
Optimal Financial Contracting and the Effects of Firm's Size

Working Paper 2018-13

SANDRO BRUSCO, GIUSEPPE LOPOMO, EVA ROPERO
AND ALESSANDRO T. VILLA

December, 2018



Stony Brook
University

Optimal Financial Contracting and the Effects of Firm's Size*

Sandro Brusco[§] Giuseppe Lopomo[¶] Eva Ropero^{||} Alessandro T. Villa^{**}

December 2018

Abstract

We consider the design of optimal dynamic financing for a firm subject to moral hazard problems. With respect to the existing literature we enrich the model by introducing durable capital. The existence of durable capital allows us to analyze the role of firm's size, separately from age and financial structure. We find that a higher level of capital decreases the probability of liquidation, increases future size and reduces the average return and volatility on the assets of the firm. Although analytical results are not available, we show through simulations that the rate of growth of capital is decreasing with size and that its variability first increases with size and eventually it declines. These results are broadly in agreement with the empirical results on the effects of firm size.

1 Introduction

We analyze the financing problem of a long-lived firm that uses durable capital and is subject to a moral hazard problem. Our goal is to characterize theoretically the links between the size of the firm, as measured by its capital, and variables such as the probability of survival, the rate of growth, and the rate of return on assets.

There is a large empirical literature on firm size which has uncovered many regularities. Cooley and Quadrini [6] summarize the empirical results on the effect of size as follows:

“Conditional on age, the dynamics of firms (growth, volatility of growth, job creation, job destruction, and exit) are negatively related to the size of the firm.”

Similarly, citing studies by Evans [13], [14] and Dunne, Roberts and Samuelson [10], [11], Sutton [30] notes in his survey the following statistical regularities:

1. *Size and Growth*: (a) the probability of survival increases with firm (or plant) size;
(b) the proportional rate of growth of a firm (or plant) conditional on survival is decreasing in size.

*This paper supersedes an earlier paper by Brusco and Ropero titled “Financing Constraints and Firm Dynamics with Durable Capital”. We thank Marco Celentani, Gianluca Clementi, Vicente Cuñat, María Gutiérrez, Álvaro Lario, Eliseo Navarro, Josep Tribò and seminar participants at IE Business School Madrid, Universidad Carlos III de Madrid, the University of Miami and the University of Bath for useful comments. We remain responsible for errors and obscurities.

[§]Department of Economics and College of Business, Stony Brook University

[¶]Fuqua School of Business, Duke University

^{||}Department of Economics, Universidad Europea de Madrid

^{**}Department of Economics, Duke University

2. *The Life Cycle*: For any given size of firm (or plant), the proportional rate of growth is smaller for older firms (or plants), but the probability of survival is greater.

The theoretical literature has lagged behind, despite the recent rapid development of the dynamic contracting literature. Analyzing the impact of size requires introducing frictions which make the current level of capital relevant for the choice of optimal policies. We do this in a simple way, by assuming partial irreversibility of the investment. In particular, the firm can sell existing (used) capital at a price that is lower than the price of new capital. The solution to the dynamic contracting problem yields policy functions that depend on the existing capital stock. This allows us to compare our theoretical predictions with the empirical evidence. These effects are independent and separate from the effects of an increase in the value of equity, discussed for example in Clementi and Hopenhayn [4] and Quadrini [25].

Our theoretical results are broadly consistent with the empirical evidence on the effects of firm size, and can be summarized as follows. In general, optimal second-best policies prescribe inefficient liquidation and underinvestment with positive probability. However, these effects are smaller for larger firms. More precisely, other things equal, at higher levels of capital the optimal probability of liquidation is smaller, and future levels of capital are higher. We also find that the returns on assets of larger firms have both lower expected value and lower volatility. In our model the returns on the firm's assets are determined by a bilateral contract, rather than being the result of an equilibrium trade-off between risk and return determined in financial markets. Our agents are risk neutral, and thus in the absence of incentive problems they would push investment to the point at which the expected return equals the risk-free rate. Instead, our model predicts a higher rate of return only when the firm is implementing a second-best optimal policy for incentive reasons. The second-best policy typically prescribes underinvestment, when compared to the first best policy. Due to the concavity of the production function, lower levels of capital are associated to higher expected returns. Firms of bigger size are more likely to implement policies close to the first best, which also generate lower expected returns.

We do not have explicit theoretical results on the relationship between firm size, rate of growth and the variability of the rate of growth. To get some insights we have run simulations, which clearly show that the rate of growth is negatively related to size, in agreement with the empirical evidence. In the model, the assumption of decreasing returns to scale implies that, absent incentive constraints, the firm should jump immediately to its optimal size and from that point forward the rate of growth should be zero. The presence of incentive constraints makes capital growth more incremental, and this in turn generates a smooth downward relationship between size and growth.

In our simulations, the *variability* of the rate of growth has an inverse U shape. Thus variability is decreasing only above a certain threshold of size. The increasing part is due to the fact that smaller firms face more severe incentive problems, which limits their growth potentials. As capital grows, the incentive constraints bind less often and this allows for a higher rate of growth. As the firm approaches its optimal size, the rate of growth and its variability tend to decline toward zero.

The rest of the paper is structured as follows. In section 2 we discuss the literature. In section 3 we introduce the model and characterize the optimal investment policy in absence of agency problems. In section 4 we study the optimal financial contract in the presence of asymmetric information. Section 5 is dedicated to defining the empirical counterparts of our theoretical model. We next focus in section 6 on the sensitivity to size of the probability of survival, the investment policy and the rate of growth. Finally, section 7 analyzes the implications of size for financial variables, such as the rate of return and the capital structure. Section 8 contains concluding remarks. Appendix I collects the proofs and Appendix II provides a description of the methodology that we have used in generating the simulations.

2 Discussion of the Literature

The impact of financial constraints on firms' investment policies has long been an important issue both in financial economics and industrial organization. One line of research has taken the financial constraints as given and analyzed their implications for firms' dynamics. Cooley and Quadrini [6], for instance, show that exogenous borrowing constraints and persistent shocks can generate relationships between age and growth of the firm that we are consistent with empirical evidence. Lorenzoni and Walentin [21] reach similar conclusions in a model with limited enforcement in which a continuum of firms are subject to aggregate economy-wide shocks and are exogenously restricted to issue state-dependent securities. More recently, Bolton et al. [3] have analyzed the optimal policies with exogenous financial constraints in a continuous time model.

In this line of research financial restrictions are essentially exogenous. This begs the question of where financial constraints come from. Motivated by this question, another line of research has analyzed the design of optimal dynamic financial contracts in the presence of asymmetric information or limited ability to enforce contracts.

Optimal Dynamic Financial Contracts. In the presence of agency problems the first-best investment policy cannot be pursued, because it violates either incentive compatibility or participation constraints for at least one of the parties. The main theme of this literature is that financing constraints are imposed to provide incentives for the borrower. These incentives typically take the form of threats to either transfer control (including liquidation) or to reduce future financing.

In our model asymmetric information plays a central role.¹ We build on the work of Quadrini [25] and, more closely, Clementi and Hopenhayn [4]. Their main focus is on the importance of financing constraints as determinants of firms' dynamics, especially size, growth and survival. In their models, the firm's revenue depends on capital investment and on random shocks,² and they show that in an optimal financial contract both the amount of investment and the value of equity increase with high revenue shocks and decrease with low ones. The sensitivity of equity value to revenue shocks provides incentives to the entrepreneur to reveal the true value of the shock. Financing constraints tend to disappear when the value of equity becomes sufficiently large, which in turn happens as the firm approaches the optimal size.

In Clementi and Hopenhayn [4] the capital invested in each period depreciates completely at the end of the period. Therefore size, defined as the amount of capital invested in the firm, has to be decided in every period and it is not a state variable. Clementi and Hopenhayn [4] use the value of equity as a proxy for size. This can be justified by the fact that in the optimal contract higher equity values lead to larger investments. However the model cannot make predictions on the effects of size which are distinct from predictions related to the financial structure of the firm.

Size as a State Variable. A firm's size can have an impact on its decisions only in the presence some adjustment costs. If capital can be bought and sold at the same price and there are no other adjustment costs, as it is the case in Quadrini [25], then in an period the firm can choose the optimal amount of capital independently of the current level of capital.

¹See Albuquerque and Hopenhayn [1] for a model of optimal financial contracts and firm dynamics with limited enforcement. One important issue in these models is the punishment which can be inflicted on a defaulting entrepreneur, in particular whether or not the entrepreneur can be excluded from financial markets. Cooley, Marimon and Quadrini [5] consider a model with limited enforcement and optimal financial contracts, assuming that a defaulting entrepreneur retains access to financial markets and can have a fresh start in the next period. Another important issue is that with limited enforcement durable capital may become more valuable since it is more easily sized in case of bankruptcy, see Rampini and Viswanathan [27].

²In Clementi and Hopenhayn [4] the sequence of productivity shocks is i.i.d.. Fu and Krishna [17] generalize the model allowing for Markovian shocks. DeMarzo et al. [9] also consider the case of persistent shocks, although their model is technically different.

In this paper we enrich the Clementi and Hopenhayn [4] model by introducing durable capital, so that size becomes a relevant variable in the firm’s decision problem. In our model capital depreciates at rate $d \in [0, 1]$. The Clementi and Hopenhayn [4] model corresponds to the special case $d = 1$. In each period the existing capital can be increased by new investment, or reduced by selling any part of it at a price that is lower than the investment’s cost. Size is defined as the amount of capital existing at the beginning of the period, determined by past investment decisions and depreciation.

The assumption of decreasing returns to scale is crucial in generating our main results. Recent work in dynamic financial contracting (De Marzo and Fishman [7], DeMarzo and Fishman [8] and DeMarzo et al. [9]) has analyzed the dynamic financial contracting problem under the assumption of constant returns to scale and convex adjustment costs. These models generate interesting and important predictions on the optimal investment policy but are ill-suited to evaluate the effect of size, as the optimal policy depends on the ratio between firm’s value and size (i.e. the value per unit of capital) rather than on the two variables separately. In the words of DeMarzo et al. [9]:

“Because our model is based on a constant returns to scale investment technology ... it is not well suited to address questions relating firm size and growth. Indeed, if we control for past performance or financial slack, size does not matter in our framework.”

Our model has decreasing returns to scale and the optimal policy is based on both on the value of the firm and its size – not on their ratio or any other single-valued function of the two. This is the minimal requisite for making predictions that identify the impact of size on a firm’s decisions.

Evidence on Used Capital. A central assumption in our paper is that used capital goods sell at a discount with respect to new capital goods. There is ample empirical support for this feature of the model. For example, Ramey and Shapiro [26], in their study of the aerospace industry, state that ‘even after age-related depreciation is taken into account, capital sells for a substantial discount relative to replacement cost’. The relevance of the market for used capital has been analyzed in many recent papers, see e.g. Eisfeldt and Rampini [12], Gavazza [18] and Lanteri [23].

One feature that is absent in the our model paper is the dependence of the used capital price on the state of the economy. Our model could easily be generalized to make the price of used capital stochastic and dependent on the state of the economy. We could also allow the state of the firm to be correlated with the state of the economy. These complications however do not seem to produce any particularly surprising insight, at least as far as the optimal policy of the single firm is concerned. They should be taken into account when looking at the aggregate consequences of investment and divestment behavior, something we do not do in this paper.

3 The Model

Time is discrete and the horizon is infinite. At time 0 an entrepreneur has an idea for a project. The project requires funding to first acquire an enabling asset (e.g. a patent or a license), and then buy capital at any subsequent time. The entrepreneur has insufficient funds and thus needs financing from a lender. Both the lender and the entrepreneur are risk neutral and have a common discount factor $\delta \in (0, 1)$. To rule out trivial cases, we assume that the cost of the enabling asset A is positive, but lower than the value generated under the optimal financing contract that we characterize in Section 4.

Once the enabling asset is acquired, the firm can buy capital and use it to generate cash flow at any time $t > 0$. At the beginning of any period the project can be continued or liquidated. If it is liquidated, the existing assets are sold at their ‘scrap’ value, which will be specified shortly. If instead the project continues, the firm makes a capital adjustment decision, i.e. either sells any fraction of its existing capital or buys new capital and adds it to the existing stock. The adjusted capital stock K_t then generates a cash flow of $\theta_t \cdot R(K_t)$, where θ_t is

the realization of a random variable with support $\{0, 1\}$ and R is a production function. We make the following assumptions about the production process.

Assumption 1 *The random variables $\{\tilde{\theta}_t\}_{t=1}^{+\infty}$ are i.i.d., with $p \equiv \Pr(\tilde{\theta}_t = 1) = 1 - \Pr(\tilde{\theta}_t = 0)$. The function $R(\cdot)$ is defined on $[0, +\infty)$, bounded, continuously differentiable, strictly increasing, strictly concave and such that $R(0) = 0$ and $p \cdot R'(0) > 1$.*

Capital depreciates at rate $d \in [0, 1]$. Denoting ΔK_t the change in capital level, through investment or sale, at time t , the law of motion of capital is

$$K_0 = 0 \quad \text{and} \quad K_t = (1 - d)K_{t-1} + \Delta K_t, \quad t = 1, 2, \dots$$

Buying new capital costs 1 per unit, while selling ‘used’ capital yields $q \in (0, 1)$ per unit. Thus the cost of adjusting capital is given by the function

$$I(K_t, K_{t-1}) = \begin{cases} K_t - (1 - d)K_{t-1}, & \text{if } K_t \geq (1 - d)K_{t-1} \\ q(K_t - (1 - d)K_{t-1}), & \text{if } K_t < (1 - d)K_{t-1}. \end{cases}$$

We make the following assumption on the liquidation (‘scrap’) value $S(K_{t-1})$.

Assumption 2 *The liquidation value at any time t depends linearly on capital, i.e.*

$$S(K_{t-1}) = S + q(1 - d)K_{t-1}, \quad 0 < S < A$$

The variable S denotes the liquidation value of the enabling asset. If K_{t-1} is the amount of capital that is available at the beginning of period $t - 1$, then the amount of capital left at the end of period $t - 1$ is $(1 - d)K_{t-1}$. Selling that amount of capital at the beginning of period t at a unit price q generates $q(1 - d)K_{t-1}$ in revenue. The inequality $S < A$ implies that it cannot be optimal to buy the enabling asset only to liquidate the firm.

As in Clementi and Hopenhayn [4], we assume that the entrepreneur is protected by limited liability, which implies that the monetary transfer from the firm to the lender in period t cannot exceed the cash flow $\theta_t \cdot R(K_t)$. We also assume that the cash flow is privately observed by the entrepreneur and not verifiable. Therefore the financing contract will have to provide incentives for the entrepreneur to report correctly the realization of $\tilde{\theta}_t$ in every period, rather than hide and divert the entire cash flow to a private account.

As a benchmark, it is useful to look at the case in which the realized cash flow is verifiable, and thus there is no agency problem. Let $W^{sym}(K)$ denote the present value of the project under symmetric information, when the capital stock at the beginning of the previous period was K . Since we assume that A is lower than the value generated under the second best policy, we have $A < W^{sym}(0)$, which together with Assumption 2 ($S < A$) implies $S < W^{sym}(0)$. This inequality in turn implies

$$S + q(1 - d)K < W^{sym}(0) + q(1 - d)K \tag{1}$$

Since the firm can always sell all the existing capital $(1 - d)K$ at unit price q and then follow the optimal investment policy starting at $K = 0$, we also have

$$W^{sym}(0) + q(1 - d)K \leq W^{sym}(K) \tag{2}$$

Combining the last two inequalities yields

$$S + q(1 - d)K < W^{sym}(K) \tag{3}$$

which establishes that *continuation is always optimal*. Therefore the problem with no asymmetric information can be restated as

$$\max_{\{K_t\}_{t=0}^{\infty}} E \left[\sum_{t=0}^{\infty} \delta^t \left(\tilde{\theta}_t R(K_t) - I_t(K_t, K_{t-1}) \right) \right] \quad (4)$$

Proposition 1 establishes that the optimal policy is characterized by the amount of capital K^* determined by the equation

$$p R'(K^*) = 1 - \delta(1 - d). \quad (5)$$

Proposition 1 *With no asymmetric information, it is optimal to buy K^* in period 1 and dK^* in every subsequent period.*

The left-hand side of equation (5) is the marginal benefit of increasing capital in any period, i.e. the expected marginal product of capital. The right-hand side is the net marginal cost of new capital – any unit of bought in period t costs one, but reduces by $(1 - d)$ the amount of capital that will be purchased in period $t + 1$.

The optimal policy follows immediately from the first order condition (5), once it is established that it is optimal to invest a strictly positive amount in every period. The intuition for the latter property is that a positive investment followed by a negative investment can never be optimal, since the firm can do better by reducing investment at time t , thus avoiding overinvestment. Since investment is positive at time 1, it has to be positive in each period.

Along the optimal path, the firm’s size remains constant at K^* . Therefore, in the problem without asymmetric information, *a higher current size does not imply a larger future size*. If, by accident, the firm reaches any level $\hat{K} > K^*$, capital will “revert to the mean” K^* , at a speed that depends on the resale price q . The reversion would take a single period if $\hat{K} < K^*/(1 - d)$ and multiple periods otherwise.

4 Optimal Contracts under Asymmetric Information

In this section we study the optimal contract under asymmetric information. We first describe the set of feasible contracts and then discuss optimality.

4.1 Feasible Contracts

A financing contract specifies a monetary transfer from the lender to the borrower at $t = 0$, and for each subsequent period, a probability of liquidation, an investment level, and payments between the two parties, as functions of the history.

At $t = 0$ the only possible activity is the acquisition of the enabling asset at price A . At any $t \geq 1$, if the firm is still active, the sequence of events is as follows. First, the firm is liquidated with probability $\lambda_t \in [0, 1]$. If the firm is liquidated, the borrower is paid Q_t , the lender keeps the remaining part of the scrap value $S(K_{t-1}) - Q_t$, and the contractual relationship is dissolved. If the firm remains active, an investment decision is made. Let K_t denote the amount of capital available at the beginning of period t . The borrower observes the production outcome $\theta_t R(K_t)$ privately, reports³ $\hat{\theta}_t \in \{0, 1\}$ to the lender and pays an amount τ_t which may depend on the current message, the level of capital K_t chosen and the history.

An outcome a_t at time $t \geq 0$ is given by $a_t = \{K_t, \hat{\theta}_t\}$ where $\hat{\theta}_t$ is the message issued by the borrower. A *history up to time t* is a collection $h_t = \{a_s\}_{s=0}^t$ and H_t is the set of all possible histories up to time t .

³By the revelation principle, allowing for other sets of feasible messages is inconsequential.

A *financing scheme* specifies a vector $(\lambda_t(h_{t-1}), Q_t(h_{t-1}), \kappa_t(h_{t-1}), \tau_t(h_t))$ for each possible history, where κ_t denotes a probability distribution⁴ over \mathbb{R}_+ . The scheme is *feasible* if for each history h_t we have $\lambda_{t+1}(h_t) \in [0, 1]$, $\text{supp } \kappa_{t+1}(h_t) \subset [0, +\infty)$, $Q_{t+1}(h_t) \geq 0$ and $\tau_t(h_t) \leq \hat{\theta}_t R(K_t)$, the last two inequalities being the consequence of borrower's limited liability.

The contract is *individually rational* if the present expected value of payments is larger than the amount invested at time 0 by both parties, where the expected value is computed over all possible histories and under the assumption that the borrower reports the true value θ_t at each period.

Let $\hat{r}_t : H_t \rightarrow \Theta$ be a reporting strategy at time t for the borrower and $\hat{\mathbf{r}} = \{\hat{r}_t\}_{t=0}^{+\infty}$ be a reporting strategy for all periods. Denote by \mathbf{r} the truth-telling strategy, that is $r_t(h_{t-1}, (K_t, \theta_t)) = \theta_t$ for each $(h_{t-1}, (K_t, \theta_t))$. Let $V_t^{\hat{\mathbf{r}}}(h_{t-1})$ be the present value of the expected payment to the entrepreneur from time t on, given history h_t and reporting strategy $\hat{\mathbf{r}}$. A contract is *incentive compatible* if

$$V_t^{\mathbf{r}}(h_t) \geq V_t^{\hat{\mathbf{r}}}(h_t)$$

for each history h_t and reporting strategy $\hat{\mathbf{r}}$. Since we look at contracts inducing truth-telling, it will be convenient to simplify notation by setting $V_t(h_t) \equiv V_t^{\mathbf{r}}(h_t)$.

4.2 Efficient Contracts

In order to analyze the incentive-efficient contract, we adopt a recursive formulation⁵ that extends the one used in Clementi and Hopenhayn [4], where the only state variable is the entrepreneur's promised utility. In our model, since $d < 1$, the current level of capital must also be treated as a state variable. We introduce the following functions:

1. At the beginning of period t , given the known history h_{t-1} , we denote by $V(h_{t-1})$ the expected continuation payoff for the entrepreneur.
2. If the firm is not liquidated, then a level of capital K_t is decided as realization of the random variable $\kappa_t(h_{t-1})$. After K_t has been decided, the entrepreneur observes θ_t and the level of production $\theta_t R(K_t)$ and reports $\hat{\theta}_t$. The expected continuation payoff for the entrepreneur after history h_{t-1} , capital choice K_t , shock realization θ_t and report $\hat{\theta}_t$ is given by

$$\hat{V}(\theta_t, \hat{\theta}_t, K_t, h_{t-1}) = \theta_t R(K_t) - \tau(\hat{h}_t) + \delta V(\hat{h}_t), \quad (6)$$

where $\hat{h}_t = (h_{t-1}, (K_t, \hat{\theta}_t))$.

The functions V and \hat{V} are linked by following the 'promise keeping' constraints

$$V(h_{t-1}) = \lambda_t Q_t + (1 - \lambda_t) E_{\kappa_t} \left[p \hat{V}(1, 1, K_t, h_{t-1}) + (1 - p) \hat{V}(0, 0, K_t, h_{t-1}) \right] \quad \forall h_t \in H_t \quad (7)$$

where λ_t , Q_t and κ_t are functions of h_{t-1} . Incentive compatibility requires

$$\hat{V}(\theta_t, \theta_t, K_t, h_{t-1}) \geq \hat{V}(\theta_t, \hat{\theta}_t, K_t, h_{t-1}) \quad \forall K_t \in \text{supp } \kappa, \quad \forall \theta_t, \hat{\theta}_t = 0, 1 \quad (8)$$

and limited liability implies

$$\tau(h_{t-1}, (K_t, 0)) \leq 0 \quad (9)$$

⁴Using a probability distribution over K_t , rather than simply a value, is useful because it ensures the concavity of the value function. This will be made clearer in the discussion after Proposition 2.

⁵The idea of using an agent's continuation utility was first introduced by Spear and Srivastava [28]. Atkeson and Lucas [2] provide a justification of the recursive approach in dynamic moral hazard problems.

Notice that there must be histories such that the entrepreneur pays a strictly positive amount when $\theta_t = 1$. If not, the participation constraint of the lender would be violated. Thus, the incentive constraint is typically binding when $\theta_t = 1$.

In order to convince the entrepreneur to report $\hat{\theta}_t = 1$ when this is the true state of the world, the incentive compatibility constraint

$$\widehat{V}(1, 1, K_t, h_{t-1}) \geq R(K_t) + \delta V(h_{t-1}, (K_t, 0)) \quad (10)$$

has to be satisfied. Notice that we have assumed $\tau(h_{t-1}, (K_t, 0)) = 0$, since negative values (equivalent to giving extra cash to the entrepreneur when $\theta = 0$ is announced) can never be optimal as they only worsen the incentive problem. Notice further that since

$$\widehat{V}(1, 1, K_t, h_{t-1}) = R(K_t) - \tau(h_{t-1}, (K_t, 1)) + \delta V(h_{t-1}, (K_t, 1)),$$

the inequality in (10) is equivalent to

$$\tau(h_{t-1}, (K_t, 1)) \leq \delta [V(h_{t-1}, (K_t, 1)) - V(h_{t-1}, (K_t, 0))]. \quad (11)$$

We will use this form of the incentive compatibility constraint in our analysis of the efficient contract.

4.3 The Value of Equity and the Value of the Firm

Since the lender has unlimited wealth, any non-negative value V can be attained simply by liquidating the firm and giving $Q = V$ to the borrower. On the other hand, negative values cannot be implemented due to the limited liability of the borrower. Thus, the set of possible values for V is $[0, +\infty)$. From now on, without loss of generality we will assume that the policy depends on the history h_{t-1} only via the promised equity value V and the accumulated level of capital K . Introducing additional variations based on histories observed along the equilibrium path cannot increase the value of the firm.

At each state (V, K) , let $W(V, K)$ be the value of the firm. The problem of finding the value function $W(V, K)$ can be decomposed in two parts. First, we can compute the value of the firm when continuation is imposed. Second, once that value has been obtained, we can use the continuation value and the liquidation value to compute the optimal liquidation policy.

Let $W_c(V_c, K)$ be the highest value of the firm that can be achieved when continuation is imposed, the level of capital is K and the entrepreneur is promised an equity value of V_c . To simplify notation, let $V^H(K')$ ($V^L(K')$) be the level of utility promised to the entrepreneur when the level of capital K' is chosen and the report is $\hat{\theta}_t = 1$ ($\hat{\theta}_t = 0$). Therefore $W_c(V_c, K)$ is obtained solving the problem

$$\begin{aligned} W_c(V_c, K) = & \max_{\kappa, \tau(\cdot), V^H(\cdot), V^L(\cdot)} E_\kappa [pR(K') - I(K', K)] \\ & + \delta E_\kappa [pW(V^H(K'), K') + (1-p)W(V^L(K'), K')] \end{aligned} \quad (12)$$

subject to

$$V_c = E_\kappa [p(R(K') - \tau(K')) + \delta (pV^H(K') + (1-p)V^L(K'))] \quad (13)$$

$$\tau(K') \leq \delta (V^H(K') - V^L(K')), \quad \tau(K') \leq R(K') \quad (14)$$

$$V^H(K') \geq 0, \quad V^L(K') \geq 0 \quad \text{each } K' \in \text{supp } \kappa \quad (15)$$

Notice that $W_c(V_c, K)$ is computed taking as given the function $W(V, K)$. Once we have the function W_c , we can rewrite the maximization problem as:

$$W(V, K) = \max_{\lambda, Q, V_c} \lambda S(K) + (1-\lambda) W_c(V_c, K) \quad (16)$$

subject to

$$\begin{aligned} V &= \lambda Q + (1 - \lambda) V_c \\ Q &\geq 0, \quad 1 \geq \lambda \geq 0. \end{aligned}$$

Standard results in dynamic programming imply that the solution to the functional equation (16) is unique (see Quadrini [25] for details). The remaining task is to characterize the properties of the functions W_c and W and of the optimal policy.

Inspecting problem (16) we can make a few simple observations. First, it is obvious that whenever the optimal policy requires $\lambda(V, K) = 0$ then we have $W(V, K) = W_c(V, K)$. Second, if $V = 0$ then the optimal policy prescribes $K' = 0$ in each period in which the firm remains active with positive probability and $Q = 0$ in case of liquidation, since the entrepreneur must be given at least $E_\kappa[pR(K')]$ in each period in which the firm is active (this value can be obtained simply announcing $\hat{\theta} = 0$ no matter what). Thus, the only efficient way to give $V = 0$ is to liquidate immediately. The next proposition provides further properties.

Proposition 2 *The functions $W(V, K)$ and $W_c(V, K)$ are non-decreasing in both arguments. For each K the functions $W(\cdot, K)$ and $W_c(\cdot, K)$ are concave and the partial derivatives with respect to V are defined almost everywhere. For each K there is a value $V_{(K)}$ such that the function W is linear in V on the interval $[0, V_{(K)}]$. The first best policy is achievable if and only if $V \geq \frac{pR(K^*)}{1-\delta}$, where K^* is the first-best level of capital defined by equation (5).*

The proposition shows that the results of Clementi and Hopenhayn [4] and Quadrini [25] continue to hold (with some obvious modifications) when capital is durable: For a given value of K the value of the firm is non-decreasing and concave in the value of equity. The linear part of the value function corresponds to the case in which liquidation occurs with positive probability: When $V < V_{(K)}$ the firm is liquidated with probability $\lambda = 1 - \frac{V}{V_{(K)}}$, while when $V \geq V_{(K)}$ the firm continues with probability 1. The important difference is that the threshold values for which liquidation occurs now depend on K .

Notice that one implication of the proposition is that the initial level of capital is not important in order to achieve the first best. Independently of the level of capital the first best can be achieved provided that the value of equity is high enough. Thus, if at any given moment in time the value V reaches $\frac{pR(K^*)}{1-\delta}$ then from that point on the optimal contract will implement the first best policy.

Another implication is that the first best cannot be achieved at period 0, i.e. at least for some time the optimal second best policy must be different from the first best policy. Since $\frac{pR(K^*)}{1-\delta}$ is higher than the value of the project under the first best policy, a contract assigning $V \geq \frac{pR(K^*)}{1-\delta}$ at time 0 would violate the individual rationality constraint for the lender. The value $pR(K^*)/(1-\delta)$ can be reached after a sequence of positive shocks.

5 The Empirical Distribution of Size

In the rest of the paper we will analyze the predictions of our model on the effects of size and compare the theoretical predictions to the stylized facts previously mentioned. We will look both at real and financial variables. Theoretical results will be used whenever possible and simulation will be used when we cannot provide analytical results. The first task is to establish the theoretical implications for the type of data that we actually observe.

5.1 The Joint Distribution of V and K

When a firm is created, the state is $(V_0, 0)$, where V_0 is the value assigned to the entrepreneur when the process is started. The value depends on the amount of self-financing that the entrepreneur can bring and on the bargaining power of the two parties. We can see V_0 as the realization of a random variable \tilde{V}_0 with distribution $f_0(V)$. Once V_0 is realized, the optimal policy induces a stochastic process $\left\{ \left(\tilde{V}_t, \tilde{K}_t \right) \right\}_{t=1}^{\infty}$ that has two absorbing states: $(0, 0)$, reached when liquidation occurs, and (V^{**}, K^*) , where $V^{**} \equiv \frac{pR(K^*)}{1-\delta}$ is the equity value when the efficient investment is made in each period and the agent never pays.⁶ From the stochastic process one can compute the distribution at time t :

$$f_t(V, K) \equiv \Pr(\tilde{V}_t = V, \tilde{K}_t = K).$$

If many firms are born at time 0 with the similar initial values V_0 , then the distribution of their values after t periods should be close to the theoretical distribution $f_t(V, K)$. However, what we typically have in the data is the distribution of firms of age t *conditional on survival*, given by

$$h_t(V, K) = \frac{f_t(V, K)}{1 - f_t(0, 0)} \quad \text{for each pair } (V, K) \neq (0, 0).$$

In empirical studies, we typically have a collection of firms born at different times. If T denotes the calendar date at which a population of firms is observed, and N_{T-k} the number of firms born at calendar date $T - k$, then the number of firms that are still alive at time T should be close to $N_{T-k}^* \equiv (1 - f_k(0, 0)) \cdot N_{T-k}$. The probability distribution of firms active at a date is given by a mixture of the distributions of firms of different ages, weighted by the number of surviving firms of each cohort. This is given by

$$h^*(V, K) = \sum_{k=0}^{+\infty} \frac{N_{T-k}^*}{N^*} h_k(V, K).$$

where $N^* = \sum_{k=0}^{\infty} N_{T-k}^*$, assuming that the sum converges.

We have solved the dynamic programming problem numerically, assuming $R(K) = \frac{1}{\alpha}K^\alpha$, $S(K) = a + bK$. Appendix II provides a detailed explanation of our methodology as well as all parameters' values specified in Table 2. With our parametrization, we have $K^* = 81.5$ and $V^* = 722.3$. The values of V_0 were drawn from a uniform distribution on the interval $[\underline{V}, \bar{V}]$, where $\underline{V} = V^*/50 \simeq 14.5$, and $\bar{V} = V^*/40 \simeq 18$.

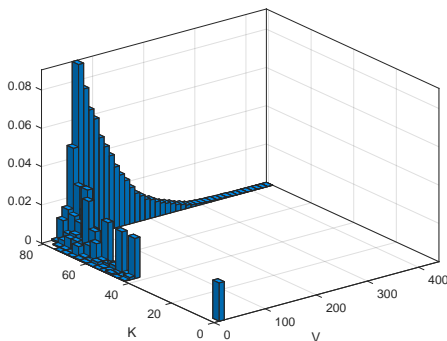


Figure 1: $h_{50}(V, K)$

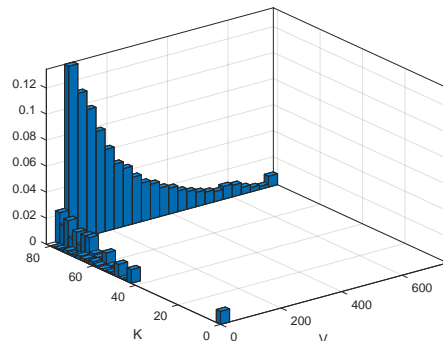


Figure 2: $h_{100}(V, K)$

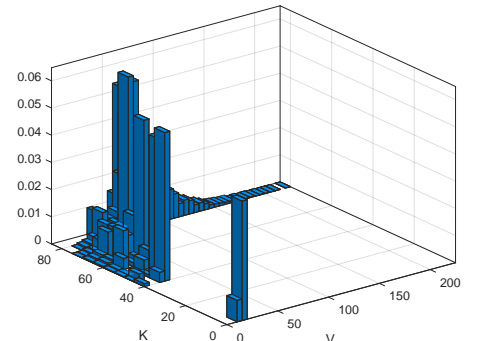


Figure 3: $h^*(V, K)$

Figures 1 and 2 show the distribution of firms surviving after 50 periods and 100 periods respectively. Figure 3 shows the mixture $h^*(V, K)$, assuming that an equal number of firms are born every 10 periods.

⁶Notice that points of the form $(0, K)$ with $K > 0$ have zero probability, since it is impossible to give $V = 0$ to the agent when K is strictly positive.

5.2 The Marginal Distribution of K

While the empirical counterpart of K is easy to identify, this is not so for V . In general V is the value given to the original entrepreneur and founder of the firm. We will later discuss a possible interpretation that follows Clementi and Hopenhayn [4], V as firm's equity and $W(V, K) - V$ as firm's debt. However, even adopting this interpretation, the market value of equity V may be difficult to observe, especially for small and medium-sized firms. For this reason we will also look at distributions over K , obtained as the marginal of the corresponding distributions over the distributions over (V, K) . In particular, let \mathcal{V} be the set of all possible values of V that can be generated by the stochastic process $\left\{ \left(\tilde{V}_t, \tilde{K}_t \right) \right\}_{t=0}^{+\infty}$. Then we can define

$$\gamma_t(K) = \sum_{V \in \mathcal{V}} h_t(V, K)$$

as the probability distribution over size of firms surviving after t periods and

$$\gamma^*(K) = \sum_{V \in \mathcal{V}} h^*(V, K)$$

as the probability distribution of firms existing at a given point in time.

Figures 4, 5 and 6 show the distributions of size for firms surviving after 10, 50 and 100 periods, respectively. Figure 7 shows the distribution $\gamma^*(K)$, obtained as the marginal of $h^*(V, K)$.

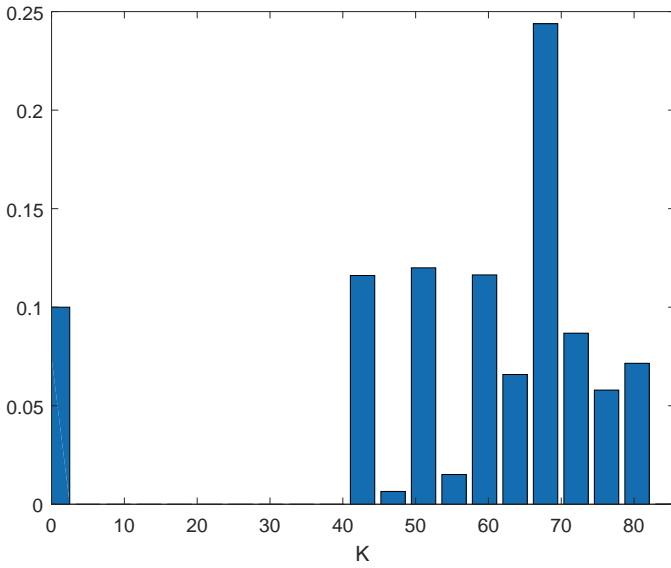


Figure 4: $\gamma_{10}(K)$

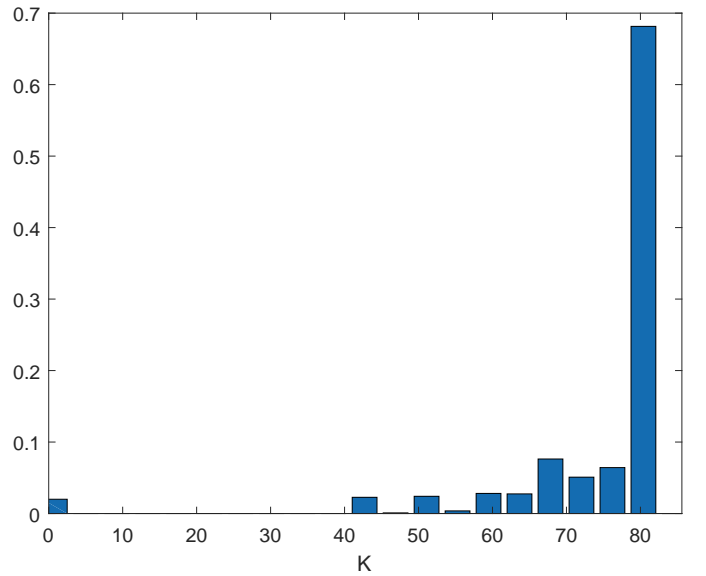


Figure 5: $\gamma_{50}(K)$

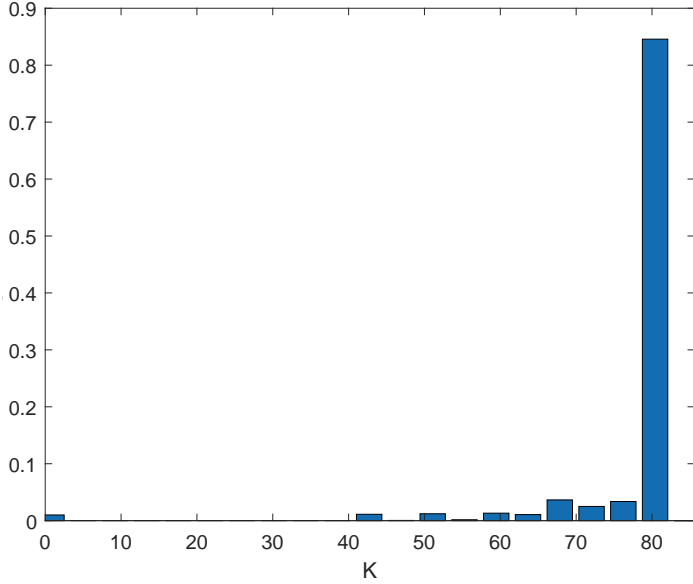


Figure 6: $\gamma_{100}(K)$

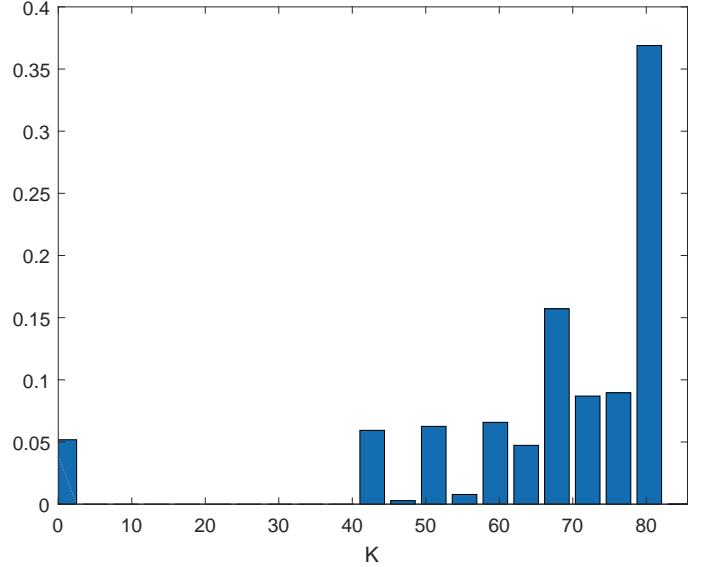


Figure 7: $\gamma^*(K)$

This is the distribution of size that we can expect for a dataset of firms born at different periods and observed at the same calendar date.

Firms that survive for a long time tend to be bigger in size. Beyond a certain level of equity \bar{V} the choice of size is at the optimal level K^* . As it is clear from the simulation, conditional on survival the distribution tends to concentrate more on K^* as age increases. Older firms tend not only to be bigger but also their size has less variance. Younger firms are smaller in size and have a larger variance. The distribution $\gamma^*(K)$ mixes these characteristics.

6 The Effect of Size on Survival and Growth

In this section we focus on the effect of size on the real variables, in particular liquidation and investment. In the next section we will look at the effect of size on financial variables.

6.1 Size and Survival

In our model the probability of liquidation is

$$\lambda(V, K) \equiv \max \left\{ 1 - \frac{V}{V_{(K)}}, 0 \right\}, \quad (17)$$

where $V_{(K)}$ is the threshold value below which liquidation occurs with positive probability. The probability of survival $1 - \lambda(V, K)$ is therefore

$$\beta(V, K) \equiv \min \left\{ \frac{V}{V_{(K)}}, 1 \right\}. \quad (18)$$

From (18) we can immediately conclude that, for each given K , the probability of survival increases in V . How to interpret this result depends on how we interpret V , and we postpone the discussion until the next section. Here we are interested in how survival depends on size. The next results establishes that, conditional on V , the probability of survival is increasing in K .

Proposition 3 *The probability of survival $\beta(V, K)$ is non-decreasing in K for each V and it is strictly increasing whenever $V < V_{(K)}$.*

This is different from the result in Clementi and Hopenhayn. In their case the capital stock is not a state variable, and it is decided in every period as part of the optimal policy. The positive correlation of size and probability of survival exists only because both size and the probability of survival are correlated to the value of equity. Firms with a high equity value are on average larger and have a higher probability of survival. Notice furthermore that the positive relation between size and equity holds only on average in the Clementi–Hopenhayn model; for some values the optimal investment decreases⁷ in V .

In our model size is measured by the existing stock of capital at the beginning of the period. It therefore makes sense to look at the effect of exogenous changes in size. The basic intuition for the result in Proposition 3 is that an increase of capital increases the value of continuation more than the value of liquidation. This is easier to see when the optimal policy prescribes strictly positive investment with probability 1 under continuation. In that case a small additional amount of capital ΔK increases the value of continuation by $(1 - d)\Delta K$ (since it reduces the necessary investment by that amount), while the value of liquidation increases only by $q(1 - d)\Delta K$ (the value at which capital can be sold).

It should be noted however that Proposition 3 provides a positive relation between size and probability of survival *conditional on V* . In empirical work it is difficult to control for V and the main existing results look at the unconditional relation. In other words, what we actually observe in empirical work is not the survival function $\beta(V, K)$ but the function

$$\beta^*(K) = E_V [\beta(V, K) | K] = \sum_{V \in \mathcal{V}} q^*(V | K) \beta(V, K),$$

where

$$q^*(V | K) = \frac{h^*(V, K)}{\gamma^*(K)}$$

Even if $\beta(V, K)$ is non-decreasing in K for each V , it may still be the case that $\beta^*(K)$ is decreasing. We may notice however that, since $\beta(V, K)$ is an increasing function of both V and K , the function $\beta^*(K)$ turns out to be increasing in K if the probability distribution $q^*(V | K)$ is increasing in K in the sense of first-order stochastic dominance.

We cannot provide a full theoretical characterization of the joint probability distribution $q^*(V, K)$. The simulations we have done support the idea that $q^*(V | K)$ is FOSD-increasing with K .

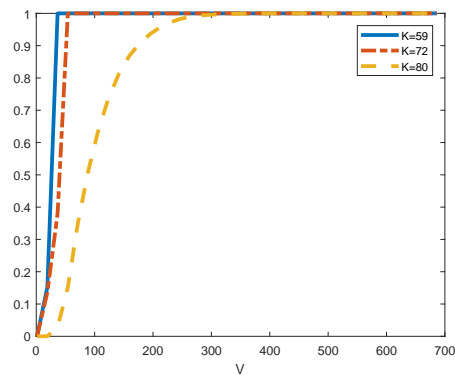


Figure 8: $Q_{50}(V | K)$

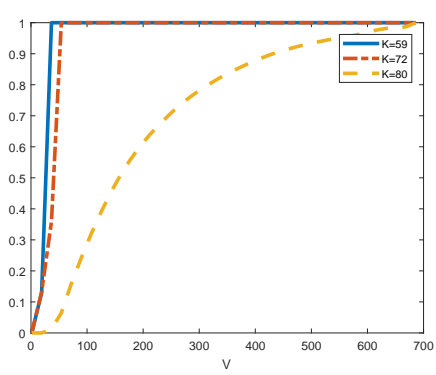


Figure 9: $Q_{100}(V | K)$

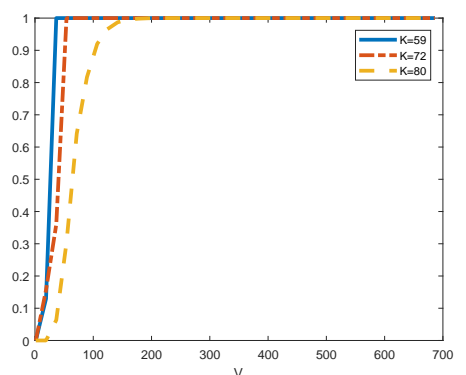


Figure 10: $Q^*(V | K)$

⁷The same is true in our model: the optimal choice K' is not necessarily monotonic in V .

Define $q_t(V|K) = \frac{h_t(V,K)}{\gamma_t(K)}$ and $Q_t(V|K)$ the corresponding cumulative distribution function. Figure 8 shows the cumulative distribution $Q_{50}(V|K)$ is FOSD-increasing with K . The same happens with $Q_{100}(V|K)$, as shown in Figure 9. Finally, Figure 10 shows the cumulative distribution function when firms of different ages are considered, $Q^*(V|K)$. Since $Q_t(V|K)$ appears to be FOSD-increasing in K for each t , the mixture $Q^*(V|K)$ has the same properties. The simulations results thus lend support to the hypothesis that $\beta^*(K)$ is increasing in K .

6.2 Size and Growth

Growth opportunities for large firms are absent in our model. Once a firm has reached the optimal size K^* the rate of growth is necessarily zero. This is a direct consequence of the assumptions on the technology and the economic environment in which the firm is operating, namely that the expected revenue function $pR(K)$ is concave and remains the same over time. Since large firms that have reached the optimal size have zero growth, it must eventually be true that the rate of growth decreases with size.

Characterizing the rate of growth when capital is lower than optimal is more complicated. Before we go there, let's see what can be said about the impact of current size on future size. Let $\tilde{K}'_{(V,K)}$ be the random variable chosen as optimal capital policy at (V, K) , with $\bar{K}'_{(V,K)}$ and $\underline{K}'_{(V,K)}$ being the supremum and the infimum of the support. The next result shows that an increase in current size, defined as existing capital stock K at the beginning of period t , causes an increase in future size, defined as the optimal capital stock K' chosen in period t . However a higher level of K in general requires less investment, since the optimal new level of capital K' is closer to existing capital.

Proposition 4 *The optimal capital choice has the following properties. For each $\Delta > 0$:*

1. $\tilde{K}'_{(V,K+\Delta)}$ (weakly) first order stochastically dominates $\tilde{K}'_{(V,K)}$.
2. If $\underline{K}'_{(V,K)} > (1-d)(K+\Delta)$ or $\bar{K}'_{(V,K)} < (1-d)K$ then $\tilde{K}'_{(V,K+\Delta)} = \tilde{K}'_{(V,K)}$.

The intuition for the first point is relatively simple. The cost of investment $I(K', K)$ is strictly decreasing in K . Thus, other things equal, it cannot be an optimal policy to choose a higher future level of capital stock when the current capital stock is lower.

The second point simply follows from the fact that increasing the level of capital when the investment is strictly positive or strictly negative does not change the constraint set and it is equivalent to adding a constant to the objective function. The optimal solution must therefore remain the same.

The expected rate of growth of a firm at the state (V, K) is given by

$$g(V, K) = \frac{E\left[\tilde{K}'_{(V,K)}\right] - K}{K}$$

where the expectation is taken using the optimally chosen probability distribution over K' . A corollary of Proposition 4 is the following.

Corollary 1 *If $\underline{K}'_{(V,K)} > (1-d)(K+\Delta)$ or $\bar{K}'_{(V,K)} < (1-d)K$ then $g(V, K+\Delta) < g(V, K)$.*

The result follows straightforwardly from the fact that the optimal policy \tilde{K}' remains the same when K increases to $K+\Delta$. Notice that the condition described in the corollary will be satisfied whenever the optimal investment

K	43	51	59	68	72	76	80
$g^*(K)$	15%	17%	7%	3%	2%	1%	0%
$\sigma_{g^*}^2(K)$	0.0004	0.0059	0.0037	0.0012	0.0021	0.0011	0.0001

Table 1: The average rate of growth $g^*(K)$ and the volatility of the rate of growth $\sigma_{g^*}^2(K)$ as a function of K .

policy is deterministic. While we cannot conclude in general that $g(V, K)$ is decreasing in K , the corollary strongly suggests that this will normally be the case.

Up to now we have focused on the role of past size in determining future size *for a given level of V* . When discussing the relationship between survival and size we have observed that empirical studies typically offer unconditional relations. The same remark applies here. In empirical studies what we typically observe is

$$g^*(K) = E_V [g(V, K) | K] = \sum_V q^*(V | K) g(V, K)$$

Thus, whether or not g^* is decreasing depends on the probability distribution over (V, K) and on how $g(V, K)$ depends on V . In Clementi and Hopenhayn [4] the relation between the capital chosen in each period K' and V is not monotonic, and we can expect the same to be true in our model. All this implies that, in general, it is difficult to establish results on the properties of $g^*(K)$. The results of our simulations show that, on average, the rate of growth is first increasing and then decreasing

The volatility of the rate of growth is another variable of interest. The only clear prediction of the model is that firms which have reached a sufficiently high level of V choose the optimal size K^* and therefore have a zero rate of growth, which in turn implies zero volatility. This must mean that, as size increases, the volatility of growth must eventually go down. In our model the main source of volatility, conditional on size, come from the fact that the investment policy depends on V . The object of interest is therefore

$$\sigma_{g^*}^2(K) = \sum_V \pi^*(V | K) (g^*(V, K) - g^*(K))^2.$$

The results of our simulations show that, on average, the volatility of the rate of growth follows a path similar to the one of the rate of growth. It is first increasing and then decreasing with K .

As previously pointed out, the model predicts that for K large enough the volatility of growth will go to zero. The simulations confirm that for lower levels of K the relationship between size and volatility of growth is not monotone.

7 The Effect of Size on Financial Variables

Our model provides predictions on the effect of size on financial variables such as the rate of return on the firms's asset and the capital structure of the firm. Given the large amount of financial data available it is tempting to use them in order to test the predictions of our model. For example, Fama and French [15], have shown that the return on equity depends negatively on the size of the firm, when controls for other variables are properly introduced.

This temptation should be resisted. Our model is one in which there is bilateral bargaining between two parties and perfect commitment. While the contract has to satisfy incentive compatibility and individual rationality, there is no assumption that in each given period the funds invested in the firm will produce a return similar to the one produced by assets with similar characteristics. Fama and French [15], and many other studies in empirical finance, look at securities which are freely traded and have a return which has to be determined in each period by equilibrium conditions. Given our assumption of risk neutrality, if the parties were not bound by long term contracts and claims to the firm's cash flow could be freely traded in the market

then the return on each security would be simply equal to the risk-free rate. Competition between lenders and borrowers determines how the expected surplus is split between the parties when the contract is signed, but after that the borrower and the lender are locked in a long-term relationship in which they are protected from (or cannot take advantage of) competitive pressure. The most accurate empirical counterpart is probably given by the long-term bilateral relationships which are sometimes set up between banks and firms or between venture capitalists and start-up firms.

7.1 Size and Rate of Return

One implication of the previous results is that the expected total return is lower for bigger firms and, under certain conditions, it is also less volatile. To see this, define the total return for the firm when the current state is (V, K) as

$$\tilde{r}(V, K) = \frac{\tilde{Y}(V, K) + \tilde{W}(V, K) - W(V, K)}{W(V, K)}, \quad (19)$$

where $\tilde{Y}(V, K)$ and $\tilde{W}(V, K)$ are the random variables determining the net income in current period and the value of the firm at the beginning of next period when the state is (V, K) . By construction we have

$$W(V, K) = E \left[\tilde{Y}(V, K) + \delta \tilde{W}(V, K) \right],$$

which in turn implies that the expected rate of return is

$$E [\tilde{r}(V, K)] = \frac{(1 - \delta) E \left(\tilde{W}(V, K) \right)}{W(V, K)}. \quad (20)$$

Since the probability distribution of $\tilde{r}(V, K)$ depends on the state (V, K) we may ask if there is any systematic impact of the state variables on the distribution. Our first result shows that the expected return decreases with size.

Proposition 5 *Let $\tilde{r}(V, K)$ be the random variable defined by (19). There is a value $\bar{\Delta} > 0$ such that*

$$E [\tilde{r}(V, K)] \geq E [\tilde{r}(V, K + \Delta)]$$

for each $\Delta \in (0, \bar{\Delta})$, with strict inequality if the optimal policy at (V, K) is such that $\underline{K}'_{(V, K)} > (1 - d)K$ or $\overline{K}'_{(V, K)} < (1 - d)K$.

Notice that the predictions are made for a fixed value of V , i.e. we consider what happens to the firm when K is increased while V remains constant. This can be done only in the presence of durable capital. To grasp the main intuition for the result, suppose that the optimal policy is single valued, there is no liquidation and investment is positive, i.e. there is a single optimal value K' and $K' > (1 - d)K$. Then the random variable $\tilde{W}(V, K)$ is

$$\tilde{W}(V, K) = \begin{cases} W(V^H, K') & \text{with prob. } p \\ W(V^L, K') & \text{with prob. } 1 - p \end{cases}$$

If we increase the amount of capital by ΔK the optimal policy does not change, as long as $K' > (1 - d)(K + \Delta K)$. Thus, $\tilde{W}(V, K) = \tilde{W}(V, K + \Delta)$. On the other hand the current value increases by $(1 - d)\Delta K$, the reduced cost of investment. Thus, $W(V, K + \Delta) = W(V, K) + (1 - d)\Delta K$. Therefore, we can write the expected return when the capital is increased to $K + \Delta$ as

$$E [\tilde{r}(V, K + \Delta)] = \frac{(1 - \delta) E \left(\tilde{W}(V, K) \right)}{W(V, K) + (1 - d)\Delta K} < E [\tilde{r}(V, K)].$$

Essentially, the point is that a larger firm has the same future values as a smaller firm, since the optimal policy does not change. However, the current value is higher because a lower investment is required, and this reduces the expected return. A similar argument can be made for the volatility of returns, at least for the case in which investment is either positive with probability 1 or negative with probability 1.

Proposition 6 *Suppose that at state (V, K) the optimal policy is such that $\underline{K}' > (1 - d)K$ or $\overline{K}' < (1 - d)K$. Then there is a value $\overline{\Delta} > 0$ such that, for each $\Delta \in (0, \overline{\Delta})$ we have*

$$\text{Var} [\tilde{r}(V, K)] > \text{Var} [\tilde{r}(V, K + \Delta)].$$

The intuition is similar to the one discussed above for the expected return. When the optimal future policy is fixed the future values are also fixed, and so is its variability. On the other hand, higher values of K reduce the need for investment and therefore increase the current value of the firm. The volatility of the return is obtained dividing the volatility of the future values by the current value; the numerator is fixed while the denominator increases when K increases, thus reducing the volatility of returns.

The results refer to the impact of size for a given value of V . Both V and K are determined jointly in response to the sequence of shocks. General results on how $E[\tilde{r}(V, K)]$ varies with respect to V are hard to come by. The next proposition provides a partial characterization.

Proposition 7 *$E[\tilde{r}(V, K)]$ is increasing in V when $V < V_{(K)}$, i.e. when the probability of liquidation is positive. $E[\tilde{r}(V, K)]$ is constant in V when $V \geq \frac{pR(K^*)}{1-\delta}$.*

The proposition provides results for low and high values of V , but not for the intermediate range, over which we do not expect a monotonic behavior.

7.2 Size and Capital Structure

We conclude our analysis with some remarks on the sensitivity of capital structure to size and investment to cash flow. While we don't have formal results, the model can be used to provide some intuition about these issues.

The financial contracts considered in this paper are quite complicated. However, we can interpret V as equity and $B(V, K) = W(V, K) - V$ as debt⁸. The evolution of the capital structure is complex, since in general K and V move together and in non-trivial ways along the optimal path. However, stretching somehow the model, assume that exogenous variations of K at any given moment are possible, e.g. because of stochastic changes in the depreciation rate. In that case the value of V remains constant, so that the conclusion is that firms with higher level of capital have a capital structure more tilted towards debt. Furthermore, if we consider the case of non-stochastic policies and use the distance $V^H - V^L$ as a measure of the volatility of equity then higher levels of K lead to higher values of K' and, since $R(K') = \delta(V^H - V^L)$, to higher volatility. Thus, in general, bigger firms are more likely to be debt-financed and have more volatile equity. The conclusion that higher debt leads to a more volatile return on equity is of course familiar from Modigliani-Miller, but here it is reached for quite different reasons (the Modigliani-Miller theorems do not hold in our model).

Consider now the sensitivity of investment to cash flow, a question that has received much attention in the literature. With durable capital the state of the firm is defined by the two variables V and K . The optimal investment policy depends separately on both variables, not just on the way in which they influence the total

⁸This interpretation follows Clementi and Hopenhayn [4]. Other interpretations are possible; for example, the firm might be entirely financed with equity, with V being the value of the 'insider equity' held by the manager and B the value of 'external equity' held by outside financiers.

value of the firm $W(V, K)$. Under some conditions (see Hayashi [20]) the investment in any given period should be explained only by the Tobin's q , i.e. the ratio $\frac{W(V, K)}{A+K}$. Quadrini [25] points out that, in models with non-durable capital and random variables $\tilde{\theta}_t$ which are independent across periods, Tobin's q is in fact a sufficient variable to explain investment. The reason is that the optimal investment is entirely explained by the value of equity V , which in turn determines the value of the firm $W(V)$. Thus, a regression of investment over Tobin's q and cash-flow should give a non-significant coefficient for cash flow.⁹ A positive coefficient for cash flow reappears when shocks are not independent, as in Fu and Krishna [17]; in that case the past value of shocks determines the future profitability of investment, so the optimal level of investment depends not only on V but also on θ_{t-1} . Since θ_{t-1} also determines the cash flow, the sensitivity of investment to cash-flow is reintroduced.

Introducing durable capital is another way in which the sensitivity of investment to cash flow may reappear. With durable capital the state of the firm is defined by two variables, V and K . The optimal investment policy depends separately on both variables, not just on the way in which they influence the total value of the firm $W(V, K)$. Thus, a regression of investment over Tobin's q and cash-flow should typically give a non-zero coefficient for cash flow.

8 Conclusions

Size has always been recognized as an important determinant of firm's behavior, but in absence of durable capital it is difficult to define 'size' properly. This paper has introduced durable capital in a model of optimal dynamic financing with moral hazard in order to develop testable predictions about the effect of firm's size on the optimal financing and investment policies. The key assumption making size relevant is that used capital is less valuable than new capital, even once depreciation has been accounted for; there is ample empirical evidence that this is in fact the case.

We show that, conditional on age, larger firms have a higher probability of survival, are more likely to be big in the future, have lower investment rates and lower average return and volatility on the assets. The results are qualitatively consistent with the empirical literature.

Further research should extend the analysis in two directions. First, it would be useful to perform a quantitative exercise, considering a parametric version of the model with realistic values for the parameters and checking whether the dynamics predicted by the model are quantitatively consistent with what has been found in empirical studies. Second, it would be interesting to move the analysis to the industry level, analyzing the endogenous determination of liquidation values as well as the impact of entry and exit.

⁹The regression's outcome depends on the exact lags which are used for the variables. Theoretically, investment at time t depends on the q at time t and nothing else. If investment is regressed over q at time $t - 1$ (as done by Fazzari et al. [16] and Gomes [19]) then even models without durable capital will exhibit the sensitivity to cash flow found in the data.

Appendix I

Proof of Proposition 1. The only difficulty in the proof comes from the fact that the function I_t is not differentiable at points where $K_t = (1-d)K_{t-1}$. However, the right derivative and the left derivative are always defined. To save notation, let $I_t \equiv I(K_t, K_{t-1})$.

We have

$$\frac{\partial I_t}{\partial K_t} = \begin{cases} 1 & \text{if } K_t > (1-d)K_{t-1} \\ q & \text{if } K_t < (1-d)K_{t-1} \end{cases}$$

and

$$\frac{\partial I_t}{\partial K_{t-1}} = \begin{cases} -(1-d) & \text{if } K_t > (1-d)K_{t-1} \\ -q(1-d) & \text{if } K_t < (1-d)K_{t-1}. \end{cases}$$

Let $\frac{\partial^+ I_t}{\partial K_t}$ be the right derivative and $\frac{\partial^- I_t}{\partial K_t}$ the left derivative. In general optimality requires

$$pR'(K_t) \geq \frac{\partial^- I_t}{\partial K_t} + \delta \frac{\partial^- I_{t+1}}{\partial K_t} \quad (21)$$

and

$$pR'(K_t) \leq \frac{\partial^+ I_t}{\partial K_t} + \delta \frac{\partial^+ I_{t+1}}{\partial K_t}, \quad (22)$$

with equality at each point at which both I_t and I_{t+1} are differentiable (i.e. points at which $K_t \neq (1-d)K_{t-1}$ and $K_{t+1} \neq (1-d)K_t$).

We now show by induction that, under the optimal policy, investment is always positive. The inequality $K_1 > (1-d)K_0 = 0$ must be satisfied for the project to be viable. To see that $K_t > (1-d)K_{t-1}$ implies $K_{t+1} > (1-d)K_t$, suppose not. Then there is a t such that $K_t > (1-d)K_{t-1}$ and $K_{t+1} \leq (1-d)K_t$. In this case $\frac{\partial I_t}{\partial K_t} = 1$ and $\frac{\partial^- I_{t+1}}{\partial K_t} = -q(1-d)$, so condition (21) becomes

$$pR'(K_t) \geq 1 - \delta q(1-d).$$

On the other hand, condition (22) for K_{t+1} becomes

$$pR'(K_{t+1}) \leq \frac{\partial^+ I_{t+1}}{\partial K_{t+1}} + \delta \frac{\partial^+ I_{t+2}}{\partial K_{t+1}}.$$

Since the highest value of $\frac{\partial^+ I_{t+1}}{\partial K_{t+1}}$ is 1 and the highest value of $\frac{\partial^+ I_{t+2}}{\partial K_{t+1}}$ is $-(1-d)q$ we have

$$\frac{\partial^+ I_{t+1}}{\partial K_{t+1}} + \delta \frac{\partial^+ I_{t+2}}{\partial K_{t+1}} \leq 1 - \delta q(1-d). \quad (23)$$

Finally, since $K_{t+1} \leq (1-d)K_t < K_t$, concavity of R implies

$$\frac{\partial^+ I_{t+1}}{\partial K_{t+1}} + \delta \frac{\partial^+ I_{t+2}}{\partial K_{t+1}} \geq pR'(K_{t+1}) > pR'(K_t) = 1 - \delta q(1-d). \quad (24)$$

Inequalities (23) and (24) are incompatible, so we have a contradiction. We conclude that strictly positive investment at time t must be followed by strictly positive investment at time $t+1$. If investment is always strictly positive the optimality condition becomes

$$pR'(K_t) = 1 - (1-d)\delta \quad (25)$$

for each t , implying a constant level of capital, given by the solution K^* to (25). ■

Proof of Proposition 2. Since we will use results from Stokey, Lucas and Prescott [29], we first recast the problem using a similar notation, so that the way in which their results are applied is clearer. Define the vector of control variables as

$$x = (\lambda_x, Q_x, K_x, \tau_x, V_x(0), V_x(1)).$$

Given a choice $y = (\lambda_y, Q_y, K_y, \tau_y, V_y(0), V_y(1))$, the return function is defined as

$$F(x, y) = \lambda_y S(K_x) + (1 - \lambda_y) [pR(K_y) - I(K_y, K_x)].$$

Define Γ as the set of vectors $(\lambda_y, Q_y, K_y, \tau_y, V_y(0), V_y(1))$ that satisfy the following:

$$\lambda_y \in [0, 1], \quad Q_y \geq 0, \quad K_y \geq 0$$

$$\tau_y \leq \min \{ \delta (V_y(1) - V_y(0)), R(K_y) \}$$

$$V_y(0) \geq 0, \quad V_y(1) \geq 0.$$

and $\Delta\Gamma(V_x(\theta))$ the set of probability distributions over Γ that satisfy

$$V_x(\theta) = E_\gamma [\lambda_y Q_y + (1 - \lambda_y) [p(R(K_y) - \tau_y) + \delta (pV_y(1) + (1 - p)V_y(0))]],$$

where $\gamma \in \Delta\Gamma(V_x(\theta))$ denotes a probability distribution over Γ .

The value function can be written as

$$W(x, \theta) = \max_{\gamma \in \Delta\Gamma(V_x(\theta))} (1 - \lambda_x) E_\gamma [(F(x, y) + \delta E[W(y, \theta')])]$$

Standard results in dynamic programming imply that $W(x, \theta)$ exists and is unique. It is also clear that neither $\Delta\Gamma(V_x(\theta))$ nor F depend on Q_x , τ_x and $V_x(\theta')$ when $\theta' \neq \theta$. Thus we can define

$$W^*(\lambda_x, K_x, V_x(\theta)) \equiv W(x, \theta).$$

Finally, we also have

$$W^*(\lambda_x, K_x, V_x(\theta)) = (1 - \lambda_x) W^*(0, K_x, V_x(\theta)).$$

Thus, function $W(V, K)$ defined as

$$W(V, K) = W^*(0, K, V)$$

is the one that we have been discussing in the text.

The function is increasing in K , because the return function is increasing in K_x and the constraint set $\Delta\Gamma(V)$ does not depend on K_x . To see that the function is increasing in V , notice that the return only depends on λ and K , but not on Q or τ . When V is increased it remain possible to use the same policies for λ and K , achieving the higher V through decreases in τ or increases in Q . Thus, increasing V expands the set of payoff-relevant policies. A similar argument establishes that W_c is increasing in V_c .

To see that $W(V, K)$ is concave in V when K is fixed, suppose that there are two values V_1 and V_2 such that

$$qW(V_1, K) + (1 - q)W(V_2, K) > W(qV_1 + (1 - q)V_2, K) \tag{26}$$

for some $q \in (0, 1)$. For the given q , consider the value $V_q = qV_1 + (1 - q)V_2$. The value V_q can be promised to the entrepreneur by offering the policy implemented at V_1 with probability q and the policy implemented at V_2 with probability $(1 - q)$; notice that such policies are clearly feasible. The expected value of the firm in that case would be the left hand side of (26). This is greater than the right hand side, contradicting the claim that

$W(qV_1 + (1 - q)V_2, K)$ is the highest value of the firm that can be achieved while giving V_q to the entrepreneur. The same argument establishes the concavity of W_c . (Notice that the argument cannot be applied to establish the concavity with respect to K given V , since K enters directly the return function.) Since $W(V, K)$ and $W_c(V, K)$ are increasing and concave in V , the partial derivatives $\frac{\partial W}{\partial V}$ and $\frac{\partial W_c}{\partial V}$ are defined almost everywhere.

The proof that $W(V, K)$ is linear in V on an interval $[0, V_{(K)}]$ is the same as in CH [4], with the only change that now the upper bound of the region over which $W(\cdot, K)$ is linear depends on K .

Finally, if the first best is implemented then an investment K^* must occur in every period independently of the history of announcements. This implies that the entrepreneur can achieve a value $\frac{pR(K^*)}{1-\delta}$ simply by announcing $\theta = 0$ in every period and stealing the output. Thus, in order to implement the first best policy we need $V \geq \frac{pR(K^*)}{1-\delta}$. When this condition is satisfied the first best can be achieved by a policy of investing K^* in every period independently of past history, paying $V - \frac{pR(K^*)}{1-\delta}$ immediately to the entrepreneur, and giving the entire output $\theta_t R(K^*)$ to the entrepreneur in each period.

■

Proof of Proposition 3. We show that $V_{(K)}$ does not increase in K and it strictly decreases if the optimal policy prescribes strictly positive investment at $(V_{(K)}, K)$. For each value V the probability of liquidation does not increase in K and it strictly decreases if $\lambda(V, K) \in (0, 1)$ and $V_{(K)}$ strictly decreases in K .

The function $W_c(V, K)$ is almost everywhere differentiable and by the envelope theorem

$$\frac{\partial W_c(V, K)}{\partial K} = (1 - d) \Pr(K' > (1 - d)K) + q(1 - d) \Pr(K' \leq (1 - d)K).$$

The liquidation value $S(K)$ is linear in K and

$$\frac{\partial S(K)}{\partial K} = q(1 - d).$$

Thus $\frac{\partial S}{\partial K} \leq \frac{\partial W_c}{\partial K}$ for each V . Since $V_{(K)}$ is the point at which the line with intercept $S(K)$ is tangent to $W_c(V, K)$, if the function W_c increases no less than the value $S(K)$ then the point $V_{(K)}$ cannot increase. In particular, if at $(V_{(K)}, K)$ we have $\Pr(K' > (1 - d)K) > 0$, i.e. the probability of a strictly positive investment is strictly positive, then $\frac{\partial W_c(V_{(K)}, K)}{\partial K} > \frac{\partial S(K)}{\partial K}$ and the value of $V_{(K)}$ strictly decreases as K increases.

The probability of liquidation can change with K only at points (V, K) at which $\lambda(V, K) < 1$. At such points we have

$$\lambda(V, K) = 1 - \frac{V}{V_{(K)}}$$

and the conclusion therefore follows from the results on $V_{(K)}$. ■

Proof of Proposition 4. Consider problem (12). We want to apply Theorem 4 in Milgrom and Shannon [24], and we will do so by showing that the objective function is quasi-supermodular in the decision variables $(\kappa, \tau(\cdot), V^H(\cdot), V^L)$ and it satisfies increasing difference in $((\kappa, \tau(\cdot), V^H(\cdot), V^L(\cdot)); K)$.

The space where $(\kappa, \tau(\cdot), V^H(\cdot), V^L(\cdot))$ is defined is the Cartesian product of the space of probability distributions κ on $[0, +\infty)$, the space of functions $\tau(x)$ such that $\tau(x) \leq R(x)$ each x and the space of non-negative functions V^H and V^L . We define the ordering on these spaces as follows:

1. $\kappa \preceq \kappa'$ if κ' first order stochastically dominates κ .
2. $\tau \preceq \tau'$ if $\tau(x) \leq \tau'(x)$ each x , and similarly for V^H and V^L .

3. $(\kappa, \tau(\cdot), V^H(\cdot), V^L(\cdot)) \preceq (\kappa', \tau'(\cdot), V^{H'}(\cdot), V^{L'}(\cdot))$ if each component of the first vector is lower than the corresponding component of the second vector.

Since the objective function does not depend on τ and it is increasing in both V^H and V^L (thus implying quasi-supermodularity) we only have to prove quasi-supermodularity with respect to κ . For convenience, we remind here the reader of some basic definitions needed to apply the Milgrom-Shannon theorem.

Given a partially ordered set X and two elements x, y in X , we define $x \wedge y$ as the largest element of X such that $x \wedge y \preceq x$ and $x \wedge y \preceq y$. Similarly, $x \vee y$ is the smallest element in X such that $x \preceq x \vee y$ and $y \preceq x \vee y$. The set X is a lattice if, given x, y in X , we have that $x \wedge y$ and $x \vee y$ are also in X . A function f defined on the lattice X is quasi-supermodular if, given two elements $x, y \in X$, whenever the inequality $f(x) \geq f(x \wedge y)$ is satisfied we also have $f(x \vee y) \geq f(y)$.

Let now consider the space of probability distribution on the positive real line endowed with the first-order stochastic dominance order. Consider two distribution κ and κ' represented by the cumulative distribution functions F and G respectively. Then $\kappa \vee \kappa'$ has cumulative distribution function $H(x) = \min\{F(x), G(x)\}$, while $\kappa \wedge \kappa'$ has cumulative distribution function $L(x) = \max\{F(x), G(x)\}$. We will prove that for any function $f(x)$, if $\int f(x) dF \geq \int f(x) dL$ then $\int f(x) dH \geq \int f(x) dG$. The first inequality can be written as

$$\int f(x) dF \geq \int_{\{x|F(x) \geq G(x)\}} f(x) dF + \int_{\{x|F(x) < G(x)\}} f(x) dG \quad (27)$$

or

$$\int_{\{x|F(x) < G(x)\}} f(x) dF \geq \int_{\{x|F(x) < G(x)\}} f(x) dG.$$

The second inequality can be written as

$$\int_{\{x|F(x) \geq G(x)\}} f(x) dG + \int_{\{x|F(x) < G(x)\}} f(x) dF \geq \int f(x) dG \quad (28)$$

or

$$\int_{\{x|F(x) < G(x)\}} f(x) dF \geq \int_{\{x|F(x) < G(x)\}} f(x) dG.$$

But this implies that whenever inequality (27) is satisfied, inequality (28) must also be satisfied. Therefore any function defined on the real numbers is quasi-supermodular when defined over the space of probability distributions over the real line.

Finally, to prove that the objective function satisfies increasing difference in $(\kappa; K)$ we have to show that the difference between the objective function at K' and the objective function computed at K is increasing in κ whenever $K' > K$. To see this, observe that such difference is given by the function

$$I(K'', K) - I(K', K') = \begin{cases} (1-d)(K' - K) & \text{if } K'' \geq (1-d)K' \\ (1-q)(K'' - (1-d)(K - K')) & \text{if } (1-d)K' > K'' \geq (1-d)K \\ q(1-d)(K' - K) & \text{if } (1-d)K > K'', \end{cases}$$

which is increasing in K'' . Therefore, $E_\kappa \left[I(\tilde{K}, K) - I(\tilde{K}, K') \right]$ is increasing in κ . This proves that the optimal policy κ is increasing in K .

Suppose now that $\underline{K}'_s > (1-d)K$, so that at s investment is always strictly positive. Suppose now that the quantity of capital is increased by a small amount Δ such that the inequality $\underline{K}'_s > (1-d)(K + \Delta)$ still holds. It must be the case that the optimal policy remains the same, so that the value of the value function increases by $(1-d)\Delta$. If this were not the case then we would have $W(V, K + \Delta) > W(V, K) + (1-d)\Delta$, but this implies that by adopting at state (V, K) the policy adopted $(V, K + \Delta)$ we would get a value strictly

higher than $W(V, K)$, a contradiction. A similar reasoning implies that the optimal policy does not change when $\bar{K}'_s < (1-d)K$ and we increase K by Δ . \blacksquare

Proof of Proposition 5. The variable $\tilde{Y}(V, K)$ is given by

$$\tilde{Y}(V, K) = \begin{cases} S(K) & \text{with prob. } \lambda \\ \tilde{\theta}R(\tilde{K}') - I(\tilde{K}', K) & \text{with prob. } 1 - \lambda \end{cases}$$

and the variable $\tilde{W}(V, K)$ is given by

$$\tilde{W} = \begin{cases} 0 & \text{with prob. } \lambda \\ W(\tilde{\theta}, V(\tilde{\theta}, \tilde{K}'), \tilde{K}') & \text{with prob. } 1 - \lambda. \end{cases}$$

Let $\tilde{Y}_\Delta, \tilde{W}_\Delta, \tilde{K}', \lambda_\Delta$ be the variables corresponding to the state $(V, K + \Delta)$ and let

$$\begin{aligned} \bar{Y} &= \lambda S(K) + (1 - \lambda) E \left[pR(\tilde{K}') - I(\tilde{K}', K) \right], \\ \bar{Y}_\Delta &= \lambda_\Delta S(K + \Delta) + (1 - \lambda_\Delta) E \left[pR(\tilde{K}') - I(\tilde{K}', K + \Delta) \right], \\ \bar{W} &= (1 - \lambda) E \left[W(V(\tilde{K}'), \tilde{K}') \right], \\ \bar{W}_\Delta &= (1 - \lambda_\Delta) E \left[W(\tilde{\theta}, V(\tilde{K}'), \tilde{K}'_\Delta) \right], \end{aligned}$$

so that

$$W(s) = \bar{Y} + \delta \bar{W} \qquad W(s + \Delta) = \bar{Y}_\Delta + \delta \bar{W}_\Delta.$$

The expected returns are

$$\begin{aligned} \bar{r} &= \frac{\bar{Y} + \bar{W} - W(s)}{W(s)} = \frac{(1 - \delta) \bar{W}}{W(s)} \\ \bar{r}_\Delta &= \frac{\bar{Y}_\Delta + \bar{W}_\Delta - W(s + \Delta)}{W(s + \Delta)} = \frac{(1 - \delta) \bar{W}_\Delta}{W(s + \Delta)}. \end{aligned}$$

Therefore, the condition $\bar{r} \geq \bar{r}_\Delta$ is equivalent to

$$\frac{W(s + \Delta)}{W(s)} \geq \frac{\bar{W}_\Delta}{\bar{W}}.$$

Suppose first that

$$\frac{\bar{Y}_\Delta}{\bar{Y}} \geq \frac{\bar{W}_\Delta}{\bar{W}}$$

This implies

$$\begin{aligned} \frac{\frac{\bar{Y}_\Delta}{\bar{W}_\Delta}}{\frac{\bar{Y}}{\bar{W}}} \geq 1 &\implies \frac{\frac{\bar{Y}_\Delta}{\bar{W}_\Delta} + \delta}{\frac{\bar{Y}}{\bar{W}} + \delta} \geq 1 \implies \frac{\bar{Y}_\Delta + \delta \bar{W}_\Delta}{\bar{Y} + \delta \bar{W}} \geq \frac{\bar{W}_\Delta}{\bar{W}} \\ &\rightarrow \frac{W(s + \Delta)}{W(s)} \geq \frac{\bar{W}_\Delta}{\bar{W}}, \end{aligned}$$

so that we are done. Next assume

$$\frac{\bar{Y}_\Delta}{\bar{Y}} < \frac{\bar{W}_\Delta}{\bar{W}}. \tag{29}$$

If this is the case we must have

$$\frac{\bar{Y}_\Delta}{\bar{Y}} < \frac{W(s+\Delta)}{W(s)} < \frac{\bar{W}_\Delta}{\bar{W}},$$

which in turn implies

$$\frac{\bar{Y}_\Delta}{W(s+\Delta)} < \frac{\bar{Y}}{W(s)}. \quad (30)$$

Let $\Delta W = W(s+\Delta) - W(s)$ and $\Delta Y = \bar{Y}_\Delta - \bar{Y}$. Then inequality (30) can be written as

$$\frac{\Delta Y}{\Delta W} < \frac{\bar{Y}}{W(s)}.$$

If $\lambda < 1$ then $\frac{\bar{Y}}{W(s)} < 1$. However, by the envelope theorem, we have

$$\lim_{\Delta K \rightarrow 0} \frac{\Delta Y}{\Delta W} = 1.$$

Thus, for ΔK sufficiently small inequality (29) can't be satisfied. ■

Proof of Proposition 6. Suppose first that at state s we have $\underline{K}'_s > (1-d)K$, i.e. investment is always positive. In this case the variance of the return is

$$\sigma_s^2 = E_{\tilde{\theta}, \tilde{K}'} \left[\left(\frac{\tilde{\theta}R(\tilde{K}') - (\tilde{K}' - (1-d)K) + W(\tilde{\theta}, V(\tilde{\theta}, \tilde{K}'), \tilde{K}') - W}{W} - \bar{r} \right)^2 \right].$$

Let

$$Z = E_{\tilde{\theta}, \tilde{K}'} \left[\tilde{\theta}R(\tilde{K}') - \tilde{K}' + W(\tilde{\theta}, V(\tilde{\theta}, \tilde{K}'), \tilde{K}') \right].$$

Then we can write

$$\bar{r} = \frac{Z + (1-d)K - W}{W}$$

and

$$\sigma_s^2 = E \left[\left(\frac{\tilde{\theta}R(\tilde{K}') - \tilde{K}' + W(\tilde{\theta}, V(\tilde{\theta}, \tilde{K}'), \tilde{K}') - Z}{W} \right)^2 \right]. \quad (31)$$

If $\underline{K}' > (1-d)K$ then a slight increase ΔK in the level of capital does not change the optimal policy. Thus, the numerator in (31) remains constant, while the denominator increases to $W + (1-d)\Delta K$. The new variance is therefore

$$\sigma_{s+\Delta}^2 = E \left[\left(\frac{\tilde{\theta}R(\tilde{K}') - \tilde{K}' + W(V(\tilde{K}'), \tilde{K}') - Z}{W + (1-d)\Delta K} \right)^2 \right] < \sigma.$$

The proof for the case $\bar{K}'_{s+\Delta} < (1-d)K$ is similar. ■

Proof of Proposition 7. If $V < V_{(K)}$ then the optimal policy is to liquidate with probability $1 - \frac{V}{V_{(K)}}$ and to continue with probability $\frac{V}{V_{(K)}}$. In case of continuation, the optimal policy is the one obtained solving problem 12 under the constraints 13, 14, 15 with $V_c = V_{(K)}$. Thus, conditional on continuation the random variable

$\widetilde{W}(V, K)$ is the same for each value $V \in (0, V_{(K)})$, which in turn implies that $E \left[\widetilde{W}(V, K) \right]$ is constant with respect to V over the range $V \in (0, V_{(K)})$. Since

$$W(V, K) = \left(1 - \frac{V}{V_{(K)}}\right) S(K) + \delta \frac{V}{V_{(K)}} W_c(V_{(K)}, K),$$

the expected rate of return can be written as

$$\begin{aligned} E[\tilde{r}(V, K)] &= \frac{(1 - \delta) \frac{V}{V_{(K)}} E \left[\widetilde{W}(V, K) \right]}{\left(1 - \frac{V}{V_{(K)}}\right) S(K) + \delta \frac{V}{V_{(K)}} W_c(V_{(K)}, K)} = \\ &= \frac{(1 - \delta) E \left[\widetilde{W}(V, K) \right]}{\left(\frac{V_{(K)}}{V} - 1\right) S(K) + \delta W_c(V_{(K)}, K)}. \end{aligned} \tag{32}$$

Since $E \left[\widetilde{W}(V, K) \right]$ is constant the expression in (32) is increasing in V .

To see that $E[\tilde{r}(V, K)]$ is constant when $V > \frac{pR(K^*)}{1-\delta}$ simply observe that in that range the optimal policy of the firms is equal to the first best policy. ■

Appendix II

The computational work in this paper consists of two main steps. First, we have computed the value function and the global optimal policies. Second, we have constructed the empirical joint distribution of firms' sizes and values, and their growth rates.

The next table shows the parameters values that we have used.

Parameter	Description	Value
p	Probability of good outcome	0.8
a	Curvature of R ($R(K) \equiv \frac{K^a}{a}$)	0.5
q	Resale Price	0.8
S_0	Scrap Value	100
d	Depreciation Rate	0.07
δ	Discount Factor	0.98

Table 2: With this parametrization $K^* = 81.5$ and $V^* = 722.3$.

In the first step the value function and the global optimal policies (for λ, K', V_L and V_H) have been computed using a value function iteration algorithm. We did not use projection methods, due to the presence of occasionally binding constraints (incentive compatibility and non-negativity of promised utilities). Moreover, since the optimal policies for λ and K' have kinks (see Figures 11 and 12, respectively), we have used a grid-based¹⁰ non-parametric approach with interpolation, instead of polynomial-based approximations. The four dimensional control space (λ, K', V_L, V_H) generated an additional source of complexity. Even after eliminating one control using the promise keeping equality constraint, the optimization problem at each step of the value function iteration algorithm remains three-dimensional.

Once the optimal policy has been computed, we moved to the second step. For each time horizon $T = 10, 20, \dots, 100$, we have simulated the life of 2000 firms. After discarding all firms that are liquidated before period T , we have built:

-
- i. the joint distribution of the firms' sizes and values and all associated distributions conditional on size (Figures 1 – 10); and
- ii. the expected growth rates, and standard deviations, conditional on firm size (Table 1).

¹⁰The grid for V has 40 nodes, uniformly spaced between 0 and $0.95 \cdot V^*$. The grid for K has 30 nodes, uniformly spaced between 0 and $1.5 \cdot K^*$. We have included values well above K^* to verify that the these values are never reached in our simulations, as predicted by the theory.

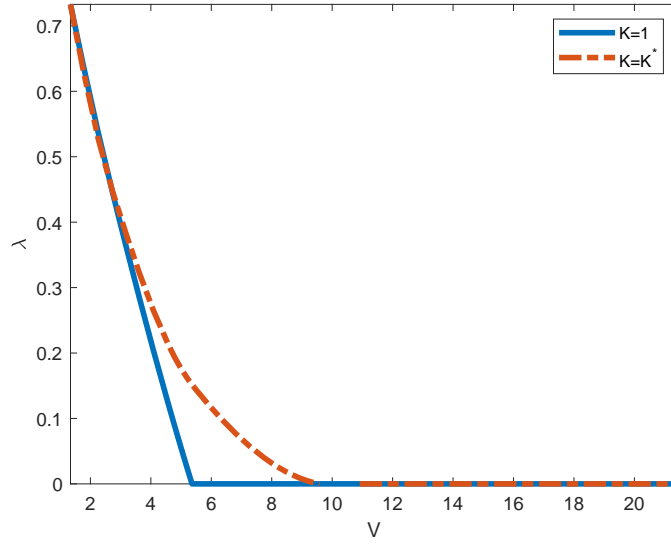


Figure 11: Two slices (for $K = 1$ and $K = K^*$) of the λ policy for a range of relatively small values of V

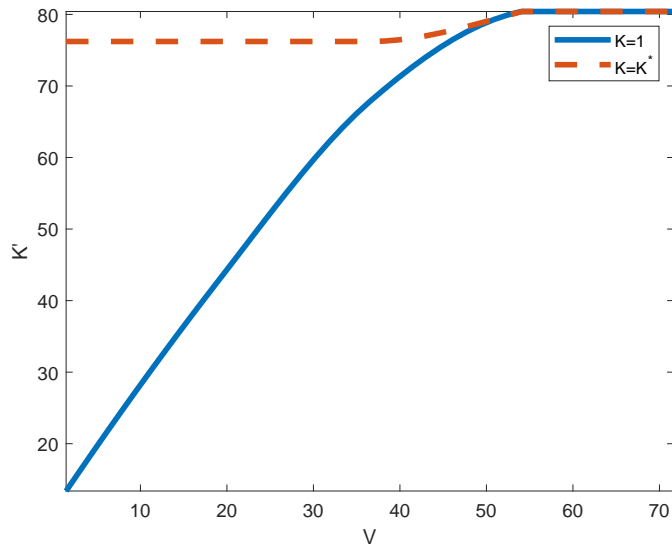


Figure 12: Two slices (for $K = 1$ and $K = K^*$) of the K' policy for a range of relatively small values of V

References

- [1] Rui Albuquerque, R. and H. Hopenhayn (2004), ‘Optimal Lending Contracts and Firm Dynamics’, *Review of Economic Studies*, **71**(2): 285–315.
- [2] Atkeson, A. and Robert Lucas (1992), ‘On Efficient Distribution with Private Information’, *Review of Economic Studies*, **59**(3): 427–453.
- [3] Bolton, P., H. Chen, Z and N. Wang (2011), ‘A Unified Theory of Tobin’s q , Corporate Investment, Financing and Risk Management’, *Journal of Finance*, **LXVI**: 1545–1578.
- [4] Clementi, G.L. and H. Hopenhayn (2006), ‘A Theory of Financing Constraints and Firm Dynamics’, *Quarterly Journal of Economics*, **121**: 229–266.
- [5] Cooley, T., R. Marimon and V. Quadrini (2004), ‘Aggregate Consequences of Limited Contract Enforceability’, *Journal of Political Economy*, **112**: 817–847.
- [6] Cooley, T. and V. Quadrini (2001), ‘Financial Markets and Firm Dynamics’, *American Economic Review*, **91**: 1286–1310.
- [7] De Marzo, P. and M. Fishman (2007), ‘Agency and Optimal Investment Dynamics’, *Review of Financial Studies*, **20**: 151–188.
- [8] De Marzo, P. and M. Fishman (2007), ‘Optimal Long-Term Financial Contracting’, *Review of Financial Studies*, **20**: 2079–2128.
- [9] De Marzo, P., M. Fishman, Z. He and N. Wang (2012), ‘Dynamic Agency and the q Theory of Investment’, *Journal of Finance*, **LXVII**: 2295–2340.
- [10] Dunne, T. Roberts, M. and L. Samuelson (1988), ‘Patterns of Firm Entry and Exit in U.S. Manufacturing Industries’, *Rand Journal of Economics*, **19**(4): 495–515.
- [11] Dunne, T. Roberts, M. and L. Samuelson (1989), ‘The Growth and Failure of U.S. Manufacturing Plants’, *Quarterly Journal of Economics*, **104**(4): 671–698.
- [12] Eisefeldt, A. and A. Rampini (2007), ‘New or used? Investment with credit constraints’, *Journal of Monetary Economics*, **54**: 2656–2681.
- [13] Evans, D. (1987), ‘The Relationship between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries’, *Journal of Industrial Economics*, **35**(4): 567–81.
- [14] Evans, D. (1987), ‘Tests of Alternative Theories of Firm Growth’, *Journal of Political Economy*, **95**(4): 657–74.
- [15] Fama, E. and K. French (1992), ‘The Cross-Section of Expected Stock Returns’, *Journal of Finance*, **47**(2): 427–465.
- [16] Fazzari, S., G. Hubbard, B. Petersen, A. Blinder and J. Poterba (1988), ‘Financing Constraints and Corporate Investment’, *Brookings Papers on Economic Activity*, **1**: 141–206.
- [17] Fu, S. and R. V. Krishna (2017), ‘Dynamic Financial Contracting with Persistent Private Information’, working paper, University of Rochester.
- [18] Gavazza, A. (2011), ‘Leasing and Secondary Markets: Theory and Evidence from Commercial Aircrafts’
- [19] Gomes, J. (2001), ‘Financing Investment’, *American Economic Review*, **91**: 1263–1285.

- [20] Hayashi, F. (1982), ‘Tobin’s Marginal q and Average q: A Neoclassical Interpretation’, *Econometrica*, **50**: 213–224.
- [21] Lorenzoni, G. and K. Walentin (2007), ‘Financial Frictions, Investment and Tobin’s Q’, NBER working paper n^o 13092.
- [22] Kiyotaki, N. and J. Moore (1997), ‘Credit Cycles’, *Journal of Political Economy*, **105**: 211–248.
- [23] Lanteri, A. (2018), ‘The Market for Used Capital. Endogenous Irreversibility and Reallocation over the Business Cycle’, forthcoming, *American Economic Review*.
- [24] Milgrom, P. and C. Shannon (1994), ‘Monotone Comparative Statics’, *Econometrica*, **62**: 157–180.
- [25] Quadrini, V. (2004), ‘Investment and Liquidation in Renegotiation–Proof Contracts with Moral Hazard’, *Journal of Monetary Economics*, **51**: 713–751.
- [26] Ramey, V. and M. Shapiro (2001), ‘Displaced Capital: A Study of Aerospace Plant Closings’, *Journal of Political Economy*, **109**: 958–992.
- [27] Rampini, A. and S. Viswanathan (2013) ‘Collateral and Capital Structure’, *Journal of Financial Economics*, **109**: 466–429.
- [28] Spear, S., and Srivastava, S. (1987), ‘On Repeated Moral Hazard with Discounting’ *The Review of Economic Studies*, **54(4)**, 599-617
- [29] Stokey, N., R. Lucas and E. Prescott (1989), *Recursive Methods in Economic Dynamics*, Harvard University Press, Cambridge, Mass.
- [30] Sutton, J. (1997), ‘Gibrat’s Legacy’, *Journal of Economic Literature*, **35**: 40–59.