Essays on dynamic agency, inequality and optimal taxation

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Thomas Michael Phelan

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Christopher Phelan, Adviser

August 2018

Contents

List of Figures

Chapter 1

Introduction

The distributions of both income and wealth in the United States are heavily skewed to the right, with tails that may be well approximated by power laws. Further, business owners are disproportionately represented at the top of these distributions and are exposed to a high degree of idiosyncratic risk. Do these facts together suggest imperfect risk-sharing remediable through government policy? If so, how should a utilitarian government balance a concern for redistribution between business owners and workers with the need to provide incentives for both business creation and investment?

This thesis addresses these questions by characterizing efficient allocations and long-run inequality in two dynamic economies that separately analyze different factors affecting business output. The first chapter focuses on the role of *human* capital and the second on *physical* capital. In each case I assume that business ownership is subject to a dynamic agency problem, with the utility of each firm owner and the output of their firm depending upon actions observable only to themselves. To induce owners to increase output, their consumption must therefore depend upon the performance of their firm, and this limits the extent to which idiosyncratic risk may be shared across society. Within these environments I calculate the degree of long-run inequality consistent with maximizing average welfare, subject to the restrictions imposed by technological constraints and the presence of asymmetric information. I then explore how these allocations may be implemented when agents may trade assets in decentralized markets and taxes are imposed on various forms of income.

In the first chapter the output of each business depends solely upon the productivity (or human capital) of the owner and the amount of workers they employ. I assume this human capital grows over time in a manner that depends partly on luck and partly on the (unobservable) effort exerted by the owner. The unobservability of effort leads to the usual tension between risk-sharing and the preservation of incentives present in all agency models. However, in contrast with the standard model of dynamic agency, the effort of the agent here has *permanent* effects on firm output, and I show that this difference has important implications for both the efficient incidence of risk and long-run inequality.

Since marginal utility from consumption is diminishing, it becomes increasingly expensive to provide incentives to exert high effort as consumption rises. In a standard agency model, this increasing cost of motivating the agent eventually overwhelms the (unchanging) associated benefit and so agents with high consumption bear little to no risk. In contrast, when effort affects productivity growth, a countervailing force is present. Although it may be more expensive to motivate an owner with a history of high firm performance, the benefits of doing so are commensurably higher because such agents are more productive and employ a larger workforce. I show that in this environment with endogenous human capital formation, more productive entrepreneurs typically bear *more* risk and the associated stationary distributions of both income and firm size exhibit the right Pareto tails observed in the data. Finally, I show that despite the wide degree of inequality exhibited in the long-run distribution, the efficient allocation may be implemented with linear taxes if contracts in the private sector are sufficiently rich.

To isolate the role of human capital and the (novel) implications its presence has for risk-bearing, the first chapter omits physical capital entirely. As such the model is unable to analyze the optimal taxation of different forms of capital income. In the second chapter I complement the above exercise by characterizing efficient allocations in an environment in which dynamic agency affects the accumulation of *physical* rather than human capital.

I suppose that output depends upon physical capital that is subject to random shocks observable only to the owner, and that owners have the ability to both divert capital to private consumption and abscond with a fraction of assets. To preserve incentives to invest, the consumption of the owner must again depend upon the performance of their firm, leading to imperfect risk-sharing and inequality in the efficient allocation.

As with the case of human capital, the addition of physical capital causes both the dynamics of risk-sharing and implied degree of long-run inequality to differ from the case of the standard agency model. Since the owner may divert output to private consumption, the degree of idiosyncratic risk borne must be increasing in both the amount of capital delegated and the owner's marginal utility from consumption. I show that in the efficient allocation, owners with histories of high firm performance are assigned high levels of capital and consumption (and hence *low* marginal utility from consumption), causing the degree of risk in consumption growth to be constant across all business owners. I characterize stationary efficient allocations, show that the above forces again imply that the associated distributions of consumption and wealth exhibit the power law pattern observed in the data and provide closed-form expressions for the thickness of the upper tails.

The model of the second chapter is also sufficiently tractable to allow me to derive a number of comparative statics results illustrating how changes in technology and demographics affect efficient long-run inequality. As one example, I show that the degree of inequality is increasing in the number of workers per entrepreneur. The intuition for this result is simple. An increase in the number of workers per entrepreneur increases the benefit from accumulating physical capital. Due to the nature of the incentive problem, this increase in the capital stock must be accompanied by an increase in consumption risk, leading to an increase in long-run inequality. Similar comparative statics are derived for changes in technology that affect the relative importance of capital and labour in production.

Finally, I show that the stationary constrained-efficient allocation in the model of the second chapter may be implemented in a general equilibrium model using taxes on labour income, risk-free savings and business profits. Despite the high degree of inequality in the stationary distribution, the implementation of the efficient allocation requires only linear taxes. The tax on entrepreneurs' savings may be positive or negative, while the tax on business profits depends solely upon the degree of private information and is independent of all technological and demographic parameters.

Chapter 2

Efficient inequality with human capital formation

This paper develops a tractable model of dynamic moral hazard which endogenously generates thick right (Pareto) tails in both consumption and productivity in the efficient societal risk-sharing arrangement. In the model, agents of heterogeneous innate ability may either become entrepreneurs and exert hidden effort to improve their productivity, or become workers and be assigned to work for an entrepreneur. This paper attempts to inform the recent debate over the appropriate policy responses to rising inequality by characterizing constrained-efficient allocations in an environment capable of capturing stylized features of the data. I show how to numerically calculate efficient stationary distributions of consumption and human capital under general specification of the production technology and underlying distribution of abilities.

It has been well-documented that the distributions of income in the United States and other developed nations is skewed to the right, with a tail that may be well approximated by a Pareto distribution. For instance, Atkinson, Piketty and Saez (11) survey theoretical and empirical studies of top incomes and find that the density f of incomes in the United States in 2009 approximately satisfied

$$f(z) \sim \frac{C}{z^{1+\alpha}}$$

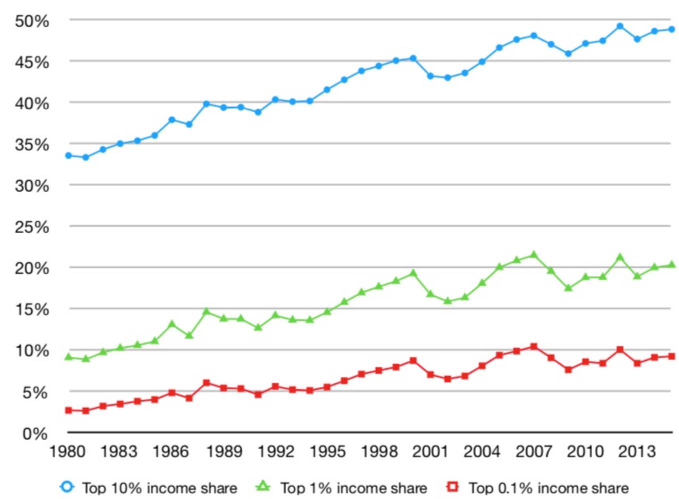

Figure 2.1: Income shares in the United States 1980-2015 including realized capital gains

as $z \to \infty$ for some constant C, with α approximately equal to 1.53. Further, numerous authors have observed a sharp rise in income inequality in the United States and other developed countries over the past few decades. Using data from the Statistics of Income division of the Internal Revenue Service, Piketty and Saez (57) estimate that the top 10% income share has risen from approximately 34% to nearly 50% over the period 1980-2015[1] and that the top 1% income share has risen from approximately 8% to over 20% over the same period. These trends are depicted in Figure 2.1.

Although the magnitudes of both levels and trends differ depending upon data sources and units of analysis, this significant rise in income inequality has also been observed in survey data[2] and in administrative-level data on labour incomes obtained from the Social Security Administration.[3]

What is perhaps less well known is that much of the recent increase in income inequality

[1]Income in Figure 2.1 is defined as market income including realized capital gains and excluding government transfers. Data and further notes available from `https://eml.berkeley.edu/~saez/#income`.

[2]See Heathcote, Perri and Violante (38), who integrate data from the Current Population Survey, the Consumer Expenditure Survey, the Panel Study of Income Dynamics and the Survey of Consumer Finances.

[3]See Guvenen, Kaplan and Song (33), Guvenen and Kaplan (32) and Song et al (34).

is due to the growing importance of business income. Smith, Yagan, Zidar and Zwick (69) and Cooper et al (18) use administrative tax data to document that business income now accounts for a greater share of the top 0.1% of income than both non-business capital income and wage income, and that the majority of this growth is due to private, pass-through entities (those not taxed at the firm level), such as partnerships and S-corporations. The importance of business income is reinforced by the observation of Guvenen and Kaplan (32) that the increase in inequality in the past decade observed in IRS data has not coincided with a similar rise in labour incomes recorded by the Social Security Administration.

Smith, Yagan, Zidar and Zwick (69) also note that top income earners are disproportionately business owners, with households in the top 1% of the income distribution fifty times more likely to receive partnership income than households in the bottom half of the income distribution. Further, these authors show business income depends on the owners' active participation rather than passive ownership by documenting that the unexpected death of an owner-manager gives rise to an average fall in profits of 54%. This may be viewed as evidence against the claim, forwarded by Piketty (56), that the recent rise in inequality is due to the increased importance of (non-human) capital income. In related work, de Backer, Panousi and Ramnath (20) document that business income is riskier than labor income: growth rates in business income are more dispersed than those for labor income and less of the dispersion is accounted for by observed heterogeneity.[4]

The appropriate policy response (if any) to this large degree of inequality and its increase over recent decades depends on the underlying economic mechanisms by which these phenomena arise. Guided largely by the findings of the seminal (static) Mirrlees (51) model, Diamond and Saez (21) survey the existing evidence on optimal taxation and argue that the top marginal tax rates may be as high as 80%, far higher than the current statutory maximum of 39.6%. However, the disproportionate role of business income amongst high

[4] In an error-components model, the authors find individual heterogeneity accounts for about 45% of the cross-sectional variance in labour income but only 25% of the cross-sectional variance in business income.

earners and the higher level of idiosyncratic risk associated with it suggests that appropriate policy responses must consider the role of entrepreneurial activity, a phenomenon absent from the static Mirrlees model.

This paper develops theory and numerical techniques to characterize constrained-efficient allocations in an environment capable of capturing the above stylized features. In doing so, I draw upon and extend the existing literature on both dynamic contracting and inequality. The recent contribution of Jones and Kim (42) characterizes competitive equilibria in an environment in which entrepreneurs are assumed unable to save and the growth rate of each agent's productivity depends partly on luck and partly on effort. I follow Jones and Kim (42) in my modeling of the underlying technology of human capital, but do not impose a particular market structure and instead require only that allocations respect underlying informational asymmetries.

Specifically, I will suppose that both initial ability (whether or not one may run a business) and effort exerted to improve productivity are privately observed by the agent. In order to preserve incentives to provide high effort, high realizations of productivity growth must be rewarded by either high future consumption or leisure, leading to imperfect risk-sharing and ex-post inequality in the efficient allocation. To analyze this dynamic moral hazard problem, I show how to extend the continuous-time techniques of Sannikov (63) to an environment in which productivity is endogenous and evolves stochastically over time. I then show how the solution to this principal-agent model may be used to derive implications for aggregate quantities and distributions by decomposing the problem of a benevolent planner facing a continuum of agents with random lifespans into a series of problems of dealing with each generation in isolation, given (shadow) prices for goods and labour which may then be varied until resources balance.

Shourideh (68), Ai, Kiku and Li (2) and Chapter 3 of this thesis all show how a Pareto distribution of consumption may emerge in the presence of asymmetric information with optimal contracting in private insurance markets. The main difference is the nature of the agency problem: in this paper entrepreneurs exert privately-observed effort to improve (publicly-observed) productivity, whereas the aforementioned papers allow entrepreneurs to divert delegated assets to private consumption.[5] I focus on the role of human capital governed by hidden effort both because it better captures the aforementioned importance of individual-specific characteristics for business income.

One virtue of this approach is that unlike Shourideh (68), Jones and Kim (42) and Ai, Kiku and Li (2), the efficient allocation in this chapter does not possess the (counterfactual) property that the Pareto exponents for income and firm size must coincide. To the best of my knowledge, the only other paper with this feature is Ai, Kiku and Li (3), where it arises due to a limited commitment problem in which the outside option of the manager is an increasing function of the firm's capital stock. Efficient allocations are quite different in these two economics. If the friction of the private markets is limited commitment on the part of managers then a government may trivially implement the first-best allocation by equalizing consumption amongst entrepreneurs (independent of firm performance) to eliminate the heterogeneity in outside options. In contrast, I will show that the efficient allocation in my model possess the stylized features noted above and only requires taxes for the implementation because of ex-ante heterogeneity between entrepreneurs and workers.

The outline of the paper is as follows: Section 2.1 conducts a review of the literature and places this paper within it; Section 2.2 characterizes the optimal contract between a single agent and principal; Section 2.3 extends this to an overlapping generations economy with a continuum of agents with heterogeneous ability; Section 2.4 characterizes and computes the stationary distributions of income in a number of example economies, Section 2.5 provides an implementation of the efficient allocation and Section 2.6 concludes. Details of the recursive techniques, numerical implementation and the welfare notions employed are relegated to the

[5] There are several other modeling differences: the welfare notion and lifecycle structure adopted here also differs from Shourideh (68), while Ai, Kiku and Li (2) pursue an equilibrium analysis.

appendix.

2.1 Literature review

A vast literature builds upon the seminal contribution of Mirrlees (51) to derive optimal taxes in informationally-constrained economies. Rothschild and Scheuer (60) conduct an optimal taxation exercise in an environment in which agents self-select into one of two industries and wages are endogenous. Scheuer (66) explicitly considers entrepreneurs and firm formation within a static model and allows for both pecuniary externalities across occupations and the possibility of occupational-specific taxation. Ales, Bellofatto, and Wang (6) adopt a span-of-control technology as in Rosen (61) and Lucas (48) and allow firm size to be endogenous. Scheuer and Werning (67) explores how optimal taxation policy must be altered the presence of 'superstar' effects in the form of assortative matching between individuals and firms. Ales and Sleet (8) conduct an optimal taxation exercise in a similar environment in which the planner has an explicit concern for the welfare of shareholders, and Ales, Kurnaz and Sleet (7) consider the role of skill-biased technical change. The above models are all static, and so in contrast with the current paper no agent bears risk or is subject to a moral hazard problem.

Stantcheva (65) conducts an optimal taxation exercise in a lifecycle model with risky human capital. However, the nature of the private-information problem is different than that considered in this paper. The wage of an agent at any moment is a function of their exogenously evolving stochastic ability and their stock of human capital. Ability is unobservable, while the stock of human capital is observable, evolves deterministically and depends on monetary costs expended by the agent. Since the law of motion of human capital is deterministic, investments may be inferred. The impediments to risk-sharing are therefore better thought of as hidden type rather than moral hazard. Similarly, in papers within the New Dynamic Public Finance literature, such as Golosov, Troshkin and Tsyvinski (31) and Farhi and Werning (25), the private information is in the form of hidden productivity and there is once again no moral hazard problem. Closer to the current paper are the works of Albanesi (5), Kapicka and Neira (45) and Best and Kleven (14), who conduct optimal taxation exercises in two-period economies with hidden effort. However, these papers do not derive implications of their framework for long-run distributions of consumption or income, nor is there an analysis of pecuniary externalities or the importance of the relative shadow values of entrepreneurial ability and worker labour.

In contrast with both the present and aforementioned papers, a related literature solves for welfare- or revenue-maximizing levels of taxation in heterogeneous-agent economies with exogenously incomplete markets. For example, Bruggemann (15) conducts an optimal taxation exercise within a model similar to Cagetti and De Nardi (16) that explicitly accounts for the role of entrepreneurs. As in Cagetti and De Nardi (16), entrepreneurship is a short-term decision, in the sense that agents may save only in a risk-free bond and decide their occupation anew each period, and the evolution of productivity is exogenous. Another closely related paper is Badel and Huggett (13), who perform an optimal taxation exercise in a model with human capital in an exogenously-incomplete markets model similar to Huggett, Ventura and Yaron (41).

For clarity, the first section of this paper analyses a principal-agent problem between a risk-averse entrepreneur and risk-neutral principal in partial equilibrium where the flow output of the entrepreneur depends solely upon their individual productivity. I characterize and compute the policy functions of the principal and show numerically that typically the agent will exert more effort and bear more risk when more productive. In the subsequent overlapping generations economy with a continuum of agents, the productivity of any entrepreneur will depend both on their productivity *and* the endogenously determined shadow price of labour, as the latter determines the societal (resource) cost of assigning workers to entrepreneurs. I view this as complementary to the above contributions, as I have a rich dynamic moral hazard component, and accordingly simplify the ex-ante heterogeneity for tractability.

2.2 Dynamic principal-agent model

For ease of exposition, I will first proceed in partial equilibrium and characterize the optimal contract between a risk-averse agent operating a risky technology and a risk-neutral principal who may trade at exogenously given prices. In the following section I will show how the problem of a benevolent planner in an overlapping generations economy may be decomposed into a series of principal-agent problems of the above form.

Individual preferences and technology

Time is indefinite and continuous. The economy consists of a single risk-averse agent and a risk-neutral principal, both of whom live forever. At any moment in time the agent may consume a flow amount c of a single good and take an action $e \in \mathcal{E}$ where $\mathcal{E} \subseteq [0,1]$. The agent has preferences over stochastic sequences of consumption $c = (c_t)_{t \geq 0}$ and effort $e = (e_t)_{t \geq 0}$ given by

$$U(c,e) = \int_0^\infty e^{-\rho t} \mathbb{E}[u(c_t, e_t)] dt \tag{2.1}$$

where the instantaneous flow utility from consumption and effort is

$$u(c,e) = \frac{c^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma} \tag{2.2}$$

for some constants $\gamma > 1$ and $\alpha > 0$. Throughout I will assume that the set $\mathcal{E}$ contains the point $e = 0$, which will be interpreted as 'retirement'. The principal will be assumed able to observe retirement, but not be able to distinguish between all other actions.

At any point in time the agent is associated with a scalar variable θ referred to as *productivity*. An agent of productivity θ inelastically produces a flow of θ units of output per time independently of their actions. The consumption and output produced by the agent are observable, while the productivity of the agent begins at the level θ_0 and evolves stochastically over time in a manner depending upon their effort. Specifically, there exists a stochastic process $Z = (Z_t)_{t \geq 0}$ defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ distributed according to standard Brownian motion, such that if the agent chooses effort levels according to a stochastic process $(e_t)_{t\geq 0}$ satisfying $e_t \in \mathcal{E}$ for all $t \geq 0$, then productivity follows the law of motion,

$$d\theta_t = \iota(e_t)1_{e_t\neq 0}\theta_t dt + \sigma 1_{e_t\neq 0}\theta_t dZ_t \tag{2.3}$$

where $\sigma > 0$ and $\iota : \mathcal{E} \to \mathbb{R}$ is given by $\iota(e) := \bar{\iota}e^{\eta} + \underline{\iota}$ for some $\bar{\iota}, \underline{\iota}, \eta > 0$. As (2.3) indicates, the productivity of the agent stops evolving upon retirement. Finally, the principal is risk-neutral and discounts at the rate of time preferences of the agent and so their preferences over output and consumption are

$$U^P(c,e) = \int_0^{\infty} e^{-\rho t}\mathbb{E}[\theta_t - c_t]dt. \tag{2.4}$$

Allocations and incentive compatibility

The actions taken by the both the agent and principal at any date may be an arbitrary function of the output observed up until that time. The following definition formalizes this mathematically.

Definition 2.2.1 (Allocations, strategies and continuation utility)**.** An allocation chosen by the principal consists of a pair of $\mathcal{F}$-adapted processes[6] $(C, e^P) = (c_t, e_t^P)_{t\geq 0}$, while an agent's strategy consists of a single $\mathcal{F}$-adapted process $e = (e_t)_{t\geq 0}$. For any allocation $(c_t, e_t^P)_{t\geq 0}$ and agent strategy $(e_t)_{t\geq 0}$, the continuation utility of the agent is given by the stochastic process $W \equiv W^{c,e}$ defined by

$$W_t := \rho \int_t^{\infty} e^{-\rho[s-t]}\mathbb{E}\left[u(c_s, e_s)\right| \mathcal{F}_t]\, ds \tag{2.5}$$

for all $t \geq 0$ almost surely.

Since the effort levels taken by the agent are private information, the principal must

[6]A sequence $(x_t)_{t\geq 0}$ is $\mathcal{F}$- adapted if and only if there exist $(\tilde{x}_t)_{t\geq 0}$ such that $\tilde{x}_t : C([0,t]) \to \mathbb{R}$ for each $t \geq 0$ and $x_t = \tilde{x}_t((Z_s)_{0\leq s\leq t})$ almost surely, for all $t \geq 0$, where for each $t \geq 0$ $\tilde{x}_t$ is measurable w.r.t. the σ-algebra generated by evaluation maps up until time t.

restrict attention to allocations that are incentive compatible. An allocation is incentive compatible if the agent wishes to adhere to the effort recommendations of the principal after every history. To formalize this notion, note that the utility of the agent under a given allocation when adhering to an arbitrary strategy amounts to evaluating the allocation under a *change of measure* associated with his own strategy. That is, the allocation of the principal specifies consumption as a function of every finite history of output, and when choosing a strategy the agent understands how his actions change the probability of each history and weights them accordingly. For any strategy $e = (e_t)_{t\geq 0}$ I will write P^e for this induced measure[7] and $\mathbb{E}^e$ the associated expectation. The utility of an agent confronted with an allocation $(c_t, e_t^P)_{t\geq 0}$ when adhering to a strategy $(e_t)_{t\geq 0}$ is then

$$U(c,e) := \rho \int_0^\infty e^{-\rho t} \mathbb{E}^e \left[u(c_t, e_t) \right] dt. \tag{2.6}$$

The formal definition of incentive compatibility is then the following.

Definition 2.2.2. An allocation $(c_t, e_t^P)_{t\geq 0}$ is incentive compatible if

$$U(c, e^P) \geq U(c,e) \tag{2.7}$$

for all strategies e. The set of all incentive compatible allocations will be denoted $\mathcal{A}^{IC}$.

The principal's problem is then the following. It is indexed by the initial productivity θ of the agent and the minimal level of promised utility W necessary for their participation.

Definition 2.2.3. Given the utility W from the outside option and productivity θ the problem of the principal is

$$\begin{aligned} V(W,\theta) &= \max_{(c,e)\in\mathcal{A}^{IC}} \int_0^\infty e^{-\rho t} \mathbb{E}[\theta_t - c_t] dt. \\ W &= \int_0^\infty e^{-\rho t} \mathbb{E}[u(c_t, e_t)] dt \\ d\theta_t &= \iota(e_t)\theta_t dt + \sigma(e_t)\theta_t dZ_t, \quad \theta_0 = \theta. \end{aligned}$$

[7] The formal construction of these measures is given in Definition A.1.2 in the appendix.

Since the increments of Brownian motion are independent, the principal's problem is naturally recursive in the variables W and θ. However, the principal's problem is difficult because the incentive compatibility constraints in Definition 2.2.2 implicitly involve a double infinity of constraints: there are infinitely many possible output paths and for each such path the agent has a continuum of actions from which to choose. However, following the techniques of Sannikov (63), I will show that incentive compatibility of an allocation is equivalent to the requirement that the associated continuation utility given in (2.5) follows a diffusion process with 'sufficiently high' volatility. Before proceeding to the general representation theorem, first note that in the deterministic case with $\sigma = 0$, the conditional expectations in (2.5) are redundant and continuation utility has a classical derivative given by

$$\frac{dW_t}{dt} = \rho(W_t - u(c_t, e_t)). \tag{2.8}$$

In the presence of Brownian shocks, one cannot simply differentiate promised utility path-by-path. However, using the martingale techniques of Sannikov (63), one may show the evolution of promised utility is simply the drift term (2.8) plus a stochastic integral with respect to the underlying noise. The proof is standard and therefore omitted.

Proposition 2.2.1. *Given any allocation $(c_t, e_t^P)_{t\geq 0}$ and agent strategy $(e_t)_{t\geq 0}$, there exists an ($\mathcal{F}$-predictable) process $S = S^{e^P,C}$ such that continuation utility W_t may be written*

$$W_t = W_0 + \rho \int_0^t (W_s - u(c_s, e_s))ds + \rho\sigma \int_0^t S_s dZ_s^e \tag{2.9}$$

for all $t \geq 0$ almost surely, where $Z_t^e(\omega) := \omega(t) - \sigma^{-1} \int_0^t \iota(\tilde{e}_s((\omega(s'))_{0\leq s'\leq s}))ds$.

The content of 2.2.1 is that the stochastic increments to continuation utility may be viewed as the increments of Z^e (the shocks to output from perspective of agent) multiplied by the *sensitivity* of promised utility to output. Such a representation exists for any allocation and strategy regardless of whether or not it is incentive compatible.

Now recall that in discrete-time environments, one may often establish a form of the one-shot deviation principle, which asserts that incentive compatibility is assured if one discourages temporary deviations from the recommended effort process. The representation obtained in 2.2.1 allows for the following continuous-time analogue of this one-shot deviation principle.

Proposition 2.2.2. *An allocation $(c_t, e_t^P)_{t\geq 0}$ is incentive compatible if and only if the sensitivity process S given in 2.2.1 satisfies*

$$S_t\iota(e_t^P) + u(c_t, e_t^P) \geq S_t\iota(e) + u(c_t, e) \tag{2.10}$$

for all $e \in \mathcal{E}$ and $t \geq 0$. The optimal choice of S_t for the planner is then $S_t = \kappa(e_t)u(C_t, e)$ when $e > 0$ and $S_t \equiv 0$ when $e = 0$, where $\kappa(e) := \alpha e^{1-\eta}[\eta\bar{\iota}(1-e)]^{-1}$.

Proof. For any strategy e and time $t \geq 0$, define a random variable $\hat{V}_t^e$ by

$$\hat{V}_t^e := \rho \int_0^t e^{-\rho s} u(c_s, e_s)ds + e^{-\rho t}W_t$$

where W_t represents the continuation utility if the agent adheres to e^P after t. From the definition of dZ^e given in 2.2.1 we have $dZ_t^{e^P} = dZ_t^e - \sigma^{-1}[\iota(e_t^P) - \iota(e_t)]$ and so by the expression (2.9),

$$\begin{aligned}
d\hat{V}_t^e &= \rho e^{-\rho t}u(c_t, e_t)dt + d\big(e^{-\rho t}W_t\big) \\
&= \rho u(c_t, e_t)dt - \rho W_t dt + \rho\Big[(W_t - u(c_t, e_t^P))dt + \sigma S_t dZ_t^{e^P}\Big] \\
&= \rho e^{-\rho t}\big(u(c_t, e_t) - u(c_t, e_t^P) + S_t[\iota(e_t) - \iota(e_t^P)]dt\big) + \rho\sigma e^{-\rho t}S_t dZ_t^e.
\end{aligned}$$

Using $\mathbb{E}^e\Big[\int_0^t e^{-\rho s}S_s dZ_s^e\Big] = 0$ for all $t \geq 0$ and effort process e, it follows that

$$\mathbb{E}^e\Big[\hat{V}_t^e\Big] = \hat{V}_0^e + \rho\mathbb{E}^e\Bigg[\int_0^t e^{-\rho s}\big(S_s\iota(e_s) + u(c_s, e_s) - [S_s\iota(e_s^P) + u(c_s, e_s^P)]\big)ds\Bigg]. \tag{2.11}$$

Since the expected utility of the agent is exactly $\mathbb{E}^e\Big[\lim_{t\to\infty}\hat{V}_t^e\Big]$, an effort process e is incentive compatible if and only if it maximises the integral in (2.11) almost surely for all

$t \geq 0$, which gives the result. $\square$

2.2.2 shows that incentive compatibility is equivalent to the responsiveness of promised utility to output being sufficiently large to outweigh the marginal benefit of a deviation. The quantity $\kappa(e)$ may be interpreted as the ratio of the benefits of deviating to the effect of a deviation on output, all as a fraction of flow utility. As expected, incentives must be more high-powered (promised utility more volatile) when deviations are hard to detect or the benefits of deviation are large. 2.2.2 also illustrates the tractability gained by the continuous-time formulation. In a discrete-time model, the principal must choose a different level of promised utility for every realization of uncertainty. In contrast, when output has continuous paths, the principal need only specify the marginal responsiveness of promised utility to output and so at any moment the choice of recommended effort and consumption determines the minimum level of volatility in promised utility necessary for incentive compatibility.

This allows us to recast the problem of the principal as an optimal control problem in the states θ and W, with flow profits as the objective, the law of motion of θ exogenously given by (2.3) and the law of motion of W given by Propositions 2.2.1 and 2.2.2. Standard results from the theory of of continuous-time optimal control then imply that the payoff to the principal as a function W and θ solves a partial differential equation. Further, the log-linearity of the evolution of productivity together with the homotheticity of preferences allow for the following reduction to a single state variable.

Lemma 2.2.3. *For all U and θ we have $V(U,\theta) = V(U\theta^{\gamma-1}, 1)\theta$ and the policy functions of the planner are functions of $U\theta^{\gamma-1}$.*

Proof. The key observation here is that incentive compatibility is unaffected if we scale consumption in every history by the same scalar. Therefore, for any U, θ and $\mu > 0$, the values of the two programs

$$\begin{aligned} &\max_{(c,e)\in\mathcal{A}^{IC}} \int_0^\infty e^{-\rho t}\mathbb{E}[\theta_t - c_t]dt & &\mu \max_{(\bar{c},e)\in\mathcal{A}^{IC}} \int_0^\infty e^{-\rho t}\mathbb{E}[\bar{\theta}_t - \bar{c}_t]dt \\ &U = \int_0^\infty e^{-\rho t}\mathbb{E}\left[u(c_t,e_t)\right]dt & &U\mu^{\gamma-1} = \int_0^\infty e^{-\rho t}\mathbb{E}\left[u(\bar{c}_t,e_t)\right]dt \\ &d\theta_t = \iota(e_t)\theta_t dt + \sigma(e_t)\theta_t dZ_t,\ \theta_0 = \mu\theta & &d\bar{\theta}_t = \iota(e_t)\bar{\theta}_t dt + \sigma(e_t)\bar{\theta}_t dZ_t,\ \bar{\theta}_0 = \theta \end{aligned}$$

coincide, as can be seen by replacing θ_t with $\mu\bar{\theta}_t$ and c_t with $\mu\bar{c}_t$. The program on the left is $V(U,\mu\theta)$ and that on the right is $\mu V(U\mu^{\gamma-1},\theta)$. $\square$

In the standard agency model of Sannikov (63), flow output depends only on current actions and promised utility is sufficient to act as a state variable. This is no longer true when actions have persistent effects on flow output as they do here. However, 2.2.3 shows that the principal's choices depend only on promised utility *per unit of output*, and suggests the following more parsimonious choice of state variable.

Definition 2.2.4. Given promised utility U and productivity θ, define normalized promised utility u by

$$u := [(1-\gamma)U]^{\frac{1}{1-\gamma}}\theta^{-1}$$

and the normalized payoff function v by

$$v(u) := V\left(\frac{u^{1-\gamma}}{1-\gamma}, 1\right).$$

Note that for fixed productivity, normalized promised utility u is simply a monotonic transformation of utility (it may be interpreted as utility in consumption units). 2.2.3 shows that optimal choices of the planner are functions only of normalized promised utility. Similar observations are made in Ai, Kiku and Li (2) and He (35), where the agency problem involves hidden diversion of resources rather than hidden effort.

Prior to formulating the problem in a manner suitable to computation, I will first determine from first principles the payoff to the principal when u is either zero or extremely high.

To this end I will first characterize the value and policy functions associated with the (sub-optimal) allocations in which the principal recommends a given effort level for the entirety of an agent's life but is unconstrained in their choice of consumption. I will denote these value functions by $v(u; e)$ for $e \in \mathcal{E}$ and refer to them as restricted-action value functions. The following shows such value functions admit a closed-form solution.

Lemma 2.2.4. *For each effort level $e \in \mathcal{E}$ and activity $e \in E$, the value function $v(u; e)$ and consumption function $c(u; e)$ are of the form*

$$v(u, e) = \frac{\rho}{\rho - \iota(e)} - \chi_1(e)u \qquad\qquad c(u, e) = \chi_2(e)u$$

for some constants $\chi_1(e)$ and $\chi_2(e)$ monotonically increasing in e and given in the appendix. Consequently, the policy function e_{rest} associated with the optimization problem $v_{\text{rest}}(u) := \max_{e \in \mathcal{E}} v(u, e)$ is decreasing in u, and the true value function satisfies

$$\lim_{u \to 0} v(u) = \frac{\rho}{\rho - \iota(\overline{e})}. \tag{2.12}$$

Proof sketch. The form of the value and consumption functions follows from guess-and-verify. The details are given in the appendix. The final claim follows from the form of $v(u, \overline{e})$ together with the obvious inequality $v(u) \leq \rho/(\rho - \iota(\overline{e}))$ for all $u \geq 0$. $\square$

The restricted-action value functions provide some insight into how optimal effort recommendations vary with normalized promised utility. The coefficients of both constant terms and the negative of the linear term in the value function are increasing in the effort recommended to the agent, which captures the fact that higher actions imply a higher growth rate of output but are also more costly. In addition, the limiting value of v given in 2.2.4 shows that the restricted value function associated with the highest action approximates the value of v near zero, so that the welfare loss from adhering to the restricted-action allocation (rather than the true efficient allocation) falls to zero as productivity rises.

Combining the above reasoning with standard results from continuous-time dynamic programming then gives the following ordinary differential equation suitable for computation. The proof amounts to first deriving a partial differential equation for $V(U,\theta)$ and then using the homogeneity in 2.2.3 to simplify, and is relegated to the appendix.

Theorem 2.2.5. *The normalized value function of the principal is a solution to the ODE*

$$0 = \max_{\substack{c\geq 0\\ e\in\mathcal{E}}} \rho(z(e)-c) + (\iota(e)-\rho)v(u) + \left(\frac{\rho}{1-\gamma}[u - c^{1-\gamma}u^{\gamma}(1-e)^{-\alpha}] - \iota(e)u\right)v'(u)$$
$$+\frac{\gamma(\rho\sigma\kappa(e))^2}{2}\left(\frac{c^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right)^2 u^{2\gamma-1}v'(u) + \frac{\sigma^2}{2}\left(\rho\kappa(e)\frac{c^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma} - u^{1-\gamma}\right)^2 u^{2\gamma}v''(u).$$

subject to the boundary and smooth-pasting conditions

$$v(0) = \frac{\rho}{\rho-\iota} \qquad v(\overline{u}) = v_{ret}(\overline{u}), \qquad v'(\overline{u}) = v'_{ret}(\overline{u})$$

where $v_{ret}(u) = \underline{z} - u$ is the flow profit per unit of θ for retiring the agent.

The solution to the differential equation given in 2.2.5 does not appear to possess a closed-form solution and numerical methods are necessary for its solution. Such numerical techniques are outlined in the appendix.

The fact that e_{rest} given in 2.2.4 is decreasing suggests that the principal will recommend high actions to individuals with *low* levels of normalized promised utility. Since individuals exerting high effort have greater incentives to deviate, the form of the restricted value function suggests that these are precisely the agents who will bear the most risk. This implies that agents bear high risk when their productivity is high *relative* to their promised utility. In the restricted-action case, promised utility typically falls over time, while productivity typically grows. It follows that it is those agents with the highest levels of promised utility are typically those with the highest levels of productivity, and hence low levels of normalized promised utility, and it is the unlucky, low promised utility agents who are retired.

This contrasts with the environments in the the seminal papers of Phelan and Townsend

(55) and Sannikov (63), in which productivity is fixed over time and the agent controls the flow of output instead of its growth. In those environments agents are typically retired when their promised utility is large because such agents are too costly to motivate. Although it remains the case here that the cost of motivating the agents increases with promised utility (simply because of diminishing marginal utility from consumption) there is also an offsetting effect: agents in the above model have high levels of promised utility precisely because their productivity has grown. The cost of motivating the agent may have grown, but so too has the benefit, and the above analysis shows that the ratio $u := [(1-\gamma)U]^{\frac{1}{1-\gamma}}\theta^{-1}$ is the key variable. The desire to smooth consumption implies promised utility typically remains stable or falls over time, whereas productivity grows exponentially on average and so normalized promised utility typically falls over time. It follows that richer and older agents will usually bear more risk. The second difference is the nature of the income effect in the preferences: the principal-agent literature usually adopts additively-separable preferences, whereas the above are multiplicatively separable for consistency with the macroeconomic literature and the positive literature that motivated the current normative analysis.

Since none of the above restricted-action value functions are strictly optimal the true value function must be calculated numerically. I will now calculate an example of the true value functions and policy functions with the following parameters:

$$\begin{array}{lllll} \rho = 0.09 & \gamma = 2 & \alpha = 0.5 & \sigma = 0.29 & z = 1 \\ \bar{\iota} = 0.07 & \underline{\iota} = 0.0 & \eta = 0.5 & \underline{e} = 0.0 & \bar{e} = 0.6. \end{array}$$

For this illustrative example, the parameters η and $\bar{e}$ are arbitrary, the parameters γ and α are within the standard range employed in the literature, the discount rate ρ is taken directly from Jones and Kim (42) (it incorporates the probability of death there), and the σ and $\bar{\iota}$ are chosen to match average growth of high earners over the period 1980-2015 as recorded in Jones and Kim (42).

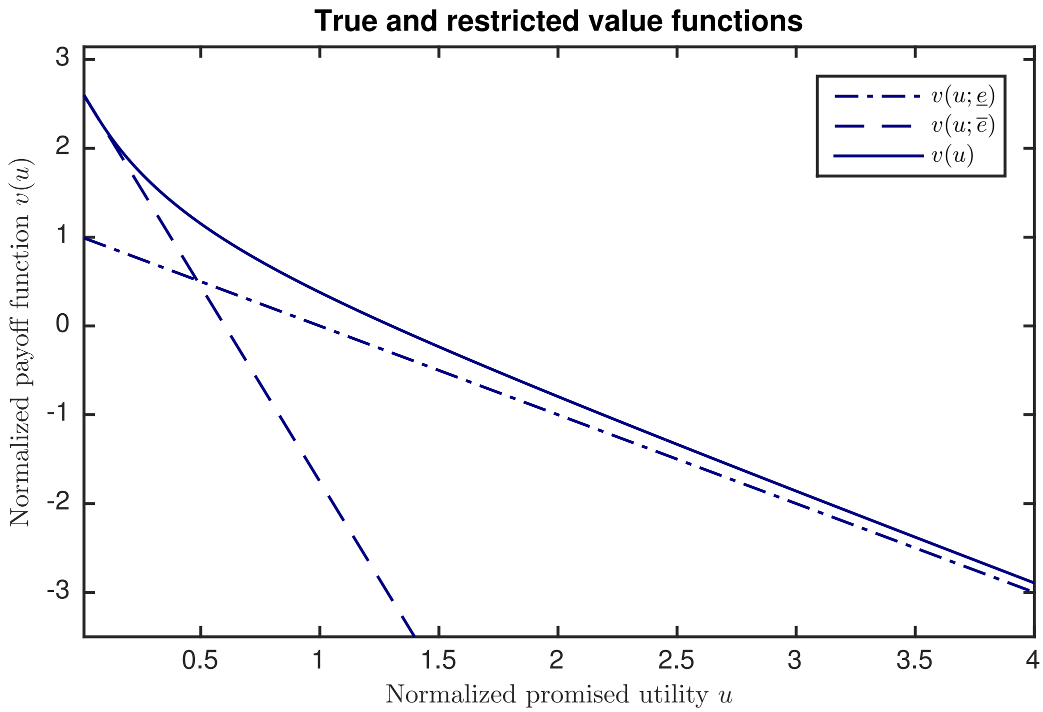

Figure 2.2: Value function and restricted value functions

Figure 2.2 plots the true value function alongside the two restricted-action value functions associated with the highest and lowest actions. Obviously the true value function must lie above the restricted-action value functions pointwise. Further, as suggested by the above discussion, the true value function is well-approximated by the restricted-action value function associated with the highest action for low values of normalized promised utility. Similarly, Figure 2.3 depicts the true consumption function alongside the two restricted-action consumption functions associated with the highest and lowest actions. As with the case of the value functions discussed above, the true consumption function is well-approximated by the restricted-action consumption function associated with the highest action for small values of normalized promised utility. Finally, Figure 2.4 depicts the recommended effort as a function of normalized promised utility. Again, as suggested by the above comments, the function $e(\cdot)$ is decreasing.

The value functions and policy functions completely solve the problem of the principal.

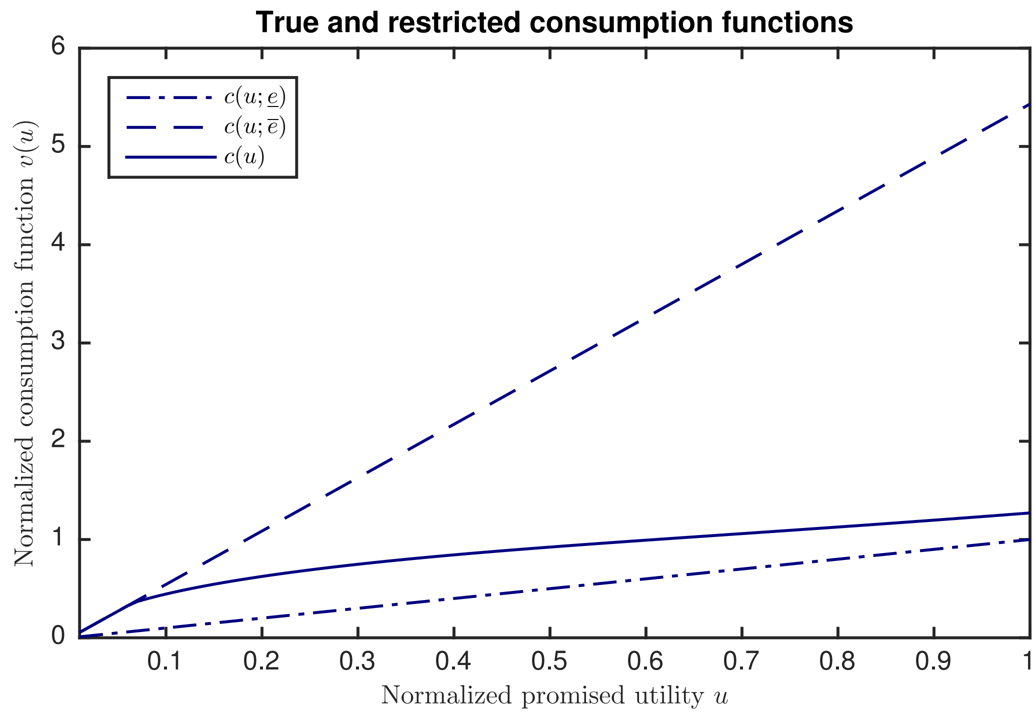

Figure 2.3: Consumption function and restricted consumption functions

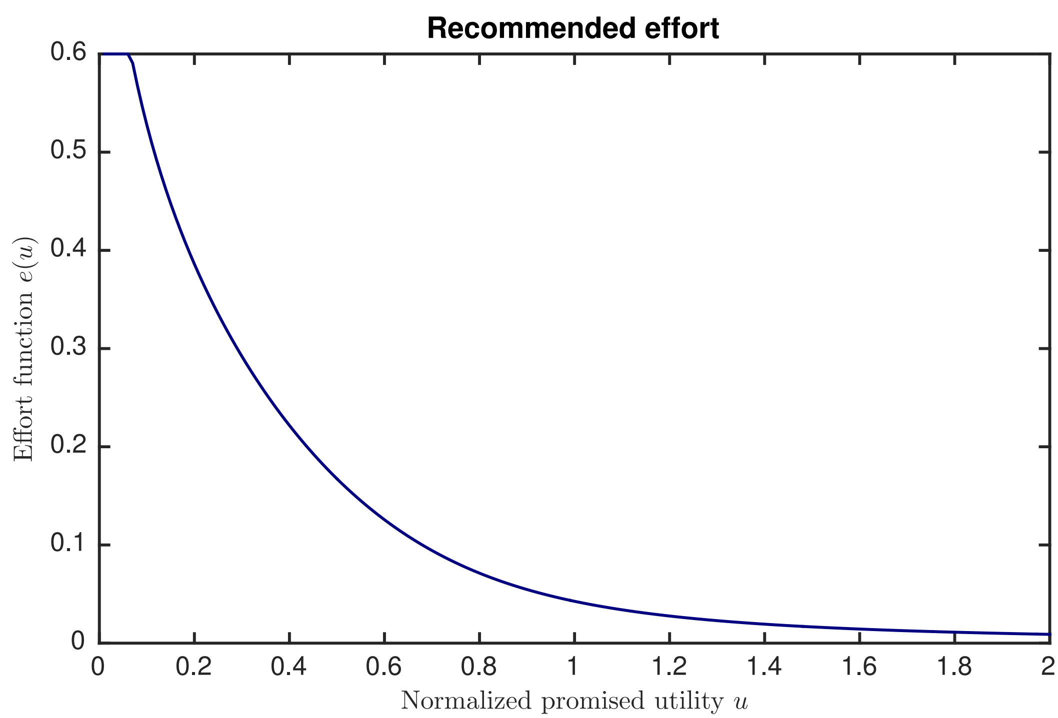

Figure 2.4: Optimal effort

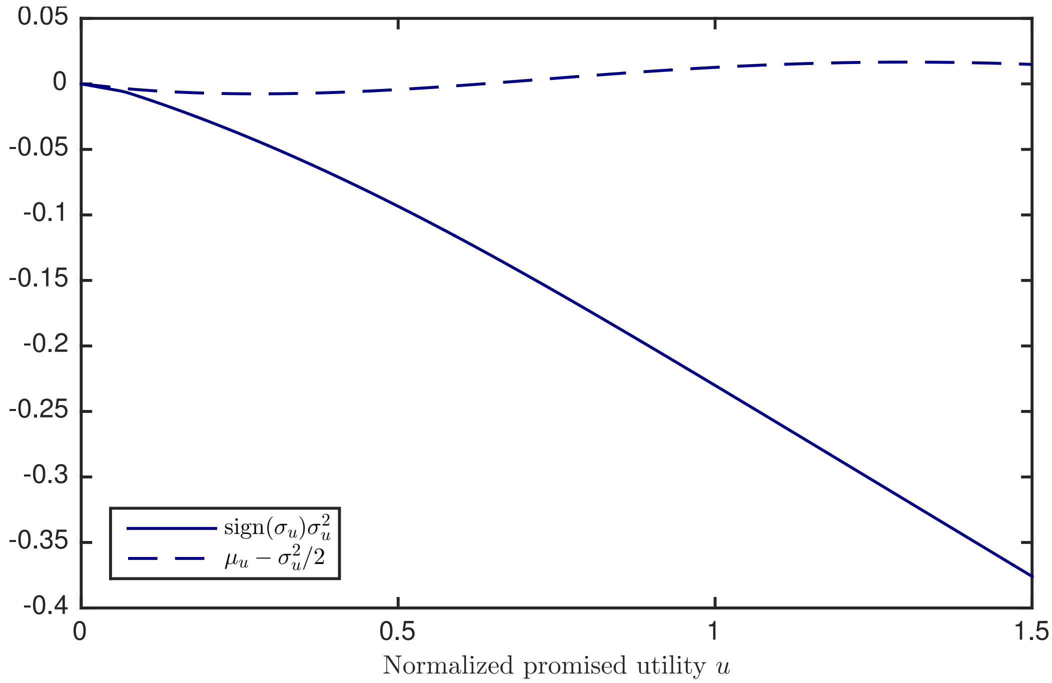

Figure 2.5: Drift and diffusion of normalized promised utility

However, the above graphs do not give any insight into the pathwise properties of efficient allocations, such as where the agent spends most of their time, or what a typical realization of the efficient contract looks like. For this we first determine the law of motion of normalized promised utility.

Lemma 2.2.6. *Normalized promised utility follows the diffusion process* $du_t = \mu_u(u)dt + \sigma_u(u)dB_t$*, where the drift and diffusion terms* μ_u *and* σ_u *are given by*

$$\mu_u(u) = \left(\frac{\rho(1-U(u))}{1-\gamma} - \iota(u) + \sigma^2\left[\frac{\gamma}{2}\left(\frac{\rho\kappa(u)U(u)}{1-\gamma}\right)^2 - \frac{\rho\kappa(u)U(u)}{1-\gamma} + 1\right]\right)udt$$
$$\sigma_u(u) = \sigma\left[\frac{\rho\kappa(u)U(u)}{1-\gamma} - 1\right]udB_t$$

and I have abbreviated $U(u) := (c(u)/u)^{1-\gamma}(1-e(u))^{-\alpha}$.

Figure 2.5 plots the diffusion σ_u^2 and the function $\mu_u - \sigma_u^2/2$, as the latter is the change in u over a period of length dt in the absence of shocks. Note that $\mu_u - \sigma_u^2/2$ may be positive or negative, but will be negative near the origin. This suggests that if the agent begins with a low level of normalized promised utility, their normalized promised utility will typically fall over time, whereas if the agent begins with a high level of normalized promised utility, their normalized promised utility may rise over time until they are retired. This observation will be important for the numerical comparative statics results derived later.

Before proceeding to the overlapping generations economy, it is instructive to compare the dynamics of consumption and risk-sharing in the optimal contract with those that obtain in economies with exogenously incomplete markets. This will serve to both highlight the forces that lead to the emergence of a power law in consumption and to explain how these forces are affected by the presence of incentive constraints and the opportunities for insurance.

The specification of technology adopted in this paper is identical to that in Jones and Kim (42), where (in contrast with the current paper) markets are exogenously incomplete and consumption simply coincides with income. Homotheticity of preferences and the exponential growth of human capital ensure that the costs and benefits of increased effort scale with current income, and it turns out that the optimal effort choice of the agent is independent of their history of income growth. The power law for income in the stationary distribution then follows from the fact that each agent experiences exponential growth in income for an exponential period of time. This result does not depend upon assumptions on the cross-sectional distribution of shocks and is even true when growth rates are deterministic: if consumption evolves according to $C'(t) = \mu_C C(t)$ for some $\mu_C > 0$ and the agent dies at rate ρ_D, then consumption C and age T in the stationary distribution satisfy

$$P(C \geq x) = P(T \geq \mu_C^{-1} \ln(x/C_0)) = e^{-\rho_D \mu_C^{-1} \ln(x/C_0)} = C_0^{\rho_D/\mu_U} x^{-\rho_D/\mu_C}$$

which is exactly a power law. The presence of idiosyncratic shocks to growth rates will (in general) alter the thickness of the tail of the stationary distribution but they are not strictly necessary for the emergence of a power law.

In contrast with Jones and Kim (42), this paper does not impose a particular market structure and instead supposes that insurance opportunities are inhibited only by the presence of asymmetric information. Section 2.3 of this paper shows that in this environment the stationary distributions of both human capital and consumption continue to exhibit the power law pattern described above. At an abstract level, the mechanism by which this occurs is the same as that given above since both consumption and human capital in this model both exhibit (approximately) exponential growth.

However, the determinants of this exponential growth are quite different across the two models. In contrast with the case of exogenously incomplete markets, the dynamics of consumption and human capital need not coincide with one another or be common across agents. 2.2.2 shows that the volatility of the the agent's consumption is an increasing function of the effort recommended by the principal and differs from the volatility of human capital. Similarly, the mean rate of consumption growth in the optimal contract depends primarily on the elasticity of intertemporal substitution and will again differ from the rate of growth of human capital. Further, in the optimal contract described above the growth rate of human capital is not constant across agents, as the principal will typically recommend higher effort to agents with higher productivity.

As such, although the distributions of human capital and consumption exhibit (asymptotically) a power law, neither corresponds with their corresponding distributions in the exogenously incomplete markets model. As Figure 2.3 demonstrates, the true law of motion of promised utility is everywhere bounded between two linear functions, which hints at the emergence of a thick right tail in consumption and promised utility. Although I cannot show this formally, this observation is consistent with the numerical examples provided in this paper.

Before proceeding to the general equilibrium context, recall the flow output of the agent in the above was simply assumed to be θ. If flow output is scaled by a constant c then the principal behaves as if confronted with an agent of productivity $c\theta$. This has the following simple (but important) consequence.

Lemma 2.2.7 (Homogeneity in productivity)**.** *For any scale parameter $c > 0$ denote the normalized value function associated with this level of productivity by $v(u; c)$. Then for all $c, u \geq 0$ we have*

$$v(u; c) = cv(u/c; 1).$$

Proof. Immediate by replacing θ in the proof of 2.2.3 with $c\theta$. □

2.2.7 will prove useful in the general equilibrium setting when the productivity of entrepreneurs depends upon an endogenously determined shadow price of labour. In what follows I will write $v(\cdot) \equiv v(\cdot; 1)$ for the normalized payoff function associated with unitary productivity.

2.3 Stationary efficient allocations

The preceding section analyzed the optimal contract between a single risk-averse agent and a risk-neutral principal. This section builds on this analysis by showing how the problem of a benevolent planner in an overlapping-generations economy may be decomposed into a series of one-on-one principal-agent problems identical in form to those considered above. This allows me to derive the effects of private information on the long-run distribution of consumption. In addition I will also allow for a very simple heterogeneity in ex-ante ability: only some agents are capable of performing the risky activity.

I will first describe the environment and specify the welfare notion adopted by the planner. The preferences and technology will be identical to those considered in the previous section. General equilibrium concerns affect both the maximum level of promised utility that may be given to a generation without violating the resource constraints, and the productivity of entrepreneurs, because the latter depends on the number of agents working for their business.

2.3.1 Environment

Time is again continuous and indefinite. At any instant there is a continuum of agents alive in the economy who die at rate ρ_D and may be one of two possible types: entrepreneurs or workers. Entrepreneurs may either operate run their own firm or work for someone else's firm, whereas workers may only work for someone else's firm. An agent enjoys utility

$$U_E(c, e) = \int_0^\infty e^{-(\rho+\rho_D)t}\mathbb{E}[u(c_t, e_t)]dt \tag{2.13}$$

from an allocation $(c_t, e_t)_{t\geq 0}$ specifying consumption and effort exerted, where flow utility is given by

$$u(c, e) := \frac{c^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}$$

where $c_t \geq 0$ and $e_t \in [0, 1]$ almost surely. For simplicity, I assume only workers may exert effort. The fraction of potential entrepreneurs will be denoted $1 - G$, where $G \in (0, 1)$. Types are permanent and unobserved but productivity evolves stochastically. Note that the discount rate in (2.13) arises from the fact that the agents discount future utility at rate ρ and die at rate ρ_D.

In addition to their productivity level θ every agent in the initial generation is indexed by a single variable $v \in V \subseteq \mathbb{R}$ that I will identify with promised utility. To each v-agent there is an associated stochastic process $Z^v = (Z^v_t)_{t\geq 0}$ distributed according to standard Brownian motion and referred to as the *noise* process for agent v. These noise processes are independent of one another and so by a law of large numbers for a continuum of agents[8] the ex-post distribution of shocks across agents will be assumed to coincide with the ex-ante distribution of shocks faced by a single agent. The noise processes of agents of subsequent generations will be indexed by the agents' dates of birth rather than promised utility and

[8]Subject to the usual measurability caveats.

so agents are only distinguished by date of birth and possibly type.

All agents have a common, fixed level of productivity $\theta_0 = \overline{\theta}$ at birth. This productivity remains constant as long as the agent is engaged in the risk-free activity. If the agent engages in the risky activity, their productivity $(\theta_t)_{t\geq 0}$ evolves according to the same law of motion given in the partial equilibrium setting

$$d\theta_t^v = \iota(e_t)1_{e_t\neq 0}\theta_t^v dt + \sigma 1_{e_t\neq 0}\theta_t^v dZ_t^v$$

where $(e_t)_{t\geq 0}$ denotes the effort exerted at any time t. In the partial equilibrium analysis the flow output of an agent with productivity θ_t at any time $t \geq 0$ was simply θ_t. In contrast, in this section the output produced by an entrepreneur will be a function of both their productivity and the total effective worker labour assigned to them. A worker simply produces a flow of $L_W(e) = e$ labour from e units of effort. If an entrepreneur of productivity θ is assigned L units of effective worker labour then the flow output produced is given by a constant-elasticity-of-substitution function of the form

$$F(\theta, L) = \left((1-\beta)[\Gamma_E\theta]^{\frac{\alpha-1}{\alpha}} + \beta[\Gamma_W L]^{\frac{\alpha-1}{\alpha}}\right)^{\frac{\alpha}{\alpha-1}}. \tag{2.14}$$

for some factor-augmenting parameters $\Gamma_E, \Gamma_W > 0$, distribution parameter $\beta \in (0,1)$ and elasticity of substitution $\alpha \in [0,\infty]$. Two special cases for the production function will be singled out. If $\alpha = \infty$ then (2.14) reduces to

$$F(\theta, L) = (1-\beta)\Gamma_E\theta + \beta\Gamma_W L. \tag{2.15}$$

I will refer to the specification (2.15) as the *perfect substitutes* case. If $\alpha = 1$, then the (limiting) value of (2.14) becomes

$$F(\theta, L) = A\theta^{1-\beta}L^{\beta} \tag{2.16}$$

where $A := \Gamma_E^{1-\beta}\Gamma_W^{\beta}$. I will refer to the specification in (2.16) as the *span-of-control* case as it is of the form considered in the span-of-control models of Lucas (48) and Rosen (61).

An allocation must specify the consumption, recommended effort exerted and effective labour assigned to every member of the initial generation throughout their (random) life as a function of initial promised utility, type and history of output, together with analogous quantities for all subsequent generations as functions of their date of birth and initial type. Implicit in this definition is the fact that effective worker labour assigned to an agent is non-zero if and only they are engaged in the risky activity. Because of the presence of the initial generation, allocations will be indexed by distributions Φ over triples (v, θ, i) of promised utility, productivity and type. The formal definition is then the following.

Definition 2.3.1 (Allocation)**.** Given a distribution Φ over promised utility, productivity and types, an allocation consists of consumption, effort recommendations and labour assignments

$$\left\{ \left(c_t^{v,\theta,i}, e_{E,t}^{v,\theta,i}, e_{W,t}^{v,\theta,i}, L_t^{v,\theta,i}\right)_{t\geq 0} \middle| \; (v,\theta,i) \in \operatorname{supp}(\Phi) \right\}$$

for the initial generation together with consumption, effort recommendations and labour assignments

$$\left\{ \left(c_t^{T,i}, e_{E,t}^{T,i}, e_{W,t}^{T,i}, L_t^{T,i}\right)_{t\geq T\geq 0} \middle| \; i = E, W \right\}.$$

for all subsequent generations.

In the above $e_{W,t}$ refers to effort supplied as a worker and $e_{E,t}$ refers to effort supplied as an entrepreneur, with the understanding that only one of these may be non-zero. To ease notation I will therefore write e_t where no ambiguity will arise. For agents not yet alive at the initial date, the superscript refers to birth date and the subscript to calendar time. I will write A for an arbitrary allocation and denote the set of all allocations by $\mathcal{A}$. Also, for any $T \geq 0$ I will write $\mathcal{A}_T$ for the set of all stochastic processes associated with agents born at time T.

Since the notion of an allocation is complicated I will define shorthand notation for aggregate consumption, labour effort and output at a given instant. Note that aggregate consumption, labour and output at any date comprise contributions from the initial generations and from subsequent generations and I will write them in this fashion for clarity. Defining $\Omega := V \times \Theta$, aggregate consumption and output at any date $t \geq 0$ are then,

$$\begin{aligned}
C_t &:= e^{-\rho_D t}\underline{C}_t + \int_0^t e^{-\rho_D[t-T]} C_t^T dT \\
\underline{C}_t &:= \int_\Omega \mathbb{E}\Big[c_t^{v,\theta}\Big]\Phi(d\omega), \qquad C_t^T := \int_\Psi \mathbb{E}\Big[c_t^{T,\psi}\Big]G(d\psi) \\
Y_t &:= e^{-\rho_D t}\underline{Y}_t + \int_0^t e^{-\rho_D[t-T]} Y_t^T dT \\
\underline{Y}_t &:= \int_\Omega \mathbb{E}\Big[F(\theta_t^{v,\theta}, L_t^{v,\theta})\Big], \qquad Y_t^T := \int_\Psi \mathbb{E}\Big[F(\theta_t^{T,\psi}, L_t^{T,\psi})\Big]G(d\psi)
\end{aligned}$$

while aggregate labour assigned to entrepreneurs and labour supplied by workers is

$$\begin{aligned}
L_t^E &:= e^{-\rho_D t}\underline{L}_t^E + \int_0^t e^{-\rho_D[t-T]} L_t^{E,T} dT \\
\underline{L}_t^E &:= \int_\Omega \mathbb{E}\Big[L_t^{v,\theta}\Big]\Phi(d\omega), \qquad L_t^{E,T} := \int_\Psi e^{-\rho_D[t-T]}\mathbb{E}\Big[L_t^{T,\psi}\Big]G(d\psi) \\
L_t^W &:= e^{-\rho_D t}\underline{L}_t^W + \int_0^t e^{-\rho_D[t-T]} L_t^{W,T} dT \\
\underline{L}_t^W &:= \int_\Omega \mathbb{E}\Big[e_{W,t}^{v,\theta}\Big]\Phi(d\omega), \qquad L_t^{W,T} := \int_\Psi e^{-\rho_D[t-T]}\mathbb{E}\Big[e_{W,t}^{T,\psi}\Big]G(d\psi).
\end{aligned}$$

I will also use the notation

$$\underline{U}_t = \int_\Omega \mathbb{E}\Big[u\Big(c_t^{v,\theta}, e_t^{v,\theta}\Big)\Big]\Phi(d\omega) \qquad\qquad U_t^T = \int_\Psi \mathbb{E}\left[u(c_t^{T,\psi}, e_t^{T,\psi})\right] G(d\psi)$$

for the total flow utility experienced by the first and subsequent generations. Note that each of the above aggregate quantities is written as the sum of two terms: the first is the contribution to the aggregate level of the initial generation, and the second is the contribution from all subsequent generations. For instance, since agents die at rate ρ_D and death is independent of the noise process, aggregate consumption of surviving members of the initial generation at date t is $e^{-\rho_D t}$ multiplied by the expected consumption of a single member of this generation. The (physical) resource constraints are then the following.

Definition 2.3.2. An allocation satisfies the goods resource constraints if

$$C_t \leq Y_t$$

for all $t \geq 0$ and satisfies the labour resource constraints if

$$L_t^E \leq L_t^W$$

for all $t \geq 0$. An allocation will be termed *resource feasible* if it satisfies both the resource constraints. The set of all resource feasible allocations will be denoted $\mathcal{A}^{RF}$.

Note that agents have preferences solely over the consumption and effort exerted during their lifetime and so there is no altruism across generations. To specify the preferences of the planner I must therefore specify how they value the flow utility of an agent experienced at any time. I will suppose that the planner places distinct weights on the agents' utility depending only on their date of birth. If $\alpha(T)$ denotes the weight placed on an agent born at time T then the preferences of the planner adopted in this paper are the following

$$\alpha(T) = e^{-\rho T} \qquad \forall T \geq 0. \tag{2.17}$$

Assumption (2.17) may be interpreted as generalized utilitarian objective over generations, since it implies that the planner values the utility experienced by an agent at any given date the sane regardless of the date of their birth. It also ensures that social preferences assume a simple form suitable for recursive analysis.

Lemma 2.3.1 (Planner's preferences)**.** *Under Assumption (2.17), the preferences of the planner over allocations A are represented by the function*

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

Proof. This simply amounts to correctly counting the mass of individuals alive at every date and interchanging the order of integration. □

2.3.2 Incentive-compatibility

The previous section specified the physical resource constraints. This section will define incentive compatibility and promise-keeping constraints on allocations.

Definition 2.3.3. Given an initial distribution Φ over promised utility and types, an allocation A satisfies promise-keeping if $U(c^{v,\theta}, e^{v,\theta}) = v$ for all (v, θ) in the support of Φ.

Notice that promise-keeping is only relevant for the first generation. Incentive compatibility requirements are of two separate types: an agent must be induced to reveal their type at birth (if not in the first generation) and an entrepreneur must be induced to follow the effort recommendations of the planner. I will refer to the first type as *type-revelation* constraints and the second constraint as *incentive compatibility* constraints. The first captures *adverse selection* and the second captures *moral hazard*.

Definition 2.3.4 (Incentive compatibility)**.** An allocation $(c_t^E, c_t^W, e_t^E, e_t^W)_{t\geq 0}$ for a particular generation is incentive compatible if

$$U_E(c^E, e^E) \geq \max\left\{U_E(c^E, e), U_W(c^W, 0)\right\}$$

for all strategies e.

Since there are two dimensions of private information, incentive compatibility must account for the possibility of double deviations, in which an agent simultaneously misreports their type and subsequently deviates from the recommended actions. However, the simple specification of hidden types in the preferences (2.13) ensures that the analysis remains tractable, because regardless of one's type, preferences over consumption and effort recommendations for agents assigned to the risky activity are common knowledge. This also ensures there is no loss in supposing all entrepreneurs and all workers are offered the same contract and so an allocation need only specify the utility to a worker and the utility of an entrepreneur if engaged in the risky activity.

It follows that the presence of hidden types simply imposes the sole additional requirement on the planner that promised be utility for entrepreneurs be sufficiently high to induce truthful revelation. I will write $\mathcal{A}^{RF} \equiv \mathcal{A}^{RF}(\Phi)$ for the set of resource feasible allocations and $\mathcal{A}^{IC} \equiv \mathcal{A}^{IC}(\Phi)$ for the set of incentive compatible allocations, viewed as functions of the initial joint distribution of types, productivity and promised utility. Finally, define

$$\mathcal{A}^{IF} := \mathcal{A}^{RF} \bigcap \mathcal{A}^{IC}$$

for the set of all *incentive feasible* allocations.

2.3.3 Recursive decomposition

The previous section defined allocations, resource feasibility, the preferences of agents and the planner over allocations and incentive compatibility. In this section I will state the planner's problem then show how it may be broken down into more tractable one-on-one principal-agent problems. The arguments employed here are a simple generalization of those employed by Farhi and Werning (24) in an endowment setting with no technological interdependence amongst agents.

Definition 2.3.5 (Planner's problem)**.** Given an initial distribution Φ over promised utility and productivity, the planner's problem is

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

The planner's problem appears intractable for an arbitrary distribution of utility, even in the case of full information and no heterogeneity in types. I will therefore focus on *stationary* solutions, in which the implied distributions of productivity and promised utility are constant over time. To solve these I will first consider the simpler problem of a planner who may trade both goods and labour intertemporally at the rate of time preference, which I will term the *relaxed* planner problem.

Definition 2.3.6 (Relaxed planner's problem)**.** Given an initial distribution Φ over types and promised utility, the relaxed problem of the planner is defined to be

$$V^R(\Phi) = \max_{A \in \mathcal{A}^{IC}} \int_0^\infty \Big(e^{-(\rho+\rho_D)t} \underline{U}_t + \int_0^t e^{-(\rho+\rho_D)[t-T]} e^{-\rho T} U_t^T dT \Big) dt$$
$$\int_0^\infty e^{-\rho t}(C_t - Y_t) dt \leq 0$$
$$\int_0^\infty e^{-\rho t}(L_t^E - L_t^W) dt \leq 0.$$

Stationary solutions to the relaxed planner's problem will solve the planner's problem.

Lemma 2.3.2. *Suppose that for some distribution Φ of initial promised utility, productivity and activities the allocation A solves the relaxed planner's problem with intertemporal price ρ and that the sequence of implied distributions of promised utility, productivity and activities $(\Phi^t)_{t\geq 0}$ is constant. Then A also solves the planner's problem with initial distribution Φ.*

Proof. If the allocation A solves the relaxed planner's problem then by definition it satisfies the promise-keeping and incentive compatibility constraints of the original problem. Further, since consumption and output at any date are solely a function of the distribution, we know that both $C_t - Y_t$ and $L_t^E - L_t^W$ are constant over time and so the discounted resource constraint implies the resource constraint. $\square$

It is easy to see that the rate of time preference is the only intertemporal price which can imply a stationary distribution, because any other price will lead to different treatment of individuals depending on their generation and therefore either a forward or backward trend in promised utility. This is in contrast with the decentralization results of Atkeson and Lucas (10) and Phelan (54), in which the rate of interest amongst competitive intermediaries must in general differ from the rate of time preference in order for the resource constraints to hold. The source of the difference is in the different welfare notions employed. For instance, in Phelan (54) the planner's objective is to minimize the resource costs of delivering a fixed level of utility to each generation. Such a planner will not wish to backload or frontload consumption indefinitely regardless of the rate of interest, and so this latter quantity may differ from the rate of time preference without violating stationarity. The virtue of the above welfare notion will become clearer below, as it ensures that the value function of the planner need only be solved for the single intertemporal rate ρ rather than for continuum of values that must be varied until resources balance.

The relaxed planning problem still takes the entire joint distribution of productivity and promised utility as its argument. However, there are now only two constraints and so the only interdependence between the decisions made by the planner regarding agents in different generations or engaged in different activities may be captured by Lagrange multipliers on resource constraints. Formally, by the general Theorem of Lagrange (Luenberger (49) chapter 8) there exist multipliers $\lambda := (\lambda_R, \lambda_L)$ such that the optimal allocation solves

$$V_\lambda(\Phi) = \max_{A \in \mathcal{A}^{IC}} \int_0^\infty e^{-\rho t}\big(U_t + \lambda_R\big[Y_t - C_t + \lambda_L\big(L_t^W - L_t^E\big)\big]\big)dt \tag{2.18}$$

where I have written

$$U_t = e^{-\rho_D t}\underline{U}_t + \int_0^t e^{-\rho_D[t-T]}U_t^T dT$$

for the total flow utility experienced at time t. For a given λ, the consumption and labour assignments for the initial generation do not affect the consumption and labour assignments of future generations. Therefore, solving (2.18) simply amounts to maximizing the components of the integral relevant to each generation *in isolation*. To see this, separate all the terms in (2.18) into the sum of contributions from each generation: recalling the aggregates

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

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

the problem (2.18) becomes

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

In this way the problem of the planner decomposes into a series of one-on-one problems similar in form to the principal-agent problems analyzed in the above partial equilibrium setting. I will refer to the problem of a planner who takes λ as given and wishes to maximize the weighted average of utility of all agents born at a given instant as a *generational planner problem.* These subproblems are similar to (but distinct from) the component planning problems introduced in the taste shocks settings of Farhi and Werning (24) and Atkeson and Lucas (10). Notice that the objects of choice in the following are the utilities, output, consumption, labour supply and labour demand of agents born at a particular date. Recall that an allocation for a particular generation is an element of $\mathcal{A}_T$ and denoted $\Big\{ (c_t^{T,\psi}, e_t^{T,\psi}, L_t^{T,\psi})_{t \geq T} \Big| \psi \in \Psi \Big\}$.

Definition 2.3.7 (Generational planner problem)**.** Given a pair of multipliers $\lambda = (\lambda_R, \lambda_L)$, the problem of the generational planner is defined to be

$$V_\lambda^G = \max_{\mathcal{A}_T^{IF}} \int_0^\infty e^{-(\rho+\rho_D)t} \Big(U_t^T + \lambda_R \Big[Y_t^T - C_t^T + \lambda_L \Big(L_t^{W,T} - L_t^{E,T}\Big)\Big]\Big) dt.$$

Now define Π to be the minimum cost of providing a weighted level of utility W to a generation of newborns when the multiplier on the labour resource constraint is λ_L,

$$\Pi(W, \lambda_L) = \min_{\mathcal{A}_T^{IF}} \int_0^\infty e^{-[\rho+\rho_D]t} \big[C_t - Y_t + \lambda_L \big(L_t^E - L_t^W\big)\big] dt$$
$$W = \overline{U}_E[1-G] + \overline{U}_W G, \quad \overline{U}_E \geq \overline{U}_W$$
$$\overline{U}_E = U_E\big(c^E, e^E\big), \quad \overline{U}_W = U_W\big(c^W, e^W\big).$$

where $\mathcal{A}_T^{IF}$ denotes the set of incentive feasible allocations for a particular generation. Notice that the only qualitative difference between the above problem and the principal-agent problem analyzed in Section 2.2 is the presence of an additional constraint requiring the utility of an entrepreneur be sufficiently high to ensure truthful revelation. Then we have the following.

Theorem 2.3.3. *The solution to the generational planner's problem when then multiplier is* $\lambda = (\lambda_R, \lambda_L)$ *is given by*

$$V^G_\lambda = \max_W \; W + \lambda_R \Pi(W, \lambda_L).$$

Definition 2.3.8. For a given pair $\lambda := (\lambda_R, \lambda_L)$ of multipliers, denote the associated stationary distribution over $\Omega := Z \times \Theta$ by Φ_λ. The stationary form of the goods and labour resource constraints then reduce to the following pair of equations

$$\begin{aligned} 0 &= \int_\Omega \mathbb{E}\Big[c^{v,\theta,\psi} - F\Big(\theta^{v,\theta,\psi}, L^{v,\theta,\psi}\Big)\Big]\Phi_\lambda(d\omega) \\ G &= \int_\Omega \mathbb{E}\Big[L^{v,\theta}\Big]\Phi_\lambda(d\omega). \end{aligned} \tag{2.19}$$

The following summarizes the relationship between the planner's problem and the component planners' problems and is the culmination of the formal analysis.

Theorem 2.3.4. *If the multiplier* λ *satisfies (2.19) then the solution to the relaxed planner's problem when the initial distribution of promised utility, productivity and types is* Φ_λ *amounts to adhering to the solutions to the generational planner's problem.*

Proof. This simply amounts to untangling definitions. □

2.3.4 concludes the discussion of the planner's problem and the manner in which is it related to simpler component planner problems. I will now combine these decomposition theorems with the optimal contract characterized earlier in partial equilibrium to infer properties of the long-run distributions of consumption and productivity.

2.4 Examples

The previous section illustrates that solving for constrained efficient allocations ultimately amounts to minimising the expected cost, discounted at the rate of time preference, of providing a weighted sum of ex-ante utility to a generation, and then varying the shadow prices of goods and labour until the resource constraints holds in the associated stationary distribution. In this section I will solve for the constrained efficient stationary distribution in a number of example economies.

2.4.1 Only entrepreneurs

First consider the case in which all agents have the ability to be entrepreneurs and worker labour is useless in production, which may be interpreted as the perfect substitutes case with $\Gamma_W = 0$. The multiplier on the labour resource constraint is then zero and the problem facing the generational planner is

$$V_\lambda^G = \max_U \; U + \lambda_R \Pi(U, \theta_0). \tag{2.20}$$

In this case only the initial level of promised utility must be varied until the resource constraint is satisfied. The problem of a generational planner confronted with a generation of newborns is then

$$\max_{\substack{A \in \mathcal{A}^{IC} \\ U = U_E(c,e)}} \; U + \lambda_R \int_0^\infty e^{-(\rho+\rho_D)t} \mathbb{E}\left[Z\theta_t - C_t\right] dt.$$

Conditional on the choice of U this problem is identical to the principal-agent problem analyzed in the partial equilibrium setting. For each initial level U of promised utility we find the associated stationary distributions of consumption and production and then vary U until the resource constraint is satisfied. This would seem to require we solve for a two-dimensional distribution for each initial level of promised utility. However, just as the hometheticity of preferences and log-linearity of the law of motion allowed for a simplification of the principal's problem, so too do the linear policy functions imply that when calculating aggregate quantities we need only restrict attention to a one-dimensional distribution.

Definition 2.4.1 (Summary measure)**.** Given a stationary distribution Φ over productivity and normalized promised utility, the summary measure is defined to be

$$m(B) = \int_B \int_0^\infty \theta\Phi(\theta, u) d\theta du$$

for all $B \subseteq [0, \infty)$.

The homogeneity of the policy functions ensures aggregate quantities may be expressed in terms of this *summary* measure. For instance, the consumption function may be written $C(\theta, u) = c(u)\theta$, so aggregate consumption is

$$\overline{C}(U) = \int_0^\infty \int_0^\infty C(\theta, u)\Phi(\theta, u) d\theta du = \int_0^\infty \int_0^\infty c(u)\theta\Phi(\theta, u) d\theta du = \int_0^\infty c(u)m(u)du$$

while aggregate output is simply $\int_0^\infty m(u)du$. The following shows that this summary measure solves a version of the Kolmogorov forward equation for the *single variable u*. It is taken from Ai (4).

Lemma 2.4.1. *Suppose that the two-dimensional process $(u_t, \theta_t)_{t\geq 0}$ evolves according to the law of motion $(du_t, d\theta_t) = (\mu(u)dt + \sigma_u(u)dZ_t, \iota(u)\theta_t dt + \sigma_\theta(u)\theta_t dZ_t)$ for some functions μ, σ_u, ι and σ_θ. Then the quantity $m(u) := \int_0^\infty \theta\Phi(\theta, u)d\theta$ solves the ordinary differential equation in the single variable u*

$$0 = -(\rho_D - \iota(u))m(u) - [(\mu(u) + \sigma_\theta(u)\sigma_u(u))m(u)]' + \frac{1}{2}[\sigma_u^2(u)m(u)]''. \tag{2.21}$$

Proof. Simply write the forward equation for the two variables (u, θ) and integrate over θ. See appendix for details. □

The appropriate boundary conditions for the ODE (2.21) are not obvious *a priori*, but need not be determined in order to find the stationary distribution. Instead, we can interpret (2.21) as the stationary distribution of a one-dimensional diffusion process and then solve for the stationary distribution of a finite-state Markov chain 'close' in distribution to the original process, following the finite-state Markov chain method of Kushner and Dupuis (47).[9] To solve for the stationary distribution we then vary the initial value of promised utility until resources balance. For any initial normalized promised utility $\overline{u}$, denote the implied density by $m_{\overline{u}}$ and write

$$\begin{aligned} M(\overline{u}) &:= \int_0^\infty m_{\overline{u}}(u)du \\ C(\overline{u}) &:= \int_0^\infty c(u)m_{\overline{u}}(u)du. \end{aligned} \tag{2.22}$$

for the average productivity and consumption of agents in the associated stationary distribution. We then seek a root of the equation

$$M(u) = C(u). \tag{2.23}$$

Now suppose that technological and preference parameters are the same as those given in the earlier example:

$$\begin{array}{lllll} \rho = 0.09 & \gamma = 2 & \alpha = 0.5 & \sigma = 0.29 & z = 1 \\ \bar{\iota} = 0.07 & \underline{\iota} = 0.0 & \eta = 0.5 & \underline{e} = 0.0 & \bar{e} = 0.6. \end{array}$$

I showed earlier that the restricted value and policy functions associated with the highest action are good approximations to the true value and policy functions for low levels of normalized promised utility. Since the drift in normalized promised utility is negative, the behaviour of the planner towards agents with high promised utility is close to that of the restricted policy function for the highest effort level. Prior to calculating the true stationary distribution, it is useful to consider the stationary distribution associated with these restricted policy and value functions as they may be calculated in closed form. As shown in 2.2.4, the restricted normalized value function and policy function of the planner

[9] Further details of this procedure are given in Appendix A.2.2.

have the explicit solutions

$$v(u;e) = \frac{z[\rho + \rho_D]}{\rho + \rho_D - \iota(e)} - \chi_1(e)u, \qquad c(u;e) = \chi_2(e)u \tag{2.24}$$

for some $\chi_1(e)$ and $\chi_2(e)$. Since $C(U, \theta) = c(u)\theta$ and $u = [(1-\gamma)U]^{\frac{1}{1-\gamma}}\theta^{-1}$, we have

$$C(U_t, \theta_t) = c(u_t; e)\theta_t = \chi_2(e)u_t\theta = \chi_2(e)[(1-\gamma)U]^{\frac{1}{1-\gamma}}.$$

In the original variables, the laws of motion for utility and human capital in the restricted-action case are given by

$$(dU_t, d\theta_t) = (\mu_U(e)U_t dt + \sigma_U(e)U_t dZ_t, \iota(e)\theta_t dt + \sigma(e)\theta_t dZ_t)$$

where for each of notation I have written the drift and diffusion of utility as

$$\mu_U(e) = (\rho + \rho_D)\big(1 - \chi_2(e)^{-1}(1-e)^{-\alpha}\big) \qquad \sigma_U(e) = (\rho + \rho_D)\sigma\chi_2(e)^{1-\gamma}\kappa(e).$$

The stationary distributions of utility in this case may be simply characterized because there is no interplay between promised utility and productivity. Put another way, promised utility follows a Markov process (by itself). It is well known[10] that the stationary distribution of a variable whose logarithm follows Brownian motion with drift, begins at a fixed $\overline{U}$ and dies at a constant rate is double-Pareto. When combined with 2.2.4 this leads to the following explicit characterizations of the distributions associated with the restricted action allocations.

Theorem 2.4.2. *For each $e \in (0,1)$ the stationary distribution of consumption associated with the restricted-action allocation for the effort level e is double-Pareto with the exponent of the upper tail given by*

$$\alpha_-(e, \gamma) = -\frac{\gamma}{2} - \sqrt{\frac{\gamma^2}{4} + 2\rho_D\left(\frac{J(e,\gamma)}{\Sigma(e)}\right)^2}$$

[10] See e.g. Luttmer (50).

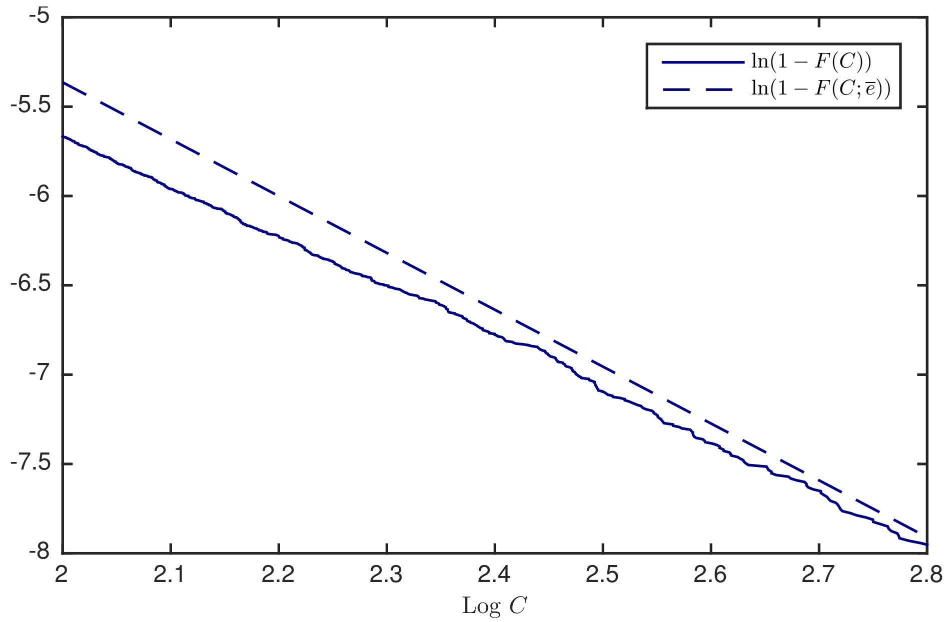

Figure 2.6: Tail of consumption distribution and restricted-action consumption distribution

where

$$J(e, \gamma) = \frac{\rho^{-1}(2\gamma - 1)(\gamma - 1)\Sigma(e)^2}{\sqrt{1 + 2\rho^{-1}(2\gamma - 1)(\gamma - 1)\Sigma(e)^2} - 1} \qquad \Sigma(e) = \frac{\rho\sigma\overline{\alpha}e^{1-\eta}}{\eta\bar{\iota}(1 - e)}.$$

The tails of the stationary distribution of consumption and the consumption distribution implied by the restricted-action policy associated with $\overline{e}$ are given by Figure 2.6. The corresponding tails of the stationary distribution of productivity and the productivity distribution implied by the restricted-action policy associated with $\overline{e}$ are given by Figure 2.7.[11] As can be seen in Figure 2.6 and Figure 2.7, the distributions of consumption and productivity have asymptotically Pareto tails (the lines are straight on a log-log scale) with slopes approximately equal to those of their restricted-action counterparts.

The maximal degree of inequality associated with the restricted-action allocations depends upon the value of $\overline{e}$, the highest effort level the agent may take. As this variable is

[11] The restricted-action distributions given in Figure 2.6 and Figure 2.7 have been shifted parallel for ease of comparison, as the purpose of this exercise is simply to compare slopes rather than absolute positions.

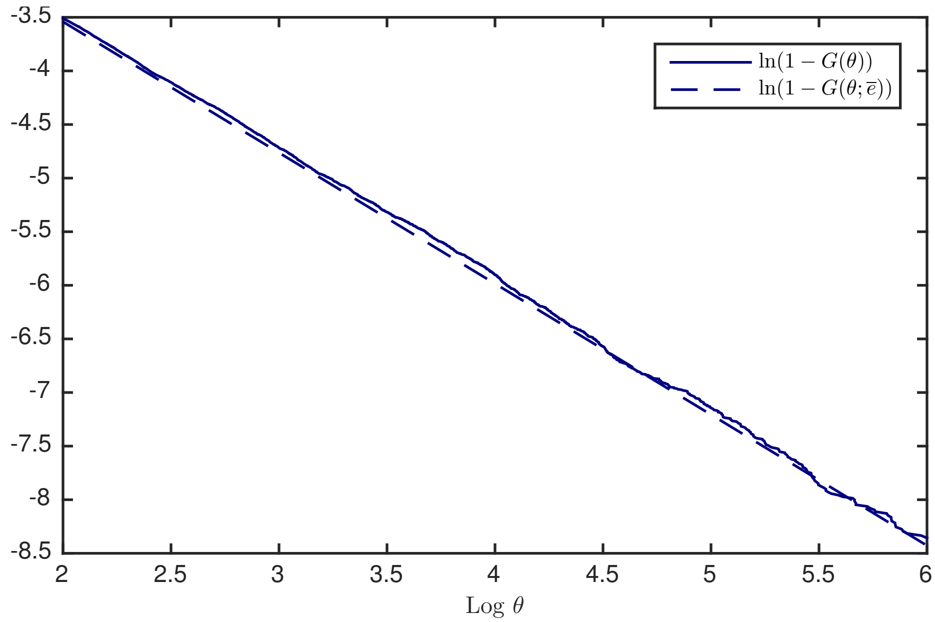

Figure 2.7: Tail of productivity distribution and restricted-action productivity distribution

difficult to read directly off the data, it is instructive to use 2.4.2 to derive the following uniform bounds on consumption inequality.

Corollary 2.4.3. *For all $\gamma \geq 1$ and $e \in (0,1)$ we have*

$$\alpha_-(e,\gamma) \leq -\frac{\gamma}{2}\left(1 + \sqrt{1 + \frac{4\rho_D}{\rho}(2 - 1/\gamma)(1 - 1/\gamma)}\right)$$

with equality in the limit as $e \to 1$.

Corollary shows that in the limiting case of logarithmic utility ($\gamma = 1$) the Pareto parameter of the upper tail for the restricted-action case satisfies the following limit,

$$\lim_{e \to 1} \alpha_-(e,1) = -1 \tag{2.25}$$

which is the lowest Pareto parameter (or thickest tail) consistent with a finite mean for consumption. By comparing with the results of Jones and Kim (42), (2.25) shows that (at least in the restricted-action case) the upper tail that obtains in the constrained-efficient allocation may be thicker than the tail in the case of exogenously incomplete markets.

2.4.2 Workers and entrepreneurs

I will now relax the assumption of no technological dependence amongst agents and suppose that only a fixed fraction $1 - G < 1$ of agents may be entrepreneurs. The output of any entrepreneur now depends on the total amount of worker labour that has been assigned to them. In contrast with the case with only entrepreneurs, a generational planner facing a population of newborns must internalize the effect that the effort levels recommended to entrepreneurs have on the shadow price of labour. I will first analyze the span-of-control specification (Cobb-Douglas), before turning to the more the general case of constant-elasticity of substitution.

Span-of-control

Suppose that the flow output per unit of time of an entrepreneur of productivity θ who has been assigned L units of labour is given by

$$F(\theta, L) = Z\theta^{1-\beta}L^{\beta}$$

for some $\beta \in (0, 1)$. Before solving the generational planner problem, notice that at any moment the optimal assignment of workers to an entrepreneur depends solely upon the shadow price of labour and the entrepreneur's productivity. It is the purely static problem

$$Y(\theta, \lambda_L) := \max_{L\geq 0}\ Z\theta^{1-\beta}L^{\beta} - \lambda_L L$$

with solution requiring only elementary algebra.

Lemma 2.4.4. *Given the shadow price λ_L the assignment of labour to an entrepreneur of productivity θ is*

$$L(\theta) = [Z\beta/\lambda_L]^{\frac{1}{1-\beta}}\theta. \tag{2.26}$$

Flow output is $Z^{\frac{1}{1-\beta}}(\beta/\lambda_L)^{\frac{\beta}{1-\beta}}\theta$, and flow output net of labour resource costs $\lambda_L L$ is given by $Y(\theta, \lambda_L) = \overline{Z}\lambda_L^{-\frac{\beta}{1-\beta}}\theta$ where $\overline{Z} = Z^{\frac{1}{1-\beta}}\beta^{\frac{\beta}{1-\beta}}(1-\beta)$.

Workers have no preferences over the identity of the entrepreneurs to whom they are assigned and so the above assignment choice may be made independently of all other decisions. The key point of Lemma 2.4.4 is that changes in the shadow price for labour simply translates into changes in the productivity of every entrepreneur, with payoffs remaining proportional to θ. Further, flow output from each entrepreneur is decreasing in the shadow value of labour. The problem of the generational planner when $\gamma = 2$ is then

$$\Pi(\lambda) = \max_{U \leq 0} \; U + \lambda_R \big(V(U, 1; \lambda_L)[1 - G] + \big(\lambda_L + U^{-1}\big) G\big).$$

It is convenient to define

$$x := \overline{Z}^{-1} \lambda_L^{\frac{\beta}{1-\beta}} [(1-\gamma)U]^{\frac{1}{1-\gamma}} \theta^{-1} = \overline{Z}^{-1} \lambda_L^{\frac{\beta}{1-\beta}} u$$

for utility in consumption units normalized by the endogenous productivity level $\overline{Z}\lambda_L^{-\frac{\beta}{1-\beta}}$ (the coefficient of θ in the above expression Y for flow output of an entrepreneur). In this way the problem may be rewritten

$$\Pi(\lambda) = \lambda_R \lambda_L G + \max_{x \geq 0} \; -\left(\overline{Z}\lambda_L^{-\frac{\beta}{1-\beta}} x\right)^{-1} + \overline{Z}\lambda_L^{-\frac{\beta}{1-\beta}} \lambda_R \left([1 - G] v(x) - G x\right)$$

so that for any pair of multipliers $\lambda = (\lambda_R, \lambda_L)$, the problem of the planner reduces to choosing an initial level of normalized promised utility x. As with the case without workers, for any such choice there will be an associated stationary density of normalized promised utility $m_x(du)$ that may be computed using only the policy functions associated with the value function v (corresponding to *unitary* productivity). The optimal choice of x depends on the multipliers λ_L and λ_R. For each choice of x we determine excess labour demand excess consumption as functions of the multipliers and find conditions under which these both vanish. Note that the labour resource constraint will be satisfied if and only if

$$G = [1 - G] M(x) (Z\beta/\lambda_L)^{\frac{1}{1-\beta}} \tag{2.27}$$

where $M(x)$ is given by (2.22). To see this, note that

RHS of (2.27) = (average θ per entrepreneur in stationary distribution)

$\times$ (mass of entrepreneurs) $\times$ (labour demand per unit of productivity (2.26))

LHS of (2.27) = labour supply by workers.

Further simplification then gives the following theorem.

Theorem 2.4.5. *The stationary level of x is a solution to the equation*

$$\left(\frac{1-G}{1-\beta}\right)M(x) = (1-G)C(x) + Gx \tag{2.28}$$

with the associated level of u given by $u = Z(1-\beta)x([1/G-1]M(x))^{-\beta}$.

Proof. From 2.4.4 we see that the output of the firm of each entrepreneur may be written

$$Z\theta^{1-\beta}L(\theta)^{\beta} = Z^{\frac{1}{1-\beta}}(\beta/\lambda_L)^{\frac{\beta}{1-\beta}}\theta.$$

Integrating over all entrepreneurs and using (2.28), output in the stationary distribution is

$$Y(x) = Z^{\frac{1}{1-\beta}}(\beta/\lambda_L)^{\frac{\beta}{1-\beta}}(1-G)M(x) = \overline{Z}\lambda_L^{-\frac{\beta}{1-\beta}}\left(\frac{1-G}{1-\beta}\right)M(x). \tag{2.29}$$

Aggregate consumption is simply the sum of consumption from both entrepreneurs and workers, or

$$\overline{Z}\lambda_L^{-\frac{\beta}{1-\beta}}[1-G]C(x) + \overline{Z}\lambda_L^{-\frac{\beta}{1-\beta}}Gx = \overline{Z}\lambda_L^{-\frac{\beta}{1-\beta}}([1-G]C(x) + Gx). \tag{2.30}$$

Dividing (2.29) and (2.30) by $\overline{Z}\lambda_L^{-\beta(1-\beta)^{-1}}$ then gives the left- and right-hand sides of (2.28), respectively. Rearranging (2.27) gives $\lambda_L^{-\beta(1-\beta)^{-1}} = ([1/G-1]M(x))^{-\beta}(Z\beta)^{-\beta(1-\beta)^{-1}}$, and so the implied level of normalized promised utility is equal to

$$u = \overline{Z}x\lambda_L^{-\frac{\beta}{1-\beta}} = \overline{Z}x([1/G-1]M(x))^{-\beta}(Z\beta)^{-\frac{\beta}{1-\beta}}$$

which reduces to the claimed expression. □

Note that to determine the constrained-efficient stationary distribution the value function need only be calculated *once* even though there is a continuum of agents and two resource constraints. 2.4.5 now allows for simple derivations of comparative statics for the initial level of x.

Corollary 2.4.6. *The stationary level of x is decreasing in G, increasing in β and independent of Z.*

Proof. Define $H(x, G) := M(x)(1-\beta)^{-1} - C(x) - xG(1-G)^{-1}$ and note that the stationary x is a root of $H(\cdot, G)$. The first two claims follow easily from the fact that H is decreasing in G and increasing in β. The last claim simply follows from the fact that Z appears nowhere in the equation defining x. □

The third claim in Corollary 2.4.6 may be viewed as a neutrality result. It shows (Hicks-neutral) changes in total factor productivity (constant Z in the above) have no effect on the stationary value of x. They therefore have no effect on inequality in the associated stationary distributions of promised utility and consumption, as these quantities are simply scaled for all agents after every history by the same proportion.

General CES production technology

The above analysis extends easily to the more general case with a constant-elasticity-of-substitution production function. Given factor-augmenting technology parameters $\Gamma_E, \Gamma_W > 0$, distribution parameter $\beta \in (0, 1)$, and elasticity of substitution $\alpha \in [0, \infty)$, define

$$F(\theta, L) = \left((1-\beta)[\Gamma_E \theta]^{\frac{\alpha-1}{\alpha}} + \beta[\Gamma_W L]^{\frac{\alpha-1}{\alpha}}\right)^{\frac{\alpha}{\alpha-1}}.$$

The following theorem follows from arguments similar to those given above.

Theorem 2.4.7. *The stationary level of x is the largest solution to the equation*

$$0 = \frac{\beta}{1-\beta}[(1/G-1)M(x)]^{1/\alpha}(\Gamma_W/\Gamma_E)^{1-1/\alpha} + (M(x) - C(x))[1/G-1] - x. \tag{2.31}$$

We may now derive comparative statics from Theorem 2.4.7 in a unified way. The abstract referred to changes in the mass of potential entrepreneurs and to changes in technology that primarily benefits potential entrepreneurs. This corresponds to increases in G or increases in Γ_W/Γ_E.

Theorem 2.4.8. *For any value of the elasticity of substitution, the stationary x is a decreasing function of G. The stationary level of x is decreasing in Γ_E/Γ_W when $\alpha > 1$ and increasing in Γ_E/Γ_W when $\alpha < 1$.*

Proof. For the first claim, for any G_1 and G_2 with $G_2 > G_1$ denote by $x_1 \equiv x(G_1)$ and $x_2 \equiv x(G_2)$ the associated stationary levels of normalized utility. Define the function

$$H(\Gamma_W, \Gamma_E, G, x) = \frac{\beta}{1-\beta}(\Gamma_W/\Gamma_E)^{1-1/\alpha}[M(x)(1/G-1)]^{1/\alpha} \\ + (M(x) - C(x))[1/G - 1] - x$$

and note that since H is decreasing in G, $H(\Gamma_W, \Gamma_E, G_1, x) > H(\Gamma_W, \Gamma_E, G_2, x)$ for all $x, \Gamma_W, \Gamma_E \geq 0$, from which $x_1 > x_2$ follows. The second claim follows from similarly, because H is decreasing in Γ_E/Γ_W when $\alpha > 1$ and increasing in Γ_E/Γ_W when $\alpha < 1$. □

The above shows how the ex-ante level of normalized promised utility x is affected by exogenous changes in technology. However, ultimately we are interested in the implied distributions of consumption, leisure and utility rather than the transformed variable x. Intuitively, the lower is x, the greater are the incentives of the planner to recommended high actions and so any change in technology that decreases in stationary x will tend to increase the risk borne by entrepreneurs and hence the efficient degree of inequality. I have found this intuition to be reinforced by the results of numerical simulations. I leave a quantitative exploration of these effects to future work.

2.5 Implementation

The foregoing analysis has focused on the forces that shape the long-run constrained-efficient levels of inequality in an economy with repeated moral hazard, with no discussion of a market structure. In this final section I will address how the constrained efficient allocations may be implemented by the planner with taxes when agents may write contracts of arbitrary complexity with private financial intermediaries. I will show that despite the rich history dependence of the optimal contract, a constant linear tax on the consumption of entrepreneurs together with a constant linear subsidy on the consumption of workers will suffice to implement the constrained-efficient allocation.

Throughout I will suppose that a fraction $1 - G$ of agents are potential entrepreneurs and that a fraction G may only be workers, with the entrepreneurs obtaining utility $\overline{\psi}U$ from the entrepreneurial activity when given utility U from consumption and effort. I will also suppose that intermediaries may borrow and lend at the rate of time preference. I will first define the problem of an intermediary faced with an agent to whom they have promised a certain amount of promised utility, given prices for labour and intertemporal trade. This problem has a similar form to that of the principal considered in the above partial equilibrium setting. For ease of notation, in what follows I will write

$$U_E((1-\tau^E)c, e) := (\rho + \rho_D)\int_0^\infty e^{-(\rho+\rho_D)t}\mathbb{E}\left[\frac{[(1-\tau_t^E)c_t]^{1-\gamma}[1-e_t]^{-\alpha}}{1-\gamma}\right] dt$$

for the utility of an entrepreneur alive at the first date from the sequences $(c_t, e_t)_{t\geq 0}$ when facing taxation $(1-\tau_t^E)_{t\geq 0}$, together with analogous quantities for all future generations.

Definition 2.5.1 (Intermediaries' problem)**.** Given sequences of wages $w = (w_t)_{t\geq 0}$, interest rates $r = (r_t)_{t\geq 0}$ and consumption taxes $(\tau_{C,t}^E, \tau_{C,t}^W)_{t\geq 0}$ on entrepreneurs and workers, respectively, the problem of a financial intermediary facing an entrepreneur with outside option U and productivity θ is defined to be

$$\Pi_E(U,\theta;w,r,\tau^E) = \max_{(c,e)\in\mathcal{A}^{IC}} \int_0^\infty e^{-\int_0^t (r_s+\rho_D)ds}\mathbb{E}\left[\overline{Z}(\beta,e)w_t^{-\frac{\beta}{1-\beta}}\theta_t - c_t\right]dt.$$
$$U = U_E((1-\tau^E)c,e)$$
$$d\theta_t = \iota(e_t)\theta_t dt + \sigma(e_t)\theta_t dZ_t, \quad \theta_0 = \theta.$$

The problem of a financial intermediary facing an agent operating the risk-free technology with outside option U and productivity θ is defined to be

$$\Pi_W(U,\theta;w,r,\tau^W) = \max_{(c,e)\in\mathcal{A}^{IC}} \int_0^\infty e^{-\int_0^t (r_s+\rho_D)ds}\mathbb{E}\left[w_t e_t \theta - c_t\right]dt.$$
$$U = U_W((1-\tau^W)c,e).$$

Definition 2.5.2. Given sequences $(\tau^E,\tau^W) := (\tau_t^E,\tau_t^W)_{t\geq 0}$ of linear taxes on the consumption of entrepreneurs and workers, a competitive equilibrium consists of sequences of wages $(w_t)_{t\geq 0}$ and interest rates $(r_t)_{t\geq 0}$, together with an allocation

$$A := \left\{(c_t^{v,\theta}, e_{E,t}^{v,\theta}, e_{W,t}^{v,\theta}, L_t^{v,\theta})_{t\geq 0}\middle|\,(v,\theta)\in\Omega\right\},\ \left\{(c_t^{T,\psi}, e_{E,t}^{T,\psi}, e_{W,t}^{T,\psi}, L_t^{T,\psi})_{t\geq T\geq 0}\right\}_{\psi\in\Psi}$$

and contracts $\left\{(c_{E,t}^{v,\theta}, e_{E,t}^{v,\theta}, c_{W,t}^{v,\theta}, e_{W,t}^{v,\theta})_{t\geq 0}\middle|\,(v,\theta)\in\Omega\right\}$ and $(c_{E,t}^T, e_{E,t}^T, c_{W,t}^T, e_{W,t}^T, L_t^T)_{t\geq T\geq 0}$ such that the following hold:

- Given the levels of promised utility

$$v = U_E\left((1-\tau^E)c_E^{v,\theta}, e_E^{v,\theta}\right) \qquad v = U_W\left((1-\tau^W)c_W^{v,\theta}, e_W^{v,\theta}\right)$$
$$U_E^T = U_E\left((1-\tau^E)c_E^{T,\theta}, e_E^{T,\theta}\right) \qquad U_W^T = U_W\left((1-\tau^W)c_W^{T,\theta}, e_W^{T,\theta}\right)$$

  for the first and future generations implied by the allocation A, the contracts chosen by the intermediaries minimize the costs of attaining v, U_E^T and U_W^T for the prices w and r. This requires that for all $(v,\theta)\in\Omega$, the contracts $(c_E^{v,\theta}, e_E^{v,\theta}, L^{v,\theta})$ and $(c_W^{v,\theta}, e_W^{v,\theta})$ solve $\Pi_E(v,\theta;w,r)$ and $\Pi_W(v,\theta;w)$, respectively, and for all $T\geq 0$, the contracts

$(c^{T,E}, e^{T,E}, L^T)$ and $(c^{T,W}, e^{T,W})$ solve $\Pi_E(U_E^T, \theta_0; w, r)$ and $\Pi_W(U_W^T, \theta_0; w, r)$, respectively.

- Financial intermediaries make zero discounted expected profits from all contracts written at or after the initial date, so that $\Pi_E\left(U_E^T, \theta_0; w, r, \tau^E\right) = \Pi_W\left(U_W^T, \theta_0; w, r, \tau^W\right) = 0$ for all $T \geq 0$. Further, there is no contract that is not offered but would make positive profits if offered.

- All agents not in the first generation choose the contract that maximizes their utility: for all $T \geq 0$ this requires

$$U_E\left((1-\tau^E)c_E^{T,\theta}, e_E^{T,\theta}\right) \geq U_W\left((1-\tau^W)c_W^{T,\theta}, e_W^{T,\theta}\right).$$

- Labour markets and goods markets clear every instant.

There is no need to impose the requirement that the government budget balance because of Walras' law. In keeping with focus of this paper on long-run inequality, I will restrict attention to *stationary* equilibria in which wages, interest rates and the distributions of consumption and productivity are constant.

Note that the competitive equilibrium without taxes need not coincide with the utilitarian constrained-efficient allocation characterized above. Competition amongst firms to provide insurance to workers will ensure that the level of promised utility is set at the point at which firms make zero expected discounted profits, and nothing forces this level to coincide with the constrained-efficient level of ex-ante utility assigned to entrepreneurs. However, the form of the preferences ensures that it is easy to describe the effects of a constant linear consumption tax on utility. Specifically, if consumption is taxed at rate τ and $(c, e) := (c_t, e_t)_{t\geq 0}$ denotes the consumption and recommended effort levels specified in the contract between the firm and the entrepreneur, then the utility of the agent is given by

$$
\begin{aligned}
U((1-\tau)c, e) &= (\rho + \rho_D) \int_0^\infty e^{-(\rho+\rho_D)t} \mathbb{E}\left[\frac{((1-\tau)c_t)^{1-\gamma}(1-e_t)^{-\alpha}}{1-\gamma}\right] dt \\
&= (1-\tau)^{1-\gamma} U(c, e).
\end{aligned} \tag{2.32}
$$

The multiplicative form of the preferences therefore implies that the imposition of a linear tax simply scales utility experienced by the entrepreneur state-by-state, leaving incentive compatibility unchanged. From the perspective of the intermediary, with the introduction of a linear tax on entrepreneurs' tax τ, the profit from contracting with an agent with normalized promised u simply changes from $v(u)$ to $v(u[1-\tau]^{-1})$. The level of normalized promised utility $u(\tau)$ that obtains in the competitive equilibrium is then simply

$$u(\tau) = (1-\tau)x$$

where x solves $v(x) = 0$. This leads to the following.

Theorem 2.5.1. *Denote by U^E and U^W the levels of promised utility associated with a newborn entrepreneur and worker, respectively, in the constrained-efficient utilitarian allocation characterized above. Then the constrained-efficient allocation may be implemented as a competitive equilibrium with interest rate $r = \rho$, constant linear taxes τ^E and τ^W, where τ^E satisfies $u(\tau^E) = [(1-\gamma)U^E]^{\frac{1}{1-\gamma}}$, and τ^W satisfies*

$$\Pi_W\left(U_W^T, \theta_0; \lambda_L, \tau^W\right) = 0,$$

where λ_L is the multiplier on the labour resource constraint found in the planning problem.

Proof. The basic idea behind the proof is that for the above choices of taxes and interest rates, the problems of the intermediaries facing entrepreneurs and workers are basically identical to the problem faced by a generational planner in the efficient allocation. The constrained-efficient allocation will depart from the competitive equilibrium only because in the latter competition amongst intermediaries will push the level of promised utility to entrepreneurs above the constrained-efficient level, and the imposition of a linear tax suffices to 'correct' this.

First note that by the observation in (2.32), a linear tax on entrepreneurial income will not affect incentives to shirk because it simply amounts to scaling down utility by the same factor along every history. When the interest rate is constant at $r = \rho$ and a constant tax $1 - \tau^E$ is levied on entrepreneurial income, the problem of an intermediary faced with an entrepreneur to whom he has promised U units of utility is identical to that of a generational planner faced with an agent of promised utility $U(1-\tau^E)^{\gamma-1}$. The linear tax specified in the statement of the theorem is exactly that level necessary to ensure that the level U^E in the constrained-efficient allocation coincides with the level such that $\Pi_E\big(U^E, \theta_0; \lambda_L, \rho, \tau^E\big) = 0$.

The final observation is that stationary level of the wage must coincide with the Lagrange multiplier found in the constrained-efficient utilitarian allocation. This is once again essentially true by construction: in the constrained efficient allocation, the multiplier λ_L is the 'price' that the generational planner faces when making the static labour-hiring decision. $\square$

In light of the decentralization results of Prescott and Townsend (58), Atkeson and Lucas (10) and Golosov and Tsyvinski (29), the fact that the desire of the taxation authority is reduced when agents are unrestricted in the contracts they may write with intermediaries is not surprising. However, it is worth noting that Theorem 2.5.1 shows that when one introduces ex-ante heterogeneity into the environment, the redistributive role of the government does not imply progressivity of taxation.

2.6 Conclusion

This paper develops theory and numerical techniques to determine the constrained efficient levels of inequality in an economy with dynamic moral hazard and random growth in human capital. I characterize the optimal contract in a principal-agent setting and show how this may be used to solve the problem of a benevolent planner facing a continuum of agents with heterogeneous abilities in an overlapping generations economy. For a particular class (those with constant effort) of incentive compatible allocations the implied stationary distributions admit closed-form expressions that possess a power law in the upper tail, and I show in numerical examples that this property carries over (at least asymptotically) to the constrained efficient allocation.

The analysis of this paper also suggests a number of possible directions for further work. I have shown that the number of workers per entrepreneur and changes in technology affect the efficient amount of risk borne by any given entrepreneur. This suggests it may be worthwhile to conduct a quantitative analysis of the significance of such changes in affecting long-run inequality. The efficient allocation in this economy typically exhibits substantial history-dependence. It would interesting to determine the welfare loss incurred when the allocation is required to be stationary, or restricted to a particular parametric class. Similarly, when considering the implementation of the efficient allocation, I have restricted attention to a benchmark case in which agent may write contracts of arbitrary complexity with a competitive sector of private intermediaries. It may be interesting to explore optimal taxation policy in the complementary case in which such contracting opportunities are absent and so that the government engages in risk-sharing rather than being purely redistributive.

Chapter 3

Efficient differential capital taxation under dynamic agency

The previous chapter characterized efficient allocations in an environment in which business income depended primarily on owner-specific characteristics. To isolate the importance of agency moral hazard in the accumulation of human capital, I assumed that physical capital played no role in production. The model was therefore unable to analyze the appropriate taxation of income from various forms of physical assets. In this chapter I conduct a complementary exercise and characterize efficient allocations in an environment in which agency frictions affect the accumulation of (physical) capital but human capital plays no role. I then show how the presence of such agency frictions provides a justification for the imposition of differential asset taxation.

Specifically, I consider a perpetual youth environment in which agents may either run their own business (be an entrepreneur) or work for another agent's business (be a worker). Only some agents have the ability to run a business and this is private information. In addition, entrepreneurship is subject to two agency frictions: physical capital is subject to (privately-observed) output shocks and may be diverted to (privately-observed) consumption; and entrepreneurs may abscond with a fixed fraction of delegated capital. An allocation in this environment must specify the occupation of every agent and the amount of capital and labour delegated to each business. The ability of entrepreneurs to divert capital to private consumption will imply that consumption must depend upon business performance, which leads to ex-post inequality in the efficient allocation. The ability of entrepreneurs to abscond with a fraction of their assets will limit the amount of capital that may be delegated to them.

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

An extensive literature, surveyed in Chari and Kehoe (17), has analyzed optimal capital and labour taxation in environments in which agents face no idiosyncratic risk and the government is assumed to have access only to linear taxes on various forms of income and consumption. Recent contributions to this tradition, such as Panousi and Reis (53) and Evans (23), extend this analysis to consider optimal linear taxation in economies with a continuum of agents subject to idiosyncratic risk with exogenously incomplete markets. In contrast, the analysis of this chapter builds upon the seminal contribution of Mirrlees (51) and the New Dynamic Public Finance literature beginning with Golosov, Kocherlakota and Tsyvinski (28) to characterize efficient allocations when the only restrictions placed upon government policy are those implied by informational asymmetries. However, the majority of this literature has focused on the implications of private-information in labour productivity on the structure of optimal tax schedules and has not explicitly accounted for entrepreneurial activity. Three notable exceptions are Albanesi (5), Shourideh (68) and the model of the first chapter of this thesis, each of which explicitly model entrepreneurial activity in the presence of agency frictions. It is therefore instructive to outline how the modeling assumptions and findings of these papers differ from those presented here.

Albanesi (5) considers a two-period model in which there are no workers, initial wealth is exogenous and common across entrepreneurs, and the returns to entrepreneurial activity depend upon unobserved effort. She finds that in general the efficient intertemporal wedge differs from the case with unobserved labour productivity and may assume either sign. Further, decentralization of the efficient allocation with an exogenous market structure may require double taxation at both the firm and individual level. I recover this latter result in my decentralization, but extend the model to an infinite-horizon setting in which wealth is endogenous and so am able to illustrate how private information affects the efficient level of long-run inequality.

More closely related with the current chapter is Shourideh (68), who also analyzes an agency model in which entrepreneurs may divert assets to private consumption. I reformulate the agency problem in continuous-time and adopt a welfare notion and lifecycle structure that leads to simpler characterizations of both efficient allocations and their decentralization.[1] The modeling of the agency problem also qualitatively changes the nature of efficient intertemporal distortions. In contrast to the findings of both Albanesi (5) and Shourideh (68), I show that the inverse Euler equation of Rogerson (59) and Golosov, Kocherlakota and Tsyvinski (28) continues to hold in the presence of production risk for a wide range of parameter values. Further, the increased tractability allows me to extend the literature in several ways.

First, in addition to characterizing a stationary efficient allocation, I show how it may be implemented in a general equilibrium model with exogenously incomplete markets. The decentralization requires only linear, time-independent taxes and so optimal policy may be

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

completely specified by five scalars, each of which admits a closed-form representation in terms of the solution to a single non-linear equation. Further, the finding that the optimal profit tax is independent of all technological parameters relies on general equilibrium effects and appears to have no antecedent in the literature on optimal taxation with endogenously incomplete markets.

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

Third, the model also allows for a sharper characterization of the long-run distributions of utility and consumption and their determinants. Indeed, in this chapter the stationary distribution of consumption admits a closed-form density of the 'double-Pareto' form[2] with the degree of inequality in the upper tail determined by the amount of risk borne by each entrepreneur. In addition to the aforementioned role played by the number of workers per entrepreneur, this allows me to show how inequality and taxes depend upon the severity of agency frictions, the amount of exogenous uncertainty, and the returns-to-scale of capital in production.

Finally, the first chapter of this thesis considered an environment in which the productivity (or human capital) of entrepreneurs grows randomly over time and depends (partly)

[2] Piecewise polynomial on the positive halfline.

on unobserved effort. The focus there was on the characterization of efficient allocations in an environment with a novel agency problem involving human capital rather than physical capital, with no discussion of how such an allocation may be implemented with exogenously incomplete markets. Although these two chapters both characterize efficiency in dynamic environments with agency frictions, the modeling assumptions, scope and results are quite different. Rather that analysing a novel agency problem, this chapter instead shows how a variation on a previously explored agency problem leads to both increased tractability and novel results for the decentralization of the efficient allocation in a general equilibrium model in which agents may only trade a risk-free bond.

The outline of this chapter is as follows: Section 3.1 analyzes a principal-agent model in which both the productivity of the agent and the interest rate are exogenous; Section 3.2 then embeds this into a macroeconomic model and characterizes stationary constrained-efficient allocations when productivity is endogenously determined by aggregate physical and labour resource constraints, and derives a number of comparative statics results; and Section 3.3 decentralizes these stationary efficient allocations in a general equilibrium model with linear taxes on labour income of workers and the (risk-free) savings and profits of entrepreneurs and plots a numerical example.

I omit technical details (such as derivations of Hamilton-Jacobi-Bellman equations, martingale techniques, measurability concerns and so on) wherever they are similar to analogous aspects of the model of the previous chapter. Technical proofs and a discrete-time version of the environment that relates the findings of the main text with those of the related literature are outlined in the appendix.

3.1 Principal-agent model

This section characterizes the optimal risk-sharing arrangement between a risk-averse agent and a risk-neutral principal in an environment where the agent may operate a risky production technology, their consumption is private-information and they may abscond with a fraction of the physical assets under their control. The environment is a slight variation on that considered in Sannikov and Di Tella (64).[3] This principal-agent problem will later be embedded into a macroeconomic model in which flow payoffs to the principal are endogenously determined by aggregate resource constraints. Most of the technical details will be relegated to the appendix.

Formal setup

Time is continuous and indefinite. The economy consists of a single risk-averse agent and a risk-neutral principal, both of whom live forever. The preferences of the consumer over stochastic sequences of consumption $c := (c_t)_{t\geq 0}$ are represented by the utility function

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

The agent has the ability to operate a constant-returns-to-scale technology subject to random shocks to productivity. The principal may delegate capital to the agent so that it may be invested in their technology. When capital delegated follows the process $K := (K_t)_{t\geq 0}$, output (net of depreciation and borrowing costs) $Y := (Y_t)_{t\geq 0}$ evolves according to

$$dY_t = [\Pi - \delta - r]K_t dt + \sigma K_t dB_t \tag{3.1}$$

where $(B_t)_{t\geq 0}$ is distributed according to standard Brownian motion and defined on a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$. The (exogenous) constant Π in (3.1) may be interpreted as the marginal product of capital and will be made endogenous in Section 3.2, while δ denotes the rate of capital depreciation and r the rate at which the principal discounts. The principal is risk neutral and so their preferences over allocations

[3] The problem considered here is simpler in one respect because savings are observable. However, as Sannikov and Di Tella (64) observe, the principal's problem often gives infinite profits in the absence of hidden savings. Rather than allowing for hidden savings I instead assume the agent may abscond with delegated capital and thereafter trade only a risk-free bond.

$(K, c) := (K_t, c_t)_{t\geq 0}$ are represented by the function

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

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

$$dY_t = [\Pi - \delta - r - a_t]K_t dt + \sigma K_t dB_t. \tag{3.2}$$

The agent may only consume a fraction ϕ of the diverted output $a_t K_t$, where $\phi \in (0, 1)$ is an exogenous constant. The parameter ϕ may be thought of as a measure of the severity of the agency problem and will play an important role in the decentralization. The specification in (3.2) may be interpreted as the continuous-time limit of the following discrete-time environments: the principal delegates resources to the agent, investment is *publicly* observed but output is subject to idiosyncratic shocks that are *privately* observed.[4] In addition to the unobservability of consumption described above, I will also assume that the consumer may at any time take a fraction ι of the capital delegated to him and abscond, and after doing so may only trade the same risk-free bond to which the principal has access.

Since consumption is unobservable, allocations (K, c) must be incentive compatible. An allocation is incentive compatible if the agent wishes to set $a_t = 0$ for all $t \geq 0$ and all histories and at no point wishes to abscond with the delegated capital. To define the notion of incentive compatibility precisely, let the underlying probability space be $(C[0, \infty), \mathcal{F}_t, P)$, where $\mathcal{F} := (\mathcal{F}_t)_{t\geq 0}$ is the σ-algebra generated by the evaluation maps[5] $(x_t)_{t\geq 0}$ and P is Wiener measure.

Definition 3.1.1. An allocation chosen by the planner is a pair (K, c) of $\mathcal{F}$-adapted processes on $C[0, \infty)$. An agent's strategy is a single $\mathcal{F}$-adapted process a defined on $C[0, \infty)$.

When the agent varies a they alter the law of motion of output and so change the

[4]Further discussion on the relation with discrete-time models is given in Appendix B.1.
[5]Defined by $x_t(\omega) := \omega(t)$ for all $\omega \in C[0, \infty)$ and $t \geq 0$.

measure used to evaluate the random variables $(K_t, c_t)_{t\geq 0}$. Denote the measure associated with strategy a by P^a and the corresponding expectation operator by $\mathbb{E}^a$ and note that the utility from adhering to such a strategy is

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

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

$$W_t := \rho \int_t^\infty e^{-\rho(s-t)}\mathbb{E}^a[\ln(c_s + \phi a_s K_s)|\mathcal{F}_t]ds. \tag{3.3}$$

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

Lemma 3.1.1. *When the agent absconds with K units of capital the utility from having access to a bond market with return rate r is given by $W = \ln K + \ln \rho + [r - \rho]/\rho$.*

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

$$k_t \leq [\iota\rho]^{-1} \exp(1 - r/\rho) \exp W_t =: \omega \exp W_t \tag{3.4}$$

for all $t \geq 0$ almost surely. An allocation is incentive compatible if the agent neither wishes to abscond with the delegated capital nor divert output to private consumption. The formal definition is as follows.

Definition 3.1.2. An allocation (K, c) is incentive compatible if

$$U^A(K, c; a) \geq U^A(K, c; 0)$$

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

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

Definition 3.1.3. Given the utility from the agent's outside option W, the marginal product of capital Π and interest rate r, the problem of the principal is given by

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

As is well-known from the theory of optimal contracting, the principal's problem is naturally recursive in the state variable W, interpreted as promised utility. Standard arguments from the continuous-time contracting literature[6] then ensure that promised utility follows a diffusion process with volatility at least as large as the marginal benefit of diverting output to private consumption. Specifically, the requirement that the allocation (K, c) be incentive compatible may be replaced by the explicit law of motion

$$dW_t = \rho(W - \ln c_t)dt + \rho\phi\sigma(k_t/c_t)dB_t. \tag{3.5}$$

Note that the drift term in (3.5) is the law of motion of W_t that would obtain in the absence of any uncertainty, as can be seen by simply differentiating the expression (3.3) with respect to time. The value function for the principal then solves the following Hamilton-Jacobi-Bellman equation

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

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

Theorem 3.1.2. *The problem of the principal is finite-valued for all sufficiently small* Π.

[6] See e.g. Sannikov (63) and Di Tella and Sannikov (64).

Further, if the no-absconding constraint does not hold with equality and $\Pi < \delta + \rho + \sqrt{\rho}\phi\sigma/2$, *then the value function of the principal is of the form* $V(W) = -\Omega(\Pi)\exp W$ *where*

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

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

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

and for brevity of notation I have defined

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

Proof. See Appendix B.3. □

Corollary 3.1.3. *When the no-absconding constraint does not hold with equality, the volatility of the consumption of the agent is given by* $\sqrt{\rho}h(\Pi)$.

Proof. This is immediate from the policy functions given in Theorem 3.1.2 and the expression (3.5) for the law of motion of promised utility. □

Corollary 3.1.3 establishes the intuitive claim that the principal is willing to assign more risk to the agent when the marginal product of capital is high. This hints at the results of Section 3.2 relating technological parameters with long-run inequality, as the latter is partly determined by the amount of risk to which each entrepreneur is exposed. Forces that increase the marginal value to society of an additional entrepreneur will therefore also tend to increase inequality in the efficient allocation.

As mentioned in the introduction, a large literature has analyzed constrained-efficient allocations in dynamic environments with privately-observed labour productivity shocks. An important result in this literature, established by Rogerson (59) in a principal-agent setting and extended to a dynamic Mirrleesian environment by Golosov, Kocherlakota and

Tsyvinski (28), is the inverse Euler equation for intertemporal distortions. The following shows that this result carries over to the model of the current environment as long as the no-absconding constraint does not hold with equality.

Corollary 3.1.4. *The stochastic process $(e^{(\rho-r)t}c_t)_{t\geq 0}$ follows a martingale in any efficient allocation when the no-absconding constraint does not hold with equality. In other words, the inverse Euler equation holds.*

Proof. From 3.1.2 the law of motion of W_t is

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

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

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

Using the expressions for $\underline{c}$ and $\underline{k}$ found in 3.1.2, the drift μ_C of C_t becomes

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

as desired. □

The proof of Corollary 3.1.4 clearly relies on knowledge of the explicit form of the policy functions derived in Theorem 3.1.2. To gain an intuitive understanding for the emergence of the inverse Euler equation it is therefore instructive to relate Corollary 3.1.4 with the more direct arguments of Rogerson (59) and Golosov, Kocherlakota and Tsyvinski (28). Constant-returns-to-scale and homotheticity of preferences imply policy functions of the form $K(W) = \underline{k}\exp W$ and $C(W) = \underline{c}\exp W$ for some $\underline{k}, \underline{c} > 0$. The Hamilton-Jacobi-Bellman equation then reduces to the single equation for the scalar Ω

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

By the expression (3.5) for the evolution of promised utility, each choice of $\underline{c}$ and $\underline{k}$ completely specifies the allocation of consumption and capital for any history of output. Given any (not necessarily optimal) scalars $\underline{c}$ and $\underline{k}$, the associated cost Ω per unit of utility $\exp W$ in consumption terms is given by

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

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

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

Note that the allocations $(\underline{c}(u), \underline{k}(u))$ are analogous to the perturbations considered in Rogerson (59) and Golosov, Kocherlakota and Tsyvinski (28) in the sense that they increase flow utility this instant by u utils whilst preserving both promise-keeping and incentive compatibility. By the assumed optimality of $(\underline{c}, \underline{k})$ and the assumption that the no-absconding constraint does not hold with equality, we must have $\overline{\Omega}'(0) = 0$, where $\overline{\Omega}(u)$ is defined to be

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

Now note that

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

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

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

which is exactly the continuous-time form of the inverse Euler equation. Note that the assumption that the no-absconding constraint does not hold with equality is necessary for the above argument because otherwise the perturbation $(\underline{c}(u), \underline{k}(u))$ will not be incentive compatible for $u > 0$. Indeed, the inverse Euler equation may fail to hold in this case.[7] However, note that even when the no-absconding constraint holds with equality it remains true that $\overline{\Omega}'(0) \geq 0$ and so we still have the inequality $\mu_C \leq r - \rho$.

Section 3.3 shows how a class of stationary efficient allocations may be decentralized in a general equilibrium model using a particular set of taxes and transfers. Such a characterization is necessarily specific to the choice of Pareto weights attached to different generations, the set of instruments available to the government and the assumed market structure. To isolate the role of agency frictions independently of general equilibrium effects, it is instructive to first analyze efficient distortions by comparing the solution to the above principal-agent problems with the allocations that arise when the agent may invest in either capital or the risk-free bond available to the principal. To motivate the definition of intertemporal wedges adopted below, first note that if an agent may invest in an asset with (gross) return R over the interval $[t, t+\Delta]$, then intertemporal optimization implies

$$u'(c_t) = \exp(-\rho\Delta)\mathbb{E}[Ru'(c_{t+\Delta})|\mathcal{F}_t]. \tag{3.6}$$

The intertemporal wedges defined in Definition 3.1.4 measure the extent to which the relation (3.6) fails for an arbitrary stochastic return.

Definition 3.1.4. Given the consumption process $(c_t)_{t\geq 0}$ in the principal-agent problem, for each asset A with return process $(R^A_t)_{t\geq 0}$ define the associated wedge ν^A implicitly by

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

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

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

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

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

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

coincides with the solution to the principal-agent problem. As such, they represent the extent to which the presence of private information forces the technological returns on each asset to differ from the returns accruing to the agent. Combining Corollary 3.1.4 with Corollary 3.1.3 shows that consumption may be written explicitly as

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

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

Lemma 3.1.5. *The intertemporal wedges for capital and the risk-free bond are given by*

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

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

Proof. See Appendix B.3. □

Figure 3.1 depicts the intertemporal wedges for both risky capital and the risk-free bond as a function of the marginal product Π, given the parameters:

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

As noted in Lemma 3.1.5, the wedge on risky capital is everywhere below that on the risk-free bond and may in fact be negative. However, I will show in Section 3.3 that these wedges

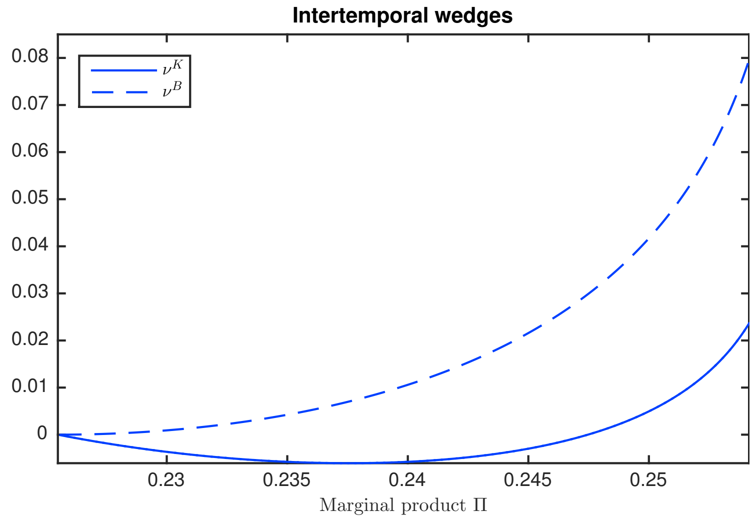

Figure 3.1: Intertemporal wedges

do not translate immediately into taxes in the decentralization of the efficient allocation. Indeed, although the wedge on the return on the risky asset everywhere exceeds that of the safe return, it does not follow that the tax on savings must exceed the tax on profits.

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

Lemma 3.1.6. *The no-absconding constraint will hold as a strict inequality whenever* $\underline{k}(\Pi) < \omega$ *and*

$$0 > \max_{c \geq 0} \ (\rho \ln c + r)\Omega(\Pi) - c + (\Pi - \delta - r)\omega - \frac{1}{2}\Omega(\Pi)[\rho\phi\sigma\omega]^2 c^{-2}. \tag{3.7}$$

Proof. See Appendix B.3. □

Note that the condition (3.7) is necessarily violated for sufficiently large ω. For instance, if one sets $c = (\Pi - \delta - r)\omega$ then the right-hand side of (3.7) becomes

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

which diverges to infinity as ω becomes arbitrarily large. It is also very important to note that although there is a range of parameters such that the no-absconding constraint does not hold with equality, such a constraint is *always* necessary to ensure that the problem of the principal is finite-valued.[8] This is easy to see by setting $\underline{k}(\omega) = \omega$ and $\underline{c} = \underline{c}(\omega)$ solving

$$\rho \underline{c}(\omega)^2 \ln \underline{c}(\omega) = \frac{1}{2}(\rho\phi\sigma\underline{k})^2 = \frac{1}{2}(\rho\phi\sigma\omega)^2 \tag{3.8}$$

and noting that the associated cost coefficient is

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

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

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

3.2 Stationary efficient allocations

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

[8]This was observed by Di Tella and Sannikov (64), who show that without the no-absconding constraint the principal's problem is well-defined if and only if the elasticity of intertemporal substitution exceeds 2 (and here it is 1).

previous chapter, and for this reason some details will be omitted.

Formal setup

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

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

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

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

where $B := (B_t)_{t\geq 0}$ is a standard Brownian motion, $A > 0$ and $\alpha \in (0,1)$ are exogenous constants and δ the depreciation rate. An allocation is now indexed by an entire initial distribution Φ over pairs (v, i) of promised utility and type, where $i \in \{E, W\}$ denotes whether or not an agent is born with the ability to be an entrepreneur or worker. The formal definition is then the following.

Definition 3.2.1. Given a distribution Φ over promised utility and types, an allocation A

consists of consumption, capital assignments and labour assignments

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

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

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

Definition 3.2.2. Given a distribution Φ over types and promised utility v, an allocation A satisfies promise-keeping if $U\left(c^{v,i}\right) = v$ for all $(v,i) \in \text{supp}(\Phi)$. An allocation is incentive compatible if it satisfies promise-keeping and the incentive compatibility conditions of the previous section are satisfied.

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

Definition 3.2.3. An allocation A is resource feasible if $C_t(A) \leq Y_t(A)$ and $L_t(A) \leq L_W$ for all $t \geq 0$. The set of such allocations will be denoted $\mathcal{A}^{RF}$. An allocation is incentive feasible given Φ if it is both resource feasible and incentive compatible given Φ. The set of all such allocations will be denoted $\mathcal{A}^{IF}(\Phi)$.

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

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

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

Definition 3.2.4. Given an initial distribution Φ, the problem of the planner is given by

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

The problem defined in Definition 3.2.4 is intractable for an arbitrary initial distribution so I will restrict attention to solutions in which aggregate consumption and output are constant over time. I will characterize such distributions using the method outlined in Farhi and Werning (24) and employed in the first chapter of this thesis, and consider, in succession, *relaxed* and *generational* planner problems. The relaxed planner problem has the same objective and state variable as in the above planner problem, but allows for intertemporal trade in goods and labour at the subjective rate of discount ρ.

Definition 3.2.5 (Relaxed planner's problem)**.** Given an initial distribution Φ over promised utility and types, the relaxed planner's problem is defined to be

$$V^R(\Phi) = \max_{A \in \mathcal{A}^{IC}(\Phi)} U^P(A)$$
$$\int_0^\infty e^{-\rho t}[C_t(A) - Y_t(A)]dt \leq 0.$$
$$\int_0^\infty e^{-\rho t}[L_t(A) - L_W]dt \leq 0.$$

Note that if an allocation solves the relaxed planner problem and the associated implied distributions of promised utility and types are constant over time at Φ, then this allocation also solves the original planner problem given the distribution Φ. Further, it is easy to see that the subjective rate of discount ρ is the only intertemporal price for which such stationarity may arise, for all other prices would induce an increasing or decreasing trend in utility across generations. Lagrange's Theorem implies that there exists a pair of multipliers $\lambda := (\lambda_R, \lambda_L)$ such that the allocation A that solves the relaxed planner's problem maximizes

[9] Again formal definitions are given in Appendix B.4.

the Lagrangian

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

where the triples $(\underline{C}_t, \underline{Y}_t, \underline{L}_t^E)$ and $(C_t^T, Y_t^T, L_t^{E,T})$ refer to consumption, output and labour assignments of initial and Tth generations, respectively, at date $t \geq 0$.[10] Although the state variable is still an entire distribution, the objective $V_\lambda(\Phi)$ may be maximized pointwise because all interdependence across agents is captured by the multipliers. One may then treat each generation in isolation and vary the multipliers λ_L and λ_R until the resource constraints hold in the implied stationary distribution. I will refer to the problem of dealing with a single generation of newborns as the *generational planner problem.*

Definition 3.2.6. The problem of a generational planner given multipliers $\lambda := (\lambda_R, \lambda_L)$ is defined to be

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

Now note that the choice of assigning labour to entrepreneurs is purely static and depends solely upon the multiplier λ_L. Conditional on assigning a newborn to be an entrepreneur, the problem of the generational planner is equivalent to the principal-agent problem analyzed Section 3.1 with the marginal product of capital now a function of the multiplier on the labour resource constraint.

Lemma 3.2.1. *Given the multipliers $\lambda := (\lambda_R, \lambda_L)$, the problem of the generational planner may be written*

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

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

[10] Detailed definitions are found in Appendix B.4.

I will assume that the weight on the utility of workers is sufficiently large that the type-revelation constraint $W_E \geq W_W$ holds with equality. Each choice of λ_R then corresponds to a level of utility promised to newborns and each choice of λ_L corresponds to a level of the marginal product of capital $\Pi \equiv \Pi(\lambda_L)$ in the principal-agent problem. One may then solve for the stationary distributions associated with each pair of multipliers by using the policy functions found in Section 3.1. Theorem 3.2.2 is the first main result of this chapter. It shows that determining the multipliers for which stationarity obtains reduces to finding an appropriate level of the marginal product of capital Π.

Theorem 3.2.2. *The stationary level of* Π *is the solution to the equation*

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

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

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

Proof. We wish to characterize the values of λ_L and λ_R such that the labour and goods resource constraints are satisfied in the stationary distributions implied by the solutions to the generational planner problem. First note that by 3.1.4 the process $(\exp W_t)_{t\geq 0}$ has no drift, and so its mean is simply $\exp W$, where W is the initial level of utility. It follows that aggregate capital may be written

$$K = L_E\underline{k}(\Pi(\lambda_L))\exp W. \tag{3.9}$$

Combining (3.9) with the static labour assignment function $l(\lambda_L) = [A(1-\alpha)]^{1/\alpha}\lambda_L^{-1/\alpha}$ the labour and goods resource constraints become

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

Simplifying the above resource constraints gives the desired equation for Π. The expressions for the multiplier and the stationary level of the capital stock then follow by combining (3.9) with the static labour assignment function. □

Note that the simplicity of the characterization given in Theorem 3.2.2 is due partly to the assumption of logarithmic preferences and partly to the welfare criterion adopted in this chapter that weights flow utility of an agent the same independently of their birth date. Other papers in the literature on dynamic economies with private information, such as Atkeson and Lucas (10) or Phelan (54), consider component planner problems similar to the above generational planner problems, but adopt a welfare criterion with either zero discounting or place weight solely upon the first generation. Such an approach necessitates solving a component planning problem for an arbitrary interest rate which is then varied until resources are balanced. In contrast, with the welfare criterion adopted here is it immediate that the only intertemporal rate for which stationarity may arise is the subjective discount rate ρ of the agents, for all other intertemporal rates would induce a trend in utility. Together with the assumption of logarithmic utility and the implication derived in Corollary 3.1.4 that consumption then follows a martingale, this implies changes in technology have no effect on the trend in consumption and affect the constrained-efficient allocation only insofar as they alter the marginal product of capital.

Now, the homogeneity of the planner's policy functions for both capital and consumption allows for a simple characterization of the stationary distribution of consumption associated with the above constrained-efficient allocation. In general, the stationary distribution of a killed geometric Brownian motion will depend on the average growth rate, the volatility and the hazard rate of death. Since the rate of death is exogenous, and Corollary 3.1.4 shows the growth rate of consumption to be zero independent of all technological parameters, it follows that the stationary distribution is solely determined by the risk $\sqrt{\rho}h(\Pi)$ borne by entrepreneurs. When combined with the defining equation for Π given in Theorem 3.2.2 this in turn allows us to determine how changes in technological parameters affect efficient long-run inequality. First note that by (3.9) and the expression for capital in Theorem 3.2.2 the initial level of the consumption of entrepreneurs is given by

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

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

Corollary 3.2.3. *The stationary distribution of the consumption of entrepreneurs associated with the constrained-efficient allocation has density given by*

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

with the exponents $\beta_\pm$ given by

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

where $\Pi = \Pi(\lambda_L)$ is the profit level given in 3.2.2, and D_1 and D_2 are determined by the requirement that the density integrate to unity and be continuous.

Figure 3.2 depicts the stationary distribution of consumption for the following parameters:

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

In this example, the exponent α, discount factor ρ, depreciation δ and fraction of entrepreneurs $L_E/(L_E + L_W)$ are taken from Cagetti and De Nardi (16), while the death

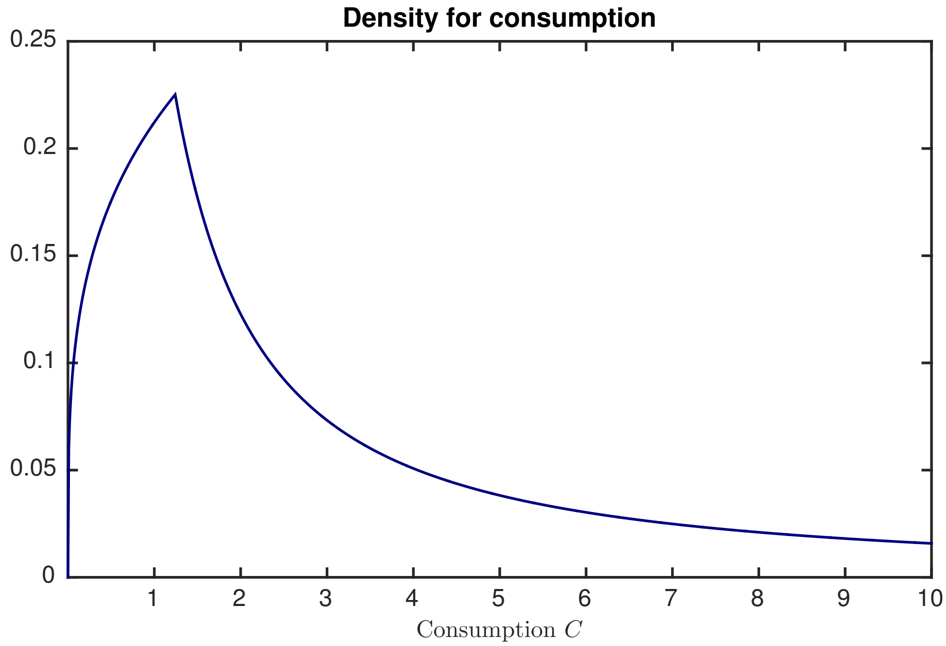

Figure 3.2: Density of consumption

rate ρ_D is chosen such that the average lifespan is 45 years (interpreted as working lifespan). The volatility term σ^2 is the average of the two values considered by Angeletos (9). Note that the upper tail of the stationary distribution in this example has Pareto exponent $\beta_- \approx -1.47$, which is close to the exponent $\beta_- \approx -1.54$ of the wealth distribution US estimated by Gabaix, Moll, Lasry and Lions (27) from the 2010 wave of the Survey of Consumer Finances.

Analysis of the defining equation for Π given in 3.2.2 allows us determine how both aggregate capital and the degree of inequality vary in response to changes in technological parameters.

Corollary 3.2.4. *The marginal product of capital* Π *and the Pareto exponent* β_- *of the upper tail in the stationary distribution are both increasing in the degree of the agency problem* $\phi\sigma$*, the ratio* L_W/L_E*, and the factor share* α*.*

Proof. Using the change of variables

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

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

$$L_E + L_W \exp\left(\frac{1}{2} - \left(1 - \sqrt{1 - x^2}\right)x^{-2}\right) = \frac{1}{\alpha}\left(1 - \sqrt{1 - x^2}\right)\left(\frac{1}{2} + \frac{(1-\alpha)\delta + \rho}{\sqrt{\rho}\phi\sigma x}\right)L_E. \tag{3.10}$$

It is easy to check that the left- and right-hand sides of (3.10) are decreasing and increasing in x, respectively. The claim that Π increases with $\phi\sigma$ then follows from the fact that the right-hand side of (3.10) is decreasing in $\phi\sigma$. Further, since the volatility $\sqrt{\rho}h(\Pi)$ may be written as an increasing function of $x(\Pi)$, this also shows that the Pareto exponent β_- is increasing in $\phi\sigma$.

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

$$0 = \frac{\underline{c}(\Pi) + L_W/L_E}{\underline{k}(\Pi)} - \frac{\Pi}{\alpha} + \delta. \tag{3.11}$$

The right-hand side of (3.11) is decreasing in Π and increasing in both L_W/L_E and α, which shows that the associated roots are also increasing in these variables. The remaining claims then follow from the fact that the volatility $\sqrt{\rho}h(\Pi)$ is increasing in Π. □

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

To illustrate the importance of agency frictions, Figure 3.3 plots the Pareto parameter β_-

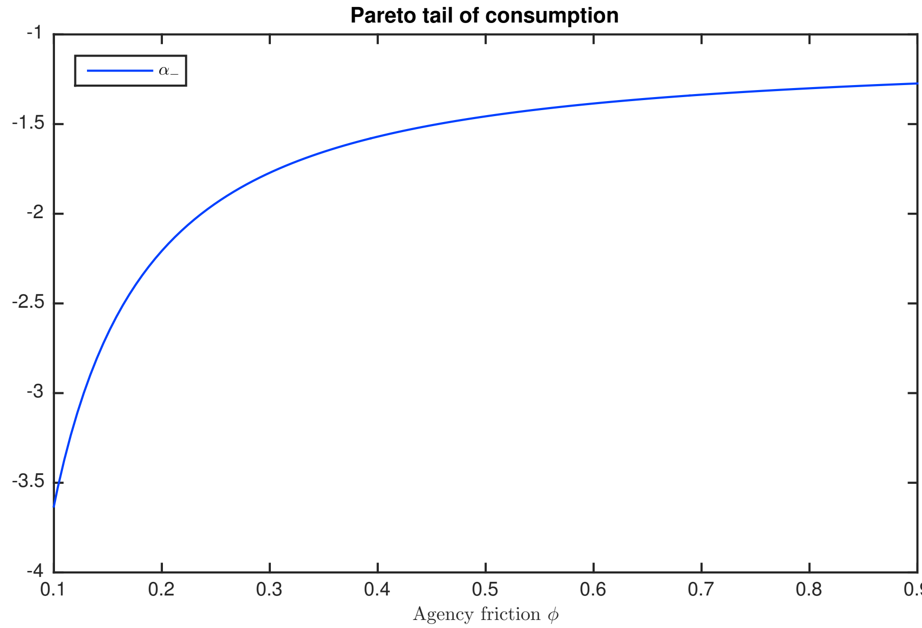

Figure 3.3: Pareto parameter as function of agency problem

as a function of the parameter ϕ, with all other parameters given by the example distribution plotted in Figure 3.2. The domain has been restricted to [0.1 0.9] because the decay in consumption diverges to $-\infty$ as $\phi \to 0$ and the no-absconding constraint begins to bind around 0.9. Similarly, Figure 3.4 plots the Pareto parameter of the upper tail as a function of the ratio L_W/L_E indicating the number of workers per entrepreneur. As with Figure 3.3, all other parameters are fixed at those given in the example depicted in Figure 3.2.

Before turning to the question of decentralization it is useful to summarize the main points of the above characterization. The efficient allocation is completely described by the following requirements: all newborns attain W units of utility, workers and entrepreneurs have zero drift in consumption, and consumption of entrepreneurs has volatility equal to $\sqrt{\rho}h(\Pi)$, where Π is given in Theorem 3.2.2. The task of the next section is to characterize the taxes that ensure these properties arise in a stationary competitive equilibrium.

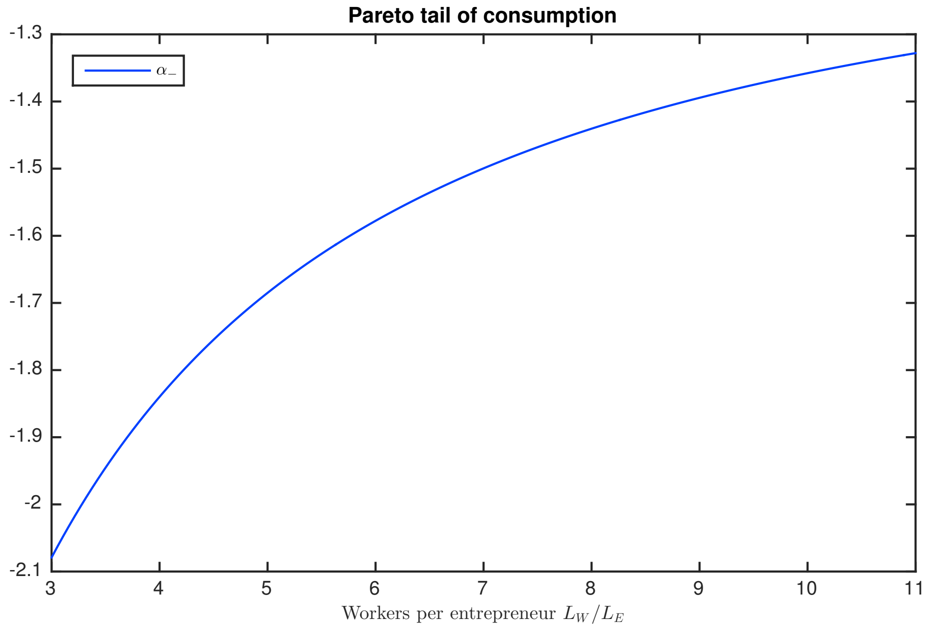

Figure 3.4: Pareto parameter as function of worker per entrepreneur

3.3 Decentralization

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

Market structure and equilibrium characterization

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

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

Definition 3.3.1. Given taxes $\tau_{LW}, \tau_{sW}, \tau_{sE}$ and τ_Π on labour income, risk-free savings and profits, the wage w, and risk-free rate r, the problem of an entrepreneur is given by

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

while the problem of a worker is given by

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

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

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

$$\Delta\left(\max_{l\geq 0} Ak_n^{\alpha}l^{1-\alpha} - wl\right) + (1-\delta\Delta)k_n - (1+\Delta r)k_n = \Delta[\Pi - \delta - r]k_n. \tag{3.12}$$

The law of motion in Definition 3.3.1 is then obtained by supposing that in addition to the above, production is also subject to proportional i.i.d. shocks, with the profits tax being levied on output net of rental costs and depreciation.

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

Definition 3.3.2. Given taxes $(\tau_{LW}, \tau_{sW}, \tau_{sE}, \tau_{\Pi})$ on labour income, risk-free savings and profits of entrepreneurs and workers, respectively, a stationary competitive equilibrium consists of inheritance levels η_E and η_W for entrepreneurs and workers, an aggregate capital stock K, wage rate w and risk-free rate r such that agents maximize, the markets for labour, capital and goods clear, and the government budget constraint is satisfied.

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

Lemma 3.3.1. *Given the wage w, risk-free rate r and depreciation δ, and taxes $\tau_{LW}, \tau_{sW}, \tau_{sE}$ and τ_{Π}, the value functions for entrepreneurs and workers are given by*

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

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

$$\mu_E(w,r) = (1-\tau_{sE})r - \rho + \left(\frac{\Pi(w)-\delta-r}{\sigma}\right)^2 \qquad \sigma_E(w,r) = \frac{\Pi(w)-\delta-r}{\sigma} \tag{3.13}$$

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

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

and the policy function for capital is

$$\frac{k_E(a)}{a} = \frac{\Pi(w)-\delta-r}{\sigma^2(1-\tau_\Pi)} =: \overline{k}(w,r). \tag{3.14}$$

Proof. See Appendix B.5. □

Lemma 3.3.1 implies that the laws of motion of wealth for entrepreneurs and workers are given by

$$\begin{aligned} da_t^E &= \mu_E(w,r)a_t^E dt + \sigma_E(w,r)a_t^E dB_t \\ da_t^W &= \mu_W(r)\left(a_t^W + \frac{(1-\tau_{LW})w}{(1-\tau_{sW})r}\right)dt. \end{aligned} \tag{3.15}$$

Theorem 3.3.2 characterizes stationary equilibrium prices and aggregate quantities given arbitrary levels of taxes and inheritance of entrepreneurs.

Theorem 3.3.2. *When the inheritance levels for workers and entrepreneurs are given by η_W and η_E, respectively, and taxes are $(\tau_\Pi, \tau_{sE}, \tau_{sW}, \tau_{LW})$, the wage w and risk-free rate r that prevail in the associated stationary equilibrium solve the pair of equations*

$$\begin{aligned} 1 &= \frac{\rho_D \eta_E L_E \overline{k}(w,r)}{\rho_D - \mu(w,r)} \\ \frac{(1-\tau_{sW})r}{1-\tau_{LW}} &= \frac{\rho(1-\alpha)\Pi(w)}{\Pi(w)-\alpha(\delta+\rho)}. \end{aligned} \tag{3.16}$$

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

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

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

where μ_E is the drift in entrepreneurial wealth given by (3.13). The stationary form of the three market clearing conditions for labour, capital and goods may then be written

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

respectively. Using the above expression for κ, the capital market-clearing equation reduces to the first equation in (3.16). Using this, the goods market clearing condition simplifies to

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

and hence

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

which reduces to the second equality in (3.16). Since $\phi_l(w) = [(1-\alpha)A/w]^{1/\alpha}$, the labour market clearing condition becomes $[(1-\alpha)A/w]^{1/\alpha}K = L_W$ and so the wage is simply the marginal product of labour, $w = (1-\alpha)A[K/L_W]^{\alpha}$, which gives the desired expression for the aggregate capital stock. The government budget constraint is automatically satisfied by Walras' law. $\square$

Decentralization

We now turn to the characterization of those taxes and inheritance share that decentralize the efficient allocation. First recall the constrained-efficient allocation is characterized by six conditions: workers have zero drift in consumption; entrepreneurs have zero drift in consumption; the volatility of the consumption of entrepreneurs is equal to $\sqrt{\rho}h(\Pi)$; workers and entrepreneurs both have lifetime utility given by W in Theorem 3.2.2; and the level of capital (equivalently, marginal product of capital) in the competitive equilibrium coincides with that given in the stationary efficient allocation.

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

$$\begin{aligned} w &= (1-\alpha)A^{\frac{1}{1-\alpha}}[\alpha/\Pi]^{\frac{\alpha}{1-\alpha}} \\ K/L_W &= [A\alpha/\Pi]^{\frac{1}{1-\alpha}}. \end{aligned} \tag{3.17}$$

The following is the second main result of this chapter, after the characterization of the stationary efficient allocation given in Theorem 3.2.2. It characterizes the inheritance level and taxes that decentralize the stationary constrained-efficient allocation found in the previous section. Note the remarkable fact that the tax on profits is independent of all technological and demographic parameters and depends solely upon the parameter ϕ.

Theorem 3.3.3. *The stationary constrained-efficient allocation coincides with the stationary competitive equilibrium allocation with the following policy choices: inheritance levels η_E and η_W are given by*

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

where Π denotes the marginal product of capital that obtains in the stationary efficient allo-

cation and characterized in Theorem 3.2.2, the profits tax is given by $\tau_\Pi = 1 - \phi$*, the tax on the labour income of workers is given by*

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

and the taxes on the savings of workers and entrepreneurs satisfy

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

Finally, the associated wage and interest rate are

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

Proof. Decentralizing the efficient allocation with taxes and inheritance level ultimately amounts to solving a system of eight equations in eight unknowns. The eight equations are the six conditions characterizing the efficient allocation and the two market-clearing conditions from Theorem 3.3.2, while the eight unknowns are the four taxes, two inheritance levels, the interest rate r and wage rate w.

The first observation is that the marginal product of capital must coincide with the level Π in the constrained-efficient allocation, which immediately gives the desired expression for the wage w. From Lemma 3.3.1 we know that the volatility of entrepreneurs' wealth is independent of taxes and given by

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

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

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

as claimed, and for entrepreneurs this implies

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

which reduces to the claimed expression for τ_{sE}. Now recall that by Theorem 3.2.2 the utility attained in the constrained-efficient allocation is given by

$$W = \ln\left(\frac{K/L_E}{\underline{k}(\Pi)}\right) = \ln\rho + \ln\left(\frac{[K/L_E]\phi\sigma}{\sqrt{\rho}h(\Pi)}\right) - \frac{h(\Pi)^2}{2}. \tag{3.18}$$

From Lemma 3.3.1 if $\mu_E(w, r) = 0$ and $\sigma_E(w, r) = \sqrt{\rho}h(\Pi)$ then the utility entrepreneurs experience from $\eta_E K$ units of wealth is

$$W_{CE} = \ln\rho + \ln(\eta_E K) + \rho^{-1}\left(\mu_E(w, r) - \frac{\sigma_E(w, r)^2}{2}\right) = \ln\rho + \ln(\eta_E K) - \frac{h(\Pi)^2}{2} \tag{3.19}$$

which coincides with (3.18) if and only if η_E is given by the above expression. Now note that if the drift in the wealth of entrepreneurs vanishes, then the capital market clearing condition from Theorem 3.3.2 simplifies to $1 = \eta_E L_E \overline{k}(w, r)$. Using the above expression for the interest rate and the fact that $\sigma(1 - \tau_\Pi)\overline{k}(w, r) = \sigma_E$, this gives

$$\sqrt{\rho}h(\Pi) = \frac{\sigma(1 - \tau_\Pi)}{\eta_E L_E} = \left(\frac{\sigma(1 - \tau_\Pi)}{\phi\sigma}\right)\sqrt{\rho}h(\Pi) \tag{3.20}$$

which implies that $1 - \tau_\Pi = \phi$. By again using the fact that $(1 - \tau_{sW})r = \rho$, the goods market-clearing condition from Theorem 3.3.2 becomes

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

which gives the claimed expression for the labour tax. Finally, we show that for the above taxes entrepreneurs and workers attain the same level of promised utility. First note that by (3.17), the present discounted value of (after-tax) wages simplifies to

$$\frac{(1-\tau_{LW})w}{(1-\tau_{sW})r} = \rho^{-1}(1-\tau_{LW})(1-\alpha)A^{\frac{1}{1-\alpha}}[\alpha/\Pi]^{\frac{\alpha}{1-\alpha}}. \tag{3.21}$$

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

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

and so the value function for the worker simplifies to

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

Using the expression for aggregate capital given in (3.17), individual workers are endowed at birth with $(1-\eta_E L_E)K/L_W = (1-\eta_E L_E)[A\alpha/\Pi]^{\frac{1}{1-\alpha}}$, and so their inherited wealth plus the present discounted value of labour income is

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

Using the defining equation for Π given Theorem 3.2.2 this simplifies to

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

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

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

which coincides with the expression (3.18) for $V_E(\eta_E K)$ given above. □

It is instructive to check that the government budget constraint is automatically satisfied in the equilibrium characterized in Theorem 3.3.3. Summing the tax revenue collected from entrepreneurs and workers gives

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

Using the expressions for the savings taxes given in Theorem 3.3.3 we have

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

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

and so the expression for aggregate taxes becomes

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

as desired. As another plausibility check, consider the behaviour of taxes as the agency problem vanishes $\phi \to 0$. In this case, inspection of the defining equation for Π given in Theorem 3.2.2 shows that $\Pi \approx \delta + \rho$ and $h(\Pi) \approx 0$. The above taxes are then approximately

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

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

It is important to note that the interest rate calculated in the competitive does not coincide with the interest rate given in the relaxed planner problem. As noted in the discussion following Theorem 3.2.2, the only intertemporal rate for the relaxed planner consistent with stationarity is the subjective rate of discount ρ, as all other rates induce a

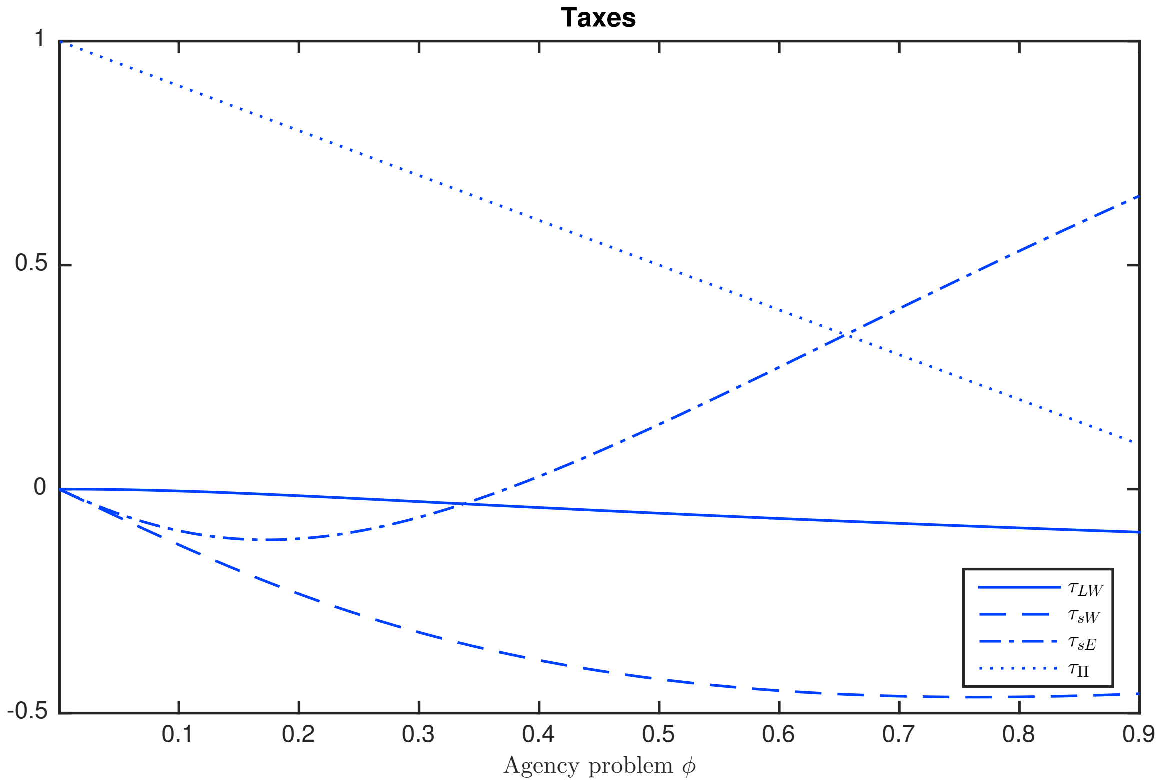

Figure 3.5: Taxes as function of agency problem

trend in consumption across different cohorts. The introduction of an intertemporal price in Section 3.2 is simply a mathematical device to characterize the efficient allocation and does not represent the return on an asset available to any agent. Indeed, using the expression for the interest rate found in Theorem 3.3.3, we can show the following.

Corollary 3.3.4. *The interest rate in the stationary competitive equilibrium that decentralizes the efficient allocation is always lower than the subject rate of discount.*

Proof. See Appendix B.5. □

One consequence of Corollary 3.3.4 is that although the intertemporal wedge in the principal-agent is always positive, it does not necessarily follow that the entrepreneurs must face a positive tax on savings in the decentralization. Figure 3.5 plots the taxes on savings, labour income and profits as a function of the parameter ϕ affecting the degree of the agency

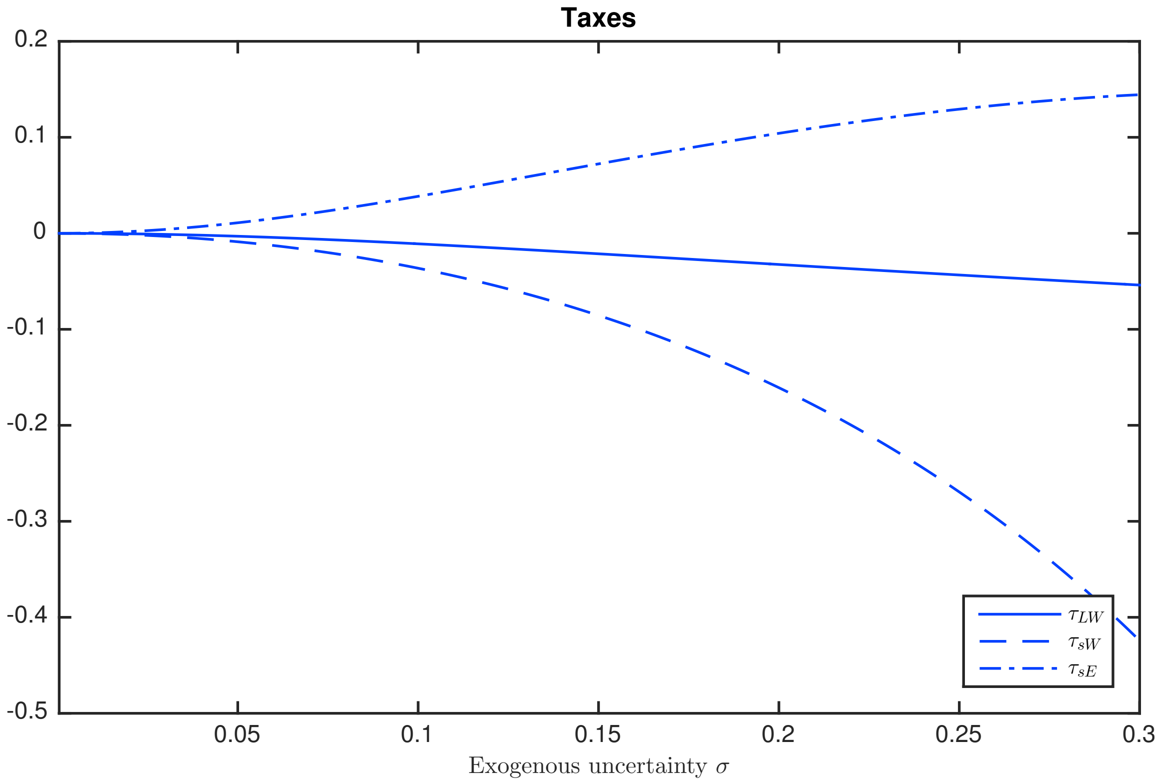

Figure 3.6: Taxes as function of exogenous uncertainty

problem, for the parameters employed in Figure 3.2 (apart from ϕ of course). Notice that for low levels of ϕ entrepreneurs actually face a subsidy on their risk-free savings. Also note that the taxes on labour income and worker savings are negative for all values in the domain.

In contrast, Figure 3.6 plots the taxes on labour income and savings as a function of the level of exogenous uncertainty σ for the same set of parameters as in Figure 3.2. Once again the taxes on workers are everywhere negative. Comparison of Figure 3.6 and Figure 3.5 shows that although changes in ϕ and σ have identical effects on the efficient allocation, they do not have identical effects on taxes.

3.4 Conclusion

This chapter characterizes a class of stationary constrained-efficient allocations in a dynamic economy with physical capital, entrepreneurs, workers and repeated moral hazard. I characterize the inheritance levels and taxes necessary for decentralization in a general equilibrium model. The decentralization is noteworthy both for its simplicity and the invariance of the profits tax to technological changes. All taxes are linear, time-independent, and admit closed-form representations in terms of the root of a single non-linear equation. For simplicity, I have restricted attention to parameters for which the no-absconding constraint holds as a strict inequality and to a particular set of functional forms for both preferences and production.

As with the case of the model in Chapter 2, I have restricted attention to the case of two (exogenous) types and focused exclusively on stationary distributions. In future work I intend to explore the extent to which the findings of this model extend to more a general environment that allow for the emergence long-run growth and incorporates a non-trivial margin for entry into entrepreneurship.

References

[1] Achdou, Yves, et al. Income and wealth distribution in macroeconomics: A continuous-time approach. No. w23732. National Bureau of Economic Research, 2017.

[2] Ai, Hengjie, Dana Kiku, and Rui Li. "Quantifying the impact of moral hazard: Evidence from a structural estimation." (2016).

[3] Ai, Hengjie, Dana Kiku, and Rui Li. "A Mechanism Design Model of Firm Dynamics: The Case of Limited Commitment." (2016).

[4] Ai, Hengjie, 2012, Aggregation in Economies with Homogenous Decision Rules, personal communication.

[5] Albanesi, Stefania. Optimal taxation of entrepreneurial capital with private information. No. w12419. National Bureau of Economic Research, 2006.

[6] Ales, Laurence, Antonio Andres Bellofatto, and Jessie Jiaxu Wang. "Taxing Atlas: Executive compensation, firm size, and their impact on optimal top income tax rates." Review of Economic Dynamics (2017).

[7] Ales, Laurence, Musab Kurnaz, and Christopher Sleet. "Technical change, wage inequality, and taxes." The American Economic Review 105.10 (2015): 3061-3101.

[8] Ales, Laurence, and Christopher Sleet. "Taxing top CEO incomes." The American Economic Review 106.11 (2016): 3331-3366.

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

[10] Andrew Atkeson, Robert E. Lucas, Jr.; On Efficient Distribution With Private Information, The Review of Economic Studies, Volume 59, Issue 3, 1 July 1992, Pages 427–453.

[11] Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez. "Top incomes in the long run of history." Journal of economic literature 49.1 (2011): 3-71.

[12] Autor, David, et al. "Concentrating on the Fall of the Labor Share." American Economic Review 107.5 (2017): 180-185.

[13] Badel, Alejandro, and Mark Huggett. "Human Capital and Economic Opportunity Global Working Group." (2014).

[14] Best, Michael Carlos, and Henrik Jacobsen Kleven. "Optimal Income Taxation with Career Effects of Work Effort." (2013).

[15] Bruggemann, Bettina. "Higher Taxes at the Top: The Role of Entrepreneurs." 2016 Meeting Papers. No. 332. Society for Economic Dynamics, 2016.

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

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

[18] Cooper, Michael, et al. "Business in the United States: Who Owns It, and How Much Tax Do They Pay?." Tax Policy and the Economy 30.1 (2016): 91-128.

[19] Cvitanic, Jaksa, Xuhu Wan, and Jianfeng Zhang. "Optimal compensation with hidden action and lump-sum payment in a continuous-time model." Applied Mathematics & Optimization 59.1 (2009): 99-146.

[20] DeBacker, Jason, Vasia Panousi, and Shanthi Ramnath. A Risky Venture: Income Dynamics within the Non-Corporate Private Business Sector. mimeo, 2014.

[21] Diamond, Peter, and Emmanuel Saez. “The case for a progressive tax: from basic research to policy recommendations." The Journal of Economic Perspectives 25.4 (2011): 165-190.

[22] Di Tella, Sebastian, and Yuliy Sannikov. “Optimal asset management contracts with hidden savings." (Working Paper, Stanford University) (2016).

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

[24] Farhi, Emmanuel, and Ivan Werning. “Inequality and social discounting." Journal of political economy 115.3 (2007): 365-402.

[25] Farhi, Emmanuel, and Ivan Werning. “Insurance and taxation over the life cycle." Review of Economic Studies 80.2 (2013): 596-635.

[26] Fleming, Wendell H., and Halil Mete Soner. Controlled Markov processes and viscosity solutions. Vol. 25. Springer Science & Business Media, 2006.

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

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

[29] Golosov, Mikhail, and Aleh Tsyvinski. “Optimal taxation with endogenous insurance markets." The Quarterly Journal of Economics 122.2 (2007): 487-534.

[30] Golosov, Mikhail, and Aleh Tsyvinski. “Policy Implications of Dynamic Public Finance." Annual Review of Economics 7 (2015): 147-171.

[31] Golosov, Mikhail, Maxim Troshkin, and Aleh Tsyvinski. “Redistribution and social insurance." The American Economic Review 106.2 (2016): 359-386.

[32] Guvenen, Fatih, and Greg Kaplan. Top Income Inequality in the 21st Century: Some Cautionary Notes. No. w23321. National Bureau of Economic Research, 2017.

[33] Guvenen, Fatih, Greg Kaplan, and Jae Song. The glass ceiling and the paper floor: Gender differences among top earners, 1981-2012. No. w20560. National Bureau of Economic Research, 2014.

[34] Song, Jae, et al. Firming up inequality. No. w21199. National Bureau of Economic Research, 2015.

[35] He, Zhiguo. "Optimal executive compensation when firm size follows geometric brownian motion." The Review of Financial Studies 22.2 (2008): 859-892.

[36] He, Zhiguo. "A model of dynamic compensation and capital structure." Journal of Financial Economics 100.2 (2011): 351-366.

[37] Heathcote, Jonathan, and Hitoshi Tsujiyama. "Optimal income taxation: Mirrlees meets Ramsey." (2015).

[38] Heathcote, Jonathan, Fabrizio Perri, and Giovanni L. Violante. "Unequal we stand: An empirical analysis of economic inequality in the United States, 1967-2006." Review of Economic dynamics 13.1 (2010): 15-51.

[39] Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. "Optimal tax progressivity: An analytical framework." The Quarterly Journal of Economics 132.4 (2017): 1693-1754.

[40] Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. Optimal tax progressivity: An analytical framework. No. w19899. National Bureau of Economic Research, 2014.

[41] Huggett, Mark, Gustavo Ventura, and Amir Yaron. "Sources of lifetime inequality." The American Economic Review 101.7 (2011): 2923-2954.

[42] Jones, Charles I., and Jihee Kim. A Schumpeterian model of top income inequality. No. w20637. National Bureau of Economic Research, 2014.

[43] Kapicka, Marek. “Efficient allocations in dynamic private information economies with persistent shocks: A first-order approach." The Review of Economic Studies (2013): rds045.

[44] Kapicka, Marek. Pareto Efficient Taxation with Learning by Doing. Working Paper, 2014.

[45] Kapicka, Marek, and Julian Neira. “Optimal taxation in a life-cycle economy with endogenous human capital formation.” University of California, Santa Barbara WP (2013).

[46] Krueger, Dirk, and Alexander Ludwig. “On the optimal provision of social insurance: Progressive taxation versus education subsidies in general equilibrium." Journal of Monetary Economics 77 (2016): 72-98.

[47] Kushner, Harold, and Paul G. Dupuis. Numerical methods for stochastic control problems in continuous time. Vol. 24. Springer Science & Business Media, 2013.

[48] Lucas Jr, Robert E. “On the size distribution of business firms." The Bell Journal of Economics (1978): 508-523.

[49] Luenberger, David G. Optimization by vector space methods. John Wiley & Sons, 1969.

[50] Luttmer, Erzo GJ. “Selection, growth, and the size distribution of firms." The Quarterly Journal of Economics 122.3 (2007): 1103-1144.

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

[52] Oberman, Adam M. “Convergent difference schemes for degenerate elliptic and parabolic equations: Hamilton–Jacobi equations and free boundary problems." SIAM Journal on Numerical Analysis 44.2 (2006): 879-895.

[53] Panousi, Vasia, and Catarina Reis. “A unified framework for optimal taxation with undiversifiable risk." (2016).

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

[55] Phelan, Christopher, and Robert M. Townsend. "Computing multi-period, information-constrained optima." The Review of Economic Studies 58.5 (1991): 853-881.

[56] Piketty, Thomas. "Capital in the 21st Century." (2014).

[57] Piketty, Thomas, and Emmanuel Saez. "Income inequality in the United States, 1913?1998." The Quarterly journal of economics 118.1 (2003): 1-41.

[58] Prescott, Edward C., and Robert M. Townsend. "Pareto optima and competitive equilibria with adverse selection and moral hazard." Econometrica: Journal of the Econometric Society (1984): 21-45.

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

[60] Rothschild, Casey, and Florian Scheuer. "Redistributive taxation in the roy model." The Quarterly Journal of Economics 128.2 (2013): 623-668.

[61] Rosen, Sherwin. "Authority, control, and the distribution of earnings." The Bell Journal of Economics (1982): 311-323.

[62] Saez, Emmanuel, and Stefanie Stantcheva. A Simpler Theory of Optimal Capital Taxation. No. w22664. National Bureau of Economic Research, 2016.

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

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

[65] Stantcheva, Stefanie. Optimal taxation and human capital policies over the life cycle. No. w21207. National Bureau of Economic Research, 2015.

[66] Scheuer, Florian. “Entrepreneurial taxation with endogenous entry." American Economic Journal: Economic Policy 6.2 (2014): 126-163.

[67] Scheuer, Florian, and Iván Werning. "The Taxation of Superstars." (2016).

[68] Shourideh, Ali. “Optimal taxation of wealthy individuals." 2013 Meeting Papers. No. 261. Society for Economic Dynamics, 2013.

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

Appendix A

Appendix to Chapter 2

A.1 Recursive analysis

In this section I will elaborate on the characterization of incentive compatibility given in Chapter 2. I will then derive the Hamilton-Jacobi-Bellman equations for both the true value function and restricted value functions, show that the latter have closed-form solutions, and derive the Kolmogorov forward equation for the summary measure. Throughout this section I will take the underlying filtered probability space to be $(\Omega, (\mathcal{F}_t)_{t\geq 0}, P)$, where $\Omega = C([0,\infty))$, P is Wiener measure and $\mathcal{F}$ is the σ-algebra generated by the evaluation maps.

A.1.1 Incentive compatibility

Definition A.1.1 (Shifted process)**.** For each effort strategy $(e_t)_{t\geq 0}$ (equivalently, choice of deterministic functions $(\tilde{e}_t)_{t\geq 0}$ defined on $C([0,t])$) the process $Z^e := (Z^e_t)_{t\geq 0}$ is defined pointwise

$$Z^e_t(\omega) := \omega(t) - \sigma^{-1}\int_0^t \iota(\tilde{e}_s((\omega(s'))_{0\leq s'\leq s}))ds$$

for each $\omega \in \Omega := C([0,\infty)$. Note that $dZ^e_t := dZ_t - \sigma^{-1}\iota(\tilde{e}_t)dt$.

In the main text I noted that a change in an agent's strategy may be viewed as a change in the measure they use to evaluate the probability of the occurrence of certain paths of output. This is made formal in the following.

Definition A.1.2 (Effort process and induced measure)**.** An effort process for the agent is a progressively measurable process $e = (e_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, P)$. This is equivalent to the existence of functions $(\tilde{e}_t)_{t\geq 0}$ such that $\tilde{e}_t : C([0,t]) \to \mathbb{R}$ for each $t \geq 0$ and $e_t = \tilde{e}_t((Z_s)_{0\leq s\leq t})$ almost surely, for all $t \geq 0$, where the $(\tilde{e}_t)_{t\geq 0}$ are measurable with respect to the filtration on Ω generated by the evaluation maps. The measure on $\Omega := C([0,t])$ implied by the process e is then defined by declaring

$$P^e(\omega(t_i) \in B_i, i = 1, \ldots, N) = P\left(\int_0^{t_i} \tilde{e}_s((\omega(s'))_{0\leq s'\leq s})ds + \sigma\omega(t_i) \in B_i, i = 1, \ldots, N\right)$$

for any points $t_1, \ldots, t_N \in [0,t]$ and Borel measurable sets $B_1, \ldots, B_N \subseteq \mathbb{R}$.

Definition A.1.2 appears complicated but simply asserts that P^e is the probability measure used by the agent to evaluate the probability of occurrences of different output levels. Note that by the Girsanov theorem, Z^e is a Brownian motion on $(\Omega, \mathcal{F}, P^e)$.

Definition A.1.3 (Time zero utility)**.** For any allocation (c, e^P) chosen by the planner and agent strategy e, define a stochastic process $V^{c,e} := (V_t^{c,e})_{t\geq 0}$ by

$$V_t := \rho \int_0^\infty \mathbb{E}^e\left[\frac{c_t^{1-\gamma}(1-e_t)^{-\alpha}}{1-\gamma}\,\middle|\, \mathcal{F}_t\right] dt. \tag{A.1}$$

One may interpret $V_t^{c,e}$ as the time zero utility based on time t information.

The law of iterated expectations ensures that for any allocation (c, e^P) and strategy e the process $V^{c,e}$ is a martingale on $(\Omega, \mathcal{F}, P^e)$. Now, for the given underlying Brownian motion $Z := (Z_t)_{t\geq 0}$ define θ to be a strong solution to the SDE $d\theta_t = \sigma\theta_t dZ_t$ and note that if $(e_t)_{t\geq 0}$ is of the form $e_t = \tilde{e}_t((Z_s)_{0\leq s\leq t})$ almost surely then we have

$$d\theta_t = \sigma\theta_t dZ_t = \sigma\theta_t\left(\sigma^{-1}\iota(\tilde{e}(\theta.))dt + dZ_t - \sigma^{-1}\iota(\tilde{e}(\theta.))dt\right) = \iota(\tilde{e}(\theta.))\theta_t dt + \sigma\theta_t dZ_t^e$$

where the process Z^e is defined in Definition A.1.1. It follows that (θ, Z^e, P^e) is *always* a *weak* solution to

$$dX_t = \iota(\tilde{e}(X.))X_t dt + \sigma X_t dW_t \tag{A.2}$$

and that choosing a strategy $\tilde{e}$ amounts to choosing the distribution of the process given in (A.2). The Brownian motion $(W_t)_{t\geq 0}$ is the noise process associated with the strategy e.

Proof of 2.2.1. The basic idea of the proof is that for *any* strategy $e := (e_t)_{t\geq 0}$, if we define a random variable X by

$$X := \rho \int_0^\infty e^{-\rho t} \frac{c_t^{1-\gamma}(1-e_t)^{-\alpha}}{1-\gamma} dt \tag{A.3}$$

then the process $V^e := (V_t^e)_{t\geq 0}$ defined in (A.1) is a martingale on $(\Omega, \mathcal{F}, P^e)$ by the law of iterated expectations. The strengthening of the martingale representation theorem given in Lemma 3.1 of Cvitanic, Wan and Zhang (19) then asserts existence of $Y = Y^{e^P, C}$ such that

$$V_t^e = \rho\sigma \int_0^t Y_s dZ_s^e$$

for all $t \geq 0$ almost surely, which gives the result upon rearrangement. □

Remark A.1.1. The strengthening of the martingale representation proved in Cvitanic, Wan and Zhang (19) and invoked above is necessary because $\mathcal{F}$ is the filtration generated by the evaluation maps ω and so is not necessarily equal to the natural filtration associated with Z^e, which is the assumption of the usual martingale representation theorem.

A.1.2 Hamilton-Jacobi-Bellman equation

In this section I will use the preceding recursive analysis to derive the Hamilton-Jacobi-Bellman equation employed in the main text. When preferences are of the form $u(c, e) = c^{1-\gamma}(1-e)^{-\alpha}(1-\gamma)^{-1}$ then from 2.2.2 the sensitivity function is of the form $S_t = \kappa(e_t)u(C_t, e)$ for κ given by

$$\kappa(e) = \frac{\alpha e^{1-\eta}}{\eta\bar{\iota}(1-e)} \tag{A.4}$$

for $e > 0$ $\kappa(0) = 0$ (because the planner may distinguish between $e = 0$ and all other actions). The original Hamilton-Jacobi-Bellman equation is given by

$$\rho V(U,\theta) = \max_{\substack{C\geq 0\\ e\in\mathcal{E}}} \rho(z(e)\theta - C) + \iota(e)\theta\frac{\partial V}{\partial\theta} + \rho\left(U - \frac{C^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right)\frac{\partial V}{\partial U} + \frac{\sigma^2\theta^2}{2}\frac{\partial^2 V}{\partial\theta^2}$$
$$+ \frac{(\rho\sigma\kappa(e))^2}{2}\left(\frac{C^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right)^2\frac{\partial^2 V}{\partial U^2} + \rho\sigma^2\kappa(e)\theta\left(\frac{C^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right)\frac{\partial^2 V}{\partial U\partial\theta}.$$

As shown in 2.2.3, the principal only conditions upon the quantity $u := [(1-\gamma)U]^{\frac{1}{1-\gamma}}\theta^{-1}$. Note that under this transformation, the planner's payoff function is asymptotically linear for both high and low levels, and that the payoff function to the planner confronted with a retired agent of productivity θ with promised utility U may be written

$$V_{\text{ret}}(U,\theta) = \underline{z}\theta - [(1-\gamma)U]^{\frac{1}{1-\gamma}} = (\underline{z} - u)\theta$$

where $\underline{z}$ is the flow payoff per unit of idiosyncratic productivity in retirement and I have used the fact that $e = 0$ is chosen in retirement. If we assume the value function is of the form $V(U,\theta) = v(u)\theta$ then the partial derivatives become

$$u = [(1-\gamma)U]^{\frac{1}{1-\gamma}}\theta^{-1} \qquad U = \frac{(u\theta)^{1-\gamma}}{1-\gamma} \qquad \frac{\partial u}{\partial U} = \frac{u}{(1-\gamma)U} = u^\gamma\theta^{\gamma-1} \qquad \frac{\partial u}{\partial\theta} = -\frac{u}{\theta}.$$

Using the homogeneity property $V(U,\theta) = v(u(U,\theta))\theta$, the first partial derivatives of the original value function are then

$$\frac{\partial V}{\partial U} = \frac{\partial u}{\partial U}v'(u)\theta = \theta^\gamma u^\gamma v'(u) \qquad\qquad \frac{\partial V}{\partial\theta} = v(u) + \frac{\partial u}{\partial\theta}v'(u)\theta = v(u) - v'(u)u$$

while the second partial derivatives are

$$\begin{aligned}
\frac{\partial^2 V}{\partial \theta^2} &= \frac{\partial}{\partial \theta}[v(u) - v'(u)u] = \frac{\partial u}{\partial \theta}[v'(u) - v'(u) - v''(u)u] = \frac{v''(u)u^2}{\theta} \\
\frac{\partial^2 V}{\partial U^2} &= \frac{\partial}{\partial U}[\theta^\gamma u^\gamma v'(u)] = \theta^\gamma \frac{\partial u}{\partial U}[u^\gamma v''(u) + \gamma u^{\gamma-1} v'(u)] = u^{2\gamma-1}\theta^{2\gamma-1}[\gamma v'(u) + u v''(u)] \\
\frac{\partial^2 V}{\partial U \partial \theta} &= \frac{\partial}{\partial U}[v(u) - v'(u)u] = \frac{\partial u}{\partial U}[-v''(u)u] = -v''(u)u^{1+\gamma}\theta^{\gamma-1}.
\end{aligned}$$

Under this change of variables, flow utility may be written

$$\frac{C^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma} = \left(\frac{c^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right)\theta^{1-\gamma}$$

and so the Hamilton-Jacobi-Bellman equation becomes

$$\begin{aligned}
\rho V(U,\theta) = \max_{\substack{C \geq 0 \\ e \in \mathcal{E}}} \; & \rho(z(e)\theta - C) + \iota(e)\theta \frac{\partial V}{\partial \theta} + \rho\left(U - \frac{C^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right)\frac{\partial V}{\partial U} + \frac{\sigma^2\theta^2}{2}\frac{\partial^2 V}{\partial \theta^2} \\
& + \frac{(\rho\sigma\kappa(e))^2}{2}\left(\frac{C^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right)^2 \frac{\partial^2 V}{\partial U^2} + \rho\sigma^2\theta\kappa(e)\left(\frac{C^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right)\frac{\partial^2 V}{\partial U \partial \theta} \\
\rho v(u)\theta = \max_{\substack{c \geq 0 \\ e \in \mathcal{E}}} \; & \rho(z(e) - c)\theta + \frac{\rho}{1-\gamma}\left(u - c^{1-\gamma}(1-e)^{-\alpha}u^\gamma\right)\theta^{1-\gamma}\theta^\gamma v'(u) \\
& + \frac{1}{2}(\rho\sigma\kappa(e))^2\left(\frac{c^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right)^2 u^{2\gamma-1}\theta[\gamma v'(u) + u v''(u)] + \iota(e)\theta[v(u) - v'(u)u] \\
& + \frac{\sigma^2\theta}{2} v''(u)u^2 - \rho\sigma^2\theta\kappa(e)\left(\frac{c^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right) v''(u)u^{1+\gamma} \\
0 = \max_{\substack{c \geq 0 \\ e \in \mathcal{E}}} & \rho(z(e) - c) + (\iota(e) - \rho)v(u) + \frac{\rho}{1-\gamma}\left(u - c^{1-\gamma}(1-e)^{-\alpha}u^\gamma\right)v'(u) \\
& + \frac{1}{2}(\rho\sigma\kappa(e))^2\left(\frac{c^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right)^2 \gamma u^{2\gamma-1} v'(u) - \iota(e)v'(u)u \\
& + \frac{\sigma^2}{2}\left((\rho\kappa(e))^2\left(\frac{c^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right)^2 - 2\rho\kappa(e)\left(\frac{c^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right)u^{1-\gamma} + u^{2-2\gamma}\right)u^{2\gamma}v''(u) \\
0 = \max_{\substack{c \geq 0 \\ e \in \mathcal{E}}} \; & \rho(z(e) - c) + (\iota(e) - \rho)v(u) + \left(\frac{\rho}{1-\gamma}[u - c^{1-\gamma}(1-e)^{-\alpha}u^\gamma] - \iota(e)u\right)v'(u) \\
& + \frac{\gamma(\rho\sigma\kappa(e))^2}{2}\left(\frac{c^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right)^2 u^{2\gamma-1}v'(u) + \frac{\sigma^2}{2}\left(\rho\kappa(e)\left(\frac{c^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right) - u^{1-\gamma}\right)^2 u^{2\gamma}v''(u)
\end{aligned}$$

as asserted in 2.2.5.

A.1.3 Kolmogorov forward equation

In general the Kolmogorov forward equation will inform us of the evolution over time of the two-dimensional density of both promised utility and productivity. However, as shown in Ai (4), the homogeneity of the policy functions and the log-linear law of motion of productivity ensure that we need only solve for the density of a single variable, referred to as the *summary measure*.

Proof of 2.4.1. The process $(\theta_t, u_t)_{t\geq 0}$ is a multidimensional diffusion process driven by the same Brownian motion and so away from the initial point u_0 the joint density Φ satisfies the Kolmogorov forward equation

$$\begin{aligned}\frac{\partial}{\partial t}[\Phi(\theta,u,t)] &= -\rho_D\Phi(\theta,u,t) - \iota(u)\frac{\partial}{\partial\theta}[\theta\Phi(\theta,u,t)] - \frac{\partial}{\partial u}[\mu(u)\Phi(\theta,u,t)] \\ &+ \frac{\sigma_\theta(u)^2}{2}\frac{\partial^2}{\partial\theta^2}[\theta^2\Phi(\theta,u,t)] + \frac{\partial^2}{\partial\theta\partial u}[\theta\sigma_\theta(u)\sigma_u(u)\Phi(\theta,u,t)] + \frac{1}{2}\frac{\partial^2}{\partial u^2}[\sigma_u^2(u)\Phi(\theta,u,t)].\end{aligned}$$

Note that for any smooth function f vanishing at zero with sufficiently rapid decay at ∞, integration by parts gives

$$\begin{aligned}\int_0^\infty \theta f'(\theta)d\theta &= [\theta f(\theta)]_{\theta=0}^\infty - \int_0^\infty f(\theta)d\theta = -\int_0^\infty f(\theta)d\theta \\ \int_0^\infty \theta f''(\theta)d\theta &= -\int_0^\infty f'(\theta)d\theta = 0.\end{aligned}$$

Recalling the definition $m(u,t) := \int_0^\infty \theta\Phi(\theta,u,t)d\theta$ and interchanging orders of integration, it follows that for all (u, t) we have the following simplifications

$$
\begin{aligned}
\iota(u)\int_0^\infty \theta\frac{\partial}{\partial\theta}[\theta\Phi(\theta,u,t)]d\theta &= -\iota(u)m(u,t)\\
-\int_0^\infty \frac{\partial}{\partial u}[\mu(u)\theta\Phi(\theta,u,t)]d\theta &= -\frac{\partial}{\partial u}[\mu(u)m(u,t)]\\
\frac{\sigma_\theta^2(u)}{2}\int_0^\infty \theta\frac{\partial^2}{\partial\theta^2}[\theta^2\Phi(\theta,u,t)]d\theta &= 0\\
\int_0^\infty \theta\frac{\partial^2}{\partial\theta\partial u}[\theta\sigma_u(u)\Phi(\theta,u,t)]d\theta &= -\frac{\partial}{\partial u}[\sigma_u(u)m(u,t)]\\
\int_0^\infty \theta\frac{\partial^2}{\partial u^2}[\sigma_u^2(u)\Phi(\theta,u,t)]d\theta &= \frac{\partial^2}{\partial u^2}[\sigma_u^2(u)m(u,t)].
\end{aligned}
$$

Interchanging the order of integration, the multi-dimensional forward equation implies

$$
\begin{aligned}
\frac{\partial}{\partial t}\int_0^\infty \theta\Phi(\theta,u,t)d\theta = -\rho_D\int_0^\infty \theta\Phi(\theta,u,t)d\theta - \iota(u)\int_0^\infty \theta\frac{\partial}{\partial\theta}[\theta\Phi(\theta,u,t)]d\theta\\
-\frac{\partial}{\partial u}\int_0^\infty \theta\mu(u)\Phi(\theta,u,t)d\theta + \frac{\sigma_\theta(u)^2}{2}\int_0^\infty \theta\frac{\partial^2}{\partial\theta^2}[\theta^2\Phi(\theta,u,t)]d\theta\\
+\sigma_\theta(u)\int_0^\infty \theta\frac{\partial^2}{\partial\theta\partial u}[\theta\sigma_u(u)\Phi(\theta,u,t)]d\theta + \frac{1}{2}\int_0^\infty \theta\frac{\partial^2}{\partial u^2}[\sigma_u(u)^2\Phi(\theta,u,t)]d\theta
\end{aligned}
$$

which is equivalent to

$$
\frac{\partial m}{\partial t} = -(\rho_D - \iota(u))m(u,t) - \frac{\partial}{\partial u}[(\mu(u)+\sigma_\theta(u)\sigma_u(u))m(u,t)] + \frac{1}{2}\frac{\partial^2}{\partial u^2}[\sigma_u^2(u)m(u,t)].
$$

Setting the partial derivative with respect to time to zero then gives (2.21). □

I will now relate the drift and diffusion terms in 2.4.1 with the Hamilton-Jacobi-Bellman equation considered in the main text. The policy functions for the law of motion of θ are given by $\iota_\theta(u) = \iota(e(u))$ and $\sigma_\theta(u) = \sigma(e(u))$ (which is σ if and only if $e \neq 0$). In order to determine the law of u I will apply Ito's lemma to the process

$$
f(U_t, \theta_t) := [(1-\gamma)U_t]^{\frac{1}{1-\gamma}}\theta_t^{-1}.
$$

Proof of 2.2.6. First note that

$$f(x,y) = [(1-\gamma)x]^{\frac{1}{1-\gamma}}y^{-1} \qquad f_1(x,y) = [(1-\gamma)x]^{\frac{\gamma}{1-\gamma}}y^{-1}$$
$$f_2(x,y) = -[(1-\gamma)x]^{\frac{1}{1-\gamma}}y^{-2} \qquad f_{11}(x,y) = \gamma[(1-\gamma)x]^{\frac{\gamma}{1-\gamma}-1}y^{-1}$$
$$f_{22}(x,y) = 2[(1-\gamma)x]^{\frac{1}{1-\gamma}}y^{-3} \qquad f_{12}(x,y) = -[(1-\gamma)x]^{\frac{\gamma}{1-\gamma}}y^{-2}.$$

For ease of notation I will suppress the dependence of κ and ι on u and write $(U_t, \theta_t) = \mu_t dt + G_t dB_t$, where $\mu_t = 2 \times 1$ and G_t is 2×2 are given by

$$\mu_t = (\rho(U_t - u(C_t, e_t)), \iota\theta) \qquad G_t = \begin{bmatrix} \rho\sigma\kappa u(C_t, e_t) & 0 \\ \sigma\theta & 0 \end{bmatrix}$$

and $B_t = (B_{1t}, B_{2t})$ is standard Brownian motion. If H_f is the Hessian of f we want to evaluate $G_t^T H_f G_t$. In general note the following

$$G_t = \begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix} \qquad H_f = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}$$
$$G_t^T H_f = \begin{bmatrix} ah_{11} + bh_{21} & ah_{12} + bh_{22} \\ 0 & 0 \end{bmatrix} \qquad G_t^T H_f G_t = \begin{bmatrix} a^2 h_{11} + 2ab h_{21} + b^2 h_{22} & 0 \\ 0 & 0 \end{bmatrix}$$

where $a = \rho\sigma\kappa C^{1-\gamma}(1-e)^{-\alpha}(1-\gamma)^{-1}$ and $b = \sigma\theta$. Since the Hessian matrix here is

$$\begin{aligned} H_f &= \begin{bmatrix} \gamma[(1-\gamma)U]^{\frac{\gamma}{1-\gamma}-1}\theta^{-1} & -[(1-\gamma)U]^{\frac{\gamma}{1-\gamma}}\theta^{-2} \\ -[(1-\gamma)U]^{\frac{\gamma}{1-\gamma}}\theta^{-2} & 2[(1-\gamma)U]^{\frac{1}{1-\gamma}}\theta^{-3} \end{bmatrix} \\ &= \begin{bmatrix} \gamma[(1-\gamma)U]^{-2} & -[(1-\gamma)U]^{-1}\theta^{-1} \\ -[(1-\gamma)U]^{-1}\theta^{-1} & 2\theta^{-2} \end{bmatrix} u \end{aligned}$$

we have

$$
\begin{aligned}
\frac{1}{2}\mathrm{Tr}[G_t^T H_f G_t] &= \frac{1}{2}[a^2 h_{11} + 2ab h_{21} + b^2 h_{22}] \\
&= \frac{u\sigma^2}{2}\left[\gamma\rho^2\kappa^2\left(\frac{C^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right)^2[(1-\gamma)U]^{-2} - 2\rho\kappa\left(\frac{C^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right)[(1-\gamma)U]^{-1} + 2\right] \\
&= \frac{u\sigma^2}{2}\left[\gamma\rho^2\kappa^2\left(\frac{[c/u]^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right)^2 - 2\rho\kappa\left(\frac{[c/u]^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right) + 2\right].
\end{aligned}
$$

Also note that if ∇f is the gradient of f, then

$$
\begin{aligned}
\nabla f^T &= \left[[(1-\gamma)U]^{\frac{\gamma}{1-\gamma}}\theta^{-1}, -[(1-\gamma)U]^{\frac{1}{1-\gamma}}\theta^{-2}\right] = [[(1-\gamma)U]^{-1}u, -u\theta^{-1}], \\
\nabla f^T \mu_t &= \left(\frac{\rho(1-[c/u]^{1-\gamma}(1-e)^{-\alpha})}{1-\gamma} - \iota\right)u \\
\nabla f^T G_t dB_t &= \left[a[(1-\gamma)U]^{-1}u - bu\theta^{-1}\right]dB_{1t} = u\left[\frac{\rho\kappa C^{1-\gamma}(1-e)^{-\alpha}}{(1-\gamma)^2 U} - 1\right]\sigma dB_{1t} \\
&= u\left[\frac{\rho\kappa(c/u)^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma} - 1\right]\sigma dB_{1t}.
\end{aligned}
$$

Writing $U(u) = (c/u)^{1-\gamma}(1-e)^{-\alpha}$ for ease of notation, the law of motion for $f(U_t, \theta_t) = u_t$ is then

$$
\begin{aligned}
df(U_t, \theta_t) &= \left(\nabla f^T \mu_t + \frac{1}{2}\mathrm{Tr}[G_t^T H_f G_t]\right)dt + \nabla f^T G_t dB_t \\
&= \left(\frac{\rho(1-U(u))}{1-\gamma} - \iota(u) + \frac{\sigma^2}{2}\left[\gamma\rho^2\kappa^2\left(\frac{U(u)}{1-\gamma}\right)^2 - 2\rho\kappa\left(\frac{U(u)}{1-\gamma}\right) + 2\right]\right)udt \\
&\quad + u\left[\frac{\rho\kappa U(u)}{1-\gamma} - 1\right]\sigma dB_{1t},
\end{aligned}
$$

as desired. $\square$

The boundary conditions for the ODE in (2.21) are not obvious *a priori*, but we do not need to know them to find the stationary distribution. I will employ the finite-state Markov chain method of Kushner and Dupuis (47) and find the stationary distribution of a finite-state Markov chain 'close' in distribution to the underlying continuous-time process.

A.1.4 Restricted value and policy functions

I will now show the restricted value and policy functions have the form asserted in 2.2.4. First note that when the recommended effort level is fixed independently of the history of output, the present discounted value of output to a planner confronted with an agent of productivity θ is

$$\rho z(e) \int_0^\infty e^{-\rho t}\mathbb{E}[\theta_t]dt = \rho z(e)\theta \int_0^\infty e^{(\iota(e)-\rho)t}dt = \frac{\rho z(e)\theta}{\rho - \iota(e)}$$

which corresponds to the constant term in the statement of 2.2.4. Recalling the notation

$$u(C,e) = \frac{\left(C(1-e)^{\overline{\alpha}}\right)^{1-\gamma}}{1-\gamma} \qquad\qquad \kappa(e) = \frac{e^{1-\eta}}{\eta\bar{\iota}(1-e)}$$

and employing arguments analogous to those given in the main text imply the appropriate Hamilton-Jacobi-Bellman equations (indexed by $e \in (0,1)$) for the cost $F(U)$ of providing U units of utility to the agent are

$$\begin{aligned} \rho F(U) = \max_{C\geq 0}\ & -\rho C + \rho\left(U - \frac{\left[C(1-e)^{\overline{\alpha}}\right]^{1-\gamma}}{1-\gamma}\right)F'(U) \\ & + \frac{(\rho\sigma\overline{\alpha}\kappa(e))^2}{2}[C(1-e)^{\overline{\alpha}}]^{2-2\gamma}F''(U). \end{aligned} \tag{A.5}$$

Into the above notation 2.2.4 asserts that the value and cost function associated with the restricted-action allocations admit closed-form expressions of the form $F(U) = -\chi_1(e)[(1-\gamma)U]^{\frac{1}{1-\gamma}}$ and $C(U) = \chi_2(e)[(1-\gamma)U]^{\frac{1}{1-\gamma}}$ for some $\chi_1(e)$ and $\chi_2(e)$.

Proof of 2.2.4. Assuming a solution of the form $F(U) = -\chi_1(e)[(1-\gamma)U]^{\frac{1}{1-\gamma}}$, the first and second derivatives are

$$F'(U) = -\chi_1[(1-\gamma)U]^{\frac{\gamma}{1-\gamma}} \qquad\qquad F''(U) = -\gamma\chi_1[(1-\gamma)U]^{\frac{\gamma}{1-\gamma}-1}.$$

Writing the policy function as $C(U) = \chi_2[(1-\gamma)U]^{\frac{1}{1-\gamma}}$ for some $\chi_2 \geq 0$ implies $C(U)^{1-\gamma}/(1-$

$\gamma) = \chi_2^{1-\gamma}U$, and so the Hamilton-Jacobi-Bellman equation (A.5) becomes

$$\begin{aligned}
-\rho\chi_1[(1-\gamma)U]^{\frac{1}{1-\gamma}} &= \max_{C\geq 0}\ -\rho C - \rho\chi_1\left(U - \frac{\left(C(1-e)^{\overline{\alpha}}\right)^{1-\gamma}}{1-\gamma}\right)[(1-\gamma)U]^{\frac{\gamma}{1-\gamma}} \\
&\quad - \frac{\chi_1\gamma}{2}(\rho\sigma\overline{\alpha}\kappa(e))^2\left[C(1-e)^{\overline{\alpha}}\right]^{2-2\gamma}[(1-\gamma)U]^{\frac{\gamma}{1-\gamma}-1} \\
-\rho\chi_1 &= \max_{\chi_2\geq 0}\ -\rho\chi_2 - \rho\chi_1\left(\frac{1-[\chi_2(1-e)^{\overline{\alpha}}]^{1-\gamma}}{1-\gamma}\right) \\
&\quad - \frac{\chi_1\gamma}{2}(\rho\sigma\overline{\alpha}\kappa(e))^2\left[\chi_2(1-e)^{\overline{\alpha}}\right]^{2-2\gamma}
\end{aligned}$$

which simplifies to

$$\chi_1 = \min_{\chi_2\geq 0}\ \chi_2 + \chi_1\left(\frac{1-\left[\chi_2(1-e)^{\overline{\alpha}}\right]^{1-\gamma}}{1-\gamma} + \frac{\gamma\Sigma(e)^2}{2\rho}\left[\chi_2(1-e)^{\overline{\alpha}}\right]^{2-2\gamma}\right). \tag{A.6}$$

The closed-form expressions for the coefficients $\chi_1(e)$ and $\chi_2(e)$ may be obtained by combining two pieces of information: the fixed-point equation (A.6) and the first-order condition from the minimization on the right-hand side of (A.6). The first of these gives

$$\chi_1(e) = \chi_2(e)\left(1 - \frac{(1-\left[\chi_2(1-e)^{\overline{\alpha}}\right]^{1-\gamma})}{1-\gamma} - \frac{\gamma\Sigma(e)^2}{2\rho}\left[\chi_2(e)(1-e)^{\overline{\alpha}}\right]^{2-2\gamma}\right)^{-1}.$$

The first-order condition for the minimization gives

$$\begin{aligned}
0 &= 1 + \chi_1(e)\left(-\chi_2(e)^{-\gamma}(1-e)^{\overline{\alpha}(1-\gamma)} + \frac{\gamma\Sigma(e)^2}{\rho}(1-\gamma)\chi_2(e)^{-1}\left[\chi_2(1-e)^{\overline{\alpha}}\right]^{2-2\gamma}\right) \\
\chi_2 &= \chi_1(e)\left(\left[\chi_2(1-e)^{\overline{\alpha}}\right]^{1-\gamma} + \frac{\gamma\Sigma(e)^2}{\rho}(\gamma-1)\left[\chi_2(1-e)^{\overline{\alpha}}\right]^{2-2\gamma}\right).
\end{aligned}$$

Equating the two expressions for $\chi_2(e)/\chi_1(e)$ implied by the above gives

$$\begin{aligned}
&-\frac{\gamma}{1-\gamma} + \frac{\left[\chi_2(e)(1-e)^{\overline{\alpha}}\right]^{1-\gamma}}{1-\gamma} - \frac{\gamma\Sigma(e)^2}{2\rho}\left[\chi_2(1-e)^{\overline{\alpha}}\right]^{2-2\gamma} \\
&= \left[\chi_2(1-e)^{\overline{\alpha}}\right]^{1-\gamma} + \frac{\gamma\Sigma(e)^2}{\rho}(\gamma-1)\left[\chi_2(1-e)^{\overline{\alpha}}\right]^{2-2\gamma}
\end{aligned}$$

which simplifies to

$$\frac{\rho}{\gamma-1}\left[\chi_2(1-e)^{\overline{\alpha}}\right]^{1-\gamma}+\Sigma(e)^2(\gamma-1/2)\left[\chi_2(1-e)^{\overline{\alpha}}\right]^{2-2\gamma}=\frac{\rho}{\gamma-1}.$$

Setting $x := \left[\chi_2(e)(1-e)^{\overline{\alpha}}\right]^{1-\gamma}$ leads to the quadratic

$$0=\Sigma(e)^2(\gamma-1/2)(\gamma-1)x^2+\rho x-\rho. \tag{A.7}$$

The positive root of (A.7) (the only sensible one, since consumption cannot be negative) is given by

$$x(e,\gamma)=\frac{\sqrt{1+2\rho^{-1}(2\gamma-1)(\gamma-1)\Sigma(e)^2}-1}{\rho^{-1}(2\gamma-1)(\gamma-1)\Sigma(e)^2}. \tag{A.8}$$

Notice that the variable x is simply the ratio of flow utility to total utility, and equals unity when $\Sigma(e)$ (as would be the case if $\sigma=0$). The expression (A.8) implies that both the value function and consumption function admit the claimed forms with

$$\begin{aligned}\chi_2(e)&=x(e,\gamma)^{\frac{1}{1-\gamma}}(1-e)^{-\overline{\alpha}}\\ \chi_1(e)&=\chi_2(e)\left(\frac{2\gamma-1}{2\gamma-x(e,\gamma)}\right)=x(e,\gamma)^{\frac{1}{1-\gamma}}(1-e)^{-\overline{\alpha}}\left(\frac{2\gamma-1}{2\gamma-x(e,\gamma)}\right)\end{aligned}$$

as desired. □

None of the restricted payoff functions turn out to be fully optimal when the planner may alter the recommended throughout the course of an agent's life. However, in numerical experiments the pointwise maxima of the above functions appear to be quite close to the true optimal function, and so serves both as a good initial guess for the numerical approximation and a guide for our intuition.

For convenience I record the following lemma that shows killed diffusion processes are distributed according to a double-Pareto distribution. The proof is standard and therefore omitted. I then specialize to the restricted-action versions of consumption and productivity.

Lemma A.1.1. *The stationary distribution of a stochastic process evolving according to $dX_t = \mu_X X_t dt + \sigma_X X_t dZ_t$ that dies at rate δ_X and is injected at some point $\overline{X} > 0$ has density given by*

$$f(x) = \begin{cases} Ax^{\alpha_+ - 1} & \text{if } x \leq \overline{X} \\ Bx^{\alpha_- - 1} & \text{if } x \geq \overline{X} \end{cases} \tag{A.9}$$

where

$$\alpha_\pm = \overline{\mu}_X/\sigma_X^2 \pm \sqrt{(\overline{\mu}_X/\sigma_X)^2 + 2\delta_X/\sigma_X^2} \tag{A.10}$$

the constants A and B are chosen such that the density is continuous and integrate to unity, and I have written $\overline{\mu}_X = \mu_X - \sigma_X^2/2$ for brevity.

Combining A.1.1 with the explicit policy functions for consumption given in 2.2.4 gives the closed-form representations for the stationary distributions associated with the restricted-action allocations asserted in 2.4.2.

Proof of 2.4.2. In the restricted-action case, the consumption function satisfies

$$C(U,\theta) = c(u;e)\theta = \chi_2(e)u\theta = \chi_2(e)[(1-\gamma)U]^{\frac{1}{1-\gamma}}, \qquad \frac{C(U,\theta)^{1-\gamma}}{1-\gamma} = \chi_2(e)^{1-\gamma}U$$

and so the law of motion of utility is

$$dU_t = (\rho + \rho_D)\big(1 - \chi_2(e)^{1-\gamma}(1-e)^{-\alpha}\big)U_t dt + (\rho+\rho_D)\sigma\kappa(e)\chi_2(e)^{1-\gamma}U_t dZ_t. \tag{A.11}$$

It follows that the law of motion of utility is of the form $dU_t = \mu_U(e)U_t dt + \sigma_U(e)U_t dZ_t$ given in A.1.1, where $\delta_X = \rho_D$ and

$$\begin{aligned} \mu_U(e) &= \rho\Big(1 - \big(\chi_2(e)(1-e)^{\overline{\alpha}}\big)^{1-\gamma}\Big) = \rho(1 - x(e,\gamma)) \\ \sigma_U(e) &= \Sigma(e)(1-\gamma)\big(\chi_2(e)(1-e)^{\overline{\alpha}}\big)^{1-\gamma} = \Sigma(e)(1-\gamma)x(e,\gamma) \end{aligned}$$

where $x(e,\gamma)$ is given in the proof of 2.2.4. Since the defining quadratic for $x(e,\gamma)$ may be rearranged to give

$$\frac{\rho(1-x(e,\gamma))}{x(e,\gamma)^2} = (\gamma - 1/2)(\gamma - 1)\Sigma(e)^2,$$

the above expressions for $\mu_U(e)$ and $\sigma_U(e)$ imply

$$(1-\gamma)\left(\frac{\mu_U(e)}{\sigma_U(e)^2} - \frac{1}{2}\right) = \frac{\rho(1-x(e,\gamma))}{(1-\gamma)\Sigma(e)^2 x(e,\gamma)^2} - \frac{1-\gamma}{2} = -\frac{\gamma}{2}$$

$$\frac{2\rho_D}{\sigma_U(e)^2} = \frac{2\rho_D(2\gamma-1)^2\Sigma(e)^2}{\left(\sqrt{\rho^2 + 2\rho(2\gamma-1)(\gamma-1)\Sigma(e)^2} - \rho\right)^2}.$$

The upper tail is then

$$\begin{aligned}\alpha_-(e,\gamma) &= (1-\gamma)\left(\frac{\mu_U(e)}{\sigma_U(e)^2} - \frac{1}{2} + \sqrt{\left(\frac{\mu_U(e)}{\sigma_U(e)^2} - \frac{1}{2}\right)^2 + \frac{2\rho_D}{\sigma_U(e)^2}}\right)\\ &= -\frac{\gamma}{2} - \sqrt{\frac{\gamma^2}{4} + 2\rho_D\left(\frac{J(e,\gamma)}{\Sigma(e)}\right)^2}\end{aligned}$$

as claimed. □

Proof of Corollary 2.4.3. From the expression

$$\alpha_-(e,\gamma) = -\frac{\gamma}{2} - \sqrt{\frac{\gamma^2}{4} + 2\rho_D\left(\frac{J(e,\gamma)}{\Sigma(e)}\right)^2}$$

we want to determine $\lim_{e\to 1} J(e,\gamma)/\Sigma(e)$. Since

$$\frac{J(e,\gamma)}{\Sigma(e)} = \frac{C\Sigma(e)}{\sqrt{1+2C\Sigma(e)^2} - 1}$$

where $C = \rho^{-1}(2\gamma-1)(\gamma-1)$, this limit is

$$\lim_{e\to 1}\frac{J(e,\gamma)}{\Sigma(e)} = \lim_{x\to\infty}\frac{Cx}{\sqrt{1+2Cx^2}-1} = \lim_{x\to\infty}\frac{\sqrt{1+2Cx^2}}{2x} = \sqrt{C/2}.$$

It follows that the appropriate bound is

$$\begin{aligned}\alpha_-(e,\gamma) &\leq -\frac{\gamma}{2} - \sqrt{\frac{\gamma^2}{4} + \rho_D\rho^{-1}(2\gamma-1)(\gamma-1)} \\ &= -\frac{\gamma}{2}\left(1 + \sqrt{1 + \frac{4\rho_D}{\rho}(2-1/\gamma)(1-1/\gamma)}\right)\end{aligned}$$

as desired. □

A.2 Numerical schemes

This section outlines how I numerically solve the ordinary differential equations derived in the main text.

A.2.1 Hamilton-Jacobi-Bellman equation

The principal-agent problems considered in the main text may solved in one of two ways: the first is the explicit finite-difference method of Oberman (52) and the second is the implicit method employed in Achdou, Han, Lasry, Lions and Moll (1) in the context of an incomplete markets model. The first is guaranteed to converge but slow, while the second is fast but not guaranteed to converge. I will adopt the implicit finite-scheme approach, but choose forward and backward differences based on arguments employed in Oberman (52), which dealt only with explicit schemes. Recall the Hamilton-Jacobi-Bellman equation is

$$\begin{aligned}0 = \max_{\substack{c\geq 0\\ e\in\mathcal{E}}}\ &\rho(z(e)-c) + (\iota(e)-\rho)v(u) + \left(\frac{\rho}{1-\gamma}[u - c^{1-\gamma}u^{\gamma}(1-e)^{-\alpha}] - \iota(e)u\right)v'(u) \\ &+ \frac{\gamma(\rho\sigma\kappa(e))^2}{2}\left(\frac{c^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right)^2 u^{2\gamma-1}v'(u) + \frac{\sigma^2}{2}\left(\rho\kappa(e)\frac{c^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma} - u^{1-\gamma}\right)^2 u^{2\gamma}v''(u).\end{aligned}$$

In the notation of Oberman (52) I seek a solution to the functional equation

$$F(u, v(u), v'(u), v''(u)) \equiv 0,$$

where F is defined by

$$F(u,v,v',v'') := \min_{\substack{x\geq 0\\ e\in\mathcal{E}}} \ \rho(c-z(e)) - (\iota(e)-\rho)v - \frac{\rho}{1-\gamma}[u - c^{1-\gamma}u^{\gamma}(1-e)^{-\alpha}]v' + \iota(e)uv'$$
$$- \frac{\gamma(\rho\sigma\kappa(e))^2}{2}\left(\frac{c^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma}\right)^2 u^{2\gamma-1}v' - \frac{\sigma^2}{2}\left(\rho\kappa(e)\frac{c^{1-\gamma}(1-e)^{-\alpha}}{1-\gamma} - u^{1-\gamma}\right)^2 u^{2\gamma}v''.$$

Oberman (52) shows that for sufficiently small time-step, the operator that solves the time-dependent Hamilton-Jacobi-Bellman by moving backwards in time is a contraction as long as the finite-differences are chosen so that F is increasing both in the values and the first differences of the approximate solution. I will adopt his guidance for the use of forward and backwards difference but use an implicit rather than an explicit scheme. The general convergence results in Oberman (52) do not carry over to the implicit case, but the method seems to work well in practice and is significantly faster than the analogous explicit scheme. Given $(v_1^n, \dots, v_I^n)$, the vector $(v_1^{n+1}, \dots, v_I^{n+1})$ is defined implicitly by

$$0 = \frac{v_i^{n+1} - v_i^n}{\Delta_T} + F_i^{n,n+1} \tag{A.12}$$

where $F^{n,n+1}$ means that the F scheme uses the $(n+1)$th values for the derivatives but the optimal choices from the nth step. Notice that as n increases in (A.12) we are moving *backwards* in time. We want a scheme such that the right-hand side of the above is increasing in spatial differences. For $\gamma = 2$ this gives

$$\frac{v_i^n - v_i^{n+1}}{\Delta_T} = \rho(c_i^n - z) + \rho[v_i^{n+1} + u_i v_i^{n+1-}] + \bar{\iota}[e_i^n]^{\eta}[u_i v_i^{n+1-} - v_i^{n+1}]$$
$$+ \rho[c_i^n]^{-1}(1-e_i^n)^{-\alpha}u_i^2 v_i^{n+1+} + u_i^3\left(\frac{\rho\sigma\alpha}{\eta\bar{\iota}}\right)^2\left[\frac{[c_i^n]^{-1}[e_i^n]^{1-\eta}}{(1-e_i^n)^{1+\alpha}}\right]^2 v_i^{n+1+}$$
$$+ \frac{\sigma^2 u_i^2}{2}\left(1 - u_i\left(\frac{\rho\alpha}{\eta\bar{\iota}}\right)\frac{[c_i^n]^{-1}[e_i^n]^{1-\eta}}{(1-e_i^n)^{1+\alpha}}\right)^2\left(\frac{v_i^{n+1+} + v_i^{n+1-}}{\Delta_u}\right)$$

where for each n and i c_i^n and e_i^n solve

$$
\begin{aligned}
&\min_{c,e}\ \rho c + \bar{\iota} e^{\eta}[u_i v_i^{n-} - v_i^n] + \rho c^{-1}(1-e)^{-\alpha} u_i^2 v_i^{n+} + u_i^3 \left(\frac{\rho\sigma\alpha}{\eta\bar{\iota}}\right)^2 \left[\frac{c^{-1}e^{1-\eta}}{(1-e)^{1+\alpha}}\right]^2 v_i^{n+} \\
&\quad + \frac{\sigma^2 u_i^2}{2}\left(u_i^2\left(\frac{\rho\alpha}{\eta\bar{\iota}}\right)^2\left[\frac{c^{-1}e^{1-\eta}}{(1-e)^{1+\alpha}}\right]^2 - 2u_i\left(\frac{\rho\alpha}{\eta\bar{\iota}}\right)\frac{c^{-1}e^{1-\eta}}{(1-e)^{1+\alpha}}\right)\left(\frac{v_i^{n+}+v_i^{n-}}{\Delta_u}\right) \\
&= \min_{c,e}\ \rho c - \bar{\iota} e^{\eta}[v_i^n - u_i v_i^{n-}] + \rho c^{-1}(1-e)^{-\alpha} u_i^2 v_i^{n+} - \frac{\rho u_i^3 \sigma^2 \alpha c^{-1} e^{1-\eta}}{\eta\bar{\iota}(1-e)^{1+\alpha}}\left(\frac{v_i^{n+}+v_i^{n-}}{\Delta_u}\right) \\
&\quad + u_i^3\left[\frac{\rho\sigma\alpha c^{-1}e^{1-\eta}}{\eta\bar{\iota}(1-e)^{1+\alpha}}\right]^2\left[v_i^{n+} + \frac{u_i}{2}\left(\frac{v_i^{n+}+v_i^{n-}}{\Delta_u}\right)\right].
\end{aligned}
$$

Note that this is of the form

$$
H(c,e) := \rho c - C_1 e^{\eta} + C_2 c^{-1}(1-e)^{-\alpha} + C_3 c^{-1} e^{1-\eta}(1-e)^{-1-\alpha} + C_4 c^{-2} e^{2-2\eta}(1-e)^{-2-2\alpha}
$$

where C_1, C_2, C_3 and C_4 are given by

$$
\begin{aligned}
C_1 &= \bar{\iota}[v_i^n - u_i v_i^{n-}] & C_3 &= -u_i^3\left(\frac{\rho\sigma^2\alpha}{\eta\bar{\iota}}\right)\left(\frac{v_i^{n+}+v_i^{n-}}{\Delta_u}\right) \\
C_2 &= \rho u_i^2 v_i^{n+} & C_4 &= u_i^3\left(\frac{\rho\sigma\alpha}{\eta\bar{\iota}}\right)^2\left[v_i^{n+} + \frac{u_i}{2}\left(\frac{v_i^{n+}+v_i^{n-}}{\Delta_u}\right)\right].
\end{aligned}
$$

The first partial derivatives of H are then

$$
\begin{aligned}
H_1 &= \rho - C_2 c^{-2}(1-e)^{-\alpha} - C_3 c^{-2} e^{1-\eta}(1-e)^{-1-\alpha} - 2C_4 c^{-3} e^{2-2\eta}(1-e)^{-2-2\alpha} \\
H_2 &= -\eta C_1 e^{\eta-1} + \alpha C_2 c^{-1}(1-e)^{-\alpha-1} \\
&\quad + C_3 c^{-1}[(1-\eta)e^{-\eta}(1-e)^{-1-\alpha} + (1+\alpha)e^{1-\eta}(1-e)^{-2-\alpha}] \\
&\quad + 2C_4 c^{-2}\big[(1-\eta)e^{1-2\eta}(1-e)^{-2-2\alpha} + (1+\alpha)e^{2-2\eta}(1-e)^{-3-2\alpha}\big].
\end{aligned}
$$

The second partial derivatives are

$$
\begin{aligned}
H_{11} &= 2C_2c^{-3}(1-e)^{-\alpha} + 2C_3c^{-3}e^{1-\eta}(1-e)^{-1-\alpha} + 6C_4c^{-4}e^{2-2\eta}(1-e)^{-2-2\alpha} \\
H_{12} &= -\alpha C_2c^{-2}(1-e)^{-\alpha-1} - C_3c^{-2}[(1-\eta)e^{-\eta}(1-e)^{-1-\alpha} + (1+\alpha)e^{1-\eta}(1-e)^{-2-\alpha}] \\
&\qquad - 4C_4c^{-3}\left[(1-\eta)e^{1-2\eta}(1-e)^{-2-2\alpha} + (1+\alpha)e^{2-2\eta}(1-e)^{-3-2\alpha}\right] \\
H_{22} &= \eta(1-\eta)C_1e^{\eta-2} + \alpha(1+\alpha)C_2c^{-1}(1-e)^{-\alpha-2} + M(c,e)
\end{aligned}
$$

where

$$
\begin{aligned}
M(e,c) &= C_3c^{-1}\frac{d}{de}[(1-\eta)e^{-\eta}(1-e)^{-1-\alpha} + (1+\alpha)e^{1-\eta}(1-e)^{-2-\alpha}] \\
&+ 2C_4c^{-2}\frac{d}{de}\left[(1-\eta)e^{1-2\eta}(1-e)^{-2-2\alpha} + (1+\alpha)e^{2-2\eta}(1-e)^{-3-2\alpha}\right] \\
&= C_3c^{-1}[-\eta(1-\eta)e^{-1-\eta}(1-e)^{-1-\alpha} + 2(1-\eta)(1+\alpha)e^{-\eta}(1-e)^{-2-\alpha} \\
&+ (2+\alpha)(1+\alpha)e^{1-\eta}(1-e)^{-3-\alpha}] + 2C_4c^{-2}[(1-2\eta)(1-\eta)e^{-2\eta}(1-e)^{-2-2\alpha} \\
&+ 4(1-\eta)(1+\alpha)e^{1-2\eta}(1-e)^{-3-2\alpha} + (3+2\alpha)(1+\alpha)e^{2-2\eta}(1-e)^{-4-2\alpha}].
\end{aligned}
$$

The above finite-difference scheme

$$
\begin{aligned}
\frac{v_i^n - v_i^{n+1}}{\Delta_T} &= \rho(c_i^n - z) + \rho[v_i^{n+1} + u_iv_i^{n+1-}] + \bar{\iota}[e_i^n]^\eta[u_iv_i^{n+1-} - v_i^{n+1}] + \rho[c_i^n]^{-1}(1-e_i^n)^{-\alpha}u_i^2v_i^{n+1+} \\
&+ u_i^3\left(\frac{\rho\sigma\alpha}{\eta\bar{\iota}}\right)^2\left[\frac{[c_i^n]^{-1}[e_i^n]^{1-\eta}}{(1-e_i^n)^{1+\alpha}}\right]^2 v_i^{n+1+} + \frac{\sigma^2u_i^2}{2}\left(1 - u_i\left(\frac{\rho\alpha}{\eta\bar{\iota}}\right)\frac{[c_i^n]^{-1}[e_i^n]^{1-\eta}}{(1-e_i^n)^{1+\alpha}}\right)^2\left(\frac{v_i^{n+1+} + v_i^{n+1-}}{\Delta_u}\right)
\end{aligned}
$$

may be rearranged to give

$$
\begin{aligned}
\frac{v_i^n}{\Delta_T} + \rho(z - c_i^n) &= \left[\frac{1}{\Delta_T} + \rho - \bar{\iota}[e_i^n]^\eta\right]v_i^{n+1} + [(\bar{\iota}[e_i^n]^\eta + \rho)u_i + N_i^n]v_i^{n+1-} \\
&+ \left[\rho u_i^2[c_i^n]^{-1}(1-e_i^n)^{-\alpha} + u_i^3\left(\frac{\rho\sigma\alpha}{\eta\bar{\iota}}\right)^2\left(\frac{[c_i^n]^{-1}[e_i^n]^{1-\eta}}{(1-e_i^n)^{1+\alpha}}\right)^2 + N_i^n\right]v_i^{n+1+}.
\end{aligned}
$$

where for ease of notation I wrote

$$
N_i^n = \frac{\sigma^2}{2\Delta_u}\left(1 - u_i\left(\frac{\rho\alpha}{\eta\bar{\iota}}\right)\frac{[c_i^n]^{-1}[e_i^n]^{1-\eta}}{(1-e_i^n)^{1+\alpha}}\right)^2 u_i^2
$$

This is of the form $b_i^n = D_i^n v_i^{n+1} + D_i^{n-} v_i^{n+1-} + D_i^{n+} v_i^{n+1+}$, which may be expanded to

$$b_i^n = \left(D_i^n + \frac{D_i^{n+} + D_i^{n-}}{\Delta_u} \right) v_i^{n+1} - \frac{D_i^{n+}}{\Delta_u} v_{i+1}^{n+1} - \frac{D_i^{n-}}{\Delta_u} v_{i-1}^{n+1} \tag{A.13}$$

where

$$b_i^n = \frac{v_i^n}{\Delta_T} + \rho(z - c_i^n), \quad D_i^n = \frac{1}{\Delta_T} + \rho - \bar{\iota}[e_i^n]^\eta, \quad D_i^{n-} = (\bar{\iota}[e_i^n]^\eta + \rho)u_i + N_i^n$$
$$D_i^{n+} = \rho u_i^2 [c_i^n]^{-1}(1 - e_i^n)^{-\alpha} + u_i^3 \left(\frac{\rho\sigma\alpha}{\eta\bar{\iota}} \right)^2 \left(\frac{[c_i^n]^{-1}[e_i^n]^{1-\eta}}{(1 - e_i^n)^{1+\alpha}} \right)^2 + N_i^n.$$

The equation (A.13) may be written $b^n = B^n v^{n+1}$, so that $v^{n+1} = [B^n]^{-1} b^n$. We iterate on the map $v^n \mapsto v^{n+1}$ until convergence.

A.2.2 Kolmogorov forward equation

It is well known that the distribution of a variable evolving according to a diffusion process is given by the Kolmogorov forward (or Fokker-Planck) equation. I refer the reader to Achdou, Han, Lasry, Lions and Moll (1) or Fleming and Soner (26) for details. As noted in the main text, I will find the summary measure of normalized promised utility by using the finite-state Markov chain method described in detail in Kushner and Dupuis (47). If $X = (X_t)_{t \geq 0}$ follows a diffusion process of the form

$$dX_t = \mu(X_t)dt + \sigma(X_t)dZ_t \tag{A.14}$$

then X may be approximated arbitrarily well by a Markov chain X^Δ indexed by a pair $\Delta := (\Delta_T, \Delta_X)$ assuming values in the finite set $\Sigma := \{\underline{x}, \underline{x} + \Delta_X, \ldots, \overline{x} - \Delta_X, \overline{x}\}$ and subject to the transition probabilities

$$
\begin{aligned}
p(x, x+\Delta_X) &= \frac{\Delta_T}{\Delta_X^2}\left(\frac{\sigma^2(X)}{2} + \Delta_X \mu(X)^+\right) \\
p(x, x-\Delta_X) &= \frac{\Delta_T}{\Delta_X^2}\left(\frac{\sigma^2(X)}{2} + \Delta_X \mu(X)^-\right) \\
p(x, x) &= 1 - p(x, x+\Delta_X) - p(x, x-\Delta_X)
\end{aligned}
$$

where $x^+ := \max\{0, x\}$ and $x^- := \max\{0, -x\}$. It is shown on page 325 of Fleming and Soner (26) and Kushner and Dupuis (47) that X^Δ converges weakly to X as $\Delta \to 0$. Now the trick is to interpret the summary measure as the stationary distribution of a diffusion process $(M_t)_{t\geq 0}$ with drift μ_M and diffusion σ_M given by

$$
\mu_M(u) = \mu(u) + \sigma_\theta(u)\sigma_u(u) \qquad \sigma_M(u) = \sigma_u(u), \tag{A.15}
$$

which dies at the rate $\rho_D - \iota(u)$ and is injected at the point $\overline{u}$ at rate ρ_D. Solving for the stationary distribution of X^Δ simply amounts to solving a particular linear system $Ax = b$, where A is constructed from the probabilities specified above and b depends upon the initial guess for normalized promised utility. The operation therefore has negligible computation time.

However, once we have found the level of ex-ante utility U associated with the stationary solution, we want to recover the implied stationary distribution of utility, and hence consumption. Unfortunately, the finite-state Markov chain method of Kushner and Dupuis (47) is difficult to apply in this case; the presence of non-negligible off-diagonal terms in the covariance matrix makes it difficult to construct an approximating Markov chain. Instead, the distributions of consumption, utility and productivity will be found via simulation using MATLAB's Financial Toolbox. This process is slow but need only be performed once for each choice of parameters, and not once for every guess of initial promised utility.

Appendix B

Appendix to Chapter 3

B.1 Discrete-time formulation

The purpose of this section is to outline a discrete-time environment that approximates the continuous-time model given in Chapter 3. It is intended to aid the reader and also allow clearer comparison with existing discrete-time environments with private information.

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

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

The appearance of Δc_n rather than c_n in (B.1) is simply a normalization.[1] The principal possesses a constant-returns-to-scale technology that only the agent has the ability to operate. At each time $n\Delta$ the principal chooses how much physical capital will be installed in the technology for the interval $[n\Delta, (n+1)\Delta]$. If K_n is the amount of capital installed at time $n\Delta$ then the amount $\Delta(\Pi + x_n)K_n$ is produced at time $(n+1)\Delta$, where $(x_n)_{n=0}^{\infty}$ is an

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

i.i.d. sequence of of mean zero random variables assuming $\pm\overline{x}$ with probability 1/2. The capital stock is subject to depreciation, with the fraction of the stock K_n remaining at time $(n+1)\Delta$ equal to $e^{-\Delta\delta}$. At time $(n+1)\Delta$ the principal also chooses additional investment I_{n+1} and so the total amount of installed capital at time $(n+1)\Delta$ is

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

The present discounted value of output minus investment is then

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

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

$$\Big(\Delta(\Pi + x_n) + e^{-\Delta\delta} - e^{\Delta r}\Big)K_n \approx \Delta(\Pi - \delta - r)K_n + \Delta x_n K_n. \tag{B.2}$$

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

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

Using (B.2) the objective of the planner is approximately

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

for the principal's objective, and

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

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

I will assume that delegated capital is publicly observable but that both output and consumption are privately observable by the agent. Consumption and delegated capital must therefore be functions only of the reported output shocks $(x_n)_{n=0}^{\infty}$. For each $n \geq 0$, write $x^n := (x_0, \ldots, x_n)$ for the history of realizations of the output shocks up to and including date n and denote by $\mathcal{X}_n$ the set of all such histories.

Definition B.1.1 (Allocations and strategies)**.** An allocation consists of a stochastic sequence of consumption and capital delegation $(K, c) = (K_n, c_n)_{n=0}^{\infty}$ where for each $n \geq 0$ we have $K_{n+1}, c_{n+1} : \mathcal{X}_n \to \mathbb{R}_+$. A strategy of the agent is a sequence of reports $X = (X_n)_{n=1}^{\infty}$ where for each $n \geq 0$ we have $X_{n+1} : \mathcal{X}_n \to \mathbb{R}_+$.

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

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

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

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

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

Definition B.1.2. An allocation (K, c) is incentive compatible if $U^A(c, K; 0) \geq U^A(c, K; X)$ for all agent strategies X and if $K_n \leq \omega \exp W_n(c)$ almost surely for all $n \geq 1$. The set of all incentive compatible allocations will be denoted $\mathcal{A}^{IC}$.

Since the output shocks assume only two values it is without loss to suppose that the reporting strategies assume only two values (either report truth or report the other possible shock), since all other deviations will be detected immediately. As is well-known it suffices to impose temporary incentive compatibility constraints that dissuade one-shot deviations. The principal then announces two possible future values $W^{\pm}$ for promised utility. The temporary incentive compatibility constraints are then

$$\begin{aligned} \left(1-e^{-\Delta\rho}\right)\ln C + e^{-\Delta\rho}W^{+} &\geq \left(1-e^{-\Delta\rho}\right)\ln(C+2\Delta\overline{x}K) + e^{-\Delta\rho}W^{-} \\ \left(1-e^{-\Delta\rho}\right)\ln C + e^{-\Delta\rho}W^{-} &\geq \left(1-e^{-\Delta\rho}\right)\ln(C-2\Delta\overline{x}K) + e^{-\Delta\rho}W^{+}. \end{aligned} \tag{B.3}$$

The first constraint in (B.3) is truth-telling for the high-shock, while the second is truth-telling for the low shock. Rearrangement of (B.3) then gives

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

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

$$\begin{aligned} W &= \left(1-e^{-\Delta\rho}\right)\ln C + e^{-\Delta\rho}\left[W^{-}+W^{+}\right]/2 \\ W^{+} &= W^{-} + \left(e^{\Delta\rho}-1\right)\left[\ln(C+2\Delta\overline{x}K)-\ln C\right]. \end{aligned} \tag{B.4}$$

Also note that simplification of (B.4) gives

$$W^{\pm} \approx e^{\Delta\rho}W - \left(e^{\Delta\rho}-1\right)\ln C \pm \left(e^{\Delta\rho}-1\right)\left[\frac{\ln(C+2\Delta\overline{x}K)-\ln C}{2}\right]. \tag{B.5}$$

Using the fact that $\ln(C+2\Delta\overline{x}K)-\ln C \sim 2\Delta\overline{x}K/C$ as $\Delta \to 0$, the expressions (B.5) may be written

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

which in turn may be written as

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

where $X_{\Delta,t}$ has mean zero, is independent over time and assumes the values $\pm\Delta$, which approximates the law of motion of promised utility in the continuous-time model.

B.2 Relation with other agency models

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

$$\begin{aligned} U^A((\Delta c_n)_{n=0}^{\infty};\Delta) &= (1-\overline{\beta})\sum_{n=0}^{\infty}\beta_n\mathbb{E}[\ln c_n] \\ U^P(K,c;\Delta) &= \Delta\sum_{n=1}^{\infty}e^{-(n+1)\Delta r}\mathbb{E}[(\Pi_n-\delta-r)K_n-C_n] \end{aligned} \tag{B.6}$$

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

$$W_n := (1-\overline{\beta})\sum_{N=n}^{\infty}(\beta_N/\beta_n)\mathbb{E}[\ln c_N]. \tag{B.7}$$

Note that under the specification (B.7) we have a recursive relation involving continuation utilities

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

The temporary incentive constraints are then

$$
\begin{aligned}
(1-\overline{\beta})\ln C_n + (\beta_{n+1}/\beta_n)W^+_{n+1}(C) &\geq (1-\overline{\beta})\ln(C_n + 2\overline{x}K_n) + (\beta_{n+1}/\beta_n)W^-_{n+1}(C) \\
(1-\overline{\beta})\ln C_n + (\beta_{n+1}/\beta_n)W^-_{n+1}(C) &\geq (1-\overline{\beta})\ln(C_n - 2\overline{x}K_n) + (\beta_{n+1}/\beta_n)W^+_{n+1}(C).
\end{aligned}
\tag{B.8}
$$

It is easy to see that only the first constraint in (B.8) will bind and so the problem of the principal is then the following.

Definition B.2.1. The principal's problem is given by

$$
\begin{aligned}
V(W) = \Delta\max_{C,K}\sum_{n=0}^{\infty}e^{-\Delta nr}\mathbb{E}[(\Pi_n - \delta - r)K_n - C_n]& \\
W = (1-\overline{\beta})\sum_{n=0}^{\infty}\beta_n\mathbb{E}[\ln C_n]& \\
(\beta_{n+1}/\beta_n)W^+_{n+1}(C) \geq (1-\overline{\beta})\ln(1 + 2\overline{x}K_n/C_n) + (\beta_{n+1}/\beta_n)W^-_{n+1}(C)&
\end{aligned}
$$

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

The promise-keeping and incentive compatibility constraints are given by

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

Simplification gives

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

Solving for $W^\pm_{n+1}$ then gives

$$W^\pm_{n+1} = (\beta_n/\beta_{n+1})W_n + (\beta_n/\beta_{n+1})(1-\overline{\beta})\Big(-\ln C_n \pm \frac{1}{2}\ln(1+2\overline{x}K_n/C_n)\Big). \tag{B.9}$$

Note that if $C_n = \underline{c}_n \exp W_n$ and $K_n = \underline{k}_n \exp W_n$ then (B.9) implies

$$W^\pm_{n+1} = (\beta_n/\beta_{n+1})\Big[\overline{\beta}W_n + (1-\overline{\beta})\Big(-\ln \underline{c}_n \pm \frac{1}{2}\ln(1+2\overline{x}\phi_{k,n}/\underline{c}_n)\Big)\Big]. \tag{B.10}$$

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

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

This discussion is summarized in the following.

Lemma B.2.1 (Homogeneity of principal's problem)**.** *The value function of the principal is of the form $V(W) = -\Omega\exp W$ for some $\Omega > 0$ and the policy functions for consumption and capital are of the form*

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

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

$$W_{n+1}^{\pm} = \frac{\overline{\beta} W_n}{\beta_{n+1}/\beta_n} - \left(\frac{1-\overline{\beta}}{\beta_{n+1}/\beta_n}\right) \ln \underline{c}_n \pm \frac{1}{2}\left(\frac{1-\overline{\beta}}{\beta_{n+1}/\beta_n}\right) \ln(1 + 2\Delta \overline{x} \underline{k}_n / \underline{c}_n). \tag{B.11}$$

I will now contrast the analysis of this model with that given in Shourideh (68). The timing in this agency problem may be summarized as follows:

1. The agent begins period n with utility (or outside option) W_n.
2. The principal assigns K_n units of capital and C_n units of consumption to the agent.
3. Output produced within period is $\Delta(\Pi + x_n)K_n$.
4. Agent reports x_n and consumes C_n plus any diverted output.
5. Fraction $1 - e^{-\Delta\delta}$ of capital depreciates during the period.
6. Principal assigns utility W_{n+1} for next period depending upon reported level of output.

In Shourideh (68) agents live for two periods and the timing is as follows:

1. Principal assigns K_n units of capital and C_n units of consumption to the agent.
2. Agent consumes C_n plus any capital diverted.
3. Output tomorrow is publicly observed and equal to $\Delta(\Pi + x_n)\underline{k}_n$ where $\underline{k}_n$ is amount of capital actually invested and x_n is random and exogenous.
4. Principal assigns consumption in second period. Agent eats and the world ends.

The above agency problems are obviously similar and so it is instructive to outline why the associated intertemporal distortions differ. Since Shourideh (68) adopts a different specification of shocks it is difficult to directly compare the two models. However, if one adopts the two-period lifecycle structure of Shourideh (68) (and the above timing) but assume shocks take only two values with equal probability, then the resulting model coincides with that given in this section with discount rates and productivities given by

$\beta_0 = e^{-\Delta\rho}, \beta_n = 0, \Pi_0 = \Pi$ and $\Pi_n = 0$ for all $n \geq 1$. In contrast, the model of this corresponds to that given in the previous section with $\beta_n = e^{-\Delta n\rho}$ and $\Pi_n = \Pi$ for all $n \geq 1$.

Now consider two successive periods, n and $n+1$, and define the following perturbation: scale $\underline{c}_n$ and $\underline{k}_n$ by $\exp u$ and $\underline{c}_{n+1}$ and $\underline{k}_{n+1}$ by $\exp(-u\overline{\beta}\beta_n/\beta_{n+1})$, for some arbitrary u. To motivate this perturbation, note that by (B.11) in B.2.1, if we scale $(\underline{c}_n, \underline{k}_n)$ by $\exp u$ and $(\underline{c}_{n+1}, \underline{k}_{n+1})$ by $\exp \overline{u}$ then the change in W_{n+1} will be $\Delta W_{n+1} = -(\beta_n/\beta_{n+1})\big(1-\overline{\beta}\big)u$ and so the change in W_{n+2} will be $\Delta W_{n+2} = (\beta_{n+1}/\beta_{n+2})\big[\overline{\beta}\Delta W_{n+1} - (1-\overline{\beta})\overline{u}\big]$. It follows that $\Delta W_{n+2} = 0$ if and only if $\overline{u} = \overline{\beta}\Delta W_{n+1}/(1-\overline{\beta}) = -\big(\overline{\beta}\beta_n/\beta_{n+1}\big)u$. This implies that the above perturbation only affects quantities in periods n and $n+1$, with all other periods unaffected. The associated change in the utility from consumption at date $t+1$ is then

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

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

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

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

$$(\beta_{n+1}/\beta_n)e^{\Delta r}\underline{c}_n \exp W = \left(\frac{[\Pi_{n+1} - \delta - r]\underline{k}_{n+1} - \underline{c}_{n+1}}{[\Pi_n - \delta - r]\underline{k}_n - \underline{c}_n}\right)\underline{c}_n \mathbb{E}\big[\exp W'\big]. \tag{B.12}$$

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

$$(\beta_{n+1}/\beta_n)e^{\Delta r}\underline{c}_n \exp W = \underline{c}_{n+1}\mathbb{E}\left[\exp W'\right]. \tag{B.13}$$

Combining (B.12) and (B.13) shows that the inverse Euler equation holds if and only if

$$\frac{\underline{c}_n}{\underline{c}_{n+1}} = \frac{[\Pi_n - \delta - r]\underline{k}_n - \underline{c}_n}{[\Pi_{n+1} - \delta - r]\underline{k}_{n+1} - \underline{c}_{n+1}}. \tag{B.14}$$

I hope that expression (B.14) clarifies things for the reader. In my infinite-horizon setting, we have $\beta_n = \beta^n$ and $\Pi_n \equiv \Pi$ and hence $\underline{k}_{n+1} = \underline{k}_n$ and $\underline{c}_{n+1} = \underline{c}_n$ for all $n \geq 0$. The equality (B.14) then obviously holds. In Shourideh (68), the agent lives for two periods (say, $t = 0, 1$) and so $\underline{k}_1 = 0 \neq \underline{k}_0$. The right-hand side of (B.14) for $n = 0$ then becomes

$$\frac{[\Pi_0 - \delta - r]\phi_{0,k} - \underline{c}_n}{[\Pi_{n+1} - \delta - r]\underline{k}_{n+1} - \underline{c}_{n+1}} = \frac{\underline{c}_0 - [\Pi_0 - \delta - r]\underline{k}_0}{\underline{c}_1}$$

which is strictly less than the left-hand side of (B.14).

B.3 Recursive analysis of principal-agent problem

This section contains proofs of 3.1.2, 3.1.5 and 3.1.6.

Proof of 3.1.2. The Hamilton-Jacobi-Bellman equation for the principal's profit function is

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

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

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

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

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

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

$$(\rho - r)\Omega = (\Pi - \delta - r)\omega + \rho\Omega \ln(\rho\Omega)$$

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

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

$$0 = r\Omega + \max_{\substack{\underline{c} \geq 0 \\ \underline{k} \leq \omega}} \; (\Pi - \delta - r)\underline{k} - \frac{\underline{k}^2}{2} \left(\rho\phi\sigma/\underline{c}\right)^2 \Omega - \underline{c} + \rho\Omega \ln \underline{c} =: T(\Omega). \tag{B.15}$$

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

$$\underline{k} = \min\left\{\omega, \left(\frac{\Pi - \delta - r}{[\sqrt{\rho}\phi\sigma]^2}\right)\frac{\underline{c}^2}{\rho\Omega}\right\}. \tag{B.16}$$

Now define a pair of functions T_1 and T_2:

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

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

and note that equation (B.15) may be written $0 = \max(T_1(\Omega), T_2(\Omega)) = T(\Omega)$.[3] Changing variables to $y = \underline{c}/\Omega$, $T_2(\Omega)/\Omega$ simplifies to

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

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

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

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

$$0 = y(\Pi)^2 \left(\frac{\Pi-\delta-r}{\rho\phi\sigma}\right)^2 - y(\Pi) + \rho. \tag{B.17}$$

The solutions to (B.17) are

$$y(\Pi) = \frac{1 \pm \sqrt{1 - \rho(2(\Pi-\delta-r)/[\rho\phi\sigma])^2}}{2((\Pi-\delta-r)/[\rho\phi\sigma])^2} = 2\rho\left(\frac{1 \pm \sqrt{1-(2(\Pi-\delta-r)/[\sqrt{\rho}\phi\sigma])^2}}{(2(\Pi-\delta-r)/[\sqrt{\rho}\phi\sigma])^2}\right). \tag{B.18}$$

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

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

[3] One may think of T_1 and T_2 as being the maxima on the right-hand side of (B.15) under the additional restrictions $\omega \leq \frac{(\Pi-\delta-r)\underline{c}^2}{\Omega[\rho\phi\sigma]^2}$ and $\omega \geq \frac{(\Pi-\delta-r)\underline{c}^2}{\Omega[\rho\phi\sigma]^2}$ respectively.

the function y simplifies to

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

Further simplification gives

$$\begin{aligned} h(\Pi)^2+1 &= \frac{2-2\sqrt{1-(2(\Pi-\delta-r)/[\sqrt{\rho}\phi\sigma])^2}}{(2(\Pi-\delta-r)/[\sqrt{\rho}\phi\sigma])^2} = \frac{h(\Pi)}{(\Pi-\delta-r)/[\sqrt{\rho}\phi\sigma]} \\ y(\Pi) &= \frac{\rho h(\Pi)}{(\Pi-\delta-r)/[\sqrt{\rho}\phi\sigma]} = \rho(h(\Pi)^2+1). \end{aligned} \tag{B.19}$$

Using (B.17) and (B.19), we have

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

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

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

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

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

Recalling (B.16), when $\rho = r$ the policy functions are given by

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

which simplifies to the claimed expressions. □

Proof of 3.1.5. For each $t \geq 0$ and return process R I want to determine the law of motion of the process

$$Y_{t,\Delta} := \mathbb{E}\left[\exp(-\rho\Delta) R_{t,t+\Delta} \frac{u'(c_{t+\Delta})}{u'(c_t)} \,\middle|\, \mathcal{F}_t\right]. \tag{B.20}$$

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

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

It follows that for risky capital the process Y satisfies

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

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

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

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

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

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

The solution to the principal-agent problem in 3.1.2 is valid provided the local maximum $y(\Pi)$ given by (B.18) is well-defined, lies in the constraint set of T_2 and is in fact a global maximum. 3.1.6 shows how these requirements may be simplified.

Proof of 3.1.6. First note that the maximum value of y in the constraint set of the function T_2 satisfies

$$\overline{y}^2 = \frac{\rho\omega}{\Omega(\Pi)}\left(\frac{[\sqrt{\rho}\phi\sigma]^2}{\Pi-\delta-r}\right) = \sqrt{\rho}h(\Pi)\left(\frac{\omega[\rho\phi\sigma]^3}{(\Pi-\delta-r)^2}\right)\exp\left(-h(\Pi)^2/2\right). \tag{B.21}$$

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

$$[\rho\phi\sigma]^2\left(\frac{\sqrt{\rho}h(\Pi)}{\Pi-\delta-r}\right)^2 < \frac{\omega\sqrt{\rho}h(\Pi)[\rho\phi\sigma]^3}{(\Pi-\delta-r)^2}\exp\left(-h(\Pi)^2/2\right)$$
$$\sqrt{\rho}h(\Pi) < \omega\rho\phi\sigma\exp\left(-h(\Pi)^2/2\right)$$

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

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

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

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

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

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

It follows that in addition to $\underline{k}(\Pi) < \omega$, we need only impose the requirement $T_1(\Omega(\Pi)) < 0$, which is strictly weaker than the requirement (3.7). □

B.4 Aggregate resource constraints

Aggregate consumption, labour and output at any date are comprised of contributions from the initial generation and subsequent generations. I will write them in this fashion for clarity. Defining $X := \mathbb{R} \times \{E, W\}$, where $i \in \{E, W\}$ indicates whether or not an agent may be an entrepreneur or a worker, aggregate consumption and output at any date $t \geq 0$ are then,

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

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

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

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

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

B.5 Decentralization

Proof of Lemma 3.3.1. Given taxes on profits τ_Π and risk-free savings τ_{sE}, the Hamilton-Jacobi-Bellman equation for the entrepreneur's value function is

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

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

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

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

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

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

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

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

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

Assuming a solution of the form

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

substitution in the Hamilton-Jacobi-Bellman equation gives

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

as claimed. □

Proof of Corollary 3.3.4. From the expression found in 3.3.3 we have $r \leq \rho$ if and only if $\Pi - \delta - \rho \geq \sqrt{\rho}\sigma h(\Pi)$. From the definition of h this will be assured as long as

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

which is always true because $\phi \in (0,1]$ and the inequality $1-\sqrt{1-x} \leq x/2$ holds for all $x \in [0,1]$. □