline{\omega}/x)\right]$$

which is necessarily negative if $e^{x^2/2} > \overline{\omega}/x$. □

For the change of variables adopted in Section 2, the capital policy function may be written as

$$\overline{k}(S, \overline{\omega}) := \overline{c}(S, \overline{\omega}) x(S, \overline{\omega}) / (\sqrt{\rho}\phi\sigma).$$

The optimal labor-capital ratio from (61) is $l(\lambda_L) = [(1-\alpha)A/\lambda_L]^{1/\alpha}$, and output per unit of capital may be written

$$Al^{1-\alpha} - \delta = A[(1-\alpha)A/\lambda_L]^{1/\alpha - 1} - \delta = \Pi(\lambda_L)/\alpha + (1/\alpha - 1)\delta. \tag{67}$$

Proof of Proposition 3.1. To see that the solution to (12), denoted $\hat{S}$, is unique whenever it exists, note that dividing by $\overline{C}(S)$ gives

$$1 - \psi + \psi/\overline{C}(S) = (S\sqrt{\rho}\phi\sigma/\alpha + \rho_S/\alpha + (1/\alpha - 1)\delta)(1-\psi)\frac{x(S, \overline{\omega})}{\sqrt{\rho}\phi\sigma} \tag{68}$$

and the right-hand side of (68) is increasing in S while the left-hand side is decreasing in S by Lemma B.1. To establish the existence of a stationary efficient allocation when such a $\hat{S}$ exists, I must show that there exist multipliers $\lambda_R, \lambda_L > 0$ such that the stationary allocation that prevails when one solves the generational planner's problem satisfies the stationary form of the resource constraints. To this end, note that using (67), when viewed as a function of λ_L, the flow production net of depreciation from the firm of an entrepreneur with utility u is

$$(Al(\lambda_L)^{1-\alpha} - \delta)\overline{k}(S(\lambda_L), \overline{\omega})u = (\Pi(\lambda_L)/\alpha + (1/\alpha - 1)\delta)\overline{k}(S(\lambda_L), \overline{\omega})u.$$

Aggregate consumption in the stationary distribution when the initial utility level is u_0 is then given by $((1-\psi)\overline{C}(S(\lambda_L)) + \psi)u_0$ and aggregate output is $(\Pi(\lambda_L)/\alpha + (1/\alpha - 1)\delta)(1-\psi)\overline{K}(S(\lambda_L))u_0$ and so canceling u_0 and recalling $S(\lambda_L) = (\Pi(\lambda_L) - \rho_S)/(\sqrt{\rho}\phi\sigma)$, the goods resource constraint will be satisfied if λ_L satisfies $S(\lambda_L) = \hat{S}$ or

$$\Pi(\lambda_L) = \alpha A^{1/\alpha}[(1-\alpha)/\lambda_L]^{1/\alpha - 1} - \delta = \rho_S + \hat{S}\sqrt{\rho}\phi\sigma$$

where I used the formula in (61). Rearranging then gives the multiplier

$$\alpha^\alpha A[(1-\alpha)/\lambda_L]^{1-\alpha} = \left(\hat{S}\sqrt{\rho}\phi\sigma + \rho_S + \delta\right)^\alpha$$

and hence

$$\lambda_L = \alpha^{\frac{\alpha}{1-\alpha}}(1-\alpha)A^{\frac{1}{1-\alpha}}\left(\hat{S}\sqrt{\rho}\phi\sigma + \rho_S + \delta\right)^{-\frac{\alpha}{1-\alpha}}. \tag{69}$$

The associated labor-capital ratio is then

$$l(\lambda_L) = [(1-\alpha)A/\lambda_L]^{1/\alpha} = \left(\hat{S}\sqrt{\rho}\phi\sigma + \rho_S + \delta\right)^{\frac{1}{1-\alpha}}[\alpha A]^{-\frac{1}{1-\alpha}}.$$

Given this λ_L and (exogenous) stock of labor L, the labor resource constraint will be satisfied if λ_R equals the unique value such that the solution $u_0 := e^W$ to the problem (64) implies that $L/l(\lambda_L) = K = (1-\psi)\overline{K}(S(\lambda_L))u_0$, or

$$u_0 = \frac{L(\alpha A)^{\frac{1}{1-\alpha}}}{(1-\psi)\overline{K}(\hat{S})}\left(\hat{S}\sqrt{\rho}\phi\sigma + \rho_S + \delta\right)^{-\frac{1}{1-\alpha}}$$

which completes the characterization. □

Proof of Proposition 3.2. Notice that the parameter $\overline{\omega} := \sqrt{\rho}\phi\sigma/(\rho\iota)$ remains fixed as we vary the agency frictions in the manner described prior to the statement of the proposition. Rearranging the resource constraint (12) then gives

$$\alpha\sqrt{\rho}\sigma(\psi/\overline{C}(S) + 1 - \psi) = ((\rho_S + (1-\alpha)\delta)/\phi + S\sqrt{\rho}\sigma)(1-\psi)x(S,\overline{\omega}).$$

Both claims in the proposition then follow from Proposition 2.1 together with the fact that the right-hand side is decreasing in ϕ and diverges as $\phi \to 0$. □

C Decentralization

C.1 General implementation

Proof of Lemma 4.1. The Hamilton-Jacobi-Bellman equation for the entrepreneur is

$$\rho V(a) = \max_{\substack{c,k\geq 0 \\ k\leq \overline{\omega}_{\mathrm{d}}(a+h)}} \rho \ln c + (r_{sE}(a+h_E) - c + (1-\tau_\pi)(\Pi - r_b)k)V'(a) + \frac{\sigma^2}{2}(1-\tau_\pi)^2k^2V''(a).$$

Substituting $V(a) = \ln(a+h) + D$ and writing $\overline{c} = c/(a+h_E)$ and $\overline{k} = k/(a+h_E)$ gives

$$\rho D = \max_{\substack{\overline{c},\overline{k}\geq 0 \\ \overline{k}\leq \overline{\omega}_{\mathrm{d}}}} \rho \ln \overline{c} + r_{sE} - \overline{c} + (1-\tau_\pi)(\Pi - r_b)\overline{k} - \frac{\sigma^2}{2}(1-\tau_\pi)^2\overline{k}^2$$

which gives both claims upon substitution. □

Proof of Proposition 4.2. I will first show that for the given after-tax returns, transfers and collateral constraint, all agents obtain the same utility at birth and the law of motion of all agents' consumption coincides with their efficient counterparts. I will then show that the allocation and prices actually constitute a competitive equilibrium as defined in Definition 4.2. To this end, first note that by Lemma A.10, the two requirements on r_{sE} and r_b may be written

$$\begin{aligned} r_{sE} &= \rho - \hat{\nu}^B = \rho - \rho\hat{x}^2 + \hat{\mu}_c \\ r_b &= \rho_S + \hat{\nu}^K - \hat{\nu}^B = \hat{\Pi} - \sqrt{\rho}\sigma\hat{x}. \end{aligned} \tag{70}$$

For the interest rate, profits tax, and constant in the collateral constraint appearing in the statement of the proposition, Lemma 4.1 implies that the constant defining the capital policy function is

$$\overline{k}_{\mathrm{d}} = \frac{\sqrt{\rho}\hat{x}}{\phi\sigma} \tag{71}$$

because the collateral constraint $\overline{k}_{\mathrm{d}} = \sqrt{\rho}\hat{x}/(\phi\sigma) = \iota^{-1}\hat{x}/\overline{\omega} \leq \overline{\omega}_{\mathrm{d}}$ holds automatically from the definition $\overline{\omega} = \sqrt{\rho}\phi\sigma/(\rho\iota)$. For the after-tax returns (70) and $\tau_\pi = 1 - \phi$, Lemma 4.1 implies that the drift in entrepreneurs' consumption satisfies

$$\mu_{c,\mathrm{d}} = r_{sE} - \rho + (1-\tau_\pi)(\hat{\Pi} - r_b)\overline{k}_{\mathrm{d}} = \rho(1-\hat{x}^2) + \hat{\mu}_c - \rho + \phi\sigma\sqrt{\rho}\hat{x}\frac{\sqrt{\rho}\hat{x}}{\phi\sigma} = \hat{\mu}_c$$

and the coefficient of the diffusion term is

$$\sigma_{c,\mathrm{d}} = (1-\tau_\pi)\sigma\overline{k}_{\mathrm{d}} = (1-\tau_\pi)\sigma\sqrt{\rho}\hat{x}/(\phi\sigma) = \sqrt{\rho}\hat{x} = \hat{\sigma}_c$$

while workers' consumption is constant. Given that $(\mu_{c,\mathrm{d}}, \sigma_{c,\mathrm{d}}) = (\hat{\mu}_c, \sqrt{\rho}\hat{x})$, Corollary A.7 implies that $\hat{\mu}_c - (\sqrt{\rho}\hat{x})^2/2 = \rho \max\left\{-\hat{x}^2/2, -\ln(\overline{\omega}/\hat{x})\right\}$ and so the constant in (19) simplifies to

$$\overline{\omega}_{\mathrm{d}} = \iota^{-1}e^{(\hat{\mu}_c - (\sqrt{\rho}\hat{x})^2/2)/\rho} = \iota^{-1}\max\left\{\hat{x}/\overline{\omega}, e^{-\hat{x}^2/2}\right\}$$

as claimed. By Lemma 4.1, entrepreneurs will be indifferent between revealing and not revealing their type at birth if

$$\ln\rho + \ln(\kappa_E\hat{K}) + \frac{1}{\rho}\left(\hat{\mu}_c - \rho\hat{x}^2/2\right) = \ln\rho + \ln(\kappa_W\hat{K}) \tag{72}$$

which, using Corollary A.7 again, rearranges to $\max\left\{-\hat{x}^2/2, -\ln(\overline{\omega}/\hat{x})\right\} = \ln(\kappa_W/\kappa_E)$, and hence

$$\kappa_W = \max\{e^{-\hat{x}^2/2}, \hat{x}/\overline{\omega}\}\kappa_E \tag{73}$$

as claimed. It remains to verify that the market-clearing conditions are satisfied for the prices and after-tax returns appearing in the statement of the proposition. Using (71), the capital market-clearing equation (aggregate capital equals the quantity demanded by entrepreneurs) is

$$\hat{K} = (1-\psi)\frac{\rho_D\kappa_E\hat{K}\overline{k}_{\mathrm{d}}}{\rho_D - \hat{\mu}_c} = (1-\psi)\frac{\rho_D\kappa_E\hat{K}}{(\rho_D - \hat{\mu}_c)}\frac{\sqrt{\rho}\hat{x}}{\phi\sigma} \tag{74}$$

which holds by the definition of κ_E in the statement of the proposition. To verify that the goods market-clearing equation holds, note that using (52), equation (73) implies that

$$1 = (1-\psi)\frac{\rho_D \overline{c}\kappa_W}{(\rho_D - \hat{\mu}_c)}\overline{k}_{\mathrm{d}} \tag{75}$$

Using $\hat{\Pi} = \alpha A \hat{K}^{\alpha-1}L^{1-\alpha} - \delta$ and the fact that all agents consume their total wealth at rate ρ, the goods market-clearing condition is

$$\hat{\Pi}/\alpha + (1/\alpha - 1)\delta = \rho\bigg((1-\psi)\frac{\rho_D\kappa_E}{\rho_D - \hat{\mu}_c} + \psi\kappa_W\bigg). \tag{76}$$

Simplifying and using (73) and (75), equation (76) simplifies to

$$\Big(\hat{\Pi}/\alpha + (1/\alpha - 1)\delta\Big)(1-\psi)\frac{\rho_D\overline{c}\kappa_W}{(\rho_D - \hat{\mu}_c)}\overline{k}_{\mathrm{d}} = \rho\bigg((1-\psi)\frac{\rho_D\overline{c}}{\rho_D - \hat{\mu}_c} + \psi\bigg)\kappa_W. \tag{77}$$

Using the fact that $\overline{k}/\overline{c} = \overline{k}_{\mathrm{d}}/\rho$, equation (77) coincides with (12). Finally, using the policy function of the agents once more, the goods market-clearing condition implies that the stationary total wealth of agents as a fraction of the capital stock is $(\hat{\Pi}/\alpha + (1/\alpha - 1)\delta)/\rho$. The wealth of private agents is then found by subtracting human wealth, which rearranges to give the debt position of the government, using the fact that the sum of public and private wealth equals the capital stock. □

Lemma C.1 (Walras' law)**.** *When the government debt is given by*

$$D = \Big((\hat{\Pi}/\alpha + (1/\alpha - 1)\delta)/\rho - 1\Big)\hat{K} - ((1-\psi)h_E + \psi h_W) \tag{78}$$

the government's budget constraint is satisfied at every instant.

Proof. I will decompose government revenue into four parts: the flow of transfers to newborns, the interest paid on debt, and the revenue raised from taxes on entrepreneurs and workers.

Using $\eta_i\hat{K} = \kappa_i\hat{K} - h_i$, the flow of revenue from the transfers to newborn agents is

$$R_T = -\rho_D[(1-\psi)\eta_E + \psi\eta_W]\hat{K} = -\rho_D((1-\psi)\kappa_E + \psi\kappa_W)\hat{K} + \rho_D((1-\psi)h_E + \psi h_W).$$

The revenue from the interest on government debt is

$$R_I = -rD = -r\Big((\hat{\Pi}/\alpha + (1/\alpha - 1)\delta)/\rho - 1\Big)\hat{K} + r((1-\psi)h_E + \psi h_W).$$

We then have the sum

$$\begin{aligned} R_I + R_T &= -r\Big((\hat{\Pi}/\alpha + (1/\alpha - 1)\delta)/\rho - 1\Big)\hat{K} + r((1-\psi)h_E + \psi h_W) \\ &\quad - \rho_D((1-\psi)\kappa_E + \psi\kappa_W)\hat{K} + \rho_D((1-\psi)h_E + \psi h_W) \\ &= r\hat{K} - r(\hat{\Pi}/\alpha + (1/\alpha - 1)\delta)\hat{K}/\rho + (r + \rho_D)((1-\psi)h_E + \psi h_W) \\ &\quad - \rho_D((1-\psi)\kappa_E + \psi\kappa_W)\hat{K} \\ &= r\hat{K} - (r + \rho_D)(\hat{\Pi}/\alpha + (1/\alpha - 1)\delta)\hat{K}/\rho + (r + \rho_D)((1-\psi)h_E + \psi h_W) \\ &\quad + (1-\psi)\frac{\hat{\mu}_c\rho_D\kappa_E\hat{K}}{\rho_D - \hat{\mu}_c}. \end{aligned} \tag{79}$$

where I used the resource constraint $(\hat{\Pi}/\alpha+(1/\alpha-1)\delta)/\rho = (1-\psi)\kappa_E+\psi\kappa_W+(1-\psi)\hat{\mu}_c\kappa_E/(\rho_D-\hat{\mu}_c)$.

I now use the fact that the revenue raised from each type of agent is the difference between the changes in pre- and post-tax total wealth. In the absence of taxes, the drift in the entrepreneurs' wealth is $(r+\rho_D)a_t - c_t + wL + (\hat{\Pi}-r)k_t$, or

$$(r-\rho_S)(a_t+h_E) + \left(\frac{1-\tau_{sE}}{1-\tau_{LE}} - 1\right)(r+\rho_D)h_E + (\hat{\Pi}-r)\frac{\sqrt{\rho}\hat{x}}{\phi\sigma}(a_t+h_E).$$

The revenue raised from the entrepreneurs is the integral of the above over a_t minus the integral of $\hat{\mu}_c(a_t+h_E)$ over a_t. The revenue raised from entrepreneurs, R_E, and workers, R_W, is then

$$\begin{aligned}
R_E &= (r-\rho_S-\hat{\mu}_c)\frac{(1-\psi)\rho_D\kappa_E\hat{K}}{\rho_D-\hat{\mu}_c} + \left(\frac{1-\tau_{sE}}{1-\tau_{LE}} - 1\right)(r+\rho_D)(1-\psi)h_E + (\hat{\Pi}-r)\hat{K} \\
R_W &= (r-\rho_S)\psi\kappa_W\hat{K} + \left(\frac{1-\tau_{sW}}{1-\tau_{LW}} - 1\right)(r+\rho_D)\psi h_W.
\end{aligned}$$

Using the resource constraint once more gives

$$\begin{aligned}
R_E + R_W = &-\hat{\mu}_c\frac{(1-\psi)\rho_D\kappa_E\hat{K}}{\rho_D-\hat{\mu}_c} + \left(\frac{1-\tau_{sE}}{1-\tau_{LE}} - 1\right)(r+\rho_D)(1-\psi)h_E + (\hat{\Pi}-r)\hat{K} \\
&+ (r-\rho_S)(\hat{\Pi}/\alpha+(1/\alpha-1)\delta)\hat{K}/\rho + \left(\frac{1-\tau_{sW}}{1-\tau_{LW}} - 1\right)(r+\rho_D)\psi h_W.
\end{aligned}$$

It follows that government revenue may be written

$$\begin{aligned}
R_G = &-(r+\rho_D)(\hat{\Pi}/\alpha+(1/\alpha-1)\delta)\hat{K}/\rho + \left(\frac{1-\tau_{sE}}{1-\tau_{LE}}\right)(r+\rho_D)(1-\psi)h_E + \hat{\Pi}\hat{K} \\
&+ (r-\rho_S)(\hat{\Pi}/\alpha+(1/\alpha-1)\delta)\hat{K}/\rho + \left(\frac{1-\tau_{sW}}{1-\tau_{LW}}\right)(r+\rho_D)\psi h_W.
\end{aligned}$$

Using $(1-\tau_{si})(r+\rho_D)h_i/(1-\tau_{Li}) = wL = (\hat{\Pi}+\delta)(1/\alpha-1)\hat{K}$, it follows that

$$\begin{aligned}
R_G = &-(r+\rho_D)(\hat{\Pi}/\alpha+(1/\alpha-1)\delta)\hat{K}/\rho + (\hat{\Pi}+\delta)(1/\alpha-1)\hat{K} \\
&+ \hat{\Pi}\hat{K} + (r-\rho_S)(\hat{\Pi}/\alpha+(1/\alpha-1)\delta)\hat{K}/\rho
\end{aligned}$$

which vanishes, as claimed. □

Proof of Lemma 5.3. Using the definition of S and the fact that x satisfies $Sx^2 - x + S = 0$ when the no-absconding constraint is strict, we can write

$$\hat{\Pi} - \rho_S - \sqrt{\rho}\phi\sigma\hat{x} = (\hat{S}-\hat{x})\sqrt{\rho}\phi\sigma = -\hat{S}\hat{x}^2\sqrt{\rho}\phi\sigma$$

from which the conclusion follows from Proposition 3.2. □

C.2 Specific implementations

Proof of Proposition 5.1. First note that in the absence of taxes, when $s_t = 0$ the law of motion of the entrepreneur's wealth simplifies to

$$[(r + \rho_D)a_t - c_t + wL]dt + (Ak_t^{\alpha-1}l_t^{1-\alpha} - wl_t/k_t - r - \delta)k_t dt + \sigma k_t dB_t. \tag{80}$$

When $r = \rho_S$ and $h_E = wL/\rho$, the average change in wealth in the absence of taxes per unit of time dt may be written more succinctly as $\rho(a_t + h_E) - c_t + (\hat{\Pi} - \rho_S)\overline{k}_{\mathrm{d}}(a_t + h_E)$. Using (71), the policy function $c_t = \rho(a_t + h_E)$, the definitions

$$S := \frac{\Pi - \rho - \tau_k}{\sqrt{\rho}\phi\sigma} \qquad\qquad \overline{\omega} := \frac{\sqrt{\rho}\phi\sigma}{\rho\iota} \tag{81}$$

and the fact that $\hat{S}\hat{x}^2 - \hat{x} + \hat{S} = 0$ when the no-absconding constraint holds (because the derivative in (35) must vanish) this becomes

$$(\hat{\Pi} - \rho_S)\frac{\sqrt{\rho}\hat{x}}{\phi\sigma}(a_t + h_E) = \hat{S}\hat{x}\rho(a_t + h_E) = \frac{\rho\hat{x}^2}{\hat{x}^2 + 1}(a_t + h_E). \tag{82}$$

Because consumption is a martingale under the stated assumptions, the coefficient of $a_t + h_E$ on the right-hand side of equation (82) is also the average revenue generated per unit of total wealth. To determine revenue raised as a fraction of *income*, note that the pre-tax income of entrepreneurs is equal to their after-tax income (which equals consumption when wealth is a martingale) plus the taxes paid, and so tax revenue as a fraction of income is

$$\frac{\rho\hat{x}^2/(\hat{x}^2+1)}{\rho\hat{x}^2/(\hat{x}^2+1) + \rho} = \frac{\hat{x}^2}{2\hat{x}^2 + 1}$$

as claimed. Using $\hat{\Pi} = \alpha A\hat{K}^{\alpha-1}L^{1-\alpha} - \delta$ we have $wL/\hat{K} = (1-\alpha)A\hat{K}^{\alpha-1}L^{1-\alpha} = (\hat{\Pi}+\delta)(1/\alpha - 1)$ from which the desired expression for debt follows from the general expression (78). □

Proof of Proposition 5.2. I will show that for any $\sigma > 0$ the tax on entrepreneurs' interest is negative for all sufficiently small ϕ and that for $\phi = 1$ the tax is positive for $\sigma < \sqrt{\rho}$ when the no-absconding constraint does not hold with equality (which is true for sufficiently small $\sigma > 0$). For sufficiently small ϕ, the no-absconding constraint holds as a strict inequality and the resource constraint becomes

$$(1 - \psi)x(S, \overline{\omega}) = \frac{\alpha\sqrt{\rho}\phi\sigma(\psi/\overline{c}(S, \overline{\omega}) + 1 - \psi)}{\rho_S + S\sqrt{\rho}\phi\sigma + (1-\alpha)\delta}. \tag{83}$$

By Proposition 4.2 and expressions for wedges in Lemma A.10, when $r = r_b$ and $r + \rho_D > 0$ (which holds for sufficiently small $\phi > 0$), the tax on entrepreneurs' interest is negative if and only if

$$\rho - \hat{\nu}^B = \rho(1 - \hat{x}^2) > r + \rho_D = \hat{\Pi} - \sqrt{\rho}\sigma\hat{x} + \rho_D,$$

which simplifies to give

$$0 > \rho x(\hat{S}, \overline{\omega})^2 + \hat{S}\sqrt{\rho}\phi\sigma - \sqrt{\rho}\sigma x(\hat{S}, \overline{\omega}). \tag{84}$$

Using the inequalities in Lemma A.9, a sufficient condition for the tax to be negative is then $(1-\phi)\sigma/\sqrt{\rho} > 2\hat{x}$. Using (83), it will suffice to ensure that

$$\frac{\alpha\sqrt{\rho}\phi\sigma(\psi/\overline{c}(S, \overline{\omega}) + 1 - \psi)}{\rho_S + S\sqrt{\rho}\phi\sigma + (1-\alpha)\delta} < \frac{1}{2}(1-\psi)(1-\phi)\sigma/\sqrt{\rho},$$

which holds for all sufficiently small ϕ. If $\phi = 1$, then the tax will be positive if the right-hand side of (84) is positive. Dividing by $\sqrt{\rho}\sigma x(\hat{S}, \overline{\omega})$, this is equivalent to

$$1 < \sqrt{\rho}x(\hat{S}, \overline{\omega})/\sigma + \hat{S}/x(\hat{S}, \overline{\omega}) \tag{85}$$

If $\sigma < \sqrt{\rho}$ and the no-absconding constraint does not hold as an equality, then it will suffice to show $x_{\text{loc}}(S) < x_{\text{loc}}(S)^2 + S$ for $S \in [0, 1/2]$. Using $x_{\text{loc}}(S) = (1 - \sqrt{1-4S^2})/[2S]$, this is equivalent to $(1-\sqrt{1-4S^2})/[2S] < (1-\sqrt{1-4S^2})^2/[4S^2] + S$, which is true for all sufficiently small $S > 0$. $\square$

D Robustness and extensions

In this appendix I show that the methodological approach adopted in the main text extends (under conditions specified below) to the situation in which utility exhibits constant relative risk aversion with $\gamma \geq 1$ and entrepreneurs are heterogeneous in ex-ante productivity.

I will not strive for complete generality and so I will not attempt to derive analogous statements for all claims made for logarithmic utility and ex-ante identical entrepreneurs. The purpose of this appendix is simply to show that the basic approach followed in the main text does not rely on logarithmic utility or homogeneity of entrepreneurs. In particular, when the no-absconding does not hold with equality, the inverse Euler equation holds and the constrained-efficient allocation can be decentralized by choosing taxes and transfers to match the constant drift and diffusion of log consumption in the efficient allocation.

D.1 Constant relative risk aversion

In this section I show that the qualitative claims in the main text extend to the case of constant relative risk aversion with $\gamma \geq 1$.

D.1.1 Characterization of value function

Proceeding analogously as in Appendix A.1, when characterizing efficient allocations with utility function $u(c) = c^{1-\gamma}/(1-\gamma)$, it is without loss of generality to assume that promised utility satisfies

$$dW_t = \rho\Big(W_t - c_t^{1-\gamma}/(1-\gamma)\Big)dt + \rho\phi\sigma k_t c_t^{-\gamma} dB_t.$$

Now define utility in consumption units as $u_t := [(1-\gamma)W_t]^{\frac{1}{1-\gamma}}$ and change variables to $\overline{c}_t$ and $\overline{k}_t$ defined by $c_t = \overline{c}_t[(1-\gamma)W_t]^{\frac{1}{1-\gamma}}$ and $k_t = \overline{k}_t[(1-\gamma)W_t]^{\frac{1}{1-\gamma}}$. In this notation, the law of motion of promised utility is

$$dW_t = \rho\Big(1 - \overline{c}_t^{1-\gamma}\Big)W_t dt + \rho\phi\sigma\overline{k}_t\overline{c}_t^{-\gamma}(1-\gamma)W_t dB_t =: \mu_W W_t dt + \sigma_W W_t dB_t.$$

If $f(W_t) := [(1-\gamma)W_t]^{\frac{1}{1-\gamma}}$ then $f'(W_t) := [(1-\gamma)W_t]^{\frac{\gamma}{1-\gamma}}$ and $f''(W_t) := \gamma[(1-\gamma)W_t]^{\frac{\gamma}{1-\gamma}-1}$, and so using Ito's lemma the law of motion for u_t is

$$\begin{aligned}
df(W_t) &= (\mu_W W_t f'(W_t) + (\sigma_W^2/2)W_t^2 f''(W_t))dt + f'(W_t)\sigma_W W_t dB_t \\
&= \rho\left(\frac{(1-\overline{c}_t^{1-\gamma})}{1-\gamma}[(1-\gamma)W_t]f'(W_t) + (\sqrt{\rho}\phi\sigma\overline{k}_t\overline{c}_t^{-\gamma})^2[(1-\gamma)W_t]^2 f''(W_t)/2\right)dt \\
&\quad + \rho\phi\sigma\overline{k}_t\overline{c}_t^{-\gamma}[(1-\gamma)W_t]f'(W_t)dB_t
\end{aligned}$$

which may be written as

$$du_t = \rho\left(\frac{1-\overline{c}_t^{1-\gamma}}{1-\gamma} + \gamma x_t^2/2\right)u_t dt + \sqrt{\rho}x_t u_t dB_t \tag{86}$$

where $x := \sqrt{\rho}\phi\sigma\overline{k}\overline{c}^{-\gamma}$. As with the case of logarithmic utility, a homogeneity argument combining the constant-returns-to-scale objective of the principal with the law of motion (86) shows that the value function is linear in u_t whenever it is finite-valued, and so it is of the form $v(u) = \overline{v}u$, where

$$\overline{v} = \overline{v}(S, \overline{\omega}) = \sup_{(\overline{c},x)\in D(\overline{\omega})} \frac{Sx\overline{c}^{\gamma} - \overline{c}}{\rho - \mu_c(\overline{c}, x)} \tag{87}$$

is a candidate coefficient of the value function, where again $S = (\Pi - \rho - \tau_k)/(\sqrt{\rho}\phi\sigma)$ and for any $\overline{\omega} > 0$, the set $D(\overline{\omega})$ is defined by

$$D(\overline{\omega}) := \big\{(\overline{c}, x) \in \mathbb{R}^2 \mid \overline{c} > 0, x \geq 0\ x\overline{c}^{\gamma} \leq \overline{\omega},\ \mu_c(\overline{c}, x) < \rho\big\} \tag{88}$$

and the coefficient of the drift in consumption is

$$\mu_c(\overline{c}, x) := \rho\left(\frac{1-\overline{c}^{1-\gamma}}{1-\gamma} + \gamma x^2/2\right).$$

Using the fact that for all $(\overline{c}, x) \in D(\overline{\omega})$, we have $(1-\overline{c}^{1-\gamma})/(1-\gamma) < 1$, or $\overline{c} > \gamma^{\frac{1}{1-\gamma}}$, I define for any $\epsilon > 0$ the following compact subset of $D(\overline{\omega})$,

$$D_\epsilon(\overline{\omega}) := \left\{(\overline{c}, x) \in \mathbb{R}^2 \;\middle|\; \overline{c} \in [\gamma^{\frac{1}{1-\gamma}}, 1/\epsilon], x \geq 0\ x\overline{c}^{\gamma} \leq \overline{\omega},\ \mu_c(\overline{c}, x) \leq \rho(1-\epsilon)\right\}. \tag{89}$$

Finally, define $\overline{\overline{x}}$ to be the solution to

$$\frac{\gamma\overline{\overline{x}}^2}{2} + \frac{1-(\overline{\omega}/\overline{\overline{x}})^{1/\gamma-1}}{1-\gamma} = 1$$

which exists because $\gamma > 1$ (and so the left-hand side is increasing in $\overline{\overline{x}}$). We then have the following analogue of Assumption A.1.

Assumption D.1. $S\overline{\overline{x}}^{1/\gamma}\overline{\omega}^{1-1/\gamma} < 1$.

The following establishes one of the claims in Lemma A.3 (the only part that is necessary in what follows) for the case of CRRA preferences with $\gamma \geq 1$.

Lemma D.1. *Assumption D.1 implies that* $\overline{v}(S, \overline{\omega}) < 0$.

Proof. First note that because $\gamma \geq 1$, if there existed $(\overline{c}, x) \in D(\overline{\omega})$ with $Sx\overline{c}^{\gamma-1} \geq 1$, then (by increasing $\overline{c}$, which weakly decreases μ_c) there would also exist $(\overline{c}, x) \in D(\overline{\omega})$ in which $\overline{c} = (\overline{\omega}/x)^{1/\gamma}$ and $Sx\overline{c}^{\gamma-1} \geq 1$. In other words, there would exist $x > 0$ such that $Sx^{1/\gamma}\overline{\omega}^{1-1/\gamma} \geq 1$ and $\gamma x^2/2 + (1 - (\overline{\omega}/x)^{1/\gamma-1})/(1-\gamma) < 1$, which violates Assumption D.1 and the definition of $\overline{\overline{x}}$. To see that the supremum in (87) is attained at some point, first note that $x < \overline{\overline{x}}$ for all $(\overline{c}, x) \in D(\overline{\omega})$, because we either have $x = 0$ or

$$\frac{\gamma x^2}{2} < 1 + \frac{\overline{c}^{1-\gamma} - 1}{1-\gamma} \leq 1 + \frac{(\overline{\omega}/x)^{1/\gamma-1} - 1}{1-\gamma}.$$

It will suffice to note that for some $\epsilon > 0$, the supremand in (87) is weakly below $-1/\rho$ (which is the value associated with $(\overline{c}, x) = (1, 0)$) for all $(\overline{c}, x) \in D(\overline{\omega}) \setminus D_\epsilon(\overline{\omega})$ because this will mean that there is no loss in restricting attention to the compact set $D_\epsilon(\overline{\omega})$ in (87). Because the no-absconding constraint implies that for all $(\overline{c}, x) \in D(\overline{\omega})$, we have $\overline{c}^{\gamma-1} \leq (\overline{\omega}/x)^{1-1/\gamma}$, by defining $\delta := 1 - S\overline{x}^{1/\gamma}\overline{\omega}^{1-1/\gamma}$ we see that it will suffice to show

$$\delta\overline{c} \geq 1 - \mu_c(\overline{c}, x)/\rho \qquad \forall (\overline{c}, x) \in D(\overline{\omega}) \setminus D_\epsilon(\overline{\omega})$$

which holds for all sufficiently small $\epsilon > 0$. Because $D_\epsilon(\overline{\omega})$ is compact this shows that the supremum in (87) is attained under Assumption D.1. $\square$

The following is the analogue of Assumption A.3.

Assumption D.2. $S\overline{\omega} + \rho\big((S\overline{\omega})^{1-\gamma} + \gamma^2\overline{\omega}^{2-2\gamma}S^{-2\gamma}\big)\overline{v}(S, \overline{\omega}) \leq 0$.

By Lemma D.1, we see that $\overline{v}(S, \overline{\omega}) < 0$ for sufficiently small $S > 0$, and so Assumptions D.1 and D.2 are obviously satisfied for all sufficiently small S. The following establishes the main claims of Proposition 2.1 for the case of CRRA preferences.

Proposition D.2. *If Assumptions D.1 and D.2 are satisfied, then the value function is finite-valued and given by* $v(u) = \overline{v}u$ *for all* $u > 0$. *Further, the no-absconding constraint holds for sufficiently small positive* S, *in which case the inverse Euler equation holds.*

Proof. As per the proof of Proposition 2.1, to establish that the optimal choices in (87) solve the HJB equation

$$0 = \sup_{\substack{\overline{c}, x \geq 0 \\ x\overline{c}^\gamma \leq \overline{\omega}}} (Sx\overline{c}^{\gamma-1} - 1)\overline{c} + \rho\left(\frac{1 - \overline{c}^{1-\gamma}}{1-\gamma} + \gamma x^2/2 - 1\right)\overline{v} =: \sup_{\substack{\overline{c}, x \geq 0 \\ x\overline{c}^\gamma \leq \overline{\omega}}} H(\overline{c}, x, \overline{v}),$$

it remains to eliminate the possibility that $H(\overline{c}, x, \overline{v}) > 0$ for some pair $(\overline{c}, x)$ satisfying $\overline{c}, x \geq 0$ and $x\overline{c}^\gamma \leq \overline{\omega}$. The existence of such a pair $(\overline{c}, x)$ satisfying $(1 - \overline{c}^{1-\gamma})/(1 - \gamma) + \gamma x^2/2 - 1 < 0$ would violate the definition of $\overline{v}$, and so it will suffice to rule out the existence of a pair $(\overline{c}, x)$ such that $(1 - \overline{c}^{1-\gamma})/(1 - \gamma) + \gamma x^2/2 - 1 > 0$ and

$$(Sx\overline{c}^{\gamma-1} - 1)\overline{c} + \rho((1 - \overline{c}^{1-\gamma})/(1 - \gamma) + \gamma x^2/2 - 1)\overline{v} > 0.$$

Since $\overline{v} < 0$, this last inequality implies

$$(Sx^{1/\gamma}\overline{\omega}^{1-1/\gamma} - 1)(\overline{\omega}/x)^{1/\gamma} + \rho\Big(\frac{1 - (\overline{\omega}/x)^{1/\gamma-1}}{1-\gamma} + \gamma x^2/2 - 1\Big)\overline{v} > 0 \tag{90}$$

for some $x > \overline{\overline{x}}$. The inequality (90) cannot hold on $x \in [\overline{\overline{x}}, \overline{\omega}^{1-\gamma}S^{-\gamma}]$ from Assumption D.1, and so to establish finiteness it will suffice for the derivative of the left-hand side with respect to x to be negative for all $x \geq \overline{\omega}^{1-\gamma}S^{-\gamma}$. This is equivalent to

$$\overline{\omega}^{1/\gamma}x^{-1/\gamma} + \rho\Big(\overline{\omega}^{1/\gamma-1}x^{1-1/\gamma} + \gamma^2x^2\Big)\overline{v} \leq 0.$$

Because $\overline{v} < 0$, this will be true for $x \geq \overline{\omega}^{1-\gamma}S^{-\gamma}$ if and only if it is true for $x = \overline{\omega}^{1-\gamma}S^{-\gamma}$, which or $\overline{\omega}^{1/\gamma}(\overline{\omega}^{1-\gamma}S^{-\gamma})^{-1/\gamma} + \rho\big(\overline{\omega}^{1/\gamma-1}(\overline{\omega}^{1-\gamma}S^{-\gamma})^{1-1/\gamma} + \gamma^2(\overline{\omega}^{1-\gamma}S^{-\gamma})^2\big)\overline{v} \leq 0$ which is exactly Assumption D.2 upon simplification. The fact that the candidate value function $\overline{v}u$ is actually the value function of the principal then follows from arguments analogous to those given in the proof of Theorem A.6.

To establish the second claim, note that Ito's Lemma implies that the inverse Euler equation is equivalent to $\mu_c = (1-\gamma)\sigma_c^2/2$. For the case with $\gamma > 1$ it is convenient to define $z := x\overline{c}^{\gamma-1}$ and to write the problem as

$$\sup_{\substack{\overline{c}>0, \overline{c}z\leq\overline{\omega} \\ \frac{1-\overline{c}^{1-\gamma}}{1-\gamma}+\gamma z^2\overline{c}^{2-2\gamma}/2<1}} \frac{(Sz-1)\overline{c}/\rho}{1 + \frac{\overline{c}^{1-\gamma}-1}{1-\gamma} - \gamma z^2\overline{c}^{2-2\gamma}/2}. \tag{91}$$

We now fix z and minimize the negative of the above supremand over $\overline{c}$, which is equivalent to minimizing the problem

$$\ln\overline{c} - \ln\left(1 + \frac{\overline{c}^{1-\gamma}-1}{1-\gamma} - \frac{\gamma}{2}z^2\overline{c}^{2-2\gamma}\right) \tag{92}$$

over the set of scalars $\overline{c} > 0$ satisfying $\overline{c}z \leq \overline{\omega}$ and $(1 - \overline{c}^{1-\gamma})/(1-\gamma) + \gamma z^2\overline{c}^{2-2\gamma}/2 < 1$. Writing $y := \overline{c}^{1-\gamma}$, the problem is equivalent to maximizing

$$\ln y + (\gamma - 1)\ln\left(\frac{y-\gamma}{1-\gamma} - \gamma z^2y^2/2\right)$$

over the set of $y > 0$ such that $y \geq (\overline{\omega}/z)^{1-\gamma}$ and $(y-\gamma)/(1-\gamma) - \gamma z^2y^2/2 > 0$. This last function is concave and diverges as $(y-\gamma)/(1-\gamma)$ approaches $\gamma z^2y^2/2 > 0$, and so the optimal choice of consumption in (92) is the minimum of the solution to the first-order condition and the boundary value $\overline{c} = \overline{\omega}/z$. The first-order condition for consumption is

$$\frac{1}{\overline{c}} = \frac{\overline{c}^{-\gamma} - \gamma(1-\gamma)z^2\overline{c}^{1-2\gamma}}{\frac{\overline{c}^{1-\gamma}-\gamma}{1-\gamma} - \gamma z^2\overline{c}^{2-2\gamma}/2}.$$

Rearranging gives $(\overline{c}^{1-\gamma} - \gamma)/(1-\gamma) = \overline{c}^{1-\gamma} - (1/2-\gamma)\gamma z^2\overline{c}^{2-2\gamma}$ and

$$\overline{c}^{1-\gamma} - 1 = -(1-\gamma)(1/2-\gamma)z^2\overline{c}^{2-2\gamma} \tag{93}$$

which simplifies to a quadratic in $\overline{c}^{1-\gamma}$, $0 = 1-\overline{c}^{1-\gamma}+(1-\gamma)(\gamma-1/2)z^2\overline{c}^{2-2\gamma}$, which has one positive solution for consumption which I denote by $\overline{c}_{\text{foc}}(z)$. Note that $\overline{c}_{\text{foc}}(z)$ is necessarily increasing in z. Using (93) and the definition of $x = z\overline{c}^{1-\gamma}$ gives $(\overline{c}_{\text{foc}}^{1-\gamma} - 1)/(1-\gamma) = -(1/2-\gamma)x^2$, which is equivalent to the inverse Euler equation. In what follows I write

$$\overline{c}(z) = \min\{\overline{c}_{\text{foc}}(z), \overline{\omega}/z\} \tag{94}$$

for the optimal c given z. Returning to the original problem (91), I now define $\overline{z}$ as the solution to

$$0 = 1 - (\overline{\omega}/\overline{z})^{1-\gamma} + (1-\gamma)(\gamma-1/2)\overline{\omega}^{2-2\gamma}\overline{z}^{2\gamma}, \tag{95}$$

which is the largest $z > 0$ such that the no-absconding constraint $\overline{c}z = \overline{\omega}$ and inverse Euler equation holds. Similarly, define $\overline{\overline{z}}$ as the solution to

$$0 = 1 + \frac{(\overline{\omega}/\overline{\overline{z}})^{1-\gamma} - 1}{1-\gamma} - \frac{\gamma}{2}\overline{\overline{z}}^{2\gamma}\overline{\omega}^{2-2\gamma}$$

which is the largest z such that the no-absconding constraint holds with equality and consumption growth is less than the rate of discount. The no-absconding constraint will hold as a strict inequality if the optimal z in (91) lies in $[0, \overline{z})$, and so it remains to show that this holds for sufficiently small $S > 0$. As with the proof of Proposition 2.1, for this it will suffice to show that $Z > 1$, where

$$Z := \min_{z\in[\overline{z},\overline{\overline{z}}]} \frac{\overline{\omega}/z}{1 + \frac{(\overline{\omega}/z)^{1-\gamma}-1}{1-\gamma} - \frac{\gamma}{2}z^{2\gamma}\overline{\omega}^{2-2\gamma}} \tag{96}$$

is the analogue of (55). I first show that the minimand in (96) exceeds 1 when $z = \overline{z}$ by noting that

$$\begin{aligned}\frac{\overline{\omega}/\overline{z}}{1 + \frac{(\overline{\omega}/\overline{z})^{1-\gamma}-1}{1-\gamma} - \frac{\gamma}{2}\overline{z}^{2\gamma}\overline{\omega}^{2-2\gamma}} &= \min_{\substack{\overline{c}>0 \\ \frac{1-\overline{c}^{1-\gamma}}{1-\gamma}+\frac{\gamma}{2}\overline{z}^2\overline{c}^{2-2\gamma}<1}} \frac{\overline{c}}{1 + \frac{\overline{c}^{1-\gamma}-1}{1-\gamma} - \frac{\gamma}{2}\overline{z}^2\overline{c}^{2-2\gamma}} \\ &> \min_{\substack{\overline{c}>0 \\ \frac{1-\overline{c}^{1-\gamma}}{1-\gamma}+\frac{\gamma}{2}\overline{z}^2\overline{c}^{2-2\gamma}<1}} \frac{\overline{c}}{1 + \frac{\overline{c}^{1-\gamma}-1}{1-\gamma}} \geq 1.\end{aligned} \tag{97}$$

The desired inequality $Z > 1$ is equivalent to

$$\overline{\omega} > z + \frac{z^\gamma\overline{\omega}^{1-\gamma} - z}{1-\gamma} - \gamma z^{2\gamma+1}\overline{\omega}^{2-2\gamma}/2 \tag{98}$$

for all $z \in [\overline{z}, \overline{\overline{z}}]$, and (97) shows that this holds at $z = \overline{z}$. The derivative with respect to z of the right-hand side of (98) is then

$$1 + \frac{\gamma(\overline{\omega}/z)^{1-\gamma} - 1}{1-\gamma} - \gamma(\gamma + 1/2) z^{2\gamma}\overline{\omega}^{2-2\gamma} = \gamma\biggl(\frac{(\overline{\omega}/z)^{1-\gamma} - 1}{1-\gamma} - (\gamma + 1/2) z^{2\gamma}\overline{\omega}^{2-2\gamma}\biggr),$$

which will be negative for $z \geq \overline{z}$ if and only if it is negative for $z = \overline{z}$. At $\overline{z}$ this expression is negative if $0 > \frac{(\overline{\omega}/\overline{z})^{1-\gamma} - 1}{1-\gamma} - (\gamma + 1/2)\overline{z}^{2\gamma}\overline{\omega}^{2-2\gamma}$ and hence, using (95), if

$$0 > (\gamma - 1/2)\overline{\omega}^{2-2\gamma}\overline{z}^{2\gamma} - (\gamma + 1/2)\overline{z}^{2\gamma}\overline{\omega}^{2-2\gamma} = -\overline{\omega}^{2-2\gamma}\overline{z}^{2\gamma}$$

which is true. □

D.1.2 Efficient stationary distribution and decentralization

In this section, I will derive analogues of the characterization of stationary efficient allocations in Proposition 3.1 and the decentralization in Proposition 4.2. As mentioned above, in this appendix I do not aim for the same level of generality as for the case with logarithmic utility. I therefore restrict attention to parameters for which the no-absconding constraint does not hold with equality, so that the collateral constraints in the consumer problem do not bind. I will first solve the individual consumer problems. I will omit the problem of workers because in the efficient allocation workers have zero drift in wealth and consumption. For CRRA utility with $\gamma > 1$, the HJB equation for an entrepreneur facing borrowing costs r_b and taxes τ_{LE} on labor, τ_{sE} on interest, and τ_π on profits is

$$\begin{aligned}\rho V_E(a) = \max_{\substack{k,c\geq 0 \\ k \leq \overline{\omega}_{\mathrm{d}}(a+h_E)}} & \frac{\rho c^{1-\gamma}}{1-\gamma} + [r_{sE}a - c + (1-\tau_\pi)(\Pi - r_b)k + (1-\tau_{LE})wL]V_E'(a) \\ & + (1-\tau_\pi)^2 \frac{\sigma^2 k^2}{2} V_E''(a)\end{aligned}$$

where $\Pi = \Pi(w) := \max_{z\geq 0} Az^{1-\alpha} - wz - \delta$ is the marginal product of capital for a given wage w and $r_{sE} = (1-\tau_{sE})(r + \rho_D)$ is the after-tax safe return available to the entrepreneur.

Lemma D.3. *The entrepreneur will choose not to divert if and only if $\tau_\pi \leq 1 - \phi$. In this case, if $\overline{\omega}_{\mathrm{d}} > (\Pi - r_b)/[\gamma(1-\tau_\pi)\sigma^2]$, then the value function is well-defined and finite if $r_{sE} > 0$, in which case it is of the form $V_E(a) = \overline{V}_E(a + h_E)^{1-\gamma}/(1-\gamma)$ for some $\overline{V}_E$, and the policy function are*

$$\begin{aligned} c(a) &= \overline{c}_{\mathrm{d}}(a + h_E) = \left(\frac{1}{\gamma}[\rho - (1-\gamma)r_{sE}] - \frac{(\Pi - r_b)^2}{2\gamma^2\sigma^2}(1-\gamma)\right)(a + h_E) \\ k(a) &= \overline{k}_{\mathrm{d}}(a + h_E) = \frac{(\Pi - r_b)(a + h_E)}{\gamma\sigma^2(1-\tau_\pi)}. \end{aligned}$$

The associated law of motion of wealth is $da = \mu_c(a_t + h_E)dt + \sigma_c(a_t + h_E)dB_t$, where

$$\mu_c = \frac{1}{\gamma}[r_{sE} - \rho] + \frac{(\Pi - r_b)^2}{2\gamma^2\sigma^2}(1+\gamma) \qquad\qquad \sigma_c = \frac{\Pi - r_b}{\gamma\sigma}.$$

Proof. Upon substituting the assumed form, the HJB equation becomes

$$\frac{\rho\overline{V}_E}{1-\gamma} = \max_{\substack{\overline{c}_{\mathrm{d}},\overline{k}_{\mathrm{d}}\geq 0\\ \overline{k}_{\mathrm{d}}\leq\overline{\omega}_{\mathrm{d}}}} \frac{\rho\overline{c}_{\mathrm{d}}^{1-\gamma}}{1-\gamma} + \left[r_{sE} - \overline{c}_{\mathrm{d}} + (1-\tau_\pi)(\Pi - r_b)\overline{k}_{\mathrm{d}}\right]\overline{V}_E - \gamma(1-\tau_\pi)^2\overline{V}_E\frac{\sigma^2\overline{k}_{\mathrm{d}}^2}{2}.$$

The first-order conditions for capital and consumption give

$$\overline{k}_{\mathrm{d}} = \frac{\Pi - r_b}{\gamma(1-\tau_\pi)\sigma^2} \qquad\qquad \overline{c}_{\mathrm{d}} = (\overline{V}_E/\rho)^{-1/\gamma}$$

where the maintained assumption on $\overline{\omega}_{\mathrm{d}}$ ensures that the collateral constraint does not bind. Substituting into the HJB equation gives

$$\begin{aligned}\frac{\rho\overline{V}_E}{1-\gamma} &= \frac{\rho^{1/\gamma}\overline{V}_E^{1-1/\gamma}}{1-\gamma} + \overline{V}_E(r_{sE} - \overline{c}_{\mathrm{d}}) + \overline{V}_E\left[(1-\tau_\pi)(\Pi - r_b)\overline{k}_{\mathrm{d}} - \frac{\gamma}{2}[(1-\tau_\pi)\sigma]^2\overline{k}^2\right]\\ &= \frac{\gamma\rho^{1/\gamma}\overline{V}_E^{1-1/\gamma}}{1-\gamma} + \overline{V}_E r_{sE} + \overline{V}_E\left[\frac{(\Pi - r_b)^2}{\gamma\sigma^2} - \frac{\gamma}{2}[(1-\tau_\pi)\sigma]^2\left(\frac{\Pi - r_b}{\gamma(1-\tau_\pi)\sigma^2}\right)^2\right]\end{aligned}$$

which rearranges to

$$\frac{\rho}{1-\gamma} = \frac{\gamma(\overline{V}_E/\rho)^{-1/\gamma}}{1-\gamma} + r_{sE} + \frac{(\Pi - r_b)^2}{2\gamma\sigma^2}$$

and therefore implies

$$\overline{c}_{\mathrm{d}} = (\overline{V}_E/\rho)^{-1/\gamma} = \frac{1}{\gamma}(\rho - (1-\gamma)r_{sE}) - \frac{(\Pi - r_b)^2}{2\gamma^2\sigma^2}(1-\gamma)$$

as claimed. Under the maintained assumptions that $\gamma \geq 1$ and $r_{sE} > 0$, the coefficient $\overline{c}_{\mathrm{d}}$ is positive and the associated value function coefficient $\overline{V}_E$ is finite. The law of motion of wealth is then

$$da_t = [r_{sE}a_t + (1-\tau_{LE})wL - c_t + (1-\tau_\pi)(\Pi - r_b)k_t]dt + (1-\tau_\pi)\sigma k_t dB_t$$

and so the law of total wealth is

$$\begin{aligned}\frac{d(a_t + h_E)}{(a_t + h_E)} &= \left[r_{sE} - \overline{c}_{\mathrm{d}} + (1-\tau_\pi)(\Pi - r_b)\overline{k}_{\mathrm{d}}\right]dt + (1-\tau_\pi)\sigma\overline{k}_{\mathrm{d}}dB_t\\ &= \left[r_{sE} - \frac{1}{\gamma}(\rho - (1-\gamma)r_{sE}) + \frac{(\Pi - r_b)^2}{2\gamma^2\sigma^2}(1-\gamma) + \frac{(\Pi - r_b)^2}{\gamma\sigma^2}\right]dt + \frac{(\Pi - r_b)}{\gamma\sigma}dB_t.\end{aligned}$$

This implies $\sigma_c = (\Pi - r_b)/(\gamma\sigma)$, while μ_c simplifies to

$$\mu_c = \frac{1}{\gamma}(r_{sE} - \rho) + \frac{(\Pi - r_b)^2}{\gamma\sigma^2}\left(\frac{1}{2}(1/\gamma - 1) + 1\right)$$

as claimed. □

As with the case of logarithmic utility treated in the main text, the homogeneity of preferences implies that the efficient allocation is again characterized by three properties: the marginal product of capital coincides with the solution to the stationary form of the goods resource constraint, and the mean and standard deviation of the growth in consumption coincide with those in the efficient allocation. For the latter, when the no-absconding constraint does not hold with equality, efficiency requires that the consumption of entrepreneurs satisfies $dc_t = \mu_c c_t dt + \sigma_c c_t dB_t$ where $\mu_c = \rho(1-\gamma)x^2/2$ and $\sigma_c = \sqrt{\rho}x$ for some constant $x > 0$. Obviously, for the welfare notion adopted in the main text, the consumption of workers is constant.

The following is a simplified analogue of Proposition 3.1 (existence of stationary efficient allocation if solution to (99) exists) and Proposition 4.2 (decentralization of efficient allocation), for the case $\gamma \geq 1$.

Proposition D.4. *The marginal product of capital that obtains in an efficient stationary distribution is $\hat{\Pi} = \rho_S + \hat{S}\sqrt{\rho}\phi\sigma$, where $\hat{S}$ is any solution to the equation*

$$\frac{\rho_D \overline{c}(S,\overline{\omega})}{\rho_D - \mu_c(S,\overline{\omega})} + \frac{\psi}{1-\psi} = [S\sqrt{\rho}\phi\sigma/\alpha + \rho_S/\alpha + (1/\alpha - 1)\delta]\frac{\rho_D \overline{c}(S,\overline{\omega})^\gamma x(S,\overline{\omega})}{(\rho_D - \mu_c(S,\overline{\omega}))\sqrt{\rho}\phi\sigma} \tag{99}$$

provided that such a solution exists and satisfies $\mu_c(\hat{S},\overline{\omega}) < \rho_D$. In this case, if the no-absconding constraint does not hold with equality at this $\hat{S}$, then writing $\hat{x} = x(\hat{S},\overline{\omega})$ and $\hat{\mu}_c = \rho(1-\gamma)\hat{x}^2/2$, the efficient allocation may be decentralized in a competitive equilibrium in which $\tau_\pi = 1-\phi$, $r_b = \hat{\Pi} - \gamma\sqrt{\rho}\sigma\hat{x}$, $r_{sE} = \rho(1-\gamma^2\hat{x}^2)$ and $r_{sW} = \rho$. The endowed total wealth of entrepreneurs as a fraction of the capital stock is

$$\kappa_E = \frac{\phi\sigma(\rho_D - \hat{\mu}_c)}{\sqrt{\rho}\hat{x}(1-\psi)\rho_D}$$

and κ_W is chosen such that entrepreneurs and workers obtain the same level of lifetime utility.

Proof. The proof proceeds in an almost identical fashion to the proof of Proposition 4.2, and so I only highlight the relevant differences. The risk borne by the agent when $\tau_\pi = 1-\phi$ is $\phi\sigma\overline{k}_{\mathrm{d}}$, which from Lemma D.3 implies that the entrepreneurs bear the efficient level of risk if $\sqrt{\rho}\hat{x} = (\hat{\Pi} - r_b)/(\gamma\sigma)$, which is true for the above choice of borrowing cost r_b. In this case, the mean growth in entrepreneurs' consumption is

$$\mu_c = \frac{1}{\gamma}(r_{sE} - \rho) + \frac{(\hat{\Pi} - r_b)^2}{2\gamma^2\sigma^2}(1+\gamma) = \frac{1}{\gamma}(r_{sE} - \rho) + \rho\hat{x}^2(1+\gamma)/2.$$

In order for the inverse Euler equation to hold, we need $\mu_c = \rho(1-\gamma)\hat{x}^2/2$, which requires $\rho(1-\gamma)\hat{x}^2/2 = (r_{sE} - \rho)/\gamma + \rho\hat{x}^2(1+\gamma)/2$ and hence $r_{sE} = \rho(1-\gamma^2\hat{x}^2)$, as claimed. The expression for κ_E then follows from reasoning identical to that given in the proof of Proposition 4.2. □

D.2 Heterogeneous entrepreneurs

This appendix justifies the claim, made at the end of Section 5.2, that the decentralization in Proposition 3.1 generalizes to the case in which entrepreneurs differ in productivity ex-ante, at least if these productivity differences are permanent and observable. Appendix D.2.1 characterizes the efficient allocations and Appendix D.2.2 derives the decentralization.

D.2.1 Efficient allocations

In this section I show how the characterization of efficient allocations changes in the presence of two types of entrepreneurs under conditions that ensure that the problem of the principal is well-defined. I will maintain the assumption adopted in the main text that all agents have logarithmic utility, but I now suppose that a fraction $\zeta \in [0,1]$ of the entrepreneurs operate with the technology represented by the function $F(K,L) = A_1 K^\alpha L^{1-\alpha}$ and the remaining fraction operate with the technology represented by the function $G(K,L) = A_2 K^\alpha L^{1-\alpha}$, where $A_2 > A_1$. However, I assume that the parameters governing the agency friction are common across entrepreneurs.

The expressions $l(\lambda_L), \Pi(\lambda_L)$ and $S(\lambda_L)$ in Appendix B.3 remain of the same form but are now indexed by $j \in \{1,2\}$,

$$\begin{aligned} l_j(\lambda_L) &= [(1-\alpha)A_j]^{1/\alpha}\lambda_L^{-1/\alpha} \\ \Pi_j(\lambda_L) &= \alpha A_j^{1/\alpha}[(1-\alpha)/\lambda_L]^{1/\alpha-1} - \delta \\ S_j(\lambda_L) &= \frac{\Pi_j(\lambda_L) - \rho_S}{\sqrt{\rho}\phi\sigma}. \end{aligned} \tag{100}$$

Rearrangement implies that for any multiplier λ_L, we have

$$S_2(\lambda_L) = S_1(\lambda_L)(A_2/A_1)^{1/\alpha} + ((A_2/A_1)^{1/\alpha} - 1)(\rho_S + \delta)/(\sqrt{\rho}\phi\sigma). \tag{101}$$

Note that $\overline{\omega}$ is common to both types and so I will drop it from the following notation for brevity. The characterization of the stationary efficient allocation now proceeds much as in the proof of Proposition 3.1, except that we have to be careful about non-negativity restrictions, because it might be the case that one type of entrepreneur is not producing in the stationary efficient allocation. The average consumption and capital delegated to entrepreneurs of each type in the stationary distribution per unit of initial utility are again given by the expressions in (65), provided that we interpret $\overline{c}(S) = 1$ and $x(S) = 0$ for $S < 0$.

The planner once again places zero weight on the utility of all types of entrepreneurs, and so because types are private information, all agents receive the same level of initial utility u_0. The variable that adjusts until resources clear is the multiplier λ_L on the labor resource constraint, but just as in Proposition 3.1, it is convenient to write the resource constraint solely in terms of the variable S_1. Motivated by the equality (101), I now define the function

$$S_2^*(S_1) = S_1(A_2/A_1)^{1/\alpha} + ((A_2/A_1)^{1/\alpha} - 1)(\rho_S + \delta)/(\sqrt{\rho}\phi\sigma). \tag{102}$$

Output (net of depreciation) per unit of initial utility u_0 in the stationary allocation for a fixed λ_L and type $j \in \{1, 2\}$ is

$$J_j(S_j(\lambda_L)) := (S_j(\lambda_L)/\alpha + (\rho_S/\alpha + (1/\alpha - 1)\delta)/(\sqrt{\rho}\phi\sigma))\frac{\rho_D\overline{c}(S_j(\lambda_L))x(S_j(\lambda_L))}{\rho_D - \mu_c(S_j(\lambda_L))}.$$

and so the net output of each type of entrepreneur is

$$\begin{aligned}
J_1(S_1) &= (S_1\sqrt{\rho}\phi\sigma/\alpha + \rho_S/\alpha + (1/\alpha - 1)\delta)\overline{K}(S_1) \\
J_2(S_1) &= (S_2^*(S_1)\sqrt{\rho}\phi\sigma/\alpha + \rho_S/\alpha + (1/\alpha - 1)\delta)\overline{K}(S_2^*(S_1))
\end{aligned}$$

where I again used the abbreviation

$$\overline{K}(S_1) = \frac{\rho_D\overline{c}(S_1)x(S_1)}{(\rho_D - \mu_c(S_1))\sqrt{\rho}\phi\sigma}$$

Instead of equation (12), the equation characterizing the stationary distribution is

$$(1-\psi)(\zeta\overline{C}(S_1) + (1-\zeta)\overline{C}(S_2^*(S_1))) + \psi = (1-\psi)(\zeta J_1(S_1) + (1-\zeta)J_2(S_1)). \tag{103}$$

The analogue of Assumption 3.1 is then the following.

Assumption D.3. *There exists a solution $\hat{S}_1$ to the equation (103) such that:*

1. *The principal's value function in Section 2 with $S = S_2^*(\hat{S}_1)$ and $\tau_k = -\rho_D$ is finite-valued.*
2. *$\mu_c(S_2^*(\hat{S}_1), \overline{\omega}) < \rho_D$.*

Relative to the case considered in Proposition 3.1 in which there was a single type of entrepreneur, the determination of the aggregate capital stock now requires more algebra, because the fraction of aggregate capital assigned to each type of entrepreneur is endogenous. I first describe how all of the aggregate quantities (capital, labor, etc) depend on the solution $\hat{S}_1$ to equation (103). I therefore rewrite the system of equations in (100) so that all quantities are functions of S.

First, note that the multiplier λ_L and S are related according to $(S\sqrt{\rho}\phi\sigma + \rho_S + \delta)^\alpha = \alpha^\alpha A_1[(1-\alpha)/\lambda_L]^{1-\alpha}$, and so the multiplier and labor-capital ratio may be written as a function of S,

$$\begin{aligned}
\lambda_L = \lambda_L(S_1) &= \alpha^{\frac{\alpha}{1-\alpha}}(1-\alpha)A_1^{\frac{1}{1-\alpha}}(S_1\sqrt{\rho}\phi\sigma + \rho_S + \delta)^{-\frac{\alpha}{1-\alpha}} \\
&= \alpha^{\frac{\alpha}{1-\alpha}}(1-\alpha)A_2^{\frac{1}{1-\alpha}}(S_2^*(S_1)\sqrt{\rho}\phi\sigma + \rho_S + \delta)^{-\frac{\alpha}{1-\alpha}}
\end{aligned}$$

which are mutually consistent by (102). The labor-capital ratio for each type may then be written

$$\begin{aligned}
l_1(\lambda_L(S_1)) &= [(1-\alpha)A_1]^{1/\alpha}\lambda_L(S_1)^{-1/\alpha}\alpha^{-\frac{1}{1-\alpha}}(1-\alpha)^{-1/\alpha}A_1^{-\frac{1/\alpha}{1-\alpha}}(S_1\sqrt{\rho}\phi\sigma + \rho_S + \delta)^{\frac{1}{1-\alpha}} \\
&= (\alpha A_1)^{-\frac{1}{1-\alpha}}(S_1\sqrt{\rho}\phi\sigma + \rho_S + \delta)^{\frac{1}{1-\alpha}} \\
l_2(\lambda_L(S_1)) &= (\alpha A_2)^{-\frac{1}{1-\alpha}}(S_2^*(S_1)\sqrt{\rho}\phi\sigma + \rho_S + \delta)^{\frac{1}{1-\alpha}}.
\end{aligned}$$

For each $j \in \{1,2\}$, denote by K_j and L_j the aggregate amount of capital and labor used by entrepreneurs of type j. Using the fact that the labor-capital ratio L_j/K_j is constant within each type $j \in \{1,2\}$, the equations in (100) imply

$$\frac{L_1}{L_2} = \frac{K_1}{K_2}\frac{L_1/K_1}{L_2/K_2} = \frac{K_1}{K_2}\frac{l_1(\lambda_L)}{l_2(\lambda_L)}.$$

Combining this with $L = (L_1/L_2 + 1)L_2$, the capital employed by the second type is

$$K_2 = \frac{L/l_2(\lambda_L)}{(K_1/K_2)l_1(\lambda_L)/l_2(\lambda_L) + 1} = \frac{L}{(K_1/K_2)l_1(\lambda_L) + l_2(\lambda_L)}$$

which ultimately implies

$$K = (K_1/K_2 + 1)K_2 = \frac{(K_1/K_2 + 1)L}{(K_1/K_2)l_1(\lambda_L) + l_2(\lambda_L)}. \tag{104}$$

Finally, because all agents obtain the same level of lifetime utility, the ratio K_1/K_2 is

$$\frac{K_1}{K_2} = \frac{\zeta\overline{K}(S_1)}{(1-\zeta)\overline{K}(S_2^*(S_1))}. \tag{105}$$

Combining equation (104) with equation (105) then gives aggregate capital solely in terms of S

$$K(S_1) = \frac{(\zeta\overline{K}(S_1) + (1-\zeta)\overline{K}(S_2^*(S_1)))L}{\zeta\overline{K}(S_1)l_1(\lambda_L(S_1)) + (1-\zeta)\overline{K}(S_2^*(S_1))l_2(\lambda_L(S_1))}. \tag{106}$$

The above gives aggregate capital as a function of S. In terms of the function in (106), the amount of capital held by each type is

$$\begin{aligned} K_1(S_1) &= \frac{K(S_1)\zeta\overline{K}(S_1)}{\zeta\overline{K}(S_1) + (1-\zeta)\overline{K}(S_2^*(S_1))} \\ K_2(S_1) &= \frac{K(S_1)(1-\zeta)\overline{K}(S_2^*(S_1))}{\zeta\overline{K}(S_1) + (1-\zeta)\overline{K}(S_2^*(S_1))}. \end{aligned} \tag{107}$$

If $\hat{S}_1$ denotes a solution to equation (103), then the efficient level of the capital stock is $\hat{K} = K(\hat{S}_1)$ for the function $K(\cdot)$ given in (106), and for $j \in \{1,2\}$ the aggregate capital of type j is $\hat{K}_j = K_j(\hat{S}_1)$ for the functions $K_j(\cdot)$ given in (107).

Proposition D.5. *When Assumption D.3 is satisfied for some $\hat{S}_1$, an efficient stationary allocation exists in which the capital stock is $\hat{K} = K(\hat{S}_1)$ and the capital stock held by each type $j \in \{1,2\}$ is $\hat{K}_j = K_j(\hat{S}_1)$. The consumption of workers is constant over time, and the consumption $(c_t^j)_{t\geq 0}$ of entrepreneurs of type $j \in \{1,2\}$ satisfies*

$$dc_t^j = \mu_c(\hat{S}_j, \overline{\omega})c_t^j dt + \sigma_c(\hat{S}_j, \overline{\omega})c_t^j dB_t. \tag{108}$$

where $\hat{S}_2 = S_2^*(\hat{S}_1)$.

D.2.2 Decentralization

In this section I state and prove the analogue of Proposition 4.2 for the case of heterogeneous entrepreneurs. When productivity is observable and permanent the arguments are now essentially unchanged relative to the main text except that we must be careful about the appropriate analogue of the resource constraint (and the associated Assumption D.3).

I will distinguish type-specific quantities using an additional subscript $j \in \{1, 2\}$. For instance, the efficient value of the wedge on the bond for the type j entrepreneur will be denoted by $\hat{\nu}_j^B = \nu^B(\hat{S}_j, \overline{\omega})$, where $\hat{S}_j$ is given in Proposition D.5. I also write $\hat{x}_1 = x(\hat{S}_1, \overline{\omega})$ and $\hat{x}_2 = x(S_2(\hat{S}_1), \overline{\omega})$. I emphasize that κ_{E1}, κ_{E2} and κ_W represent the lifetime wealth of agents as a fraction of aggregate capital $\hat{K}$, not as a fraction of either $\hat{K}_1$ or $\hat{K}_2$. To state the decentralization more succinctly, I define some new notation. Denote by H the aggregate (after-tax) level of human wealth

$$H = (1-\psi)(\zeta h_{E1} + (1-\zeta)h_{E2}) + \psi h_W \tag{109}$$

and the associated efficient level of output net of depreciation is denoted

$$\hat{Y} = A_1 \hat{K}_1(\hat{S}_1)^\alpha \hat{L}_1(\hat{S}_1)^{1-\alpha} - \delta \hat{K}_1(\hat{S}_1) + A_2 \hat{K}_2(\hat{S}_1)^\alpha \hat{L}_2(\hat{S}_1)^{1-\alpha} - \delta \hat{K}_2(\hat{S}_1)$$

Proposition D.6. *When Assumption D.3 holds and $x(S_2(\hat{S}_1), \overline{\omega}) < 1$, the stationary efficient allocation can be implemented as an equilibrium of the form given in Definition 4.2 if the following conditions are satisfied:*

1. *The tax on profits is common across entrepreneurs at $\tau_\pi = 1 - \phi$.*
2. *The interest rate r and taxes τ_{sW} and $\{\tau_{iEj}, \tau_{Ij}\}_{j=1,2}$ are any values such that after-tax returns satisfy $r_{sW} = \rho$, $r_{sEj} = \rho - \hat{\nu}_j^B$, and $r_{bj} = \rho_S + \hat{\nu}_j^K - \hat{\nu}_j^B$.*
3. *The transfers and labor taxes are chosen such that κ_{E1}, κ_{E2} and κ_W satisfy*

$$1 = (1-\psi)\zeta \frac{\rho_D \kappa_{E1}}{\rho_D - \hat{\mu}_{c1}} \frac{\sqrt{\rho}\hat{x}_1}{\phi\sigma} + (1-\psi)(1-\zeta)\frac{\rho_D \kappa_{E2}}{\rho_D - \hat{\mu}_{c2}} \frac{\sqrt{\rho}\hat{x}_2}{\phi\sigma}$$
$$\kappa_W = \kappa_{E1} \max\left\{e^{-\hat{x}_1^2/2}, \hat{x}_1/\overline{\omega}\right\} = \kappa_{E2} \max\left\{e^{-\hat{x}_2^2/2}, \hat{x}_2/\overline{\omega}\right\}$$

   and the constant in the collateral constraint for the jth type is $\overline{\omega}_{\mathrm{d}j} = \iota^{-1} \max\left\{\hat{x}_j/\overline{\omega}, e^{-\hat{x}_j^2/2}\right\}$.

4. *The level of government debt is $D = \hat{Y}/\rho - \hat{K} - H$.*

Proof. The proof is almost identical to that of Proposition 4.2 and so I only highlight the differences. To see how the equations governing κ_{E1}, κ_{E2} and κ_W change, note that Lemma 4.1 implies that the constant defining the capital policy function is

$$\overline{k}_{\mathrm{d}j} = \frac{\sqrt{\rho}\hat{x}_j}{\phi\sigma} \tag{110}$$

Using (110), the capital market-clearing equation becomes

$$\hat{K} = (1-\psi)\zeta\frac{\rho_D\kappa_{E1}\hat{K}\overline{k}_{\mathrm{d}1}}{\rho_D-\hat{\mu}_{c1}} + (1-\psi)(1-\zeta)\frac{\rho_D\kappa_{E2}\hat{K}\overline{k}_{\mathrm{d}2}}{\rho_D-\hat{\mu}_{c2}} \tag{111}$$

which simplifies to the first of the two equations governing κ_{E1}, κ_{E2} and κ_W. The second equation governing these constants is the indifference relation between all agents, just as in the proof of Proposition D.6. The wealth of private agents is then found by subtracting human wealth, which rearranges to give the debt position of the government. □

References

S. Rao Aiyagari. Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting. *Journal of Political Economy*, 103(6):1158–1175, 1995. doi:10.1086/601445.

Stefania Albanesi. Optimal Taxation of Entrepreneurial Capital with Private Information. Working Paper 12419, National Bureau of Economic Research, August 2006.

Fernando Alvarez and Urban J. Jermann. Efficiency, Equilibrium, and Asset Pricing with Risk of Default. *Econometrica*, 68(4):775–797, 2000. doi:10.1111/1468-0262.00137.

George-Marios Angeletos. Uninsured Idiosyncratic Investment Risk and Aggregate Saving. *Review of Economic Dynamics*, 10(1):1–30, January 2007. doi:10.1016/j.red.2006.11.001.

George-Marios Angeletos and Vasia Panousi. Revisiting the supply side effects of government spending. *Journal of Monetary Economics*, 56(2):137–153, March 2009. doi:10.1016/j.jmoneco.2008.12.010.

Spencer Bastani and Daniel Waldenström. How Should Capital Be Taxed? *Journal of Economic Surveys*, 34(4):812–846, 2020. doi:10.1111/joes.12380.

Olivier J. Blanchard. Debt, Deficits, and Finite Horizons. *Journal of Political Economy*, 93(2): 223–247, April 1985. doi:10.1086/261297.

Robin Boadway and Kevin Spiritus. Optimal taxation of normal and excess returns to risky assets. *The Scandinavian Journal of Economics*, forthcoming, 2025. doi:10.1111/sjoe.12566.

Corina Boar and Matthew Knowles. Optimal taxation of risky entrepreneurial capital. *Journal of Public Economics*, 234:105100, June 2024. doi:10.1016/j.jpubeco.2024.105100.

Marco Cagetti and Mariacristina De Nardi. Entrepreneurship, Frictions, and Wealth. *Journal of Political Economy*, 114(5):835–870, October 2006. doi:10.1086/508032.

Christophe Chamley. Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives. *Econometrica*, 54(3):607–622, 1986. doi:10.2307/1911310.

V. V. Chari and Patrick J. Kehoe. Optimal Fiscal and Monetary Policy. In *Handbook of Macroeconomics*, volume 1, pages 1671–1745. Elsevier, January 1999. doi:10.1016/S1574-0048(99)10039-9.

Juan Carlos Conesa, Sagiri Kitao, and Dirk Krueger. Taxing Capital? Not a Bad Idea after All! *American Economic Review*, 99(1):25–48, March 2009. doi:10.1257/aer.99.1.25.

Julio Davila, Jay H. Hong, Per Krusell, and José-Víctor Ríos-Rull. Constrained Efficiency in the Neoclassical Growth Model With Uninsurable Idiosyncratic Shocks. *Econometrica*, 80(6):2431–2467, 2012. doi:10.3982/ECTA5989.

Jason DeBacker, Vasia Panousi, and Shanthi Ramnath. A Risky Venture: Income Dynamics among Pass-Through Business Owners. *American Economic Journal: Macroeconomics*, 15(1):444–474, January 2023. doi:10.1257/mac.20200134.

Sebastian Di Tella. Uncertainty Shocks and Balance Sheet Recessions. *Journal of Political Economy*, 125(6):2038–2081, December 2017. doi:10.1086/694290.

Sebastian Di Tella. Optimal Regulation of Financial Intermediaries. *American Economic Review*, 109(1):271–313, January 2019. doi:10.1257/aer.20161488.

Sebastian Di Tella and Yuliy Sannikov. Optimal Asset Management Contracts With Hidden Savings. *Econometrica*, 89(3):1099–1139, 2021. doi:10.3982/ECTA14929.

Evsey D. Domar and Richard A. Musgrave. Proportional Income Taxation and Risk-Taking. *The Quarterly Journal of Economics*, 58(3):388–422, 1944. doi:10.2307/1882847.

Sebastian Dyrda and Marcelo Pedroni. Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Income Risk. *The Review of Economic Studies*, 90(2):744–780, March 2023. doi:10.1093/restud/rdac031.

David Evans. Optimal Taxation with Persistent Idiosyncratic Investment Risk. Mimeo, 2014.

Emmanuel Farhi and Ivan Werning. Inequality and Social Discounting. *Journal of Political Economy*, 115(3):365–402, 2007. doi:10.1086/518741.

John Geanakoplos and Heracles M. Polemarchakis. Existence, Regularity, and Constrained Suboptimality of Competitive Allocations When the Asset Market Is Incomplete. Technical Report 764, Cowles Foundation for Research in Economics, Yale University, 1985.

Aart Gerritsen, Bas Jacobs, Kevin Spiritus, and Alexandra V Rusu. Optimal Taxation of Capital Income with Heterogeneous Rates of Return*. *The Economic Journal*, page ueae083, September 2024. doi:10.1093/ej/ueae083.

Mikhail Golosov and Aleh Tsyvinski. Policy Implications of Dynamic Public Finance. *Annual Review of Economics*, 7(1):147–171, 2015. doi:10.1146/annurev-economics-080614-115538.

Mikhail Golosov, Narayana Kocherlakota, and Aleh Tsyvinski. Optimal Indirect and Capital Taxation. *The Review of Economic Studies*, 70(3):569–587, 2003. doi:10.1111/1467-937X.00256.

Fatih Guvenen, Gueorgui Kambourov, Burhan Kuruscu, Sergio Ocampo, and Daphne Chen. Use It or Lose It: Efficiency and Redistributional Effects of Wealth Taxation. *The Quarterly Journal of Economics*, 138(2):835–894, May 2023. doi:10.1093/qje/qjac047.

Kenneth L. Judd. Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics*, 28(1):59–83, October 1985. doi:10.1016/0047-2727(85)90020-9.

Narayana R. Kocherlakota. Implications of Efficient Risk Sharing without Commitment. *The Review of Economic Studies*, 63(4):595–609, October 1996. doi:10.2307/2297795.

James A. Mirrlees. An Exploration in the Theory of Optimum Income Taxation. *The Review of Economic Studies*, 38(2):175–208, 1971. doi:10.2307/2296779.

Vasia Panousi. Capital Taxation with Entrepreneurial Risk. Technical Report 2010-56, Board of Governors of the Federal Reserve System, 2010.

Vasia Panousi and Catarina Reis. Optimal Capital Taxation with Idiosyncratic Investment Risk. Technical Report 2012–70, Board of Governors of the Federal Reserve System, 2012.

Vasia Panousi and Catarina Reis. A Unified Framework for Optimal Taxation with Undiversifiable Risk. *Macroeconomic Dynamics*, 25(6):1331–1345, September 2021. doi:10.1017/S1365100519000695.

Thomas Phelan. The Optimal Taxation of Business Owners, March 2023.

F. P. Ramsey. A Contribution to the Theory of Taxation. *The Economic Journal*, 37(145):47–61, 1927. doi:10.2307/2222721.

William P. Rogerson. Repeated Moral Hazard. *Econometrica*, pages 69–76, 1985. doi:10.2307/1911724.

Ali Shourideh. Optimal Taxation of Wealthy Individuals. Technical Report 261, Society for Economic Dynamics, 2013.

Matthew Smith, Danny Yagan, Owen Zidar, and Eric Zwick. Capitalists in the Twenty-first Century. *The Quarterly Journal of Economics*, 134(4):1675–1745, 2019. doi:10.1093/qje/qjz020.

Stefanie Stantcheva. Dynamic Taxation. *Annual Review of Economics*, 12(1):801–831, 2020. doi:10.1146/annurev-economics-100119-013035.

Ludwig Straub and Iván Werning. Positive Long-Run Capital Taxation: Chamley-Judd Revisited. *American Economic Review*, 110(1):86–119, January 2020. doi:10.1257/aer.20150210.