

FIRM SIZE AND OPTIMAL GROWTH RATES[†]

by

Uzi Segal^{*}

and

Avia Spivak^{**}

August 1985

Revised January 1986

UCLA Working Paper #380

[†]We are grateful to Siegfried Berninghaus, Richard Kihlstrom, Ariel Pakes, Joel Sobel, and John Sutton for their comments.

^{*}Department of Economics, University of California, Los Angeles, 405 Hilgard Ave., Los Angeles, CA 90024.

^{**}Department of Economics, Ben Gurion University, Beer Sheva, Israel

Abstract

This paper presents a theoretical model in which, due to dissolution costs, the rate of growth of small firms tends to be higher and more variable than that of larger firms. This model also predicts that for large firms the rate of growth is fixed, as claimed by Gibrat's Law.

1. Introduction

One of the major decisions a firm must make is how to allocate the profits between dividends and retained earnings. Retained earnings reinvested in the firm provide for future growth. The rate by which the firm grows is thus an endogenous decision variable, arrived at as a result of intertemporal maximization.

This consideration is entirely overlooked in discussions of growth rate and firms' size. Simon, who contributed significantly to this literature (see Simon and Bonini (1958) and the references therein), adopted Gibrat's Law, the assumption that a single firm's rate of growth is independent of its size. He proceeded to obtain the stationary size distribution consistent with this assumption. Jovanovic (1982) also assumed that rates of growth, albeit stochastic, are exogenous to the firm.

This paper seeks to close this gap by regarding the growth rate as a solution to an explicit intertemporal maximization problem. In order to avoid the agency problem, it is assumed that the firm is owned by a single shareholder who is also the manager. The owner is risk-neutral and seeks to maximize the expected discounted value of his income from the firm. The firm's growth rate is stochastic and diversified. Increasing reinvestment in the firm changes the parameters that govern the stochastic growth process so that the expected growth rate increases as well. Reinvestment of profits is thus beneficial on two counts: First, it increases the expected size of the firm and hence the total expected future profits. Secondly, bigger firms are less likely to go out of business (and to suffer dissolution costs), because the growth is diversified. Hence, investment in growth decreases the expected dissolution costs of the firm.

Our modeling thus introduces increasing returns to scale, not because of technology, which shows constant returns to scale, but because units of the firm insure each other mutually against the disappearance of the firm.

Since there are scale economies, one may suspect that the rate of growth will change with the size of the firm. This indeed is found to be true. Our major result is that smaller firms on the average grow faster than bigger firms. We also show that when the size of the firm tends to infinity, the growth rate converges to a positive number. In other words, large firms satisfy Gibrat's Law, that the firm grows at a fixed rate.

Jovanovic (1982) explained the same empirical phenomenon with a different approach. Firms enter an industry not knowing their true cost function but only the industry's average. Every period they use the observation on the noisy cost to update their estimate of cost and their optimal production decisions. So, firms that are more efficient than the average will (on the average) learn that their output will grow. The other, less efficient firms, will produce less and less and eventually will exit the industry. In this model it turns out that rates of growth for smaller firms are larger and more variable than those of bigger firms. An entirely different theoretical approach, that of Lucas (1978), determined the size and growth of firms by the optimal managerial span of control.

The empirical evidence supports our theoretical finding that for small firms growth rates slow down with size and that for large firms growth rates tend to be fixed. Mansfield (1962) found an inverse relationship between size and rate of growth for a sample of small firms. Pasigian and Hymer (1962) and Hart and Prais (1956) found for their samples of large firms that growth rates were independent of size. One must realize, though, that the sample of small firms may have an upward sampling bias of the growth rates for small firms

because of possible bankruptcy and hence exclusion from the sample of small firms with low rates of growth.

The rest of this paper is organized as follows. Section 2 presents the model. Section 3 considers optimal growth under the assumption that the rate of reinvestment in the firm does not vary throughout the firm's life. Section 4 relaxes this assumption and allows the rate of reinvestment to change at the firm's will. Section 5 concludes the paper.

2. The Model

We assume a firm whose units are independent. If a unit is successful, it will give rise to another unit. If it fails, the unit will disappear. This stylized description of diversified growth fits the way that some firms develop, especially firms that are R&D intensive: each unit ("division") of the firm produces a single product. New products are conceived in existing divisions, and then are produced in newly created divisions.¹

Formally we assume that the firm is a collection of units which are stochastically identical and independent. The number of units at time t is $N(t)$, and is assumed to be a random variable. Denote by N the initial number of units, $N = N(0)$. Each unit has a λdt probability of changing in the following time interval dt into two units and a μdt probability of disappearing, $\lambda, \mu \geq 0$. Thus, the model is assumed a continuous time model. Each unit generates an income stream of y as long as it exists. The number of units is changing at random discretely up and down by one unit at a time, the probability of two units changing simultaneously is zero. If the number of units go down to zero, the firm goes out of business.

¹See the description of the organization of Minnesota Mining & Manufacturing Co. in: "In Search of Excellence".

A major assumption of this model is that the firm's owners suffer a loss of L when the firm dissolves. These dissolution costs are the intangibles that are implied in the existence of the firm such as reputation, the ability of people to work as a team, etc. L is related to the cost of putting a new firm together and acquiring a reputation. The risk-neutral owners of the firm are interested in the expected value of the profits, discounted at a positive discount factor δ . ($\delta > \lambda - \mu$ to ensure boundedness of this value).

Initially, we assume that λ , μ , and y are given to the firm. We will explore the following question: What is the value (at $t = 0$) of a firm, comprised of N units, to its owners. Later, we will consider optimizing behavior, where the firm uses the tradeoff between λ , μ , and y .

Given λ , μ , and y , the expected discounted income stream is made up of two parts: expected discounted income and expected discounted loss.

2.1 Expected Discounted Income

We start the discussion with a firm of one unit, so $N = 1$ at $t = 0$. The size of the firm as a function of time, $N(t)$, is a random variable. The expected discounted value of the income stream is

$$Y = E \left[\int_0^{\infty} e^{-\delta t} y N(t) dt \right] = \int_0^{\infty} e^{-\delta t} y E[N(t)] dt$$

Following Harris (1963, p. 104), $E(N(t)) = e^{-t(\mu-\lambda)}$. (Note that when $\mu = \lambda$ $E[N(t)] = 1$ and when $\lambda > \mu$, $E[N(t)]$ grows exponentially with time.)

Substituting and integrating with respect to time, with the assumption $\delta + \mu - \lambda > 0$, we obtain

$$Y = \frac{y}{\delta + \mu - \lambda}$$

Since the N units grow independently, the total expected discounted income is simply NY . Thus, regarding the income aspect, we have constant returns to

the size of the firm.

2.2 Expected Discounted Loss

Let $\phi = \phi(\lambda, \mu, \delta; N)$ be the expected discounted loss. In the Appendix it is proved that

$$\phi = \begin{cases} L\delta \int_0^{\infty} e^{-\delta t} \left| \frac{1 - e^{-t(\mu-\lambda)}}{1 - \frac{\lambda}{\mu} e^{-t(\mu-\lambda)}} \right|^N dt & \text{for } \lambda \neq \mu.^2 \\ L\delta \int_0^{\infty} e^{-\delta t} \left[\frac{\lambda t}{1+\lambda t} \right]^N dt & \text{for } \lambda = \mu. \end{cases}$$

We now list some properties of this loss function.

Property 1: ϕ is decreasing in δ and λ and is increasing in μ .

When δ is higher, future loss is less important. An increase in λ decreases the death probability ξ and so does a decrease in μ .³

Property 2: ϕ is homogeneous of degree 0 in λ, μ, δ .

Proof: Multiply λ, μ, δ by α and substitute $x = \alpha t$. The same ϕ integral is obtained.

²For computational purposes, one may substitute $x = e^{-t(\mu-\lambda)}$ and obtain

$$\phi = \frac{L\delta}{\lambda-\mu} \int_1^{\infty} x^{-\frac{\delta}{\lambda-\mu}-1} \frac{[x-1]^N}{\left[\frac{\lambda}{\mu}x-1\right]^N} dx \quad \text{for } \lambda \neq \mu.$$

³Formally, $\frac{\partial \phi}{\partial \delta} = -L \int_0^{\infty} t e^{-\delta t} f_N(t) dt < 0$. $\frac{\partial \phi}{\partial \lambda} < 0$ if

$$\frac{\partial}{\partial \lambda} \left[\frac{e^{t(\lambda-\mu)}}{\frac{\lambda}{\mu} e^{t(\lambda-\mu)} - 1} \right] < 0, \quad \text{which follows from the fact that } 1 - x - e^{-x} < 0 \quad \text{for}$$

all $x \neq 0$. A similar argument shows that $\frac{\partial \phi}{\partial \mu} > 0$.

Property 2 amounts to saying that ϕ is invariant to the units by which the time is measured.

Property 3: ϕ is declining and convex in N , and $\lim_{N \rightarrow \infty} \phi = 0$. Furthermore, $\lim_{N \rightarrow \infty} N\phi = 0$.

Proof: $\frac{\partial \phi}{\partial N} = L\delta \int_0^{\infty} e^{-\delta t} \xi^N \ln \xi dt < 0$ because $\xi < 1$ for all t .

$$\frac{\partial^2 \phi}{\partial N^2} = L\delta \int_0^{\infty} e^{-\delta t} \xi^N (\ln \xi)^2 dt > 0.^4$$

To show that $\phi = \int_0^{\infty} e^{-\delta t} \xi^N dt$ and $N\phi$ tend to zero as N tends to infinity, we use the facts that for all t , $\xi < 1$, $e^{-\delta t} \xrightarrow{t \rightarrow \infty} 0$, and ξ is nondecreasing in t . It follows that the sequence of functions $e^{-\delta t} \xi^N$ converge to the zero function uniformly in t , hence $\phi \rightarrow 0$ as $N \rightarrow \infty$. A similar proof holds for $N\phi$ by noting that $N\xi^N \rightarrow 0$ as $N \rightarrow \infty$. Q.E.D.

The last property will be very useful in the sequel because it characterizes the returns to scale that the firm possesses. Note that the only assumption used here is that the probability of disappearance ξ is less than one. This assumption must be satisfied in every stochastic model. The stronger assumption, responsible for Property 3, is the independence of units. Note that even if the loss from bankruptcy was proportional to size, we would still get the same returns to scale properties.

⁴Mathematically, the function is defined for all N . Its properties can thus be found by differentiation. Economically, ϕ is defined for $N = 1, 2, 3$. We will later use the fact that for $N = 0$, $\phi = L$.

2.3 Value of the Firm to its Owners

From the above analysis we know that the value of a firm of size N to its owners equals $NY - \phi(N)$. Since $\phi(N)$ is convex and diminishing to zero, $NY - \phi(N)$ is concave and asymptotic to the straight line NY .

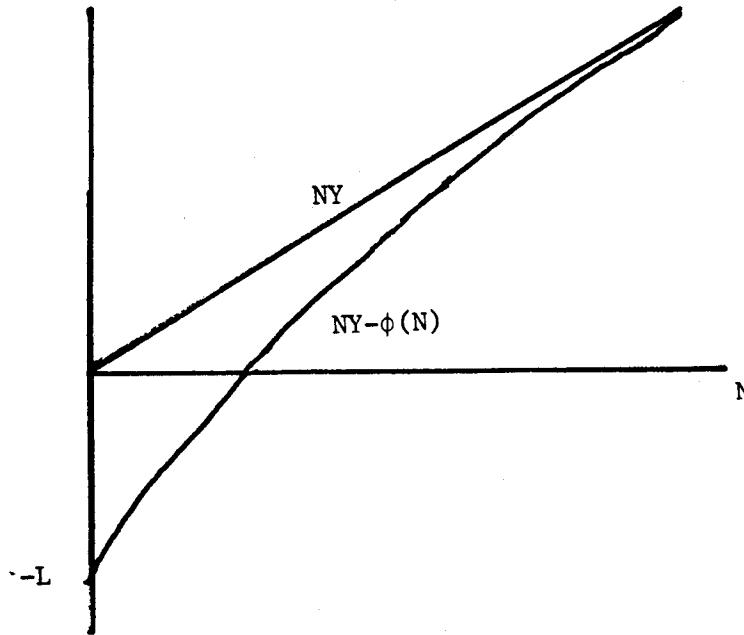


Figure 1

In Figure 1 we use the mathematical fact that for $N = 0$, $\phi = L$. The interpretation is that when $N = 0$ the firm is dissolved and must pay the cost L . We see that a sufficiently large firm has the same value to its owners in the presence or absence of dissolution cost. This suggests that the case $L = 0$ might be a useful benchmark for the behavior of firms.

3. Optimal Growth: The Naive Approach

The firm can change the parameters λ and μ by plowing back more or less of its income, thus changing the dividend stream y . We assume that each existing unit creates a total income stream of \bar{y} . The firm retains a stream of g that is reinvested in the firm to obtain growth and distributes dividends $y = \bar{y} - g$. It is assumed that the sum $b = \lambda + \mu$, which governs the

chances of one unit to change, is constant. Thus, an increase in g increases λ according to the relationship $g = g(\lambda)$ and decreases μ by the same magnitude. We assume that $g' > 0$, $g'' > 0$, $\lim_{\lambda \rightarrow 0} g'(\lambda) = 0$, and $\lim_{\lambda \rightarrow b} g'(\lambda) = \infty$. Because the units are independent, the parameters and income stream may, in principle, vary across different units. However, because of symmetry and the convexity of g , the decisions for all the units will be identical.

In this section we consider two problems. The first is the optimal growth for a firm without bankruptcy cost. The second is the case of the myopic or naive firm. This firm plans on the assumption that the optimal g will not change with N . However, when N changes the firm finds it optimal to change g again. We will show that for this case, g is inversely related to N . This will be the first proof that growth rate decreases with size.

3.1 The Zero Dissolution Cost Case

When $L = 0$, the expected discounted profits of the firm are $NY = N \frac{y}{\delta + \mu - \lambda}$. Substituting $y = \bar{y} - g(\lambda)$ and $\mu = b - \lambda$, the problem becomes

$$\text{Max}_{\lambda} N \frac{\bar{y} - g(\lambda)}{\delta + b - 2\lambda}$$

The first order conditions for maximum are

$$N \frac{2(\bar{y} - g(\lambda)) - g'(\lambda)(\delta + b - 2\lambda)}{[\delta + b - 2\lambda]^2} = 0$$

or

$$g'(\lambda) = 2 \frac{\bar{y} - g(\lambda)}{\delta + b - 2\lambda}$$

Because $\delta + b - 2\lambda = \delta + \mu - \lambda > 0$, a sufficient condition for maximum is $g''(\lambda) > 0$. This condition may take an alternative form. Denote by g^* , $\lambda^* = \lambda(g^*)$, and $\mu^* = b - \lambda^*$ the solution to the maximum problem and $a^* = \frac{\bar{y} - g^*}{\delta + \mu^* - \lambda^*}$. Then at

g^* ,

$$(1) \quad g'(\lambda^*) = 2a^*$$

As expected, the optimal growth decision is independent of N and the maximum value of the firm is simply a^*N .

3.2 Optimal Growth Rate as a Function of Size: The Myopic Approach

Let $L > 0$ and assume that the firm decides on the optimal λ when its size is N , based on the myopic assumption that this λ will be kept unchanged in the future. Since λ is constant in all subsequent periods, the loss function ϕ is well defined and the firm seeks to solve

$$\text{Max}_{\lambda} NY(\lambda) - \phi(N, \lambda) = \text{Max}_{\lambda} \left[N \frac{\bar{y} - g(\lambda)}{\delta + b - 2\lambda} - L\delta \int_0^{\infty} e^{-\delta t} \xi^N dt \right]$$

where

$$\xi = \frac{1 - e^{-t(b-2\lambda)}}{1 - \frac{\lambda}{b-\lambda} e^{-t(b-2\lambda)}}$$

We denote the solution to the myopic problem by $\tilde{\lambda}(N)$. The following theorem states that growth declines with size:

Theorem 1: The sequence $\tilde{\lambda}(N)$ is a monotonically decreasing sequence converging to the optimal no dissolution loss λ^* .

Proof: The proof is by comparative statics of the optimum conditions.

Mathematically, the function is defined for all nonnegative values of N although it makes economic sense only for nonnegative integer values. Thus, we will show that $\frac{\partial \tilde{\lambda}}{\partial N} < 0$.

The first order conditions are (the superscript \sim on λ is deleted for convenience):

$$(2) \quad H(N, \lambda) \stackrel{\text{def}}{=} N \frac{\partial Y}{\partial \lambda} - L\delta \int_0^{\infty} e^{-\delta t} N \xi^{N-1} \frac{\partial \xi}{\partial \lambda} dt = 0$$

By the implicit function theorem, $\frac{\partial \lambda}{\partial N} = -\frac{H_N}{H_\lambda}$. $H_\lambda < 0$ by second order conditions for maximum. To show that $\frac{\partial \lambda}{\partial N} < 0$, it suffices to prove that $H_N < 0$. Now,

$$H_N = \frac{\partial Y}{\partial \lambda} - L\delta \int_0^\infty e^{-\delta t} \xi^{N-1} \frac{\partial \xi}{\partial \lambda} dt - L\delta \int_0^\infty e^{-\delta t} N \xi^{N-1} \ln \xi \frac{\partial \xi}{\partial \lambda} dt$$

By (2) the sum of the first two entries is zero and

$$H_N = -L\delta \int_0^\infty e^{-\delta t} N \xi^{N-1} \ln \xi \frac{\partial \xi}{\partial \lambda} dt.$$

$\xi < 1$, hence, $\ln \xi < 0$. $\frac{\partial \xi}{\partial \lambda} < 0$ for $\lambda \neq \frac{b}{2}$ by straightforward differentiation,⁵ hence $H_N < 0$.

Since $\lim_{N \rightarrow \infty} NY(\lambda) \rightarrow 0$, $\tilde{\lambda}(\infty)$ equals λ^* , the non-bankruptcy cost solution. Q.E.D.

4. Optimal Growth: The Dynamic Programming Approach

The firm may learn from experience that the optimal growth rate changes with size. It may then take into account the dependency of λ on N in its optimization. This more sophisticated firm solves a dynamic programming problem. Intuitively, the knowledge that a decision on today's λ does not dictate tomorrow's λ allows the firm greater freedom and we expect λ to be more responsive to the change of N . Using the dynamic programming approach, we will show that λ decreases with the size of the firm.

⁵After some manipulations, one finds that $\frac{\partial \xi}{\partial \lambda} < 0$ iff $2t[1 - \frac{\lambda}{b-\lambda} e^{t(2\lambda-b)}] > [1 - e^{t(2\lambda-b)}][\frac{b}{(b-\lambda)^2} + \frac{2\lambda b}{b-\lambda}]$ or equivalently if the

function $\theta(\lambda) = 2t(b-\lambda)(b-2\lambda) - b(1 - e^{t(2\lambda-b)})$ is always positive. $\theta(\lambda)$ is a convex function with the minimum occurring at $\lambda = \frac{b}{2}$. $\theta(\frac{b}{2}) = 0$, hence $\theta(\lambda) > 0$ for all $\lambda \neq \frac{b}{2}$.

4.1 The Dynamic Program

We denote by $V(N)$ the maximum expected discounted profit that a firm of size N can attain, if it chooses a different optimal λ for different sizes of N . By definition, $V(0) = -L$. Given the choice of λ , a firm starting with N operating units will have a stream of income $Ny = N[\bar{y}-g(\lambda)]$ until the firm changes its size to either $N+1$ or $N-1$. Denoting the random date of the change by t_1 , the expected discounted value of the firm is

$$W(N, \lambda) = E \left[\int_{t_1}^{\infty} Ne^{-\delta t} [\bar{y}-g(\lambda)] dt + \frac{\lambda}{\lambda+\mu} e^{-\delta t_1} V(N+1) + \frac{\mu}{\lambda+\mu} e^{-\delta t_1} V(N-1) \right]$$

The first part of the sum is the discounted value of the income stream that the N units produce. In the second part, $\frac{\lambda}{\lambda+\mu}$ is the probability that the firm will grow in one unit, conditional upon the occurrence of change.

Because we have a continuous time model, the change affects one unit only.

The probability of more than one unit changing at exactly the same time is zero and hence this event is ruled out. The second and third parts are thus the discounted values of $N+1$ and $N-1$ units, weighted by their conditional probabilities. W depends on λ because the distribution of t_1 depends on λ , $V(N+1)$ and $V(N-1)$ are given numbers, and $\mu = b - \lambda$. If $W(N)$ is maximized, $V(N)$ is obtained. So

$$V(N) = \max_{\lambda} W(N, \lambda) = \max_{\lambda} E \left[\int_{t_1}^{\infty} Ne^{-\delta t} [\bar{y}-g(\lambda)] dt + \frac{\lambda}{\lambda+\mu} e^{-\delta t_1} V(N+1) + \frac{\mu}{\lambda+\mu} e^{-\delta t_1} V(N-1) \right]$$

We first calculate $W(N, \lambda)$ and then apply the optimization. Noting that

$$\int_0^{\infty} Ne^{-\delta t} [\bar{y}-g(\lambda)] dt = \frac{N}{\delta} (1-e^{-\delta t_1}) [\bar{y}-g(\lambda)], \text{ we obtain}$$

$$(3) \quad W(N, \lambda) = E \left[\frac{N}{\delta} (1-e^{-\delta t_1}) [\bar{y}-g(\lambda)] + \frac{\lambda}{\lambda+\mu} e^{-\delta t_1} V(N+1) + \frac{\mu}{\lambda+\mu} e^{-\delta t_1} V(N-1) \right]$$

The probability distribution of t_1 is the solution to the following well-known elementary problem: given N light bulbs whose life is distributed exponentially with a parameter b , what is the probability distribution of the time t_1 elapsing until the first bulb dies out? The density function obtained is $f(t_1) = Nbe^{-Nbt_1}$.⁶ We can now calculate

$$E_{t_1} [e^{-\delta t_1}] = \int_0^{\infty} e^{-\delta t_1} f(t_1) dt_1 = \int_0^{\infty} e^{-\delta t_1} Nbe^{-bNt_1} dt_1 = \frac{Nb}{\delta + Nb}$$

Substituting into (3) one obtains

$$(4) \quad W(N, \lambda) = \frac{N}{\delta + Nb} [\bar{y} - g(\lambda) + \lambda V(N+1) + (b - \lambda)V(N-1)]$$

The optimal λ satisfies the following first order conditions:

$$(5) \quad g'(\lambda) = V(N+1) - V(N-1)$$

$g''(\lambda) > 0$ is a sufficient condition for a maximum.

4.2 Optimal Policy for Zero Dissolution Cost Using Dynamic Programming

We illustrate the use of the optimality condition by showing that in the case of no dissolution cost, $V(N) = a^*N$, or that the best policy λ^* found above may be obtained by using our new criterion. Of course, since the best policy in the case $L = 0$ is independent of N , the use of dynamic programming is redundant and this illustration merely checks the consistency of our two different approaches. We assume that $V(N+1) = a^*(N+1)$ and $V(N-1) = a^*(N-1)$. We obtain $V(N) = a^*N$, and that the optimal policy λ^* satisfies equation (1):

⁶Formally, the cumulative distribution is $F(t_1) = 1 - [1 - \int_0^{t_1} be^{-bt} dt]^N = 1 - e^{-bNt_1}$. The density function is $f(t_1) = \frac{d}{dt_1} F(t_1) = Nbe^{-bNt_1}$.

$$(1) \quad g'(\lambda^*) = 2a^*$$

First, $g'(\lambda) = V(N+1) - V(N-1) = a^*(N+1) - a^*(N-1) = 2a^*$. Now

$$V(N) = \frac{N}{\delta + Nb} [\bar{y} - g(\lambda^*) + \lambda^*V(N+1) + (b - \lambda^*)V(N-1)]$$

Substituting $\bar{y} - g(\lambda^*) = a^*(\delta + b - 2\lambda)$ (see the definition of a^* in Section 3.1), $V(N+1) = a^*(N+1)$, $V(N-1) = a^*(N-1)$, one obtains $V(N) = a^*N$.⁷

4.3 Properties of the $V(N)$ Function

The major property that we want to establish here is the concavity of the function V in N . The concavity is likely to be present because the loss becomes less likely when size increases, and the incremental improvement eventually goes down to zero. As we will see later, the concavity implies that the growth rate decreases with size.

Claim 1: $V(N)$ is strictly increasing.

Proof: Suppose that instead of N units the company has $N+1$ units. If it operates N units with the former policy and the $N+1$ st unit with some λ such that $\bar{y} - g(\lambda) > 0$, it will have an additional expected income. Its expected loss will go down because the expected loss decreases in N . Since it has a feasible policy that creates more value, $V(N+1) > V(N)$.

Q.E.D.

Claim 2: $a^*N - \phi(N, \lambda^*) < V(N) < a^*N$. Hence, $\lim_{N \rightarrow \infty} V(N) - a^*N = 0$.⁸

⁷This "proof" depends, of course, on the assumption that $V(N+1) = a^*(N+1)$ and $V(N-1) = a^*(N-1)$. This section illustrates the dynamic programming process; it does not solve the zero dissolution cost case.

⁸The right hand inequality holds only for $L > 0$.

Proof: The left hand inequality follows from the fact that $V(N)$ is optimal for the loss L and $a^*N - \phi(N, \lambda^*)$ is available. The right hand inequality follows from optimality of a^*N for the no loss case, and from the fact that the loss decreases the income stream. The limiting results follow from

$$\lim_{N \rightarrow \infty} \phi(N, \lambda^*) = 0.$$

Q.E.D.

According to Claim 2, the V function is bounded between two concave functions (see Property 3 in Section 2.2). One is thus inclined to believe that V is itself concave. This more difficult result is proved in the next claim.

Claim 3: $V(N)$ is a concave function. It is strictly concave when $L > 0$.

Proof: We wish to show that for all $N > 1$, $V(N+1) - V(N) < V(N) - V(N-1)$ and equivalently, that $2V(N) - V(N+1) - V(N-1) > 0$. $V(N) > W(N, \lambda^*)$, thus it is sufficient to show that

$$2W(N, \lambda^*) - V(N+1) - V(N-1) > 0$$

By equation (4), $W(N, \lambda^*) = \frac{N}{\delta + Nb} (\bar{y} - g^*) + \frac{\lambda^* N}{\delta + Nb} V(N+1) + \frac{(b - \lambda^*) N}{\delta + Nb} V(N-1)$, hence

$$\begin{aligned} 2W(N, \lambda^*) - V(N+1) - V(N-1) = \\ \frac{N \{ 2(\bar{y} - g^*) + (2\lambda^* - b)[V(N+1) - V(N-1)] \} - \delta [V(N+1) + V(N-1)]}{\delta + Nb} \end{aligned}$$

When a firm with $N-1$ units receives two more units, it can always operate these two units with the strategy λ^* , creating an extra discounted expected income of $\frac{2(\bar{y} - g^*)}{\delta + b - 2\lambda^*}$. The two extra units also decrease the expected loss, hence the maximum value $V(N+1)$ satisfies

$$(6) \quad V(N+1) > V(N-1) + \frac{2(\bar{y} - g^*)}{\delta + b - 2\lambda^*}$$

By Claim 2, $V(N+1) < a^*(N+1)$ and $V(N-1) < a^*(N-1)$, hence

$$(7) \quad V(N+1) + V(N-1) < 2a^*N = 2N \frac{\bar{y} - g^*}{\delta + b - 2\lambda^*}$$

Using (6) and (7) one obtains

$$2W(N, \lambda^*) - V(N+1) - V(N-1) >$$

$$\frac{1}{\delta+Nb} \left\{ N[2(\bar{y}-g^*) + (2\lambda^*-b) \frac{2(\bar{y}-g^*)}{\delta+b-2\lambda^*}] - 2\delta N \frac{\bar{y}-g^*}{\delta+b-2\lambda^*} \right\} = 0$$

Q.E.D.

4.4 A Proof That Optimal Growth Declines With Size

We now have all the elements to show that the optimal λ declines with N .

Theorem 2: When $L > 0$, the optimal growth rate $\lambda(N)$ is monotonically decreasing with N and converging to the no loss growth rate λ^* .

Proof: By (5), $\lambda(N)$ is defined by the condition $g'(\lambda(N)) = V(N+1) - V(N-1)$ or equivalently, $\lambda(N) = (g')^{-1} [V(N+1)-V(N-1)]$. The inverse function $(g')^{-1}$ exists, is continuous, and is monotonically increasing because $g'' > 0$. By Claim 3, $V(N+1) - V(N-1)$ is monotonically decreasing with N , hence $\lambda(N)$ is monotonically decreasing. A simple application of Claim 2 yields $\lim_{N \rightarrow \infty} [V(N+1) - V(N-1)] = 2a^*$. Since $(g')^{-1}$ is continuous and since by equation (1), $(g')^{-1}(2a^*) = \lambda^*$, $\lim_{N \rightarrow \infty} \lambda(N) = \lambda^*$. Q.E.D.

Theorem 2 states that the inverse relationship between size and rate of growth becomes weaker as the size of the firm increases. In other words, for sufficiently large firms the growth rate is independent of size. This result is known as Gibrat's Law, thus the Theorem shows that Gibrat's Law holds for large firms.

5. Concluding Remarks

This paper looked at the growth process as an endogenous process where the firm chooses its optimal growth rate to maximize its expected present value. The firm is a collection of independent units whose fates are

independent. At each instant, each unit may turn into two independent units or disappear, according to a well defined stochastic process. If all the units disappear, the firm goes out of business and suffers costs of dissolution. The firm is small relative to the industry, hence its growth potential is independent of its size. The owners of the firm are risk-neutral and interested in the discounted expected value of the firm's profit, including the dissolution cost. The size of the firm is the number of units that it operates. Because the units are independent, the discounted probability of bankruptcy is smaller when the firm is larger and thus the source of the returns to scale.

The stochastic process discussed above is called branching process. It was first studied by Dalton in the context of a family's survival. We add to this classical process a discounting of the future. Farrel (1970) used branching processes without discounting to examine questions of survival of firms.

By retaining earnings and reinvesting them in the firm, the firm can change the parameters governing the stochastic process, increasing the probability of multiplication and reducing the probability of disappearance. We assume that the sum of these probabilities is fixed, thus the firm's decision variable is the size of retained earnings.⁹ For any number of units that the firm owns and operates, it will decide on the growth rate. A consistent optimization must take into account the change in growth that will follow the expected increase or decrease in the firm's size or, in other words, will solve a dynamic program.

⁹We abstract from consideration of differential tax rates on dividends and capital gains, often cited as a reason for reinvestment and growth.

This paper found that smaller firms grow faster. This finding is true both for myopic (or constrained) optimization where the firm does not plan or cannot effect a future change in its behavior, and in the non-myopic optimization, where the firm takes fully into account its future reaction of the growth rate to the size. In both cases, the effect of size on change of growth diminishes. Very large firms grow at a rate independent of their size. The model also predicts that the actual rate of growth for smaller firms varies more than that of larger firms because the rate of growth for the firm is an average of the rate of growth of its units. Since units' growth is independent, variability must decrease with size.

The results of this paper depend crucially on the existence of dissolution costs. These costs are independent of size because in our model before a company goes out of business it must dwindle to a firm of the minimal size possible, which is one unit.

Our concept of dissolution costs includes a subset of the bankruptcy costs discussed in the financial literature.¹⁰ We do not have the administrative costs of selling the firm's assets to pay for its loans because we do not allow debts. Our dissolution cost is the loss of intangibles of the firm such as reputation, organization, etc., related to the cost of putting a new firm together.

¹⁰The significance of bankruptcy cost was discussed in the financial literature because of its bearing on the debt-equity decisions (Modigliani-Miller (1963)). There is ongoing debate about the importance of bankruptcy cost. Robinchek and Myers (1966), Baxter (1967), and others underscored their importance while Haugen and Senbet (1978), Miller (1977), and Warner (1977) claimed that bankruptcy costs are rather insignificant. (See also Ang, et al. (1982)).

Appendix

We start with N units and look for a distribution function $F_N(t)$ for the event that all N units disappear exactly at or before time t . Denote by $f_N(t)$ the corresponding density function. If this function is known, then the expected discounted loss is $\phi = \int_0^{\infty} L e^{-\delta t} f_N(t) dt$. Because units are independent, $F_N(t) = \xi^N$, where ξ is the probability that one unit will disappear before or at time t . Referring again to Harris (1963, p. 104),

$$\xi = \frac{1 - e^{-t(\mu-\lambda)}}{1 - \frac{\lambda}{\mu} e^{-t(\mu-\lambda)}} \quad \text{for } \lambda \neq \mu \quad \text{and} \quad \xi = \frac{\lambda t}{1 + \lambda t} \quad \text{for } \mu = \lambda.$$

(The case $\mu = \lambda$ can be obtained by letting $\mu \rightarrow \lambda$ in the former expression, using

l'Hospital's rule.) Note that when $\mu > \lambda$ $\lim_{t \rightarrow \infty} \xi = 1$, hence ruin is certain, but when $\mu < \lambda$ $\lim_{t \rightarrow \infty} \xi = \frac{\mu}{\lambda} < 1$, so the unit has a positive probability of living forever while giving birth to other units. Since $\lim_{t \rightarrow \infty} \xi^N < 1$, the whole firm has a positive probability of eternal life. Note also that when $\lambda = 0$ and $\mu > 0$, the process is reduced to the exponential distribution.

We can now calculate the expected discounted loss function ϕ . Using integration by parts and the fact that $F_N(0) = 0$,

$$\begin{aligned} \phi &= \int_0^{\infty} L e^{-\delta t} f_N(t) dt = L e^{-\delta t} F_N(t) \Big|_0^{\infty} + L \int_0^{\infty} \delta e^{-\delta t} F_N(t) dt = \\ &= L \delta \int_0^{\infty} e^{-\delta t} F_N(t) dt = L \delta \int_0^{\infty} e^{-\delta t} \xi^N dt \end{aligned}$$

References

- Ang, James S., Chua, Jess H. and John J. McConnel. "The Administrative Costs of Corporate Bankruptcy: A Note," Journal of Finance, vol. 37, No. 1, March 1982, pp. 219-26.
- Baxter, N.D. "Leverage, Risk of Ruin and the Cost of Capital," Journal of Finance, vol. 22, September 1967, pp. 395-403.
- Du Reitz, G., "New Firm Entry in Swedish Manufacturing Industries During the Post-War Period," Stockholm, 1975.
- Farrel, M.J. "Some Elementary Selection Processes in Economics," Review of Economic Studies, April 1970, pp. 305-19.
- Harris, Theodore E. The Theory of Branching Processes, Springer-Verlag, Berlin, 1963.
- Hart, P.E. and S.J. Prais. "The Analysis of Business Concentration: A Statistical Approach," Journal of the Royal Statistical Society, vol. 119, pt. 2, 1956, pp. 150-91.
- Haugen, R.A. and L.W. Senbet. "The Insignificance of Bankruptcy Costs to the Theory of Optimal Capital Structure," Journal of Finance, vol. 23, May 1978, pp. 383-93.
- Jovanovic, Boyan. "Selection and the Evolution of Industry," Econometrica, vol. 50, No. 3, May 1982, pp. 649-70.
- Lucas, Robert E. Jr. "On the Size Distribution of Business Firms," Bell Journal of Economics, vol. 9, No. 2, Autumn 1978, pp. 508-23.
- Mansfield, E. "Entry, Gibrat's Law, Innovation and the Growth of Firms," American Economic Review, vol. 52, 1962, pp.. 1023-51.
- Miller, Merton H.. "Debt and Taxes," Journal of Finance, vol. 32, May 1977, pp. 261-75.

- Modigliani, F. and M.H. Miller. "Corporate Income Taxes and the Cost of Capital: A Correction," American Economic Review, vol. 53, June 1963, pp. 433-43.
- Pashigian, P. and S. Hymer. "Firm Size and Rate of Growth," Journal of Political Economy, vol. 52, 1962, pp. 556-69.
- Peters, T.J. and R.H. Waterman. In Search of Excellence, Harper and Row, New York, 1982.
- Robinchek, A.A. and S.C. Myers. "Problems in the Theory of Optimal Capital Structure," Journal of Financial and Quantitative Analysis, vol. 1, June 1966, pp. 1-35.
- Simon, Herbert E. and C.P. Bonini. "The Size Distribution of Business Firms," American Economic Review, vol. 48, 1958, pp. 607-17.
- Warner, J.B. "Bankruptcy Costs: Some Evidence," Journal of Finance, vol. 32, May 1977, pp. 337-47.