A Simple Theory of Why and When Firms Go Public

Sudip Gupta, *Fordham Business School*
John Rust, *Georgetown University*[*]

January 16, 2017

Abstract: We introduce a simple model of a public firm's optimal investment, dividend, and debt policy to obtain a closed form solution for the value function and optimal investment and dividend payout decision functions in certain cases. We also characterize the optimal policy of a privately held firm and show how an owner's desire to *consumption smooth* distorts its investment and dividend policy, resulting in a loss of market value relative to a publicly owned firm with a comparable level of capital and debt. We then consider the decision of the owner of a privately held firm to "take it public" via an *initial public offering* (IPO), that transforms the privately firm into a public firm with shareholders whose objective becomes maximization of its equity valuation in the stock market. We characterize the conditions under which the owner of a private firm will want to undertake an IPO relative to other options (such as borrowing from a bank or investment of retained earnings, while keeping the firm private), and if the owner decides to undertake an IPO, how much of the IPO proceeds will the original owner "cash out", how much of the proceeds will be reinvested in the firm, and how much of an ownership stake will the original owner retain in the newly formed public firm.

Keywords: IPOs, investment, fiinaincing decisions, dynamic programming
JEL classification: C13-15

[*]PRELIMINARY AND INCOMPLETE: NOT FOR CIRCULATION. Direct questions to: John Rust, Department of Economics, Georgetown University, Washington, DC phone: (301) 801-0081, e-mail: `jr1393@georgetown.edu`.

1 Introduction

We introduce a simple model of the firm to better undertand theoretically the understudied question of why and when the owner of a private firm may choose to take their firm public by holding an initial public offering (IPO). The model we consider is simple enough to provide an analytical solution and full characterization of the optimal investment and dividend policy of a "public firm" which invests in a single liquid capital good k. We use the term "public firm" to distinguish it from a "private firm" which we also analyze. The key difference is that a public firm's objective is to adopt an investment and borrowing policy to maximize the discounted stream of dividends, whereas a private firm adopts and investment and borrowing policy to maximize the discounted stream of *utility.*

Debt policy is complex but we analyze the model in the presence of "single period debt" where existing debt can be "rolled over" and refinanced with another single period loan. We consider situations where a firm uses both debt and retained earnings to finance its investment, but under our assummptions, the firms we study will not find it optimal to hold cash balances, but rather either invest all cash, or pay it out as dividends to shareholders. We show how firms that start with little initial capital have a desire to borrow in order to accelerate their accumulation of capital. We show how external capital in the form of debt significantly shortens the time it takes the firm to achieve "optimal scale" compared to a firm that faces liquidity constraints and is unable to borrow. So the debt option does significantly enhance the value of sufficiently small firms, but has little advantage for firms that are able to enter with a sufficient level of initial capital.

Consider first a firm that has no ability to borrow and which must finance any new investment out of current cash flows. We assume a "putty-clay" production technology where the firm can purchase new capital k using cash flows but once the capital is installed, it cannot be "liquidated" or partially sold to obtain more cash. The firm is constrained to invest using only the new cash flows produced by this capital stock $f(k)$.

At the start of period t, suppose the firm has a capital stock of k_t. It obtains a deterministic cash flow (return) of $f(k_t)$, where $f'(k) > 0$, $f''(k) < 0$ and $\lim_{k\downarrow 0} f'(k) = +\infty$. Using this cash flow, the firm can either pay dividends, $D \geq 0$, or invest an amount $I \geq 0$, subject to the firm budget constraint $D + I \leq f(k_t)$. The amount invested k_t is subject to a deterministic depreciation rate $\delta \in (0,1)$ and the investment is long-term and irreversible, in the sense that the only way to reduce k_t is via depreciation. However the amount invested can be increased by new investment I so that the capital stock follows the law of motion

$$k_{t+1} = k_t(1-\delta) + I. \tag{1}$$

We initially assume that the firm is liquidity constrained and cannot borrow, so it can only use its current cash flow $f(k_t)$ to finance dividend payments and new investment. The firm discounts the future at a constant rate $\beta \in (0,1)$ and its objective is to maximize the present discounted value of future dividend payments.

In section 2 we provide the analytical solution for the optimal investment and dividend policy of a public firm that does not have access to credit markets (i.e. cannot borrow). In section 3 we extend the solution to the case where the public firm has access to credit markets where the firm can borrow via a sequence of single period debt contracts at a fixed rate of interest. We show how the option to borrow increases the value of sufficiently small firms. In DPI the value function is approximated by a piece-wise linearly interpolated solution to a linear system of equations over a finite grid of points in the state space, k. In section 4 we introduce the problem of a "private firm" that is subject to borrowing constraints, where an individual invests his/her private wealth in the firm and operates the firm not to maximize the discounted stream of dividends, but rather to maximize the discounted stream of utility from consumption, where consumption includes payment of profits from the firm. We show that the investment policy of a private firm is very different from the investment policy of a public firm due to the consumption smoothing motive of the owner of a privately held firm. In section 5 we consider whether the owner of a private firm would wish to "go public" by selling off their ownership interst in their firm, converting it from a privately owned firm to a publicly owned firm whose objective is to maximize the discounted value of dividends. This decision can be viewed as a simplified model of an "initial public offering" (IPO). We show that similar to borrowing, the IPO decision is generally optimal only for firms that are sufficiently small. When the firm is sufficiently large, the owner would prefer to remain private rather than "take it public." In section 6 we consider a more realistic version of the IPO decision where the original owner may retain a partial ownership interest in the public firm after the IPO, and the owner uses the cash raised in the IPO (that he did not consume or invest elsewhere) to invest back into the newly created public firm. We show that IPOs done this way are similar to debt financing in that they enable the original owner of the firm to obtain "leverage" by reinvesting the IPO proceeds back into the firm. This reinvestment accelerates the growth of the firm similar to debt finance, but without the constraints imposed by borrowing constraints and the threat of bankruptcy that debt financing implies. In the appendix we discuss the numerical algorithm we used to compute some of the solutions in cases where analytical solutions to the problems were not available (or we were not clever enough to find them). We compare the analytical solution to the public firm problem when it does not have access to credit markets the numerical solution produced by the the method of "discrete policy iteration" (DPI) to show that our numerical solutions are highly accurate.

2 Publicly held firm without access to capital markets

We start our analysis assuming the firm does not have access to capital markets, that is, it is not allowed to borrow. The only way for the firm to finance its desired investments is via retained earnings. Let $V(k)$

denote the value of a publicly held firm when its capital investment is $k \geq 0$. Recall the term "publicly held" signals that the firm's objective function is to maximize the discounted value of dividend payments to shareholds. The Bellman equation for the firm is given by

$$V(k) = \max_{0 \leq I \leq f(k)} \left[f(k) - I + \beta V(k(1-\delta) + I) \right]. \tag{2}$$

If $f(0) = 0$, then it is clear that $V(0) = 0$, since when the firm has no capital investment, it generates no cash returns, and thus it cannot invest any more funds, and therefore will not receive any future cash flows from which it can pay out dividends in the future. Since the marginal return to investment approaches infinity as $k \downarrow 0$, it is reasonable to conjecture that the firm's optimal investment policy has three different regions: 1) an initial region $[0, \underline{k})$ where the firm pays no dividends and devotes all cash flows to investment, 2) an intermediate region $[\underline{k}, \overline{k}]$ where the firm invests and pays dividends, and 3) a final region $(\overline{k}, \infty)$ where the firm has "excess capital" and so it does not invest and pays out all cash flow in the form of dividends. In the intermediate zone where the firm invests and pays dividends, we conjecture that the firm invests just enough to achieve a target or "steady state" level of capital k^* which is the solution to

$$k^* = \underset{k \geq 0}{argmax}\, \beta \frac{f(k) - \delta k}{(1-\beta)} - k \tag{3}$$

The interpretation is that k^* is the level of capital that can be maintained in steady state that maximizes the discounted present value of the firm, net of the cost of the initial investment k. That is, if an investor were to take k in cash and invest it in the firm today, and in all future periods the firm's investment equals the replacement investment (i.e. the depreciation of this invested capital, δk), in order to maintain the capital at level k, then the optimal initial investment that maximizes the difference between the net present value of the firm after the investment (a perpetual dividend stream of $f(k) - \delta k$ that starts with a one period delay) and the initial amount of the investment k is the value k^* given in equation (3). It is not hard to see using calculus that k^* is given by

$$k^* = f'^{-1}(1/\beta - 1 + \delta). \tag{4}$$

where f'^{-1} is the inverse of the marginal return function, $f'(k)$, which is invertible due to our assumption that $f''(k) < 0$.

If we write $\beta = 1/(1+r)$ where $r > 0$ is the one period "market interest rate", then we can rewrite the first order condition determining the optimal steady state capital stock k^* as follows

$$f'(k^*) = r + \delta \tag{5}$$

and observe that this is similar to the equation for the "Golden rule" steady state capital stock in the Solow growth model, except that the population growth rate n is used in place of a "market interest rate" r. The intuition for condition (5): the marginal cash flows produced by the optimal stead state capital stock must be sufficient to cover 1) depreciation of capital, δ, and 2) the opportunity cost of capital, r. Thus, the marginal product of capital equals the sum of these, $r+\delta$, at the optimal steady state level of the capital stock.

Given this, we conjecture that the optimal investment rule $I(k)$ (which describes investment when we are not necessarily at the steady state, k^*) takes the following form

$$I(k) = \begin{cases} f(k) & \text{if } k \in [0,\underline{k}) \\ k^* - (1-\delta)k & \text{if } k \in [\underline{k},\overline{k}] \\ 0 & \text{if } k \in (\overline{k},\infty). \end{cases} \tag{6}$$

It is easy to see that $\overline{k}$ is given by

$$\overline{k} = \frac{k^*}{(1-\delta)} \tag{7}$$

and $\underline{k}$ is given by

$$f(\underline{k}) = k^* - (1-\delta)\underline{k} \tag{8}$$

These values of $\underline{k}$ and $\overline{k}$ ensure that the optimal investment function $I(k)$ is a continuous function of k. The optimal dividend function is then determined trivially as follows by assuming that the budget contraint is binding at all k

$$D(k) = f(k) - I(k). \tag{9}$$

Using equation (6) we obtain the following equation for $D(k)$

$$D(k) = \begin{cases} 0 & \text{if } k \in [0,\underline{k}) \\ f(k) + (1-\delta)k - k^* & \text{if } k \in [\underline{k},\overline{k}] \\ f(k) & \text{if } k \in (\overline{k},\infty). \end{cases} \tag{10}$$

Below we verify these conjectures are correct and derive an explicit formula for the value function $V(k)$ by making use of the Bellman equation (2), and showing that these conjectured optimal investment and dividend policies do result from the solution to the firm's Bellman equation. We summarize the main conclusions in Theorem 0 below.

Theorem 0: *Consider the optimal investment and dividend policy of a publicly held firm that does not have the option of borrowing, the solution to which is given by the Bellman equation (2). The optimal investment policy is given by the function $I(k)$ in equation (6) and the optimal dividend policy is given in equation (10) where the constants k^*, $\underline{k}$ and $\overline{k}$ are given in equations (5), (8) and (7), respectively. Under the optimal policy, the firm will reach the optimal steady state capital stock k^* in a finite number of periods starting from any positive initial capital stock k. However if $k=0$ and $f(0)=0$, the firm will remain forever in a non-investment, no-dividend absorbing state, so V(0)=0.*

We prove Theorem 0 below, starting by verifying the firm's optimal investment policy for k in the "unconstrained region" $[\underline{k},\overline{k}]$ where there is an interior solution for the optimal level of investment $I(k)$

implied by the Bellman equation (2). That is, assuming that V is differentiable in this region, then $I(k)$ must satisfy the following first order or *Euler equation*

$$1 = \beta V'(k(1-\delta)+I(k)). \tag{11}$$

Substituting the optimal investment rule $I(k)$ into the right hand side of the Bellman equation (2) and differentiating with respect to k, making use of the *envelope theorem,* we have

$$\begin{aligned} V'(k) &= f'(k)+(1-\delta)\beta V'(k(1-\delta)+I(k)) \\ &= f'(k)+(1-\delta) \end{aligned}$$

where we used the fact that the Euler equation (11) holds for $k \in [\underline{k},\overline{k}]$. The envelope equation implies that $V(k)$ is given by

$$V(k) = f(k)+(1-\delta)k+C \tag{12}$$

for some constant C when $k \in [\underline{k},\overline{k}]$. Notice that at $k = k^*$ the firm generates a perpetual dividend stream of $f(k^*)-\delta k^*$ so this implies that

$$V(k^*) = \frac{f(k^*)-\delta k^*}{(1-\beta)} \tag{13}$$

So using the other formula for $V(k)$ from equation (12), this implies that the unknown constant C is given by

$$C = \frac{\beta\,[f(k^*)+\delta k^*]}{(1-\beta)} - k^* \tag{14}$$

Thus, we can see that C equals the optimized right hand size of the net gain from initial investment in equation (3) which determined the optimal steady state capital stock value k^*. Thus, the value of the firm in the interval $[\underline{k},\overline{k}]$ is this optimized value, plus $f(k)+(1-\delta)k$. The intuition for this formula is that once the firm is in the interval $[\underline{k},\overline{k}]$, its investment $I(k) = k^* - (1-\delta)k$ will enable it to achieve the optimal steady state capital level k^* in the following period. So it follows that $V(k)$ equals the net dividends this period, $D(k) = f(k)-I(k) = f(k)-k^*+(1-\delta)k$ plus the present value of all future dividends in all subsequent periods $\beta[f(k^*)-\delta k^*]/(1-\beta)$ where this period's investment has enabled the firm to achieve the optimal steady state capital stock k^*.

Now we need to verify that the optimal investment rule $I(k)$ for $k \in [\underline{k},\overline{k}]$ really is the formula we conjectured, $I(k) = k^*-(1-\delta)k$. To show that this is correct, we need to show that this satisfies the Euler equation (11). Using the closed form solution for $V(k)$ in equation (12) we can rewrite the Euler equation as

$$1 = \beta\left[f'(k(1-\delta)+I(k))+(1-\delta)\right] \tag{15}$$

Solving this equation for $I(k)$ we can see that

$$\begin{aligned} I(k) &= f'^{-1}(1/\beta-(1-\delta))-(1-\delta)k \\ &= k^*-(1-\delta)k \end{aligned}$$

which does indeed match the formula we conjectured in equation (6).

Finally we need to derive formulas for $V(k)$ in the "constrained no dividend region" $[0,\underline{k})$ where $I(k) = f(k)$ and show that it is indeed optimal for the firm to invest all available cash flow and not pay any dividends in this region, and also for the "excess capital, no investment region" we need to derive $V(k)$ and show that it is indeed optimal for the firm to invest zero in this region and pay out all cash flows as dividends, $I(k) = 0$ and $D(k) = f(k)$.

Consider the latter region first. Consider a value of $k > \overline{k} = k^*/(1-\delta)$ that is sufficiently close to $\overline{k}$ so that after depreciation we have $k(1-\delta) \in [\underline{k}, \overline{k}]$. In particular, we have $(1-\delta)k > k^*$, so that after depreciation (assuming zero investment) the capital exceeds the optimal steady state capital level k^* in the region where the firm invests and pays dividends.

We claim that for this value of k the value of the firm is given by

$$V(k) = f(k) + \beta\left[f((1-\delta)k) + (1-\delta)^2 k + C\right], \tag{16}$$

and the optimal investment at this value k is $I(k) = 0$. To see this, we consider the value of investing a positive amount $I > 0$

$$V(k,I) = f(K) - I + \beta\left[f((1-\delta)k + I) + (1-\delta)[(1-\delta)k + I] + C\right] \tag{17}$$

Notice that for this fixed value of k, the function $V(k,I)$ is strictly concave in I due to our assumption that f is strictly concave. So it is sufficient to show that $\frac{\partial}{\partial I}V(k,I) < 0$ at $I = 0$. The concavity of $V(k,I)$ in I then implies that this partial derivative is negative for all higher values of I which implies that the optimal value of investment is zero at this value of k, $I(k) = 0$. Evaluating the partial derivative of $V(k,I)$ with respect to I at $I = 0$ we have

$$\frac{\partial}{\partial I}V(k,I) = -1 + \beta f'((1-\delta)k) + \beta(1-\delta) \tag{18}$$

However the first order condition for the optimal steady state capital stock level can be written as

$$0 = -1 + \beta f'(k^*) + \beta(1-\delta) \tag{19}$$

Since f is strictly concave and $(1-\delta)k > k^*$, equation (18) implies that $\frac{\partial}{\partial I}V(k,I) < 0$ at $I = 0$, and so we can conclude it is optimal for the firm not to invest at k.

Since we know that no investment is optimal at this point, we conclude that

$$V(k) = f(k) + \beta\left[f((1-\delta)k) + (1-\delta)^2 k + C\right] \tag{20}$$

This will hold for any value of k such that $k^* < (1-\delta)k < k^*/(1-\delta)$. Continuing inductively we can so that if $k(1-\delta) > \overline{k} = k^*/(1-\delta)$ but $k < k^*/(1-\delta)^3$, it will take 2 periods for capital to depreciate to a value $(1-\delta)^2 k \in (k^*, \bar{k})$. We can show that in this interval of k zero investment is optimal as well, using an argument similar to the one above. In fact by a formal induction proof, we can show that $I(k) = 0$ for all $k > \bar{k} = k^*/(1-\delta)$ and $V(k)$ is given by

$$V(k) = \sum_{i=0}^{n-1} \beta^i f((1-\delta)^i k) + \beta^n \left[f((1-\delta)^n k) + (1-\delta)^n k + C\right] \quad k \in \left[k^*/(1-\delta)^n, k^*/(1-\delta)^{n+1}\right). \quad (21)$$

Since the equation above satisfies the Bellman equation (2) by construction, it follows that $I(k) = 0$ and $D(k) = f(k)$ is the optimal investment and dividend policy for the firm in the region $k > k^*/(1-\delta)$ and $V(k)$ is given by the formula in (21) once we determine the smallest number of periods n that are required for the capital to depreciate down to a level $k \in (k^*, k^*/(1-\delta))$ where it becomes optimal for the firm to invest again.

Now consider the final interval $k \in [0, \underline{k})$. In this region we claim that it is optimal for the firm to invest all available cash flow and pay no dividends. That is, $D(k) = 0$ and $I(k) = f(k)$. We now verify that this conjecture is correct. Recall that $\underline{k}$ was defined as the solution to the equation

$$f(\underline{k}) + (1-\delta)\underline{k} = k^* \quad (22)$$

Consider a $k < \underline{k}$, but a value not so close to zero so that if the firm invests all cash flow and pays zero dividends, then its capital stock at the start of next period, $f(k) + (1-\delta)k$, satisfies

$$f(k) + (1-\delta)k > \underline{k}. \quad (23)$$

How do we know there is a $k < \underline{k}$ that satisfies inequality (23) above? First notice that $f(k) + (1-\delta)k$ is a strictly concave function of k and notice that at k^*, it is easy to manipulate the first order condition determining the optimal steady state capital stock in equation (3) to show that

$$f'(k^*) + (1-\delta) = 1/\beta > 0. \quad (24)$$

and hence we conclude that $f(k) + (1-\delta)k$ is strictly increasing in k for $k \leq k^*$. But since dividends must be positive at k^* we have $f(k^*) - \delta k^* > 0$ which is equivalent to $f(k^*) + (1-\delta)k^* > k^*$. Then since $\underline{k}$ is defined as the value of k that solves $f(k) + (1-\delta)k = k^*$, it follows that $\underline{k} < k^*$. If $k < \underline{k}$, then the fact that $f(k) + (1-\delta)k$ is strictly increasing in k implies that $f(k) + (1-\delta)k < f(\underline{k}) + (1-\delta)\underline{k} = k^*$.

Let $k < \underline{k}$ be such that $f(k) + (1-\delta)k > \underline{k}$. We now want to show that it is optimal for the firm to invest all cash flow, $f(k)$, and pay zero dividends. The investment-specific value function is $V(k,I)$ given in equation (17) above. We want to show that $\frac{\partial}{\partial I}V(k,I) > 0$ for all $I \in [0, f(k)]$. This is given by

$$\frac{\partial}{\partial I}V(k,I) = -1 + \beta f'((1-\delta)k + I) + \beta(1-\delta). \quad (25)$$

Noting that $V(k,I)$ is strictly concave in I it is sufficient to show that $\frac{\partial}{\partial I}V(k,I) > 0$ when I takes the maximum possible value, $I = f(k)$. In this case, the partial derivative in equation (25) reduces to

$$\frac{\partial}{\partial I}V(k, f(k)) = -1 + \beta f'((1-\delta)k + f(k)) + \beta(1-\delta). \quad (26)$$

But we know that, from the argument above, that $\frac{\partial}{\partial I}V(k^*, f(k^*)) = 0$ and that $f(k) + (1-\delta)k$ is strictly increasing in k for $k < k^*$. So this implies that $\frac{\partial}{\partial I}V(k, f(k)) > 0$ as claimed.

Let $\underline{k}^1$ be given by the solution to $f(\underline{k}^1) + (1-\delta)\underline{k}^1 = \underline{k}$, or $\underline{k}^1 = 0$ if no solution exists. Then, it is not hard to show using the same argument as above that if $\underline{k}^1 > 0$ we must have $\underline{k}^1 < \underline{k}$. Then for all $k \in [\underline{k}^1, \underline{k})$ we have $I(k) = f(k)$ and $D(k) = 0$ and

$$V(k) = \beta\left[f(k(1-\delta) + f(k)) + (1-\delta)[(1-\delta)K + f(k)] + C\right] \tag{27}$$

If $\underline{k}^1 > 0$ then we can recursively define $\underline{k}^j$, $j = 2, 3, \ldots$ by the formula

$$f(\underline{k}^j) + (1-\delta)\underline{k}^j = \underline{k}^{j-1} \tag{28}$$

until the first value of j is reached where $\underline{k}^j = 0$. Define the function $T(k)$ by

$$T(k) = k(1-\delta) + f(k) \tag{29}$$

and define the composite powers of T, T^2, T^3, etc by

$$T^2(k) = T(T(k)) = T(k)(1-\delta) + f(T(k)) \tag{30}$$

and in general

$$T^j(k) = T^{j-1}(T(k)), \quad j = 1, 2, \ldots \tag{31}$$

where we define $T^0(k) = k$. Then if $k \in [\underline{k}^j, \underline{k}^{j-1})$ (where $\underline{k}^0 = \underline{k}$) we have

$$V(k) = \beta^j\left[f(T^j(k)) + (1-\delta)T^j(k) + C\right]. \tag{32}$$

Using an induction argument, we can show that $I(k) = f(k)$ and $D(k) = 0$ for k in every interval $[\underline{k}^j, \underline{k}^{j-1})$. By construction, $V(k)$ satisfies the Bellman equation (2). We conclude that we have derived a closed form solution for the optimal investment policy in equation (6) and the optimal dividend policy in equation (10) and have an analytic (if recursive) expression for the value function in equations (12), (21) and (32) where the constant C is given by equation (14) and the optimal steady state capital stock k^* is given by equation (4).

Theorem 0 tells us that the optimal investment policy is for the firm to invest all profits back into the firm and pay no dividends while $k < \underline{k}$, where $\underline{k} < k^*$ is the lower boundary of the "linear investment region" there the firm has enough accumulated capital to jump to the optimal steady state capital stock k^* in a single period. In addition, as long as $k > \underline{k}$, the firm also has enough surplus profits to also pay dividends to its shareholders. This implies that there are zero dividends until the first period where k exceeds $\underline{k}$, then a "partial dividend" in that period equal to $f(k) + (1-\delta)k - k^*$, followed by an infinite stream of dividends equal to $f(k^*) - \delta k^*$.

Now consider a firm that starts out with an arbitrarily small initial investment in capital k_0. The firm will reinvest all the cash flow from this very small initial investment and keep doing that until the capital stock first exceeds $\underline{k}$. How long will this take as a funcion of k_0? We now wish to show that for many production functions, the time it will take to reach $\underline{k}$ will be finite, no matter how small k_0 is, provided k_0 is positive. But this implies that the right hand limit of the value function is positive

$$\lim_{k \downarrow 0} V(k) > 0. \tag{33}$$

whereas if $f(0) = 0$, we know from Theorem 0 that $V(0) = 0$. Thus, we conclude that there is a *discontinuity in the value function at* $k = 0$ and this discontinuity arises naturally from the restriction that the firm is not able to "get off the ground" until at least some arbitrarily small initial investment is made in it.

To capture this property formally, define the sequence of functions $g(k,t)$ recursively by $g(k,0) = k$ and

$$g(k,t) = g(k,t-1)(1-\delta) + f(g(k,t-1)), \quad t = 1,2,\ldots \tag{34}$$

Then $g(k,t)$ is the amount of capital the firm could accumulate if it started with initial capital stock k at period $t = 0$ and reinvested all profits in periods $t = 1,2,\ldots,t$. We now define a property we call *finite period reachability*

Definition A strictly concave production function f satisfying $f(0) = 0$ has the property of *finite period reachability* if we have

$$\lim_{k \downarrow 0} \inf_t \{t | g(k,t) > \underline{k}\} < \infty \tag{35}$$

where $\underline{k}$ is defined in equation (22).

We can verify that the finite reachability condition holds for specific choices of f such as $f(k) = \sqrt{k}$. For example we can calculate numerically, starting from $k_0 = 1 \times 10^{-250}$ that it will take $t = 19$ periods of full reinvestment of profits for the firm to reach the capital stock threshold of $\underline{k} = 24.4416$, and $V(k_0) = 30.6954$. Notice the $f(k) = \sqrt{k}$ satisfies an "Inada condition", i.e. $\lim_{k \downarrow 0} f'(k) = +\infty$. However satisfying an Inada condition is not a necessary condition for the finite period reachability condition to hold. For example consider the production function $f(k) = \sqrt{k+1} - 1$. Then $f'(0) = 1/2$, yet the finite period reachability condition is still satisfied, though it takes 1565 periods before the capital stock can grow from $k_0 = 1 \times 10^{-250}$ to exceed $\underline{k}$, and $V(k_0) = 0.5056$. Clearly, the smaller the marginal productivity of capital is for very small initial investments, the longer it will take for the firm to grow its capital stock to the threshold value $\underline{k}$, and therefore the magnitude of the discontinuity in V at $k = 0$ is smaller.

Of course, the finite period reachability condition may not be a particularly realistic assumption in practice: there may be fixed setup costs that must be incurred to get a firm "off the ground" and in such cases, we would expect that $V(k) = 0$ for all k below the minimal fixed costs that are necessary to get the firm off the ground. However we think this is an interesting example of an apparantly "continuous" dynamic programming problem where the value function has a discontinuity. Normally we expect that small changes in "initial conditions" should only lead to small changes in payoffs, but when the finite period reachability condition is satisfied, an arbitrarily small initial investment in the firm leads to a discontinuous jump in the value of the firm — i.e. it results in an abitrarily high rate of return from this initial investment.

Theorem 1 *Assume that* β *and* δ *are in the unit interval* $(0,1)$*, and that* $f(k)$ *is strictly concave and continuously differentiable. Then* $V(k)$ *is strictly concave and continuously differentiable for almost all* $k > 0$*. If* $f(0) = 0$ *and the finite period reachability condition (35) is satisfied, then the value function has a discontinuity at* $k = 0$ *where* $V(0) = 0$ *whereas* $\inf_{k>0} V(k) > 0$.

The proof that $V(0) = 0$ if $f(0) = 0$ is intuitive: if the firm has no initial capital stock, then it cannot generate any cash flow, and thus cannot invest. Without any investment, the firm's capital stock will be stuck at $k = 0$, and since $f(0) = 0$ the firm will never be able to generate positive cash flows and thus have zero value. Formally, the Bellman equation when $k = 0$ is given by

$$V(0) = \max_{0 \le I \le f(0)} \left[f(0) - I + \beta V((1-\delta)0 + I)\right]. \tag{36}$$

But if $f(0) = 0$, then clearly $I = 0$, which implies that $V(0) = \beta V(0)$, and if $\beta \in (0,1)$, this implies that $V(0) = 0$.

The proof of the differentiability and strict concavity of V involves considering V over the three different regions $k \in [0, \underline{k})$, $k \in [\underline{k}, \overline{k}]$ and $k \in (\overline{k}, \infty)$. In the middle region, $V(k) = f(k) + (1-\delta)k + C$ and strict concavity in this region follows from the assumption that $f(k)$ is strictly concave. In the upper region $(\overline{k}, \infty)$ V is given by formula (21) and it is straightforward to see that V is strictly concave in this region as well. Finally in the initial "no dividend" region $[0, \underline{k})$, the concavity follows from an induction argument. We first show by induction that for each $j \ge 1$ that $T^j(k)$ is concave. Then using the properties of compositions of concave functions, it is easy to show from equation (32) that V is strictly concave on $[0, \underline{k})$ as well. To prove the discontinuity in V at $k = 0$ we first note that $V(0) = 0$, since if the firm has no initial capital, then due to the fact that $f(0) = 0$, the firm has no cash flows, and thus cannot invest anything or pay any dividends. Thus, it will be stuck at a zero capital absorbing state and must have a value of zero. However if the firm can makes an arbitrarily small initial investment k_0, the finite period reachability condition (35) implies that after a finite number of periods of reinvesting all of the profits back into the firm to grow its capital stock at the maximal rate (which is an optimal investment strategy while $k < \underline{k}$ due to Theorem 0), then after some finite time $t < \infty$ we will have $k_t > \underline{k}$. It follows from Theorem 0 that once this condition is satisfied it is optimal for the firm to start paying dividends and the capital stock will converge to the optimal steady state capital stock k^* in a finite number of periods and dividends equal to $f(k^*) - \delta k*$ in every period after that. This implies that $V(k_0) > 0$ for any $k_0 > 0$, i.e. $V(k)$ has a discontinuity at $k = 0$.

Figure 1 plots the optimal investment and dividend rules for the case $f(k) = \sqrt{k}$. We see that optimal

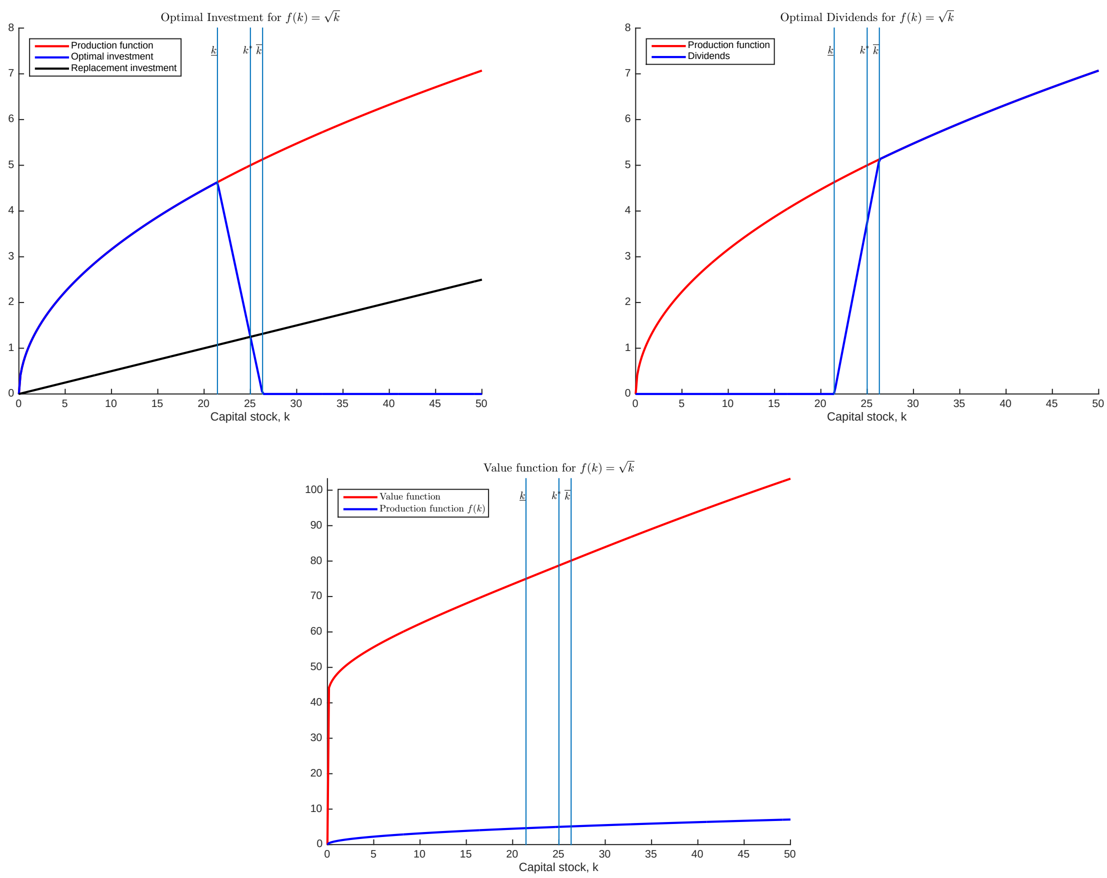


Figure 1: Optimal investment and dividend policy and value function for $f(k) = \sqrt{k}$

investment intersects the black "replacement investment" line (i.e. the line δk) exactly at k^*, the optimal steady state capital stock level, which equals 25 in this example. The level of optimal investment at the steady state is $\delta k^* = 1.25$, which of course is just enough to offset the corresponding depreciation in capital. Figure 1 also plots the value function for this problem. Notice there are no discontinuities in the value function at the various break points, $\{\underline{k}^j\}$ and $\{k^*/(1-\delta)^j\}$ above and below the cutoffs $\underline{k}$ and $\overline{k}$ defining the region where investment and dividends are positive. The value function is monotonic and strictly concave in k and satisfies $V(0) = 0$.

The dynamics of the capital stock are clear: starting from any k the capital stock converges globally to the unique optimal steady state level k^* in a finite number of periods. For $k \in [\underline{k}, \overline{k}]$ the firm undertakes investment $I(k) = k^* - (1-\delta)k$ enables it to jump to the optimal steady state value k^* in a single period. When initial capital is either below or above this region, the firm has to wait several periods for capital to accumulate above the lower $\underline{k}$ threshold, or depreciate down below the upper $\overline{k}$ threshold.

2.1 Extending the model to allow non-concave production functions

In the previous section we were able to derive essentially a closed form solution for the optimal investment strategy of the firm for a general case of *concave* cash flow production functions $f(k)$. In this section we extend the model to consider *non-concave* production functions. Figure 2 plots a pair of non-concave production functions formed by grafting logistic "S-curves" on to the basic concave production function we considered in the previous section. That is, the figure plots production functions of the form

$$f(k) = \sqrt{k} + \theta_1 \left[\frac{\exp\{(k-\theta_2)/\theta_3\}}{1+\exp\{(k-\theta_2)/\theta_3\}} \right], \tag{37}$$

where $\theta_2 = 80$ is a "location parameter", $\theta_3 = 10$ is a "scale parameter" that determines how steep the S-curve is, and θ_1 is a "height parameter" that determines the overall productivity. Figure 2 plots two production functions, one is for a "productive firm" where $\theta_1 = 10$ and the other is for a "less productive firm" with $\theta_1 = 2$.

The reason we believe non-concave production functions are potentially interesting is because they can enable us to model *growth stages* of firms. We can imagine a firm starting out with little initial capital and investing in a "first stage technology" that is concave, such as $f(k) = \sqrt{k}$. However after it makes its investment in its first stage technology and grows sufficiently large, the firm may be able to continue to invest in a "second stage" technology that could potentially be far more productive than its first stage technology. This second stage technology is represented by the second additive S-curve component in the production function in equation (37) and in figure 2. To reach this higher level of production and cash flow, the firm may need to undertake signficant, large fixed investments that initially do not have high returns (high marginal product of capital, $f'(k)$) but after sufficient investment the firm can enter an *increasing returns to scale region* where $f''(k) > 0$ before again returning to a concave region after sufficient capital has been invested and the firm has more or less fully mastered and exploited its second stage technology.

As we noted in the previous section, the simple theory with concave production functions leads the firm to grow until it reaches the "Golden rule" capital stock k^* satisfying equation (5). When the production function f is concave, it is evident that there is *only one steady state, Golden rule solution.* However it should be clear from figure 2) that if the firm's production function is no longer concave, *there is the possibility of multiple steady state "Golden rule" solutions.* Which of these Golden Rule steady states will firm end up at? In addition, will the firm's investment stil retain the form given in equation (6) if its production function is not concave?

Figure 3 provides the answer to this question. It plots the numerically calculated optimal investment strategies corresponding to the two production functions plotted in figure 2. The first observation is that despite the non-concavity of the production functions, the optimal investment rules $I(k)$ still take the three region form that we illustrated in equation (6) in the concave case. That is, there are still a pair of thresholds

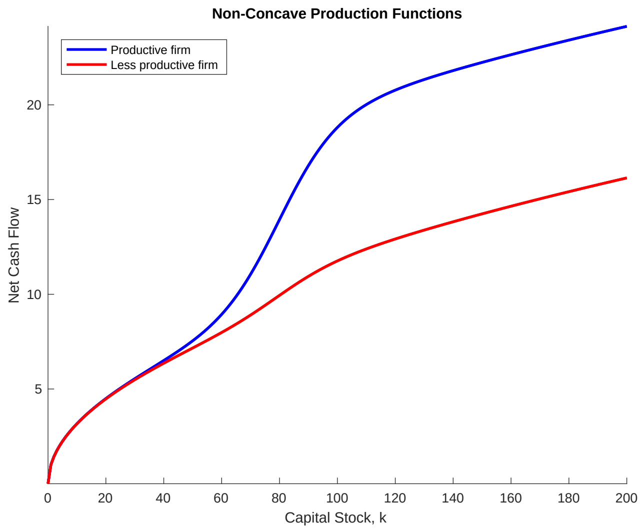


Figure 2: Non-concave production functions $f(k)$

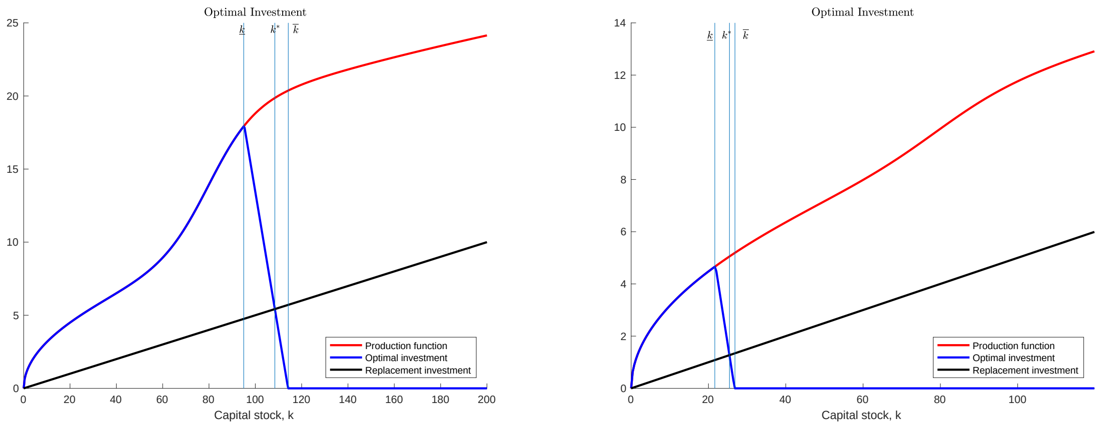


Figure 3: Optimal investment strategies for non-concave production functions $f(k)$

$(\underline{k}, \overline{k})$ such that: a) the firm reinvests all of its cash flow for $k \leq \underline{k}$, b) the firm does no further *net investment* if $k > \overline{k}$, and c) in the interval $[\underline{k}, \overline{k}]$ investment decreases linearly. Further, there is an optimal steady state capital stock, k^*, where the Golden rule relation (5 is satisfied.

The left hand panel of figure 3 plots the optimal investment rule for the firm with the productive second stage technology (i.e. $\theta_1 = 10$). In this case the optimal steady state capital stock $k^* = 108$ is in fact the largest steady state Golden rule solution. Comparing this to figure 1, we see that the steady state size of the firm is more than four times larger due to the presence of this attractive second stage investment opportunity. However the right hand panel of figure 3 shows that for the firm with the less productive second stage investment opportunity, it determines that the additional cash flow is not sufficiently high to justify investing to reach it, given that approximately the same level of capital investment would be

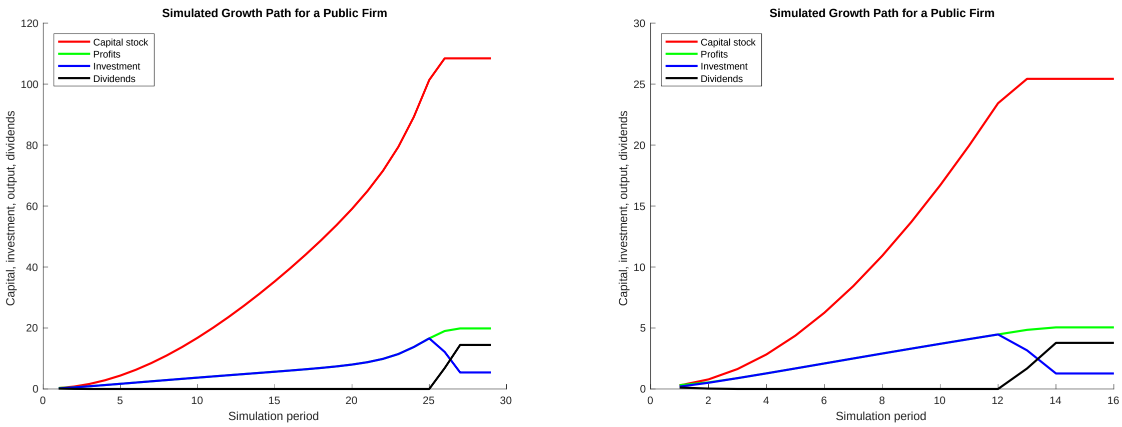


Figure 4: Simulated growth paths for firms with non-concave production functions $f(k)$

required. This firm decides to forgo the second stage investment opportunity and thus it ends up at an optimal steady state capital stock k^* that is very near the value $k^* = 25$ that was optimal for the firm with the concave $\sqrt{k}$ production function. Thus, for this firm the optimal steady state capital stock k^* is the *lowest* of the steady state Golden rule solutions.

Figure 4 presents simulated trajectories for the two firms, with each starting from an initial capital stock of $k_0 = 0.1$. The qualitative features of the growth in capital and output and the trajectories for investment and dividends are the same in both cases. Both firms pay no dividends and reinvest all cash flows back into the firm to grow the capital stock as quickly as possible when $k < \underline{k}$, and only begin to pay dividends when $k \geq \underline{k}$ as they steadily reduce investment as they approach their respective optimal steady state capital stocks, k^*. As we noted above, k^* is more than four times larger for the firm with the more productive second stage investment opportunity than the one with the less attractive opportunity, and thus it takes the former firm twice as long to reach its steady state capital stock k^*. However the delay is worth it as the the steady state output (plotted by the green curves in the figures) is more than 4 times as large, and so are its steady state dividends.

From figure 3 it is tempting to conjecture that the general form of the optimal investment rule is the same for a non-concave production function f as the general characterization for concave f that we established in Theorem 0 in the previous section. That it, the firm has two thresholds $(\underline{k}, \overline{k})$ such that the firm reinvests all cash flow and pays zero dividends if $k \leq \underline{k}$, and sets investment equal to δk^* for $k > \overline{k}$, and in the interval $[\underline{k}, \overline{k}]$ its investment is a decreasing linear function of k that enables the firm to reach its optimal steady state capital stock $k^* \in [\underline{k}, \overline{k}]$ in a single period.

Unfortunately the non-concave case things are more complicated as illustrated in figure 5 below. We see that in this case there are *two "optimal" steady state "Golden rule" capital stocks* k^*. There is a low steady state at $k^* = 24.9797$ (nearly the same as for the concave example in the previous section where

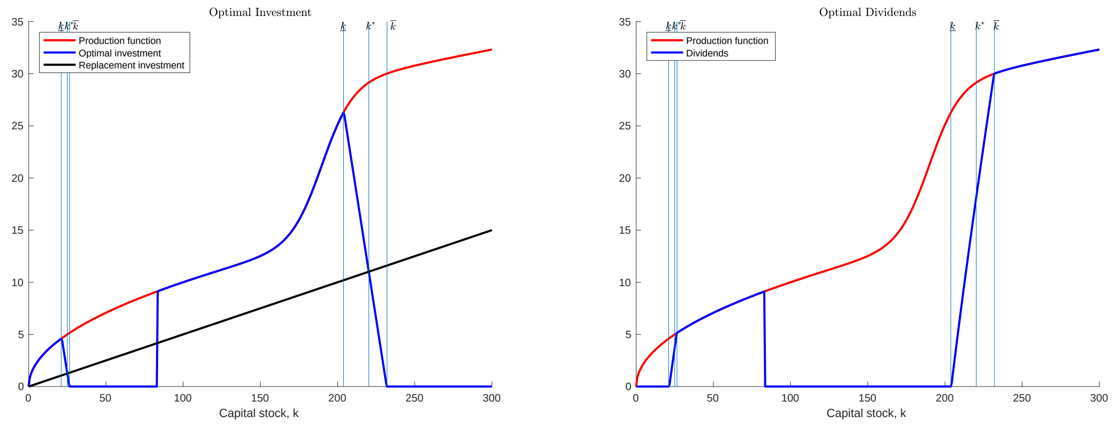


Figure 5: More complex optimal investment policy for a non-concave production function $f(k)$

$f(k) = \sqrt{k}$), as well as a high steady state $k^* = 222.24$. The domain of attraction for the the low steady state k^* is $(0, 83)$, whereas the domain of attraction for the high steady k^* is $[83, \infty)$. That is, if the firm starts out with sufficiently little initial capital, it will not be optimal for the firm to invest enough to reach the high steady state Golden rule capital stock at $k^* = 222.24$, and instead it is optimal for the firm to converge to the low steady state value at $k^* = 24.9797$. Essentially, the level of delay and investment necessary to reach the higher steady state k^* (which is also significantly more profitable with a value of $v(k^*) = 381$ compared to a value at the low steady state of $v(k^*) = 79$) makes it uneconomic for the firm to try to reinvest its retained earnings to grow enough to reach this higher steady state capital stock and profitability level.

However if the firm were endowed with sufficient initial capital, $k > 83$, then there is far less delay to reach the higher optimal steady state capital stock and it is optimal for such a firm to reinvest all of its profits and pay no dividends until it reaches this higher steady state capital stock $k^* = 222.24$. This examplei highlights an interesting paradox and constraint on growth created by the firm's inability to borrow. Suppose the firm could borrow any amount it desired at the time the firm was initially established. As we will show in section 5, the firm will want to borrow an amount that will immediately put it at one of the steady state capital stocks, k^*. Suppose that the firm can borrow at the same interest rate that its own dividends are discounted at, i.e. the market rate of interest $r_m = 1/\beta - 1 = 0.05$ (in this example). Then if the firm borrowed enough to reach the low steady state its payoff, after paying off the principal and interest on this initial debt (here for simplicity we ignore the fact that it would actually take multiple periods of rolling over single period debt contracts to fully pay off this initial debt) is

$$V(k^*) - k^*(1 + r_m) = 79 - 24.9797 * (1.05) = 52.77 \tag{38}$$

whereas if the firm borrowed 222.24 to reach the higher steady state capital stock k^* its net of debt and interest payoff would be

$$V(k^*) - k^*(1 + r_m) = 381 - 222.24(1.05) = 149.75. \tag{39}$$

Thus, this provides a stark illustration of how borrowing constraints and the lack of initial capital can lead an optimizing firm to converge to lower and "suboptimal" steady state outcome. The lower steady state *is* fully optimal for the firm that has insufficient initial capital and faces borrowing constraints, but it not optimal compared to a firm that does have the option to borrow. Roughly speaking, the inability to borrow causes the firm with limited initial capital to forgo undertaking an investment opportunity that could nearly triple the net of debt return to a firm that did have access to capital markets. In addition, we will see that even if firms can borrow, if they still face *borrowing constraints* that prevent them from borrowing as much as they would like to borrow, this can still lead them to forgo attractive investment opportunities. It will become a key part of our explanation for why firms choose to "go public" via an IPO.

It is possible to extend the model in other directions to allow for other types of realistic non-convexities in the model such as the case where investment projects are *lumpy* — i.e. they require fixed initial "down-payments" or can only be undertaken in discrete "chunks." Our model so far has assumed that capital is a perfectly divisible commodity, the firm is free to choose any level of investment it desires subject to its budget constraint. However once we allow for additional nonconvexities, we need to pay closer attention to the method by which the firm finances its investments, especially in situations where the firm must come up with a fixed downpayment such as to acquire a patent or new technology that will enable it to proceed and undertake further continuous investments in what we have called a "second stage technology" especially if the firm's level of cash flow in any given period is insufficient to pay this fixed investment fee. To handle this situation, we need to introduce an additional state variable into the model, *liquid capital* that represents cash holdings that the firm must acquire to finance large lumpy fixed investment costs. The liquid capital can be considered a type of "savings" that the firm must carry due to a *cash in advance constraint* that new investment must be financed by cash payments in situations where the firm is unable to borrow. Liquid capital is also important as a "buffer stock" against unexpected productivity shocks, so we will defer a formal extension of our model to handle liquid capital to the next section where we also consider how to extend the model fo allow for stochastic shocks.

2.2 Extending the model to allow for stochastic shocks

The next natural extension of the model is to allow for stochastic productivity shocks $\{\eta_t\}$. That is, we consider models of the form

$$y_t = f(k_t, \eta_t) \tag{40}$$

where η_t is a stochastic shock to output (cash flow) in period t. A natural and simple specification is one with lognormal, multiplicative shocks to the production functions we considered previously,

$$y_t = f(k_t)\exp\{\eta_t\} \tag{41}$$

where $\eta_t \sim N(0,\sigma^2)$ are *IID* stochastic productivity shocks. This specification implies that cash flow in any given period is lognormally distributed with parameters $(f(k_t),\sigma^2)$.

In the lognormal specification, if $f(k) \geq 0$, then there is zero probability that the firm will ever experience negative cash flows, corresponding to a *loss.* We can extend the model fo allow for this possibility by allowing for *fixed production costs F* which may also have their own stochastic components ε_t. For example, one specification might be

$$y_t = f(k_t)\exp\{\eta_t\} - (F + \varepsilon_t) \tag{42}$$

where $\varepsilon_t \sim N(0,\gamma^2)$ can be regarded as a stochastic component of fixed costs of production, including management and "front office" costs. With this specification it is easy to see there is a positive probability of losses and once we allow for this, we are also forced to think more carefully of how the firm finances its investments and dividends, while allowing for the possibility of cash flow shortfalls or even outright losses.

To this end we introduce *liquid capital* l_t which we can consider to be either cash or liquid short term securities that the firm can use to finance its expenditures. We continue to assume that the firm cannot borrow, and it must make expenditure decisions regarding new and replacement investment and dividend payments at the start of each period, but these expenditures are subject to the constraint that they cannot exceed the firm's start of period liquid capital holdings. Production takes place during the period so we assume that the cash flow y_t is not known to the firm *ex ante* at the start of period t but its value will be realized and known at the end of the period. Thus, it will contribute to the amount of liquid capital available to the firm at the start of period $t+1$.

Given that there can be positive probability of a cash flow loss under the specification that allows for fixed production costs, (42) we also have to specify what happens to the firm its liquid asset balance l_t at the start of any period is *negative* $l_t < 0$. We intepret this as a situation where the firm is unable to pay all of its suppliers or workers. This may involve a "default" or contractual penalties, but we assume that a default to workers or suppliers cannot force a liquidation of the firm as can be the case when the firm borrows via secured debt contracts. However any suppliers or workers who have not been paid may refuse to continue to work for, or supply to the firm until their back wages or invoices have been paid, possibly with some sort of "late payment fee" penalty. So we assume that in the event that $l_t < 0$ the firm cannot undertake any new investments or pay dividends until it has first paid off these past due balances and fees.

So all cash flows of the firm will go towards restoring its liquid capital to a positive level, and once this occurs, the firm is allowed to pay dividends and undertake investments again.

Thus, this discussion implies the following law of motion for liquid capital

$$l_{t+1} = \begin{cases} l_t - D - I + f(k_t, \eta_t, \varepsilon_t) & \text{if } l_t \geq 0 \\ l_t(1+r_p) + f(k_t, \eta_t, \varepsilon_t) & \text{if } l_t < 0 \end{cases} \tag{43}$$

where $f(k_t, \eta_t, \varepsilon_t)$ is just a compact way of writing the firm's cash flow production function, allowing for the possibility of fixed costs with random shocks ε_t in addition to the proportional lognormal production shocks η_t. For simplicity we assume a proportional penalty $r_p > 0$ that constitutes an interest "late fee" if the firm experiences a cash flow loss. Then the negative value of l_t that constitutes an implicit type of debt that must be paid off with interest (penalty) before the firm is able to return to normal operations and undertake investment and pay dividends again.

With this specification we can write the value of the firm as $V(k,l)$, where the firm's value now depends both on its physical capital k and its liquid capital l. The Bellman equation is given by

$$V(k,l) = \max_{\substack{D+I \leq l \\ D \geq 0, I \geq 0}} \left[D + \beta \int_\eta \int_\varepsilon V(k(1-\delta)+I, [l-D-I](1+r_m) + f(k,\eta,\varepsilon)) F(d\eta, d\varepsilon) \right] \tag{44}$$

if $l \geq 0$, and

$$V(k,l) = \beta \int_\eta \int_\varepsilon V(k(1-\delta)+I, l(1+r_p) + f(k,\eta,\varepsilon)) F(d\eta, d\varepsilon), \tag{45}$$

if $l < 0$, where $r_m \geq 0$ is the market rate of return to liquid capital and $r_p > 0$ is the penalty the firm pays for being in "default" due to incurring cash flow losses that creates an implicit form of debt that we discussed above. In the Bellman equations (44) and (45), $F(\eta, \varepsilon)$ is the joint distribution of the stochastic shocks. For example in terms of our previous discussion we can consider for concreteness that F is a bivariate normal distribution for the proportional shock to output η and the additive shock ε to fixed costs.

The presence of stochastic shocks is not only realistic, but it leads to new issues that help to clarify the difference between the behavior of public and private firms. We will assume a public firm behaves in a *risk neutral manner* and *maximizes the expected discounted value of dividends.* However a private firm *maximizes the expected discounted value of utility of consumption.* If the owner of a private firm has a concave utility function, their behavior will be *risk averse* and we will see that risk aversion can lead the owner of a private firm to invest *more conservatively* than a publicly owned firm. Even in a model without any stochastic shocks, we will see that the owner of a private firm distorts his/her investment plan out of a desire for *dividend smoothing* which ultimately derives from a desire to *consumption smooth.* Thus, an owner of a private firm will generally not want to pay zero dividends because this will lower his/her consumption in such periods. As a result, the owner of a private firm is willing to pay a penalty in the form of *slower growth* in order to dividend-smooth.

Once we have stochastic shocks, we can also introduce additional types of uncertainty, including *uncertainty over a firm's future investment opportunities.* An especially simple specification that involves

Bayesian learning is outlined below. Suppose that there are two potential future "second stage" investment opportunities as represented by two different stochastic production functions $f(k_t, \eta_t, \tau_1)$ and $f(k_t, \eta_2, \tau_2)$. Assume for simplicity that the firm is certain about its first stage production technology and suppose this is the $\sqrt{k}$ production function given in equation (37) but the firm is not sure whether its second stage technology is the more productive one characterized by parameter $\theta_1 = 10$ or the less productive one with parameter $\theta_1 = 2$. Then we can let $\tau_1 = \theta_1 = 10$ and $\tau_2 = \theta_1 = 2$ and endow the firm with a prior π_0 which represents its initial belief about whether its second stage production technology if the more productive τ_1 technology with $\theta_1 = 10$.

In addition, and for simplicity, consider the lognormal specification of stochastic production shocks given in equation (41). Then it is easy to show that at time t the firm's posterior belief that its second stage technology will be the more productive one is π_t given by the recursive Bayesian updating rule

$$\pi_t = \frac{g(y_t|\tau_1)\pi_{t-1}}{g(y_t|\tau_1)\pi_{t-1} + g(y_t|\tau_2)(1-\pi_{t-1})}, \tag{46}$$

where $g(y_t|\tau_i)$ is a lognormal density for realized output at time t given it is of type τ_i, $i \in \{1,2\}$, which has parameters $(f(k_t, \tau_i), \sigma^2)$.

It is not difficult to extend the Bellman equations (44) and (45) to allow for firm learning about its second stage production technology and for brevity we omit it. However we do note that the objects $V(k,l,\pi)$, $I(k,l,\pi)$ and $D(k,l,\pi)$ providing the firm's expected discounted profits and optimal investment and dividend policies are not functions both of the current capital stock k and liquid capital l, but also on *the firm's beliefs about its second stage investment technology* π. The addition of this extra level of uncertainty considerably enriches the model, since it will lead certain firms with sufficiently optimistic beliefs about their future investment and growth opportunities to undertake an ambitious investment path as in the left panel of figure 4 above, whereas firms that are more pessimistic about their future investment/growth prospects will invest less and not not grow as large, similar to the outcome in the right hand panel of figure 4.

We can also incorporate stochastic productivity shocks and uncertainty about the potential attractiveness of second stage technologies and investment opportunities into a model of the behavior of a privately owned firm that we will discuss in section 6. We will show that these differing beliefs (which can be interpreted as "degrees of optimism") can contribute to the explanation as to why some firms invest heavily, grow rapidly and may ultimately go public via an IPO whereas other firms may grow more slowly and never go public via an IPO. But before we do this, it is important to extend our model of a public firm to allow for debt.

2.3 Extending the model to allow debt

Note that the solution we provided above has the "boundary condition" $V(0) = 0$, i.e. if a firm has no initial capital stock, it will not have any cash flows to invest, and thus it is never able to "get off the ground" even though there may be an attractive production technology $f(k)$ that the firm "owns". We might think of f as the "entrepreneurial idea" but that idea cannot be implemented without a cash investment to get the firm going. As long as the firm has some way of getting this initial investment, it is enough to get it going and eventually a sequence of investments will lead it to reach the optimal steady state capital stock k^*, which is the same capital stock it would choose if it had sufficient capital to make a large one time investment at the optimal scale.

We now extend the model to allow the firm to borrow. We first consider the easiest case where the firm has no existing debt and makes an initial borrowing decision at period 0. Assume that the owners of a public firm can borrow an unlimited amount at an interest rate $r_b \geq r = 1/\beta - 1$ and have capital level $k < k^*$. As we noted in the previous section, the optimal capital stock k^* is the value that satisfies

$$V'(k^*) = 1/\beta = 1 + r \tag{47}$$

which is the level of investment where the return to investing an additional dollar in the capital stock equals the market rate of return on the firm, $1 + r$. Further, we showed that

$$V'(k^*) = f'(k^*) + 1 - \delta \tag{48}$$

so the effect of investing a marginal dollar to increase the capital stock in the firm equals the marginal product of capital, $f'(k^*)$, plus the residual value of that extra unit of capital after depreciation, $1 - \delta$. Combining these two expressions, we obtain the classic "Golden rule" formula for the optimal steady state capital stock k^*

$$f'(k^*) = \delta + r \tag{49}$$

i.e. at the optimal steady state capital stock k^*, the marginal product of capital must be just large enough to cover the rate of depreciation of capital δ plus the opportunity cost of capital, r.

If the firm has existing capital $k < k^*$ and cannot borrow, then it has to use retained earnings to finance a gradual buildup of its capital stock until it reaches the optimal steady state value k^*. Now consider how this changes if the firm can borrow at the same rate of interest r as its return on equity (i.e. the interest rate that its future cash flows are discounted at). Let b be the amount the firm borrows. With sufficient borrowing, the firm can finance any desired level of investment I. So we can write $I = f(k) - b$, i.e. the amount the firm can invest equals its cash flow stream $f(k)$ plus any additional amount it chooses to borrow, b, where we use the convention $b < 0$ as borrowing whereas $b > 0$ would indicate saving (e.g. investment in bonds).

Thus, we can write the firm's problem as that of choosing the level of borrowing b to maximize the value of equity in the firm

$$\max_{b \leq 0} \beta[V(k(1-\delta)+f(k)-b)+b(1+r)] \tag{50}$$

which equals the net value of equity after the firm undertake both the investment and the borrowing necessary to finance it. In equation (50) we multiply the net value of the firm's equity at time $t = 1$ by the discount factor β to discount back the net equity in the firm to time $t = 0$ when the borrowing and investment decision is actually made.

It is clear from equation (50) that the optimal level of borrowing b^* satisfies $V'(k(1-\delta)+f(k)-b^*) = 1+r$, which implies that $b^*(k) = k(1-\delta)+f(k)-k^*$. Thus, we have derived the optimal debt policy for the firm in period 0:

$$b^*(k) = \begin{cases} k(1-\delta)+f(k)-k^* & \text{if } k < k^* \\ 0 & \text{if } k \geq k^*. \end{cases} \tag{51}$$

The firm could also borrow an additional amount x to finance dividend payments at period $t = 0$ but it is clear that borrowing to pay dividends has no effect on the value of the firm: borrowing x to pay extra dividends increases the value of current equity by x but increases the present value of future liabilities by $\beta x(1+r) = x$, so we have assumed without loss of generality that the firm will not undertake an "superfluous" additional borrowing to pay dividends to its shareholder when there is no strict gain in shareholder value from doing this.

What is the optimal financial policy for the firm from period 1 onwards? To avoid default the firm will have to pay off its debt with interest $b^*(k)(1+r)$ and if $f(k^*) - \delta k^* + b^*(k)(1+r) < 0$, then the firm's net cash flows after investment are insufficient to retire its current debt and the firm will have to roll over its single period loan by taking out another single period loan. In the absence of any debt, the firm would be able to pay dividends equal to $d(k^*) = f(k^*) - \delta k^*$ to its shareholders. By rolling over the entire debt, the firm would be able to make these dividend payments in period $t = 1$. However consider what would happen to the value of equity if the firm reduced dividend payments by x in order to reduce the amount of its borrowing by x. This reduces the debt payment due at period $t = 2$ by $(1+r)x$, and the present value of this reduction in debt at $t = 1$ is $\beta x(1+r) = x$. Thus, the reduction in current dividends of x is exactly counterbalanced by the present value of the reduction in future liabilities of x, leaving the value of equity in the firm independent of financial policy as represented by the amount of additional debt to pay off, x.

This argument can be extended to show that the owners of the firm are indifferent about whether or how to pay off the accumulated debt once they achieve the optimal steady state capital stock k^*. That is, the value of equity is the same if the firm rolls over its debt perpetually, or follows any path for gradually paying its debt over time, financed by a corresponding reduction in dividends. Also, note that the Modigliani-Miller Theorem holds: the total value of the firm $V(k^*)$ is unaffected by firm financial policy about whether to pay off the firm's accumulated debt, and if show, the time path pat over which it is paid off.

On the other hand, note that the total value of the firm is *most definitely affected by the firm's initial borrowing decision!* The starkest way to illustrate this point is to consider a firm that initially has no capital to invest, $k = 0$. Then as we saw in Theorem 1, the value of the firm is $V(0) = 0$. A firm that has access to debt markets can borrow and amount $b^*(0) = k^*$ resulting in an immediate equity valuation of $\beta[V(k^*) - k^*(1+r)] > 0$. Theorem 1 below covers the less extreme case where the firm does have some positive initial capital stock k.

Theorem 2 *Consider a firm with capital stock that satisfies $f(k) + (1-\delta)k < \underline{k}$ where $\underline{k}$ is the upper boundary of the liquidity constrained region of a firm that does not have access to capital markets, given in (8). This implies that $k < \underline{k}$ and thus the borrowing-constrained firm optimally invests all of its cash flow: $I(k) = f(k)$ and pays no dividends as proved in Theorem 0. Suppose the firm is suddenly given access to capital markets and is allowed to borrow unlimited amounts at interest rate $r = 1/\beta - 1$. Let $E(k,0)$ denote the optimal value of equity in this firm after undertaking its desired investment and borrowing. We have*

$$E(k,0) = \beta V(k^*) + k(1-\delta) + f(k) - k^* > V(k). \tag{52}$$

Thus, the ability to borrow strictly increases the equity valuation of the firm as long as k is in the "borrowing constrained" region. To see why Theorem 2 is true, note that k^* is the solution to

$$\max_k [\beta V(k) - k] \tag{53}$$

and thus we have

$$\beta V(k^*) - k^* > \beta V(k(1-\delta) + f(k)) - [k(1-\delta) + f(k)] \tag{54}$$

as a result of the strict concavity of V, where the expression on the right side of inequality (54) is equal to $V(k)$ in the investment constrained region where $k \in (0, \underline{k})$. The proof is completed by noting that inequalities (54) and (52) are equivalent.

We can define a Bellman equation for the equity value of the firm more generally as

$$E(k,b) = \max \left[0, \max_{\substack{I \geq 0 \\ \underline{B} \leq b' \leq f(k) - I + b}} \left[f(k) - I + b - b' + \beta E(k(1-\delta) + I, b'(1+r_b)) \right] \right]. \tag{55}$$

where r_b is the interest rate on single period debt, and $\underline{B} < 0$ is the borrowing limit for the firm. Notice that the firm has three choice variables: investment I, dividends d and the amount of new borrowing b'. However we can eliminate d by expressing it in terms of (I, b'). We have $d = f(k) - I + b - b'$ and the constraints that dividends cannot be negative implies that $f(k) - I + b - b' \geq 0$ which implies the upper bound on b' equal to $f(k) - I + b$ in the optimization over b' in the Bellman equation (55).

The outer max expression in equation (55) can be interpreted as a *limited liability constraint* that insures that the equity valuation of the firm can never be negative. If the firm were to ever have too much debt for a given capital stock so that the value of the currently due debt exceeded the present value of future dividend payments (and thus a negative equity value), the owners of the firm can always default, allowing the bondholders to take control of the firm to recover as much of the amount due to them as they possibly can.

Bankruptcy can be accomodated for in this model as an event that occurs if the firm were ever to default on its currently due debt b. In equation (55) b' is the amount of new borrowing the firm undertakes, which is often to "roll over" the currently due debt b by taking out a new one period loan. However if the firm's current debt b is so large that even if the firm does not invest, $I = 0$, and devotes all cash flow $f(k)$ towards paying off this debt, the amount of new borrowing would violate the borrowing constraint $b' = f(k) + b \geq \underline{B}$. In this situation, the firm is not allowed to undertake enough new borrowing to pay off its existing debt and a default occurs. At this point the firm goes bankrupt and the existing bondholders are allowed to take ownership of the firm in order to collect as much of the currently due debt as possible.

Theorem 3 (Modigliani-Miller) *Suppose the firm is not bankrupt and the borrowing rate is $r_b = 1/\beta - 1$ and the firm does not face any binding borrowing constraints and $k \geq \underline{k}$ where $\underline{k}$ is the upper boundary of the borrowing-constrained region given in equation (8). Then we have*

$$E(k,b) = V(k) + b \tag{56}$$

i.e. the equity valuation of the firm equals its total value $V(k)$ less the amount of its currently due debt where V is the solution to the Bellman equation for the firm that has no access to borrowing, (2). However for $k < \underline{k}$ we have

$$E(k,b) > V(k) + b \tag{57}$$

and we can conclude that for all $k \geq 0$ we have

$$E(k,b) - b = E(k,0) \quad b \geq -E(k,0) \tag{58}$$

i.e. the total value of firm (debt plus equity), $E(k,b) - b$ does not depend on b as long as the firm's equity valuation is positive, which will be guaranteed as long as the debt does not exceed the total value of a debt-free firm, $b \geq -E(k,0)$.

To prove Theorem 3, note that the Bellman equation has a unique solution $E(k,b)$ and we can use a "guess and verify" approach to show that if $k \geq \underline{k}$, then $E(k,b) = V(k) + b$. If we substitute this conjectured formula for $E(k,b)$ into the right hand side of the Bellman equation for E, (55) and we get

$$\begin{aligned}
E(k,b) &= \max_{\substack{I \geq 0 \\ \underline{B} \leq b' \leq f(k) - I + b}} \left[f(k) - I + b - b' + \beta V(k(1-\delta) + I) + b' \right] \\
&= \max_{I \geq 0} \left[f(k) - I + \beta V(k(1-\delta) + I) \right] + b \\
&= V(k) + b
\end{aligned} \tag{59}$$

since it is evident that the optimal value of I in the middle equation of (59) is exactly what is given by in Theorem 0 and in equation (6), which implies that the first term in brackets in the middle equation of (59) is just $V(k)$. However when $k < \underline{k}$, this reasoning breaks down since in our derivation of the original Bellman equation (2) the firm is liquidity constrained in this region and had to settle for the constrained optimal solution $I(k) = f(k)$ when $k < \underline{k}$. However when the firm can borrow, it can do better and borrow enough to jump to the optimal steady state capital stock k^* in the very first period by borrowing $f(k) + k(1-\delta) - k^*$ (which implies that it pays no dividends). So in particular, as we showed in Theorem 2, $E(k,0) > V(k)$ for $k < \underline{k}$. However we can use the "guess and verify" approach and the Bellman equation for equity (55) to show that $E(k,b) = E(k,0) + b$. Clearly the equity valuation fo the firm will be non-negative as long as the amount of debt satisfies $b \leq -E(k,0)$, but otherwise it is obvious from the identity $E(k,b) - b = E(k,0)$ that the total valuation of the firm does not depend on the amount of debt it has, b, so we have verified that the Modigliani-Miller Theorem holds.

We conclude from Theorem 3 that even in the frictionless, no tax case where $r_b = r = 1/\beta - 1$, that borrowing and dividend policy *do not matter* and the Modgliani-Miller Theorem does hold. The firm can choose to borrow more to pay greater current dividends, but the increase in firm value from the larger dividend payment is exactly offset by the increase in liabilities the firm takes on to pay the higher dividend. However it is important to note that investment policy, on the other hand, *does matter* and the firm has a well defined optimal investment policy given by the function $I(k,b)$ given below

$$I(k,b) = \begin{cases} k^* - (1-\delta)k & \text{if } k \leq \overline{k} \\ 0 & \text{if } k > \overline{k} \end{cases} \tag{60}$$

where $\overline{k}$ is defined in equation (7) in section 2. Note that while we have written investment as depending potentially on both k and b, notice from the right hand side of equation (60) does not depend on b, and thus, the amount of the firm's debt has no effect of its investment decisions.

The other key difference between the optimal investment policy for a borrowing-constrained firm and one that faces no borrowing constraints, is the elimination of the investment-contrained interval $[0,\underline{k}]$ where $I(k) = f(k)$. Thus a firm that starts in period $t = 1$ with no initial debt and which is allowed to borrow as much funds as it needs can jump immediately to the optimal steady state capital stock k^* starting in period $t = 2$ after borrowing the amount it needs to reach k^*, $b'(k) = (1-\delta)k - k*$. If the firm is encumbered with some initial debt b in period $t = 1$ (say the firm needed to borrow to purchase the technological rights/patents that enables it to produce using the production function $f(k)$), then the unconstrained firm can simply borrow enough to pay back the currently due debt b plus the additional amount $k(1-\delta) - k^*$ to enable it to invest and get to the steady state capital stock k^* in period $t = 2$. The firm could also borrow even more to pay dividends to its shareholders right away, but as we showed above, doing this will not increase equity value.

Though the *Wikipedia* entry on the Modigliani-Miller Theorem states that it is a "capital structure irrelevance principle" that forms "the basis of modern thinking on capital structure" it is important to note how fragile this result is and how it breaks down when any of the key underlying assumptions is modified. In Theorem 4 we show that when the firm faces binding borrowing constraints, the Modigliani-Miller capital structure irrelevance result fails, and the firm's investment and valuation will be affected by its level of debt.

Theorem 4 (Modigliani Miller Theorem fails in the presence of borrowing constraints) *Suppose the assumptions of Theorem 3 hold except the firm can borrow at most $\underline{B} < 0$ in any period. Then the Modigliani Miller Theorem fails to hold: the total value of the firm is no longer independent of the level of its initial debt, i.e.*

$$E(k,b) \neq E(k,0) + b \tag{61}$$

for all possible values of (k,b). In particular, we have

$$E(0,\underline{B}) = 0 < E(0,0) + \underline{B} \tag{62}$$

and the firm's investment policy is no longer independent of its debt b. In particular, we have

$$I(0,\underline{B}) = 0 < I(0,0). \tag{63}$$

The proof of Theorem 4 is straighforward: if the firm faces a borrowing constraint of $\underline{B}$ and starts its life with zero capital, it is unable to borrow any additional funds to get off the ground, and hence it will not be able to invest so $I(0,\underline{B}) = 0$ and its equity valuation will be zero as well $E(0,\underline{B}) = 0$ since the firm starts its life too encumbered by debt, and essentially faces no other choice other than immediate bankruptcy. However even if the firm does have a positive level of capital, excessive initial debt can create binding liquidity constraints for the firm that limits the amount the firm can invest, and this distortion in investment penalizes the value of the equity by more than the amount of the firm's initial debt *per se.*

We illustrate this in figure 6 which plots $E(k,b)$ as a function of k for four different values of $b \in \{0,10,20,30\}$ when we assume the borrowing limit is $\underline{B} = -30$. We see that the curves for $b = 0$ and $b = 10$ are parallel so we have $E(k,-10) = E(k,0) - 10$. This means that if the firm is not encumbered with too much initial debt, then the borrowing constraint is not binding and the value of equity of a firm with initial debt of $b = -10$ equals the value of equity of a firm with no initial debt, $E(k,0)$, less 10. However the curves for $b = 20$ and $b = 30$ are no longer parallel to the green curve for the graph of $E(k,0)$. This is because the value of initial debt is sufficiently large that it constrains the firm's ability to borrow, and this distorts the optimal investment policy and is reflects by an equity value $E(k,b)$ that is lower than the value predicted by the Modigliani Miller Theorem: $E(k,0) + b$. For example the lowest black line in figure 6 plots $E(k,-30)$ and we see that $E(0,-30) = 0$ as predicted by Theorem 4. When the firm's initial debt is sufficiently high, even when the firm starts with a positive but sufficiently low level of

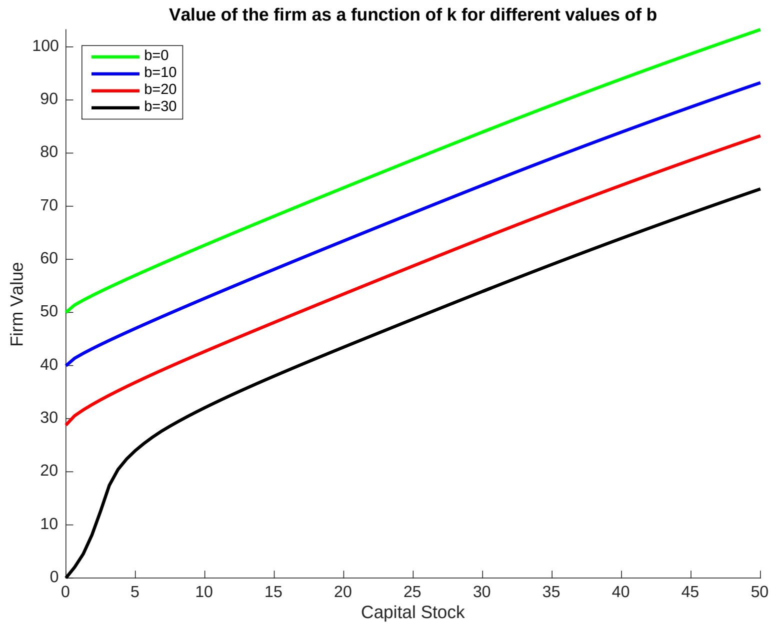


Figure 6: Illustration of the failure of the Modigliani Miller Theorem

capital the firm's equity value is penalized due to the binding borrowing constraints that prevent the firm from reaching the optimal steady state capital stock k^* as quickly as is possible when the firm starts with less initial debt and thus avoids having its investment policy distorted by binding liquidity constraints.

In the remainder of this section we consider the more realistic cases where $r_b \neq r$. The Modigliani Miller Theorem also fails whenever $r_b \neq r$. Consider first the case where $r_b < r = 1/\beta - 1$, where the firm can borrow funds at a lower interest rate than its return on equity r. It is intuitively clear that in this case, the firm would want to borrow as much as it possibly can to pay current dividends to its shareholders. Each dollar it borrows today allows it to raise dividends by a dollar, but the market valuation of debt it will have pay back next period is $\beta(1+r_b) < 1$, so the firm's valuation increases by incurring debt to pay current dividends. However it is not realistic for the firm to be able to borrow unlimited amounts at r_b. Suppose the firm faces a borrowing limit $\underline{B}$: we will now show that unless the firm has very little initial capital and too much initial debt, the firm strictly prefers to maintain a debt load of $\underline{B}$ every period, so its capital structure is no longer irrelevant. The firm will also use some of its debt to pay dividends once it capital stock k exceeds the lower threshold $\underline{k}$ defined in equation (8). We summarize these results in Theorem 5 below.

Theorem 5 (optimal capital structure when $r_b < r$) *Suppose the assumptions of Theorem 4 hold except that $r_b < r$. The optimal debt policy is for the firm to borrow the maximum amount $\underline{B}$ in all periods except in states (k,b) where k is sufficiently low and b is sufficiently close to $\underline{B}$ that the firm's cash flow is insufficient to pay off the interest on its debt if it borrowed the maximum amount allowed $\underline{B}$. In these*

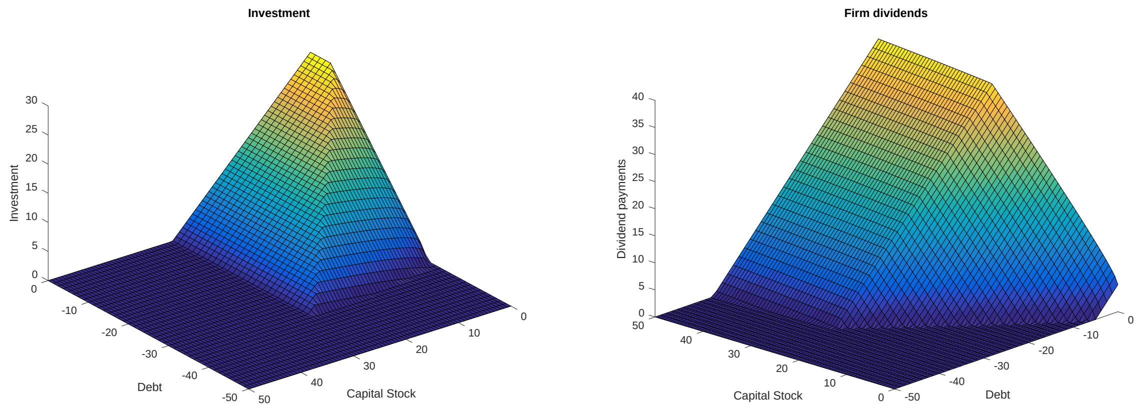


Figure 7: Optimal investment and dividend policies when $r_b = 0.03$ and $r_m = 0.05$

states the both the firm's borrowing and investment will be constrained by the borrowing constraint and the firm's desire to avoid default and bankruptcy. If the firm has less initial debt, it will borrow the maximum amount allowed and use the proceeds for investment, in order to reach the steady state capital stock k^* as quickly as possible. If the firm's initial capital stock is below $\underline{k}$, it will not pay dividends until its capital stock exceeds $\underline{k}$. If the firm has a sufficiently small initial debt b and sufficiently low initial capital, if it can borrow enough to reach the optimal steady state capital stock k^* in the first period and has residual borrowing capacity, it will use it to pay dividends to its shareholders. Once it reaches steady state, its dividend stream will be lower than for a firm without debt due to the interest its pays on its debt $\underline{B}$.

Figure 7 plots the optimal dividend and investment decision rule for the firm in the case where $r_b = 0.03$ and $r = 0.05$ and $\underline{B} = -30$. This borrowing limit is sufficiently high that a firm without any initial capital or debt can borrow the maximum amount $\underline{B}$ to reach the steady state capital stock $k^* = 25$ in the first period, and using the remaining proceeds to pay out 5 in dividends to its shareholders. In all subsequent periods the firm pays a dividend equal to 2.85 which equals the net cash flow stream $f(k^*) = 5$ less depreciation of $\delta k^* = 1.25$ and less interest on its debt of $0.90 = r_b\underline{B}$. The firm is willing to incur a lower perpetual dividend stream due to the fact that its initial borrowing in period $t = 0$ enabled it to pay its shareholders the large initial dividend of 5.

However we can see from figure 7 that if the firm is encumbered by initial debt, it may not be able to borrow enough to reach the steady state capital stock in the first period. The firm will be borrowing constrained in a way that is similar to a firm that has no ability to borrow that we analyzed in section 2, and the curved region on the optimal investment function $I(k,b)$ indicates the effects of the liquidity constraint on firm investment. For example if the firm starts at initial condition $(k,b) = (0,-20)$ it can only invest $I(k,b) = 10$ in the first period because it has to use the other 20 of its total borrowing of $\underline{B} = 30$ to pay off the existing debt. Thus, the firm is debt constrained and will not pay any dividends until its capital exceeds

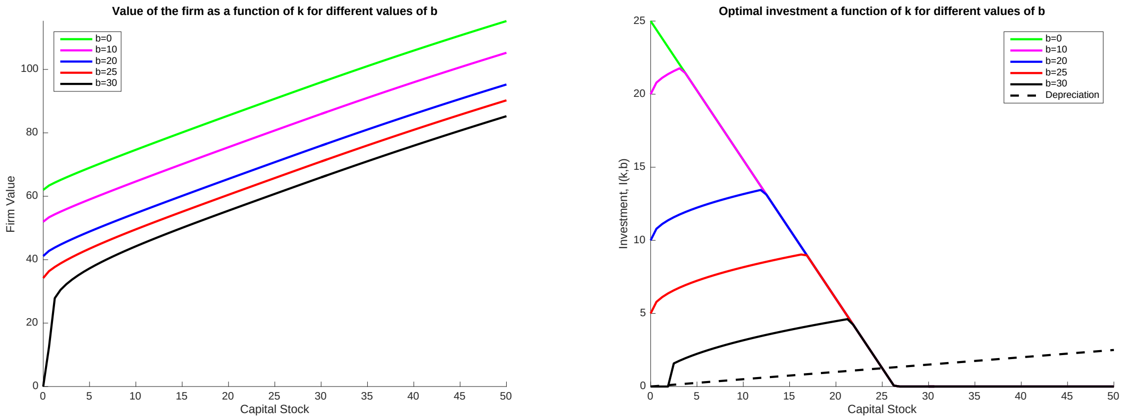


Figure 8: Effect of debt on firm value and investment

the threshold $\underline{k}$.

Figure 8 provides more detail on how initial debt encumbers the firm, reducing its value and the amount it can invest. The left panel shows slices of the equity value of the firm a function of k for different values of $b \in \{0, -10, -20, -25, -30\}$. Equity values are parallel to each other for sufficiently high values of k which implies that debt has an additively separable, linear effect on equity value in this region. However for sufficiently small initial values of k the effect is nonlinear, and increases in initial debt reduces equity value by more than the incremental value of the debt. In particular, we find that $E(0, -30) = 0$ which is consistent with equation (62) in Theorem 4. The left hand panel of figure 8 plots slices of the optimal investment function for the same five initial values of debt b and we can clearly see how larger initial debt reduces investment at low initial values of capital k. Of course the reduced investment is caused by the binding borrowing constraints imposed by the higher initial debt.

Figure 9 illustrates simulated paths for capital stock and dividends starting from five initial conditions of the form $(0, b)$ where $b \in \{0, -10, -20, -25, -28\}$. Note that when the firm starts out with no initial debt, it can invest and reach the optimal steady state capital stock $k^* = 25$ in the first period. However as its debt increases, the firm finds itself in the borrowing constrained region where it has to invest in smaller chunks over a longer period of time due to the need to roll over its existing debt and pay interest on it. The right hand panel shows the effect of the borrowing constraint on dividend payments. If the firm starts with no initial debt and no capital, it can borrow in the first period up to its borrowing limit $\underline{B} = -30$ and invest to reach the optimal steady state capital stock k^* in the first period and pay a \$5 dividend to its shareholders. However as the initial debt b rises, the firm is forced to follow a slower capital accumulation trajectory and it pays no dividends to its shareholders until its capital exceeds the threshold $\underline{k}$ given in eequation (8).

In the remainder of this section we consider the case $r_b > r$, i.e. where the interest rate that the firm can

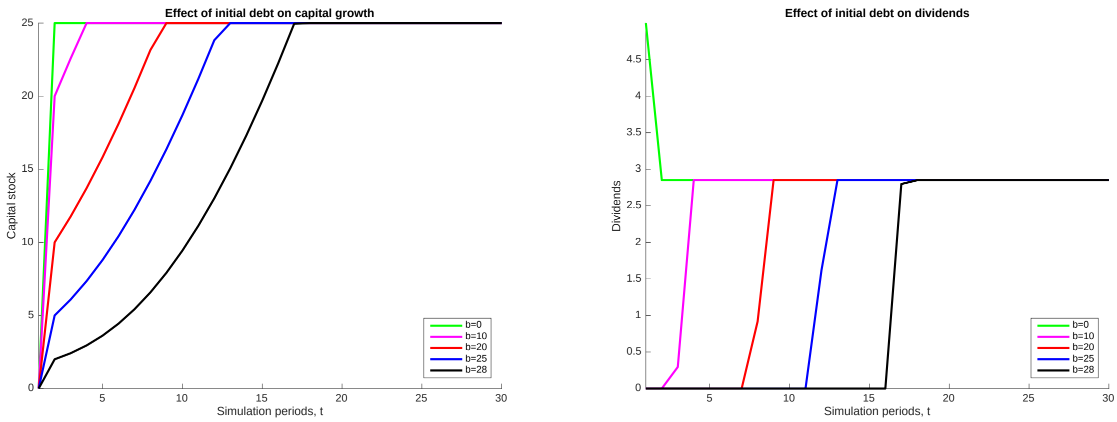


Figure 9: Effect of debt on capital accumulation and dividends

borrow is higher than the interest rate investors use to discount its dividends. Will the firm choose to carry long term debt on its books in this case? Theorem 6 shows that the answer is negative: the firm generally only uses debt for *short term investment purposes* in order to accelerate its investment path to the optimal steady state capital stock k^*, but once it has accumulated sufficient capital (and thus has sufficient retained earnings), it uses its retained earnings to pay off its debt and does not start paying dividends until its debt is entirely retired. There is a key exception to this: if the firm has too much initial debt and insufficient initial capital, it is optimal for it to undertake *strategic bankruptcy* by borrowing to pay dividends to its shareholders and not investing. This causes the firm to hit its borrowing limit and the lack of investment and initial capital results in insufficient cash flow to pay its accumulated debt obligations.

Theorem 6 (optimal capital structure when $r_b > r$) *Suppose the assumptions of Theorem 5 hold except that $r_b > r$. There are two possible steady state outcomes for the firm in this case: a) bankruptcy if the initial capital k is sufficiently low and initial debt b is sufficiently close to or greater than the borrowing limit $\underline{B}$, or b) a debt-free steady state outcome where the firm maintains the optimal steady state capital stock k^* and pays dividends. In case a) the firm is either immediately insolvent if its cash flow is not enough to pay off its existing debt even if it borrows the maximum amount allowed, $\underline{B}$, or the firm will borrow to pay dividends but not invest, which results insolvency and bankruptcy in the following period. In case b) the firm borrows during a capital accumulation phase during which it pays no dividends. Once its capital stock is sufficiently high, the firm gradually pays off its accumulated debt, balancing the need to continue to grow its capital stock with its desire to pay off its debt. Once the debt is paid off it will never borrow again and all future investment is financed by retained earnings. Thus, the firm's optimal dividend and investment policy coincide with the policy of a firm that does not have access to capital markets given in Theorem 0. In particular, when the firm's capital stock exceeds the threshold $\underline{k}$ given in equation (8), the firm starts to pay dividends, and its dividends and investment in the steady state are $f(k^*) - \delta k^*$ and δk^*,* *respectively.*

The key to the proof of Theorem 6 is Lemma 1 below:

Lemma 1 *Under the assumptions of Theorem 6 we have*

$$\frac{\partial}{\partial b}E(k,b) \geq 1. \tag{64}$$

The proof of Lemma 1 is rather straightforward: if the constraint on the level of optimal borrowing b' is not binding, then via application of the Envelope Theorem to the Bellman equation for the equity value of the firm $E(k,b)$ given in equation (55) implies that

$$\frac{\partial}{\partial b}E(k,b) = 1. \tag{65}$$

Otherwise, the constraint is binding, and it is not difficult to show that in this case we have $b' = f(k) - I(k,b) + b$ which implies that dividends are zero, and

$$\frac{\partial}{\partial b}E(k,b) = \beta\frac{\partial}{\partial b}E(k(1-\delta)+I(k,b),(f(k)-I(k,b)+b)(1+r_b))(1+r_b). \tag{66}$$

However due to the fact that the constraint on borrowing is binding, we have

$$-1+\beta\frac{\partial}{\partial b}E(k(1-\delta)+I(k,b),(f(k)-I(k,b)+b)(1+r_b))(1+r_b) \geq 0, \tag{67}$$

which together with equation (66) establishes the desired result.

Lemma 1 implies the intuitive conclusion that the firm will never want to borrow to pay dividends, since an extra dollar of borrowing to pay a dividend raises the value of the firm by a dollar today, the effect of the extra debt that is incurred to pay this extra dividend depresses the value of the firm by $\beta(1+\frac{partial}{\partial b}E(k(1-\delta)+I(k,b),(1+r_b)b'(k,b)) \geq 1$, resulting in a net reduction in the value of equity in the firm. This is why the firm wants to be debt-free in steady state, because we can run the argument just given in reverse: by sacrificing a dollar of dividends today, the firm can reduce its debt which increases its net equity valuation.

We can show that is optimal for the firm to borrow when its capital stock is sufficiently low, because the borrowing accelerates its capital accumulation and reduces the amount of time it takes to reach the optimal steady state capital stock k^*. However because debt is costly the firm faces a tradeoff between the returns to investing versus the returns to paying off its debt. This leads an optimal investment policy where the firm generally invests, but in some situations the firm may temporarily stop investing to accelerate the rate at which it can pay off its debt. We can show that whenever the firm does borrow, it will never pay dividends due to the arguments given above. However in some circumstances the firm will use its borrowing to invest, but in other cases the firm will choose not to invest and use all of its cash flow to pay off its accumulated debt as quickly as possible. The firm thus stops doing any further borrowing once it

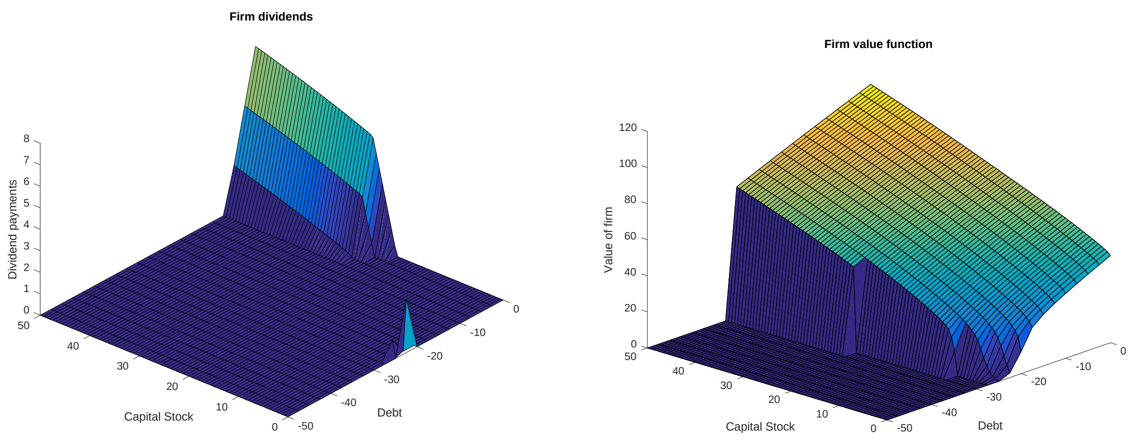


Figure 10: Optimal dividend policy and equity valuation for $r = 0.05$ and $r_b = 0.07$ with $\underline{B} = -25$

has paid off its debt, and once this happens, the firm follows the same optimal investment and dividend policy as a firm that does not have access to capital markets as we established in Theorem 0.

A striking conclusion from Theorem 6 is the possibility of strategic bankruptcy by firms that have low initial capital that are overencumbered by debt. The firm realizes that it is not sufficiently productive to make it worthwhile to pay off its initial debt, but if there are lenders willing to lend to this firm, the firm will cynically exploit them by borrowing to pay dividends, knowing full well that this will result in the firm's bankruptcy in the following period. The firm is essentially following a "take the money and run" strategy here, except that it cannot run: once the firm fails to meet its debt obligations (by reaching its borrowing limit and being able to borrow any more to remain solvent), the creditors seize the firm and the owners' equity value falls to zero. However the owners calculate that it is better for them to have one period of dividends and then go bankrupt rather than try to borrow to invest in and grow the firm and eventually become debt-free.

We see the strategic bankruptcy in the graph of optimal dividends for a firm where $r = 0.05$ and $r_b = 0.07$ with a borrowing limit of $\underline{B} = -25$ in the left hand panel of figure 10. We see a secondary peak in the dividend policy in the region near $(k,b) = (0,-25)$ and for capital in this region, strategic bankruptcies can occur. For example consider a firm with $(k,b) = (1,-25)$. The optimal policy for this firm is to borrow $b'(k,b) = -25$ and invest $I(k,b) = 0$ and pay dividends equal to $f(1) - I + b - b' = 1$ to its shareholders. The next period its debt and interest equal $b = -26.75$ and its capital has depreciated to $k = 0.95$ and so its cash flow $f(k) = 0.97$ is not sufficient to pay off the total debt due, even if it borrows the maximum amount allowed, $b' = \underline{B} = -25$. So the firm has intentionally borrowed in period $t = 1$ to pay dividends, knowing full well that it will go bankrupt in period $t = 2$.

The right hand panel of figure 10 plots the equity value $E(k,b)$. We see this function appears generally

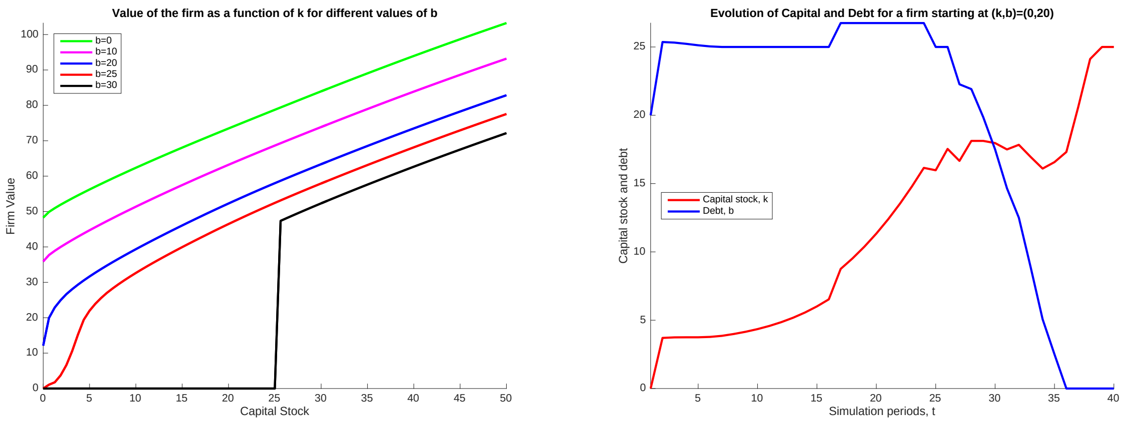


Figure 11: Effect of debt on firm value and capital accumulation for $r = 0.05$ and $r_b = 0.07$ with $\underline{B} = -25$

concave, but is essentially linear in b for large values of k, but concave in b for low values of k, especially for values of b that are close to the borrowing threshold $\underline{B}$. Equity values fall to zero when initial debt is significantly larger than $\underline{B}$ and this represents a region of insolvency where the firm's debt is too large given its cash flow and borrowing limit to enable it to pay off its existing debt b.

The left hand panel of Figure 11 provides more detail on the value function by ploting slices of the equity valuation function $E(k,b)$ as a function of k for various values of b. We see that $E(k,b)$ is actually *not* concave in k for all values of b, and in particular the red line which plots $E(k,-25)$ shows that this function starts at 0 and has an initial convex region before turning concave. Notice that the firm is bankrupt when $b = -30$ and $k < k^* = 25$. This is because its cash flow is less than $f(k) = \sqrt{25} = 5$, and even if it devoted all of this cash flow and borrowed the maximum amount $\underline{B} = -25$, the firm could not pay off its existing debt of $b = -30$. Similarly if $b > -25$ and $k = 0$, the firm cannot earn any cash flow and cannot borrow enough to roll over its existing debt, so $E(k,b) = 0$.

However if $b \geq -25$ and $k > 0$, then the firm has enough combined cash flow plus borrowing capacity to avoid immediate bankruptcy, and generally we see that the lower the amount of initial debt, the higher the value of the firm. The less the amount of initial debt, the less "encumbered" the firm is by the the interest costs necessary to to refinance its existing debt each period, so the firm can invest faster and then retire its debt faster.

The right hand panel of figure 11 show the complex nature of the capital accumulation and debt paths predicted by this simple model. The figure shows the debt and capital paths for a firm that starts out with $(k,b) = (0,-20)$. In the first period it borrows $b' = -23.70$ and uses 20 of this to pay off its existing debt and invests $I(k,b) = 3.70$. Notice that the firm chooses not to borrow all the way up to its borrowing limit $\underline{B} = -25$ because doing so would put it too close to its borrowing limit in the future periods, thereby reducing its *future borrowing potential*. We see that the firm borrows less than $\underline{B}$ for the first 15 periods while it gradually accumulates capital. Then starting in period $t = 16$ the firm has enough accumulated capital to start borrowing at the maximum amount allowed, $\underline{B}$ and we can see how this additional borrowing helps to accelerate the growth of its capital stock.

Then starting in period $t = 24$, when its capital stock is $k = 16.15$, the firm starts to pay off its accumulated debt. During this "payoff phase" the firm's investment path is more volatile, falling to zero in periods 26, 32 and 33, and positive in other periods. Then in period $t = 35$ the firm is able to pay off its debt completely, so from period $t = 36$ onward it becomes debt-free. In these periods the firm's investment policy coincides with the investment policy given in Theorem 0. In particular, the firm does not start paying dividends until $k > \underline{k}$ which occurs in period $t = 38$. In this period, the firm also invested enough retained earnings to reach the optimal steady state capital stock $k^* = 25$, and so starting in period $t = 39$ it can pay the optimal steady state dividend level of $f(k^*) - \delta k^* = 3.75$.

We conclude that the existence of costly debt (i.e. debt whose interest rate r_b exceeds the firm's discount rate r) does benefit the firm, but it also introduces distortions that affect both dividend and investment policy. In particular, during the period where the firm is borrowing, it will not be optimal for the firm to pay dividends and the firm's investment will be distorted and constrained by the firm's borrowing constraints. However once the firm's capital stock grows to a level that generates sufficient retained earnings, the firm can start to focus on paying off its accumulated debt as quickly as possible. Only once the firm has fully paid off its costly debt will it resume a rapid rate of investment to reach the optimal steady state capital stock k^* and start paying dividends.

It is possible to extend our analysis to consider the effect of stochastic shocks to firm productivity and more complex financial constraints, such as when external lenders impose a capital dependent borrowing constraint $\underline{B}(k)$. That is, $\underline{B}(k)$ is the maximum amount banks or other lenders would be willing to lend the firm when its capital stock is k. We can think of $-\underline{B}(k)$ as the "collateral value" of the firm: the value that lenders believe they could obtain in a worst case scenario where the firm defaulted and the existing capital investment in the firm would have to be sold off in a "fire sale" to outside investors. However a better case scenario is that in the event of a default, the bondholders simply gain ownership of the firm and continue to manage it as a going concern. Under this scenario we would have $-\underline{B}(k) = V(k) - C$ where C is a fixed transactions and legal cost associated with a default and a changeover in the ownership of the firm from the existing equity holders to the bondholders.

Note that it is natural to impose a constraint that $\underline{B}(0) = 0$, i.e. a firm with no initial capital investment is precluded from being able to borrow since the firm is not yet a "going concern" and there would naturally be a concern on the part of creditors whether the investor might simply "take the money and run" leaving them without any collateral or any operating firm that could provide for at least partial repayment of their loan.

We leave the details of these extensions to future work, so that we can turn to an analysis of the optimal behavior of a privately owned firm to better understand the differences in investment and dividend policies of public and private firms.

2.4 Debt Policy for Firms with Nonconcave Production Functions

There is one interesting final topic we consider that sheds light on a key motive why smaller firms might choose to go public even when they also have the option to borrow. As we showed in subsection 2.1, when a firm has a nonconcave production function but cannot borrow, a sufficiently small firm can get "stuck" at the smaller Golden Rule capital stock k^* because it is too costly and involves too much delay to finance the investments needed to grow to the larger steady state capital stock k^* via retained earnings alone. We conjectured if firms could borrow enough in a single lump sum initial investment they would prefer to borrow enough to immediately finance and invest in the larger scale plant, i.e. at the higher steady state capital stock k^*. However the simplistic calculations we did in equations (38) and (39) ignored the reality that even at the higher steady state capital stock, k^*, the firm would not be generating enough cash flow to pay off the large initial debt in just a single period.

Consider a firm with the production function illustrated in figure 5 and now allow it to borrow up to $\underline{B} = -90$ via single period debt contracts at an interest rate of $r_b = 0.07 > r_m = 0.05$. The left hand panel of figure 12 shows the gain in value to a public firm from the option to borrow. There is no gain in value from the option to borrow when the firm has sufficient capital stock, specifically when k exceeds the larger Golden Rule steady state capital stock level $k^* = 222$. This is because, as we showed above, when $r_b > r_m$, it is not optimal for the firm to carry any debt in steady state. However for firms with sufficiently low levels of initial capital, the option to borrow helps them accelerate their growth so they can reach the higher steady state capital stock faster.

The right hand panel of figure 12 shows the simulated trajectory of a public firm that starts out with initial capital $k_0 = 20$. Recall from figure 5 that if the firm could not borrow at all, it would not be optimal for it to invest its retained earnings to reach $k^* = 222$. Instead a firm that cannot borrow would undertake two much smaller investments that are sufficient to get it to the *smaller* steady state k^*, and the firm remains there forever after. This is the "growth trap" due to lack of financing we discussed earlier. When the firm can borrow up to a limit of $\underline{B} = -90$, then the firm can escape this "growth trap" and reach the higher steady state capital stock $k^=$222 after 29 periods. Interestingly, it is not optimal for the firm to *start borrowing immediately.* Instead the firm uses its retained earnings to finance its growth for the first 17 periods, and then in period 18 it borrows the maximum amount $\underline{B} = -90$. It rolls over this debt for 3 more periods andin period 22 it starts to retire the accumulated debt until by period 28, it becomes debt-free. Via a final investment of its retained earnings it reaches the higher steady state capital stock $k^* = 222$ in

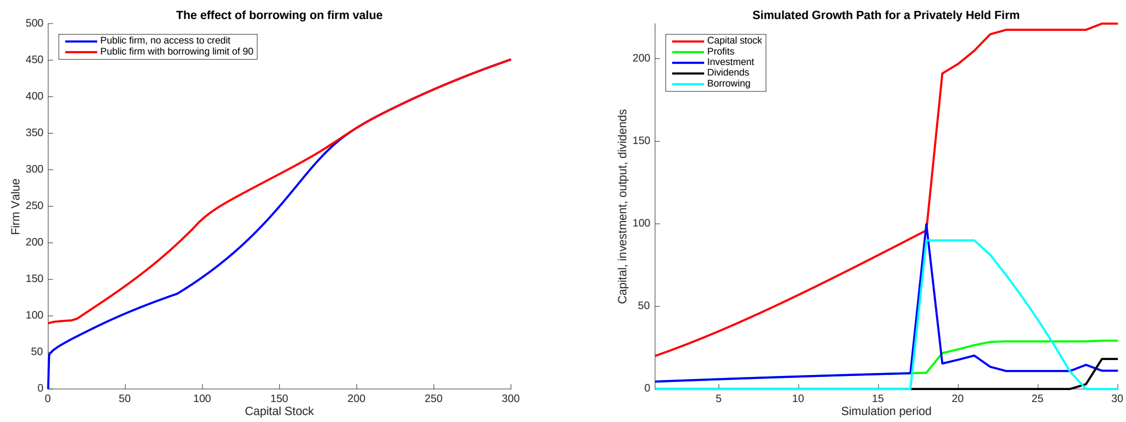


Figure 12: Gains from borrowing for a public firm with non-concave $f(k)$ and $\underline{B} = -90$

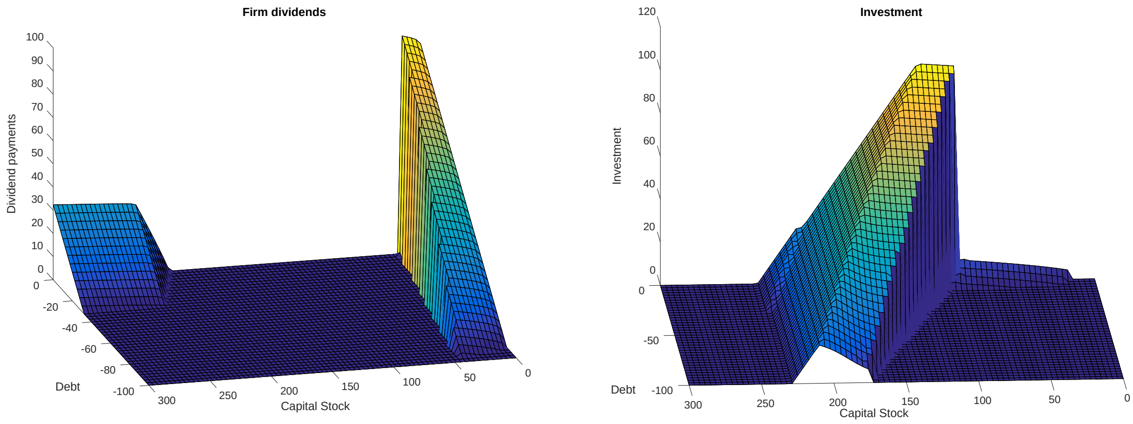


Figure 13: Optimal dividend and investment policy for a firm with non-concave $f(k)$ and $\underline{B} = -90$

period 9 where it pays out a per period dividend of 18.18 in perpetuity.

However if we were to simulate the behavior of a firm with less initial capital, say $k = 3$, quite different behavior emerges. In this case the firm borrows the maximum amount $\underline{B} = -90$ in the first period, pays a dividend of $D(k,b) = 91.73 = -\underline{B} + f(k)$ and then goes bankrupt in the following period. This is another case of "strategic bankruptcy" in our model but in this case, the firm is not encumbered with any initial debt. The firm simply makes the calculation that a one time payment of 91.73 is better than the option of investing the money and rolling over a series of one period debt contracts to try to grow and achieve the higher steady state capital stock $k^* = 222$. If lenders realize this, then it seems unlikely that they would be willing to lend much to small "startups" because of this strategy bankruptcy incentive.

The left hand panel of figure 13 illustrates the strategic bankruptcy motive. It is the high region of dividends for firms with low levels of capital. The dividends from following the strategic "take the money and run" strategy clearly dominate the steady state dividends that the firm could pay out each period in perpetuity around the higher steady state capital stock $k^* = 222$. The right hand panel of figure 13 illustrates the corresponding investment policy. The firm will not find it optimal to engage in strategic bankruptcy if it not encumbered with debt initially, and thus firms with smaller initial capital stocks will find it optimal to invest and escape the "growth trap" of remaining at the lower steady state capital stock $k^* = 25$ forever. There is a discontinuous threshold of capital where the firm finds it optimal to borrow up to the maximum allowed, $\underline{B} = -90$, which enables it to a finance a series of large investments that enable it to quickly reach the larger steady state capital stock of $k^* = 222$, and after reaching this level, investment falls back to replacement level, $I(k^*) = \delta k^* = 11.10$.

The behavior of the firm and its equity valuation are very sensitive to the magnitude of its borrowing constraints. Figures 14 and 15 illstrate what happens if the firm has a borrowing limit of $\underline{B} = -20$. From the left hand panel of figure 14 we see that, not-surprisingly, the value of the firm's equity is reduced and is virtually the same as the value of a firm that does not have the option to borrow both at low levels of capital stock and at sufficiently high levels. Where borrowing makes a difference is for a firm that has little or no initial capital and must borrow to get off the ground, and for a firm with intermediate levels of capital stock $k \in [60, 200]$ where the ability to borrow helps it accelerate the accumulation of capital necessary to reach the higher steady state capital stock $k^* = 222$.

Notice, that by lowering the amount the firm can borrow, the incentive for strategic default has been eliminated: it is no longer optimal for the firm with little initial capital to borrow up to the maximum and then go default and go bankrupt, unlike the case shown in figure 15 when the firm had a much larger borrowing limit, $\underline{B} = -90$. Of course if the firm has sufficiently small capital stock and is initially encumbered with too much debt, it will immediately go bankrupt, similar to what we illustrated above for a firm with a concave production function. However when the borrowing limit is set "correctly" the ability to borrow a moderate amount eliminate the temptation for strategic default.

For example, the right hand panel of figure 14 shows the trajectory of a firm with initial capital $k = 1$ and debt of $b = -18$. The dynamics are surprisingly complex in a model that has no stochastic shocks. This firm borrows $B(k,b) = -19.80$, nearly up to its limit, $\underline{B} = -20$ in the first period to finance an investment of $I(k,b) = 2.8$. Then it rolls over this debt, reducing it very slowly over the next 6 periods. Then in period 8 it suddenly reverses course and borrows up to its limit $\underline{B}$ to finance a spike in investment of $I(k,b) = 3.53$. Then it reduces debt gradually for 3 more periods before again borrowing up to its limit to finance another investment spike of $I(k,b) = 4.09$ in period 12, and finally there is another borrowing/investment spike at period 14. From period 15 onwards the firm steadily pays off its debt until it is debt-free in period 25. Then via a final spike in investment (now financed by retained earnings) in periods 26 and 27, the firm reaches the lower steady state capital stock level $k^* = 26.25$. The firm remains debt free and pays a dividend of 3.81 per period in perpetuity.

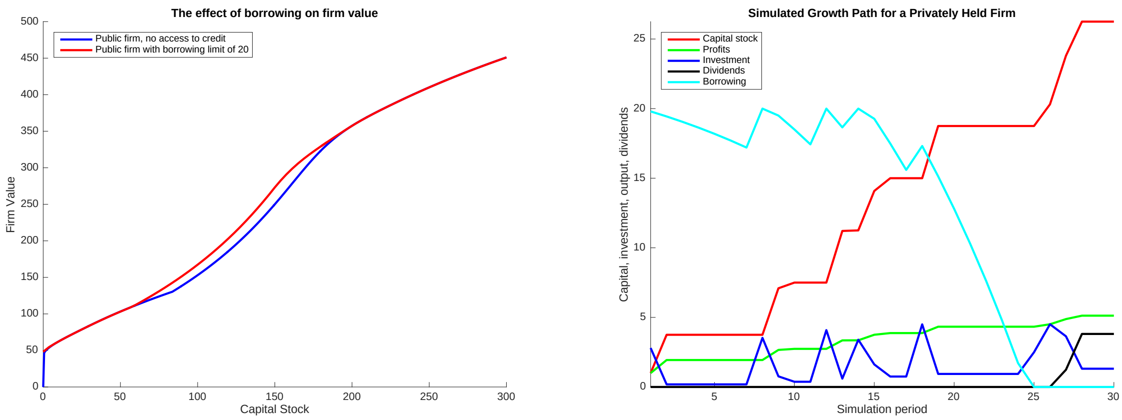


Figure 14: Gains from borrowing for a public firm with non-concave $f(k)$ and $\underline{B} = -20$

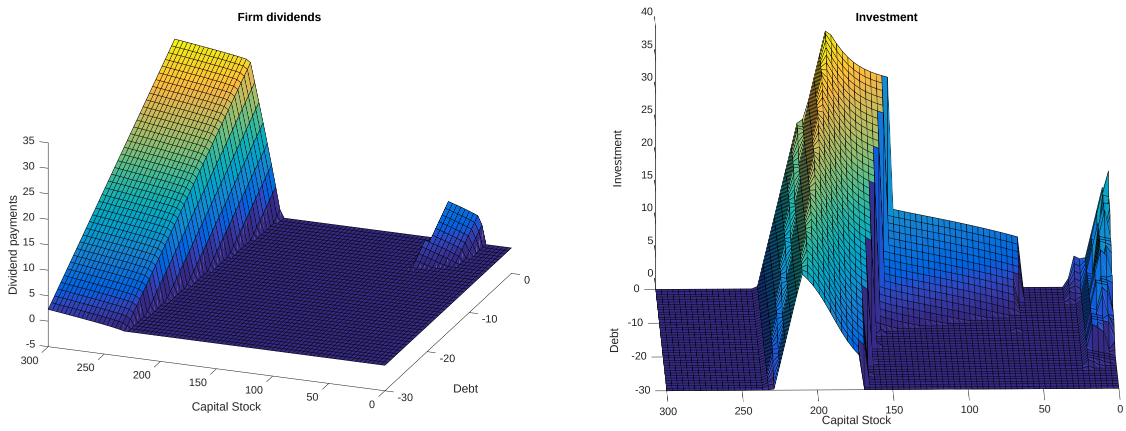


Figure 15: Optimal dividend and investment policy for a firm with non-concave $f(k)$ and $\underline{B} = -20$

The lower borrowing limit of $\underline{B} = -20$ eliminates the strategic default motive, and still generates in a monotonic increase in the value of the firm, the increase in value is significant mainly at $k = 0$ (where a firm with no ability to borrow has zero value, as shown in Theorem 1) and also for $k \in [60, 190]$, which is the region where the ability to borrow helps accelerate the firm's investment trajectory to enable it to reach the higher steady state capital stock $k^* = 222$ faster than it could if it had to finance investment solely via retained earnings. At the same time, a borrowing limit of $\underline{B} = -20$ is insufficient to eliminate the "growth trap" that causes sufficiently small firms to converge to the lower steady state capital stock. This is why the ability to borrow does not increase the equity value of the firm for $k \in (0, 55]$. We can see this in the peak in the investment policy function $I(k,b)$ for small values of k and b sufficiently near 0 in the right hand panel of figure 15. This represents investments made to reach the lower steady state capital stock $k^* = 24$ in the growth trap region of the state space. For larger values of k that are greater than the upper threshold $\overline{k}$, optimal investment falls to zero and the firm allows its capital to depreciate back to the lower steady state

k^*. It is only when capital is sufficiently large, above the threshold $\underline{k} = 60$ where we see a discontinuous jump in investment. For these larger values of initial capital k, the firm has sufficient incentive to make the necessary investments necesary to reach the higher steady state capital stock $k^* = 222$.

We can see from the left hand panel of figure 15 that the firm pays no dividends in this intermediate region of capital where it is investing to reach the higher steady state capital stock $k^* = 222$. Interestingly, the firm does not immediately use debt finance to help accelerate its growth to the higher k^*. It is not until it has accumulated enough capital via retained earnings that it is optimal for the firm to undertake borrowing, when it enters the "convex region" of its production technology that provides the higher marginal returns to capital that can help justify and pay for the interest expense on its borrowing. Then as this debt-financed peak in investment pushes the firm into the concave region of its production function, it reduces its borrowing and rapidly pays it off to become debt free just before a final wave of investments financed by retained earnings enables it to reach the higher steady state capital stock $k^* = 222$, where it remains in a debt-free state in perpetuity.

As you can see, the nature of the optimal investment and financial policy has become quite complex already, even in a relatively simple deterministic model. The main insights to take away are 1) excessive borrowing limits (or excessive initial debt) create a strong temptation for strategic default by small firms, and may explain why small firms face binding liquidity constraints, 2) even relatively generous borrowing limits do not eliminate the "growth trap" that causes small firms to prefer to invest at a small scale to reach the less ambitious lower steady state capital stock k^* rather than incur the delay and expense of using a combination of retained earnings and debt finance to reach the higher steady state capital stock where the firm is much larger and more profitable. We shall see that this leads to a key motive for public firms to issue more shares (or undertake an IPO): to generate the financing necessary to grow and reach the higher scale and steady state capital stock by avoiding the costs and delay associated with growth financed exclusively by retained earnings and debt in the face of binding liquidity constraints.

3 Optimal investment for a private firm

In this section we consider the problem faced by a privately owned firm to contrast how the optimal investment and dividend policy of a private firm differs from that of a public firm. Initially we will consider a firm that does not have the option to borrow and we will ignore the option of this firm to go public via an IPO. We will then relax these restrictions in the following sections in order to fully develop our simple theory of why and when firms go public.

Consider an individual who owns the production technology and who has private wealth w that they can invest (partially or fully) in their own firm. The individual has utility function $u(c)$ satisfying $u'(c) > 0$ and $u''(c) < 0$. The suppose the market interest rate is $r_m = r$ but the individual's personal interest rate is r_p and thus the individual discounts future utility using discount factor $\beta = 1/(1+r_p)$. If the person purchased an annuity with their initial endowment of wealth w they would receive discounted lifetime utility of $u(rw)/(1-\beta)$. Now suppose instead the person invests their wealth to buy an equivalent amount of capital $w = k$ and from each period onward the owner manages the firm to obtain dividends which he/she consumes. What is the optimal investment and dividend policy for this "privately held firm"?

Suppose that $w > k^*$ where k^* is the optimal steady capital stock of the publicly held firm given in equation (4) above. Is it optimal for the owner of the private firm to invest this amount too? Assume that after making an initial capital investment k, the private owner restricts attention to "steady state" investment policies $I(k) = (1-\delta)k$ that will maintain the capital stock of the firm at the initially invested value k forever. What is the optimal value of k that the owner would choose?

This is given by the solution k_p^* to

$$k_p^* = \underset{0 \le k \le w}{argmin} \frac{u(f(k) - \delta k - r(w-k))}{1-\beta}. \tag{68}$$

The first order condition for the optimal steady state policy is

$$f'(k_p^*) = \frac{1}{\beta} - 1 + \delta \tag{69}$$

so we see in fact that $k^* = k_p^*$: the owner of a private firm would invest to the same steady state capital stock value that a publicly held firm would choose if it were to make an initial investment and be able to borrow the funds necessary at the same interest rate as the discount rate the market uses to value the firm (i.e. to discount its future dividend stream).

However assume that $w < k^*$. For the moment, let's conjecture that the owner would choose to invest all of his/her wealth in the firm, so they will receive no annuity income after sinking all of their initial wealth as an investment in their firm. Assuming the owner cannot borrow, the Bellman equation for the privately held firm is given by

$$V(k) = \max_{0 \le I \le f(k)} \left[u(f(k) - I) + \beta V(k(1-\delta) + I) \right]. \tag{70}$$

The first order condition for optimal investment is given by

$$u'(f(k) - I(k)) = \beta V'(k(1-\delta) + I(k)). \tag{71}$$

If we were to assume an "Inada condition" i.e. that $\lim_{c \downarrow 0} u'(c) = +\infty$, then it is easy to see that the optimal investment policy will always entail paying some positive level of dividends, i.e. $I(k) < f(k)$ for all k. However it may still be the case that if the firm had sufficient capital, it may be optimal not to invest, i.e. $I(k) = 0$ for $k \ge \overline{k}$, though the value of $\overline{k}$ may be different than the value $\overline{k} = k^*/(1-\delta)$ at which a public firm stops investing.

Using the Envelope theorem, we have

$$V'(k) = u'(f(k) - I(k))f'(k) + \beta V'((1-\beta)k + I(k))(1-\delta), \tag{72}$$

but using the first order condition (71) we have

$$V'(k) = u'(f(k) - I(k))[f'(k) + (1-\delta)], \tag{73}$$

and substituting this back into the first order condition (71) we can derive the Euler equation characterizing the private investor's optimal investment policy $I(k)$

$$u'(f(k) - I(k)) = \beta u'\left(f(k(1-\delta) + I(k)) - I(k(1-\delta) + I(k))\right)\left[f'(k(1-\delta) + I(k)) + (1-\delta)\right]. \tag{74}$$

This is a non-linear functional equation for I and it is ordinarily not an easy one to solve via numerical methods. It is not clear there there is a closed form solution in this case, unlike the one we found for the optimal investment policy of a publicly held firm.

However we can show there is a unique steady state solution k_p^* to the Euler equation, and that $k_p^* = k^*$, the same steady state solution for a public firm. Note that any steady state, we have $I(k) = \delta k$ and substituting this for $I(k)$ in the Euler equation above we obtain

$$u'(f(k) - \delta k) = \beta u'\left(f(k) - I(k)\right)\left[f'(k) + (1-\delta)\right], \tag{75}$$

or $f'(k) = 1/\beta - 1 + \delta$, for which the only solution is $k = k^*$. This suggests that even if the private investor does not have sufficient initial wealth to invest in the firm at the optimal level k^*, the subsequent investment policy will lead the firm to gradually accumulate capital and converge to the optimal steady state asymptotically.

Figure 16 plots the optimal investment and dividend policy functions for a privately held firm and compares them to the ones chosen by a publicly held firm. The solutions for the privately held firm were calculated numerically using the discrete policy iteration algorithm described in the appendix. We see that both are quite different from each other. The top left panel shows the optimal dividend policies plus the level of replacement investment necessary to keep the capital stock from declining. The intersection of the optimal investment curves and the black replacement investment line defines the optimal steady state capital stock level k^* and as predicted by our analysis above, we see that it is the same for both the public and privately held firm.

Away from the steady state, investment and dividends are quite different from each other. Investment by the privately held firm is less than investment by the public firm for $k \in (0, k^*]$, but investment by the privately held firm is greater than investment by the public firm for $k > k^*$. The pattern for dividends is the opposite: the private firm pays higher dividends than the public firm for $k \in (0, k^*]$, but lower dividends

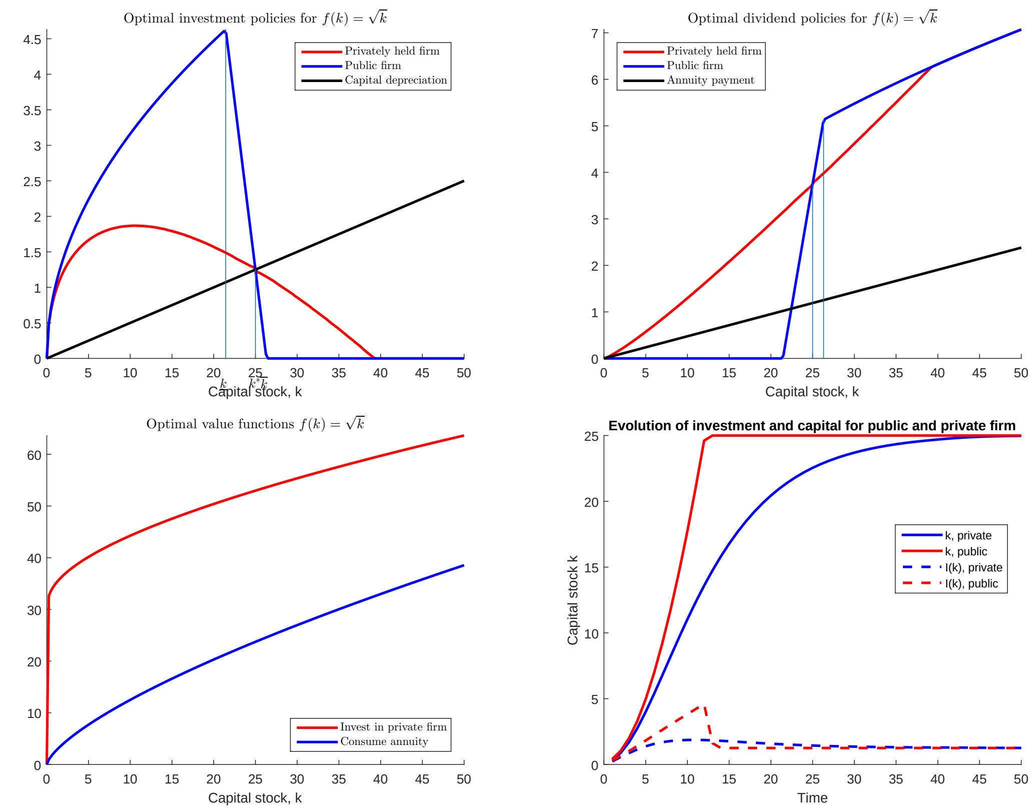


Figure 16: Optimal investment and dividend policy and value function for $f(k) = \sqrt{k}$

for $k > k^*$, unless capital is sufficiently high that both the public and private firm stop investing, and in this region the dividend payments coincide.

The lower left panel of figure 16 plots the value of the privately held firm $V(k)$ and compares it to the utility the investor would have obtained if they invested all of their wealth in an annuity earning the market rate of return. We see that at least if investment is framed as an all or nothing choice, it is always preferable for the investor to invest their wealth in the private firm rather than in an annuity. The private firm generates sufficiently greater returns to dominate the return of $r = .05$ that the person could obtain from an annuity. Another way to see this is to look at the black line in the right hand top panel of figure 16. This plots the annuity income the investor would receive each period if they invested all of their wealth into an annuity. We see that the dividend income from investing in a private firm dominates the annuity income they would receive at all levels of initial investment k.

Finally, the lower right hand panel of figure 16 compares the evolution of investment and capital stock for a public and a private firm that each begin life with an initial capital stock of $k = 1$. We see that due to the higher early investment, the public firm reaches the steady state capital stock $k^* = 25$ after only 15 periods, whereas the privately held firm approaches k^* only asymptotically.

We summarize our analysis of the private firm in Theorem 7 below.

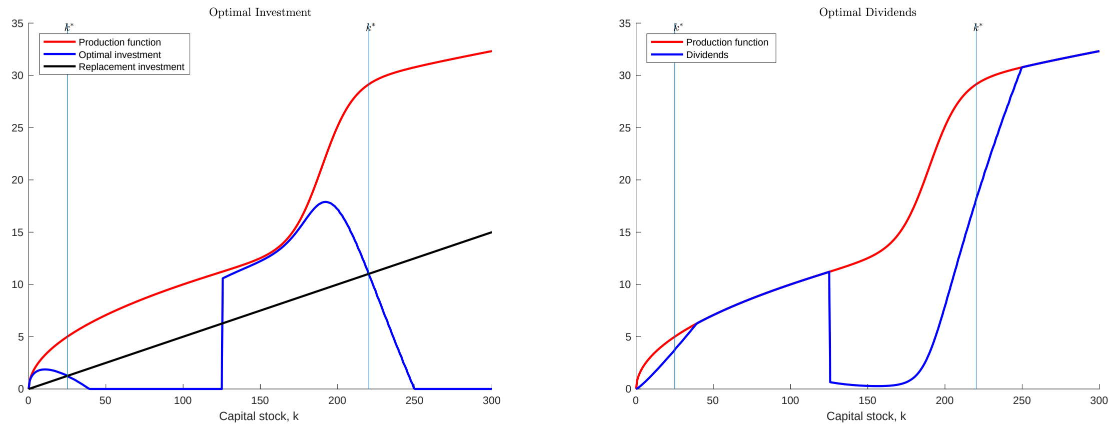


Figure 17: More complex optimal investment/dividend policy for non-concave $f(k)$

Theorem 7 (behavior of a private firm without access to capital markets) *Consider a privately owned firm where the owner has a concave utility function $u(c)$ and production function $f(k)$ where u satisfies an Inada condition. Assume the firm cannot borrow and does not have the option to do an IPO in the future. Then the optimal dividend policy is to pay positive dividends for any positive level of the capital stock. The privately owned firm adopts an inefficient investment policy (i.e. a policy that does not maximize its market valuation) due to the owner's desire to use dividends to consumption-smooth, that results in slower rate of accumulation of capital. However the steady state capital stock for a privately owned firm, k^*, is the same as a publicly owned firm, and both pay the same level of dividends in steady state, $f(k^*) - \delta k^*$.*

We conclude this section by illustrating the non-concave case, where the owner has a concave utility function but the production technology $f(k)$ is non-concave for the reasons explained in section 3. Figure 17 illustrates the optimal investment and dividend policy for the owner of a private firm who faces the same non-concave production technology that we solved for a public firm in section 3 (see figure 5). Similar to the concave case analyzed above, the owner of a private firm distorts his/her investment in order to generate positive dividends in all circumstances. We also see that the private firm has the same two Golden Rule steady state values of k^* that we computed for the public firm, but the domains of attraction to these steady states are different. The owner of a private firm would need substantially more initial capital, more than $k = 125$ to be willing to invest and grow to reach the higher steady state value of k at $k^* = 222$, whereas a public firm only needs k in excess of 80 to be willing to invest and reach this higher Golden Rule k^*.

However we can see from the right hand panel of figure 17 that the private firm's owner is willing to get by on very little dividend income in order to rapidly invest and reach the higher steady state k^* and thereafter enjoy a permanent dividend stream of $D(k^*) = 18.12$ after incurring "depreciation expenses" of $\delta k^* = 11$. However the main message is that similar to the case of a public firm, there is a potential for borrowing constraints to create a "liquidity trap" that causes the private firm to forgo attractive investment opportunities and remain a small firm. Though the private owner is behaving optimally given his/her financial constraints, the behavior is in a sense suboptimal relative to a world where the firm could borrow to reach the higher steady state k^* faster, rather than incurring the delay of having to slowly build up enough capital through investment of retained earnings. In the next section we will study how borrowing can help the firm avoid this globally suboptimal outcome.

3.1 How the Option to Borrow Affects the Behavior of a Private Firm

Now suppose the owner of a private firm has the option to borrow at an interest rate $r_b > 0$. We continue to rule out the option to "go public" in the future via an IPO and focus on this section on characterizing how the borrowing option affects the owner's behavior. As we noted in the previous section, a private firm's desire to use dividend policy to consumption-smooth results in a slower capital accumulation path, and this is the main source of inefficiency in its behavior relative to a public form.

Let $V(k,b)$ denote the discounted utility of a private firm with capital stock k and single period debt b and subject to borrowing constraint $\underline{B}$. The Bellman equation for the owner is given by

$$V(k,b) = \max_{\substack{I \geq 0 \\ \underline{B} \leq b' \leq f(k) - I + b}} \left[u(f(k) - I + b - b') + \beta V(k(1-\delta) + I, b'(1+r_b)) \right] \tag{76}$$

We consider the case where $r_b = r$ first. Recall our analysis of a public firm in this case. From Theorem 3, we know that if there are no borrowing constraints then the Modigliani-Miller Theorem holds for a public firm, and the firm's dividend and borrowing policy are not well defined. The firm will want to borrow enough in the first period to reach the optimal steady state capital stock k^* but the firm is indifferent about whether or not it should subsequently pay off this initial debt, or borrow more to pay dividends. However the Modigliani Miller Theorem no longer holds for a private firm: Theorem 8 below shows that a private firm facing no borrowing constraints will borrow enough to invest in the steady state capital stock k^* in the first period, but the owner will also borrow more than this in order to pay dividends, and the owner will maintain an optimal level of debt b^* *permanently* even though this reduces the level of dividends the owner will receive in steady state because by maintaining this debt, the owner is able to *smooth his consumption stream.* The owner is willing to pay interest on perpetually rolled over debt to achieve this flat consumption profile.

Theorem 8 (Modigliani Miller Theorem fails for a privately held firm) *Consider the owner of a privately held firm who faces no borrowing constraints and who can borrow at interest rate r_b which is the same as the owner's subjective discount rate. Then the optimal policy will be to borrow and amount equal to b^* given by*

$$b^* = \beta[f(k^*) + k^*(1-\delta)] \tag{77}$$

and continually rolls over this initial debt in all future periods. The owner invests enough to reach the *optimal steady state capital stock k^* in the first period and pays a dividend equal to $f(k^*) - (\delta + r)k^*$ every period. Thus the owner is not indifferent about the payment of dividends and financial policy matters, since the owner strictly prefers to carry an optimal level of debt b^* even though this reduces the dividends the owner will receive in steady state due to the payment of interest on the debt.*

To prove Theorem 8 consider, without loss of generality, a firm that starts with zero initial capital and no initial debt. We consider which initial level of investment and borrowing such a firm would want to undertake to maximize the discounted utility of the owner. We can show that because the owner has diminishing marginal utility, the optimal policy will involve a time-invariant or steady-state configuration of dividend and interest payments. Thus if b is the level of debt the owner will choose at $t = 0$ and k is the desired capital stock, the owner's problem is given by

$$V(0,0) = \max_{\substack{k \geq 0 \\ b \geq k}} \left[u(b-k) + \beta \frac{u(f(k) - \delta k - rb)}{1-\beta} \right] \tag{78}$$

It is easy to show that the optimal initial investment (and thus the optimal steady state capital stock) is k^* that satisfies $f'(k^*) = r + \delta$, and the optimal level of borrowing is the value b^* given in equation (77). Further it is not hard to see that

$$V(k^*, b^*) = \frac{u(f(k^*) - \delta k^* - rb^*)}{1-\beta} \tag{79}$$

and these expressions satisfy the Bellman equation (76), and thus the claimed investment, dividend and debt policy is in fact the optimal policy for this privately held firm. Further we have $b^* - k^* = f(k^*) - \delta k^* - rb^*$, so that via this optimal borrowing and investment in period $t = 0$ the owner achieves a perfectly flat consumption profile, but at the expense of paying the interest cost rb^* in every period in perpetuity.

Corollary 8.1 *The option to borrow at interest rate $r_b = r$ strictly increases the welfare of the owner of a private firm. In particular, for an owner with no initial capital stock, $k = 0$, we have*

$$V(0) = \frac{u(0)}{1-\beta} < V(0,0) = \frac{u(f(k^*) - \delta k^* - rb^*)}{1-\beta} \tag{80}$$

where $V(0)$ is the discounted utility of an owner with no initial capital who does not have the option to borrow, whereas $V(0,0)$ is the the discounted utility of an owner with no initial capital and no initial debt but who does have the option to borrow.

When the owner does face binding borrowing constraints, the owner may not be able to borrow enough in the first period to reach the optimal steady state capital stock k^*. Instead, the owner will have to borrow and gradually invest, while also making a tradeoff of how much of the borrowed funds to use for investment and how much to use to pay dividends. We can show that the owner will generally not borrow right up to the borrowing limit $\underline{B}$ when their initial capital is sufficiently low because borrowing the maximum allowed could subject the owner to the possibility of insolvency.

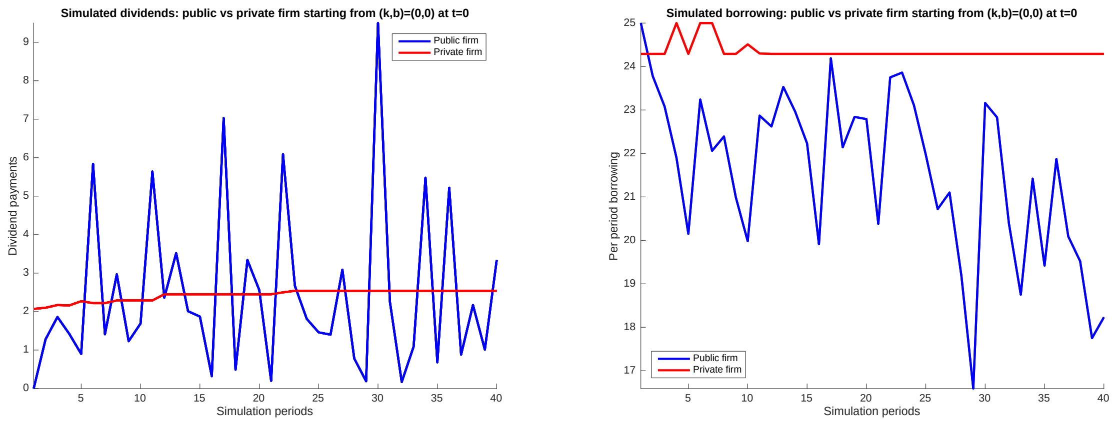


Figure 18: Simulated dividend and borrowing for a public firm (blue lines) and private firm (red lines)

However similar to the case of a public firm, we can demonstrate that an owner of a private firm may engage in strategic default if their utility function does not satisfy an Inada condition (so that $u(0) > -\infty$). For example a firm whose owner has utility function $u(c) = c^{.7}$ and who has $r = r_b = .05$ with $\underline{B} = -30$ will find it optimal to borrow and go bankrupt starting from initial condition $(k,b) = (0,-28)$. Here, as in the case of the public firm, the initial debt is too large and the owner obtains higher discounted utility from borrowing $b'(k,b) = -29.39$ in the first period, paying off the initial debt of $b = -28$ and investing $I(k,b) = 1.39$ and paying no dividends in the first period, and then in period $t = 2$ the owner borrows up to the limit $\underline{B} = -30$, invests nothing, and pays himself dividends of 0.32, and then goes bankrupt the following period. The owner calculates the alternative strategy of trying to slowly acquire capital via borrowing and investing over a long period of time to reach the optimal steady state capital stock k^* is lower than the discounted utility of the strategic bankruptcy decision, which is $V(0,-28) = \beta u(.32)$.

Figure 18 illustrates the strong incentives for "dividend smoothing" by the owner of a private firm, contrasting it with the essentially random dividend and borrowing policy of a public firm, a consequence of the Modigliani Miller Theorem and small random errors in the numerical solution of the public firm's problem that generate the seemingly random fluctations in borrowing and dividend payments. Since the firm is formally indifferent between borrowing more to pay dividends or reducing dividends to retire debt, the numerical methods we used to solve and simulate the public firm's behavior will be affected by small rounding and numerical approximation errors in what would otherwise be (if all numerical calculations were 100% exact) a completely flat objective function over its choice of borrowing and dividends.

Now consider the optimal behavior of a firm where $r_b < r$. We have already seen that in the case of a public firm the Modigliani Miller Theorem fails and the optimal policy for the firm is to borrow as much as it can (subject to its borrowing constraint) except for the caveats we noted for firms with low initial capital that are already encumbered by debt. In these cases the firm will generally borrow less than the

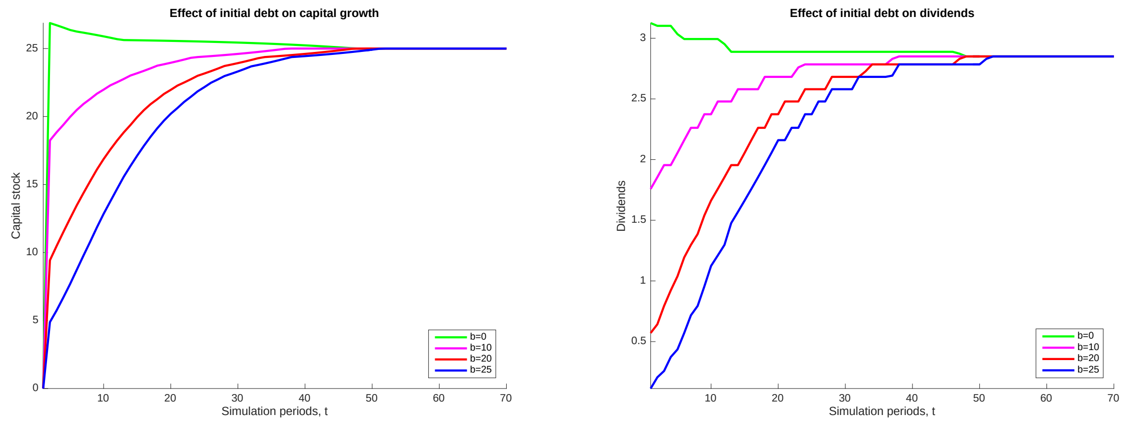


Figure 19: Effect of debt on capital accumulation and dividends

maximum allowed to avoid bankruptcy, though there may be cases where the firm also engages in strategic bankruptcy as well.

Theorem 9 (optimal borrowing for privately held firm when $r_b < r$) *Consider the owner of a privately held firm who faces a borrowing constraint $\underline{B}$ and who can borrow at interest rate $r_b < r$ where r is the owner's subjective interest rate for discounting future utility. The optimal policy is to borrow the maximum amount allowed if the capital stock is not too large, and provided the firm's initial debt b is not too large that the firm is insolvent in the first period. If the borrowing constraint $\underline{B}$ is sufficiently large to allow a firm with no initial capital and no debt to reach the optimal steady state capital stock k^* in the first period, it can be optimal for an owner with residual borrowing capacity to invest more than k^* in period 1 and pay dividends. The globally stable steady state is $(k^*, \underline{B})$, so the firm maintains the maximum allowed debt load in perpetuity and pays a dividend equal to $f(k^*) - \delta k^* - r_b \underline{B}$.*

Figure 19 displays simulated trajectories for capital stock and dividends for a privately owned firm with $r = 0.05$ and $r_b = 0.03$ starting from initial conditions $(k,b) = (0,b)$ where $b \in \{0, -10, -20, -25\}$. The green line in the left panel shows the trajectory for the firm's capital stock when it starts out with no capital and no initial debt. The owner borrows up to $\underline{B} = -30$ and chooses to to invest *more than* $k^* = 25$ in period $t = 1$ and uses the remainder of the debt proceeds to pay dividends to himself. This excessive initial investment enables the owner to pay higher than normal dividends over the next 40 periods as the ower lets the initial investment gradually depreciate to the optimal steady state capital stock $k^* = 25$.

The other curves show trajectories for initial conditions where the owner is encumbered by initial debt. We see that this initial debt limits the amount of investment the owner can undertake, as well as the amount of dividends he can pay himself. So in these cases the capital stock and dividend trajectories approach the steady state values, $k^* = 25$ and $f(k^*) - \delta k^* - r_b \underline{B} = 2.85$ from below. The higher the initial level of debt, the longer it takes for the firm to reach the optimal steady state values of capital and dividends.

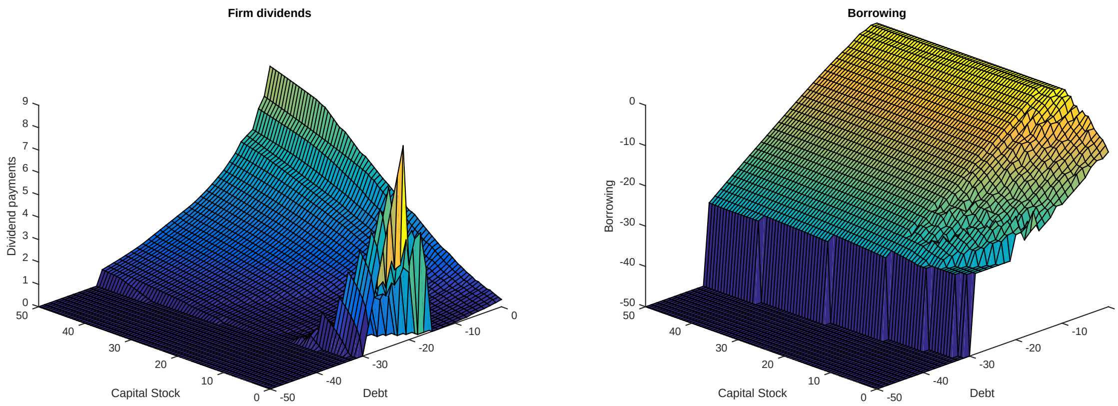


Figure 20: Optimal dividends and borrowing for a private firm with $r = 0.05$ and $r_b = 0.10$ with $\underline{B} = -30$

It is interesting to compare the evolution of capital and dividends for a private firm illustrate in figure 19 to the paths for capital and dividends for a public firm in the same circumstances (i.e. $r = 0.05$, $r_b = 0.03$ and $\underline{B} = -30$) illustrated in figure 9 above. Both the private firm and the public firm converge to the same steady state values of capital $k^* = 25$ and dividends $f(k^*) - \delta k^* - r_b\underline{B} = 2.85$, but the public firm converges to the steady state much more rapidly: in fewer than 20 periods compared with more than 50 periods for the private firm. The public firm achieves the more rapid growth in its capital stock by not paying dividends when its capital is sufficiently low. Notice also the public firm never undertakes any "excess capital investment" such as the private owner undertakes in the the initial state $(k, b) = (0, 0)$. The slower rate of capital accumulation constitutes an efficiency cost of the private owner's desire to dividend smooth, and would result in a lower valuation for the company if the owner were to take his company public and continue to follow this investment and dividend policy.

Finally we consider the case where $r_b > r$. As we already showed for a public firm, it is optimal for the firm to pay off its debt, and a public firm will not hold debt in steady state. We now show that this is also the case for a private firm as well, provided the interest rate on debt is sufficiently high.

Theorem 10 (optimal borrowing for privately held firm when $r_b > r$) *Consider the owner of a privately held firm who faces a borrowing constraint $\underline{B}$ and who can borrow at interest rate $r_b > r$ where r is the owner's subjective interest rate for discounting future utility. There are two possible steady state outcomes for the firm. One involves either immediate bankruptcy if the firm's initial debt b is too high given k and its borrowing constraint $\underline{B}$, or strategic default in which the firm may make small initial investments but borrow mostly to pay dividends for a few periods before going into default, and the other steady state outcome is $(k^*, 0)$ where the owner's investments enable the firm to reach the optimal steady state capital stock and the firm retires any initial debt used to finance initial dividends and investment, and carries no debt in steady state.*

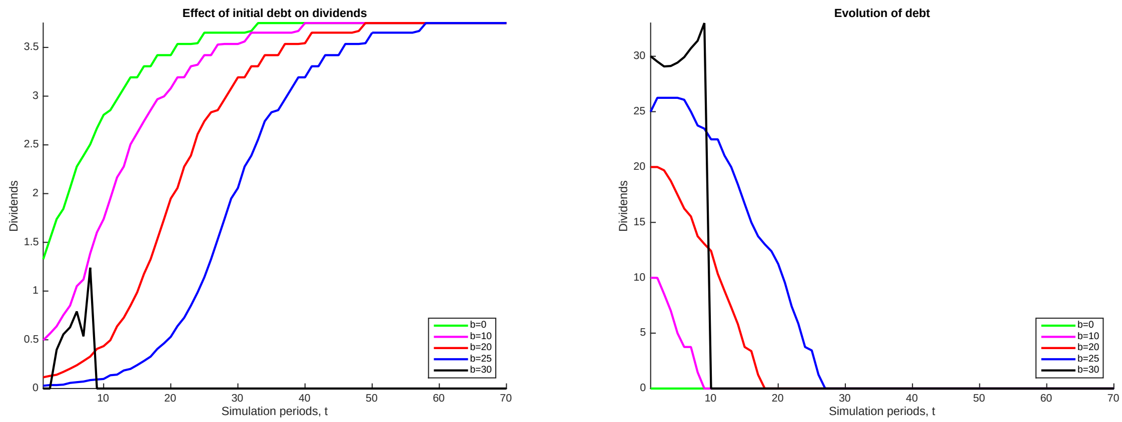


Figure 21: Debt and dividends for a private firm with initial $k = 10$, $\underline{B} = -30$ and varying levels of b

Figure 20 illustrates the conclusions of Theorem 10 by displaying the optimal dividend and borrowing policy of a privately owned firm with an owner who has a utility function $u(c) = c^{.7}$ and discount factor $\beta = 1/(1+r)$ with $r = 0.05$ and can borrow at rate $r_b = .10$ up to a maximum of $\underline{B} = -30$. The left hand panel shows the dividend policy, and the secondary peak in dividends around $k = 0$ and b between -20 and -30 represents the "strategic default zone" where the owner concludes it is not optimal to do any "serious" investment, and mainly borrows to pay dividends for several periods before going bankrupt. For example an owner who starts at $(k,b) = (0,-20)$ will borrow $b'(k,b) = -23.96$ at $t = 1$ and invest $I(k,b) = 0.62$, and pay himself dividends of 3.33. Then in period $t = 2$ the owner borrows $b'(k,b) = -30 = \underline{B}$ and pays himself a dividend of 4.44, but given the small capital stock, the owner is unable to pay off this loan in period $t = 3$ and so the firm goes bankrupt. The right panel illustrates the borrowing policy for this owner. We see that the owner generally borrows less than the maximum amount allowed, $\underline{=} - 30$, and this implies the dynamic summarized in Theorem 10 where the owner will pay off the initial borrowing and hold no debt in steady state.

We conclude this section with figure 21 that illustrated simulated paths for dividends and debt for an owner of a firm who starts with initial capital stock $k = 10$ and faces an interest rate of $r_b = .10$ and borrowing limit of $\underline{B} = -30$ for varying levels of initial debt b. The green curve shows the dividends and debt paths for an owner with no initial debt, $b = 0$. The owner decides not to do any borrowing and thus remains debt-free for its entire life. Not surprisingly, this owner is able to pay the highest dividends. The other curves plot cases where owners are encumbered by initial debt and are unable to increase their dividends as rapidly. If the the owner's initial debt b is not too large (e.g. $b = -10$ or $b = -20$) the owner pays off the debt steadily but balances the desire to pay off debt rapidly with the desire to consumption-smooth (and thus pay dividends in the interim). If the owner's initial debt is somewhat larger, $b = -25$, then the owner does some incremental additional borrowing but does not borrow up to the maximum allowed, $\underline{B} = -30$ since this would result in default.

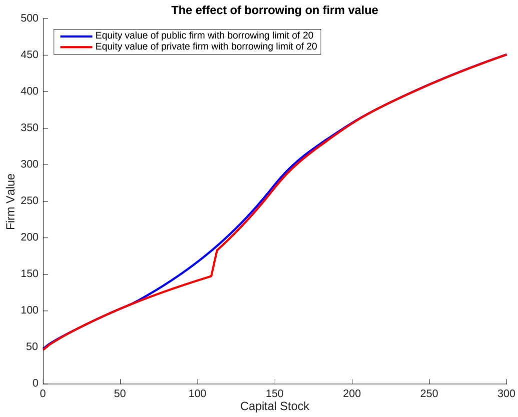
However an owner who has even more initial debt, $b = -30$ concludes that trying to pay off this initial debt and keeping the firm as a "going concern" is not worthwhile. The owner slyly borrows somewhat below the maximum amount allowed in the first three periods so debt actually initially decreases slightly. During this time the owner uses the proceeds to pay dividends to himself but does not do any investment and lets his capital stock depreciate. Then starting in period $t = 4$ the owner starts to increase his borrowing until reaching the borrowing limit $\underline{B} = -30$ in period $t = 8$. By period $t = 9$ the principal and interest are too large to repay given the dwindling cash flow stream from the depreciated capital stock, and the firm goes bankrupt.

3.2 Debt Policy for a Private Firm with a Non-Concave Production Function

Now consider the case of a private entrepreneur who has the same utility function as above, $u(c) = c^{.7}$, but who faces the non-concave production function $f(k)$ illustrated in figure 5 that we introduced in section 2.1. Figures 22 and 23 illustrate how a private owner's desire to consumption smooth distorts the financing and investment policy and reduces the growth rate of the firm.

The right hand panel of figure 22 provides a simulation of a private firm with the same initial conditions that we used to simulate the trajectory of a public firm in figure 14: $k = 1$ and $b = -18$. The consumption-smoothing motivation is immediately evident: the private firm pays dividends in *every period* even though dividends in the first five periods are very small, $D(k,b) = .02$. In contrast, the public firm pays *no dividends* until period 27, after both borrowing and devoting all retained earnings to rapidly building up its capital stock. The willingness to delay dividend payments enables the public firm to reach the steady state capital stock of $k^* = 26.25$ by period 28 whereas the private firm does not reach it until period 62. By period 28, the private firm has only accumulated a capital stock of $k = 18.42$ and pays a dividend of only $D(k,b) = 1.21$ whereas the public firm is able to pay the steady state dividend of $D(k,b) = 3.81$.

In addition to investing more slowly, the private firm relies on debt to "dividend smooth". The private firm is in debt for the first 45 periods of the simulation whereas the public firm is in debt for only the first 25 periods. Thus, even though the private firm is able to pay a small level of dividends immediately, the public firm achieves a higher present value of dividends by accelerating its investments and paying larger dividends with a delay. The gain in value achieved by the public firm is illustrated in the left hand panel of figure 22. It compares the present value of the dividend policy of the public and private firms, both of which face an identical borrowing limit of $\underline{B} = -20$ and borrowing rate of $r_b = 0.07$. The public firm (blue line) dominates the present value of dividends of the private firm, but the difference is relatively small when capital stocks are sufficiently low ($k \leq 60$) or high ($k \geq 120$). The reason the loss in value is largest in the interval $k \in (60, 120)$ is that the private firm is *less patient* than the public firm, even though both



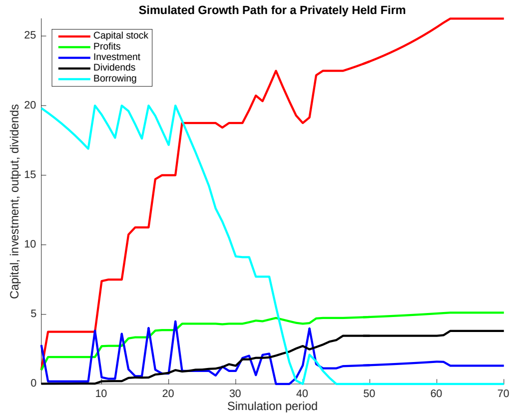


Figure 22: Gains from borrowing for a private firm with non-concave $f(k)$ and $\underline{B} = -20$

firms discount at the same rate $r = 0.05$.

For example, if we simulate both the public and private firm starting at an initial capital stock of $k = 60$, the public firm invests rapidly again to reach the *higher* steady state capital stock of $k^* = 222$ by period 27. It practices "abstinence" by not paying any dividends until period 26, just before it reaches the larger steady state capital stock k^*, and starting in period 28 onwards it pays a dividend of $D(222,0) = 18.18$. The private firm, on the other hand, does not see the benefit in an aggressive investment policy and the need to sacrifice dividend payments (current consumption) in order to attain this higher steady state capital stock. So it simply "coasts" and does not invest at all in the first 9 periods, and in periods 10 to 21 its new investments are too small to counteract depreciation, so it consciously allows its capital stock to depreciate to enable the owner to eat the dividends. By period 22 the capital stock has reached the *lower* steady state value of $k^* = 26.25$, and due to the unwillingness to make the bigger investmenst that the public firm made, the cost to the private owner is a much lower dividend of only $D(26.25,0) = 3.81$ in perpetuity.

Otherwise there is a qualitative similarity between the optimal investment and dividend policies of the public and private firms, which you can see visually by comparing figure 23 (for the private firm) with figure 15 (for the public firm). Both types of firms are willing to invest (rather than strategically default) when capital stocks are very low, but both are willing to discontinuously increase their level of investment if their initial capital stock k is sufficiently high. The main difference is that the zone of capital stocks where investment is zero is larger for the private firm than the public firm, reflecting the greater impatience and unwillingness to make significant investments to reach the higher steady state level of k^* that we discussed above. The other key difference is that the private firm pays positive dividends in *all states* whereas the public firm pays zero dividends in nearly all states except those where the capital stock $k \geq \underline{k}$ when investment switches from being monotonically increasing in k to monotonically decreasing in k.

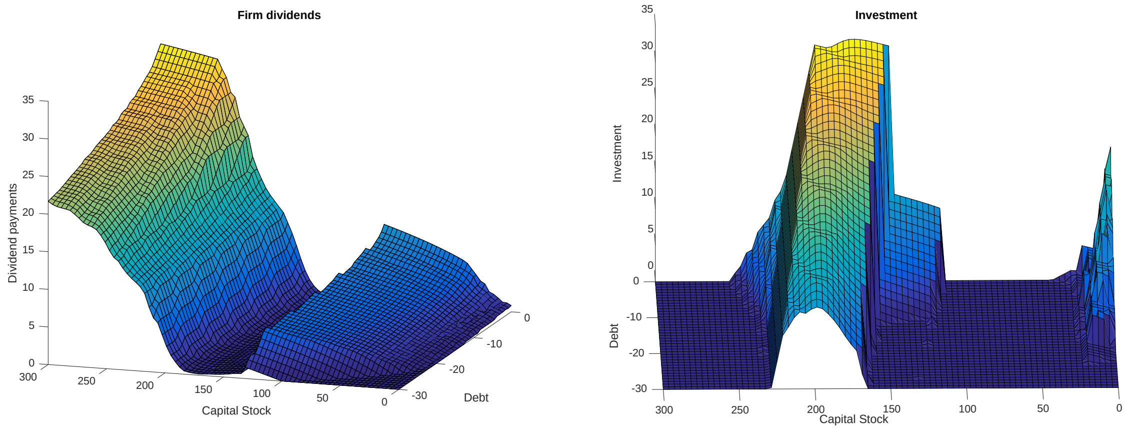


Figure 23: Optimal dividend and investment policy for a private firm with non-concave $f(k)$ and $\underline{B} = -20$

Thus, we have illustrated the sense in which the desire to consumption-smooth can lead to an inefficient investment and financial policy for a private firm relative to the value-maximizing strategy followed by a public firm. Notice that these inefficiencies appear quite clearly even though the utility function we chose, $u(c) = c^{.7}$, is rather close to the linear utility function $u(c) = c$ that corresponds to risk neutrality and results in the same optimal investment and financial policy as a public firm. Allowing for more curvature in the utility function will enhance the incentive to consumption smooth and thus result in greater inefficiencies. Figure 24 illustrates the increased gain in value from going public for an owner of a private firm whose utility function is more concave (i.e. a utility function $u(c) = c^{.3}$ vs the utility function $u(c) = c^{.7}$ we used previously in this section). The gains to going public are largest at an intermediate capital stock of approximate $k = 132$: the market valuation of a public firm is over 40% higher than the market valuation of the stream of dividends produced by the private firm, distorted by the private owner's desire to consumption-smooth.

Adding uncertainty to the model creates an additional avenue for inefficient operation of a private firm: the risk aversion of the owner can lead him/her to fail to undertake investments that they deem "too risky" relative to those that would be deemed to be attractive investments by a risk-neutral manager of a public firm. This suggests that a *separation theorem* might be possible that makes the owner strictly better off by doing an IPO: By selling out his/her firm and taking it public, the owner gains by using financial markets to consumption-smooth, while making it possible for a manager to run the company in a more efficient, value-maximizing manner which results in more wealth for the owner which can be more efficiently smoothed using financial instruments such as annuities rather than using dividend, borrowing, and investment policy as a second-best way to smooth consumption.

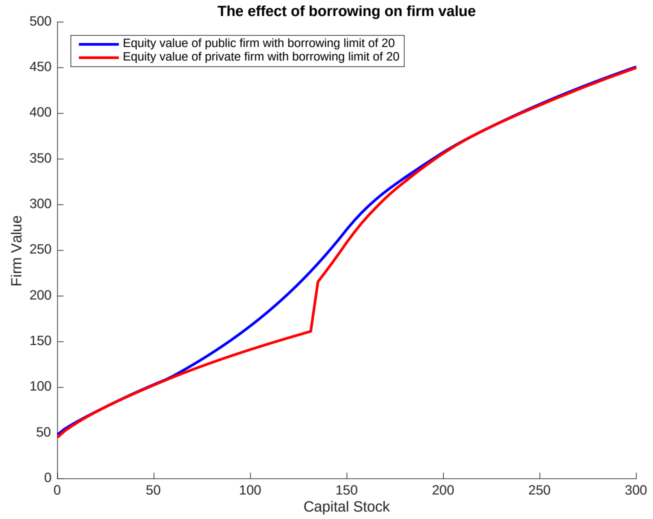


Figure 24: Increase gains to going public for an owner private firm with a more concave utility function

4 Modeling the IPO Decision

In this section we extend our model of a privately owned firm to give the owner the option of "taking the firm public" via an IPO (initial public offering). After the IPO the firm would be be run by a professional manager whose objective is to maximize the present value of dividends. By selling off a 100% stake in the firm at the IPO, the owner no longer has any operating control over the firm, but the owner can take the proceeds raised by the IPO and buy an annuity and live happily ever after on this annuity income. What will the owner decide to do: sell their firm in an IPO, or keep the firm private?

We start by considering the simplest case first, where neither the private firm nor the public firm has the option to borrow. Thus there is no motivation for the private owner to go public in order for the newly created public firm to obtain access to credit markets that the private owner did not have access to. Still, provided the costs of undertaking an IPO are not too high, the private owner owner will be better off by selling out rather than continuing to operate his firm privately.

Theorem 11 (Conditions for Going Public) *Consider a firm that is initially privately held by an owner whose personal subjective discount rate r_p is the same as the market interest rate r_m at which dividends of publicly traded firms are discounted at. If there is an actuarially fair annuity market where perpetual stream of annuity payments can be purchased at an initial price that equals the present value of these annuity payments, also discounted at the same rate $r = r_p = r_m$, then if there are no costs to doing an IPO, the owner would always prefer to sell out the private firm and use the proceeds to purchase an annuity, except if $k = 0$ or $k = k^*$, where the owner is indifferent about selling out or not. If there is a proportional fee $\rho \in (0,1)$ for doing the IPO, (so the owner only receives the share $(1-\rho)$ of the IPO proceeds), then* *only private firms with sufficiently low levels of initial capital k will find it optimal to do an IPO. If there are also up-front fixed costs F required to do an IPO, then if* ρ *and F are not too high, there will be an interval of values of the private owner's capital stock,* (k_l, k_u) *where it is optimal for the firm to go public via an IPO. An owner with capital less than* k_l *will not have sufficient size to afford the fixed costs of undertaking an IPO, and an owner will capital greater than* k_u *can afford to do the IPO but finds the transactions costs too high to make it worthwhile.*

In the case of no transactions costs to doing an IPO, the proof of Theorem 11 is intuitively clear: we showed that the private owner has a motive to consumption smooth, but this motive distorts the owner's investment policy, since his desire to pay dividends in every period slows the rate of accumulation of capital to the optimal steady state value k^*. By going public, the newly public firm will avoid these inefficiencies and the owner can continue to consumption-smooth by using the IPO proceeds to purchase an annuity. This result is in effect, a type of *separation theorem* between investment and consumption, and shows that in the absence of transactions costs, firms ought to be publicly held rather than privately held, since it is more efficient to use capital markets than investment policy to smooth out consumption streams. Note that if $k = 0$, then since neither the public or private firm can borrow, and therefore the owner cannot gain from doing an IPO. If $k = k^*$, then the firm is already at the optimal steady state capital stock, and since this results in a flat consumption stream of $f(k^*) - \delta k^*$ per period (the same as what an annuity would pay the owner if he sold out), there is no gain from doing an IPO in this case either.

Let $V_{pub}(k)$ be the *market value* of the firm (which does not have access to borrowing, similar to the privately held firm) given by the solution to the Bellman equation (2) in section 2. This represents the funds the owner would raise if this firm were to be sold in an IPO. The owner can then use these IPO proceeds to purchase an annuity equal to $(1-\beta)V_{pub}(k)$. Thus, the discounted utility to the owner from holding an IPO is given by

$$V_{ipo}(k) = \frac{u((1-\beta)V_{pub}(k))}{(1-\beta)}. \tag{81}$$

Notice that $V_{ipo}(k)$ is measured in utils, not in dollars. The owner compares this utility value to the utility value of keeping his firm private and operating it to maximize their lifetime discounted utility. Call this discounted utility value $V_{pri}(k)$: it is the solution to the private owner's Bellman equation (70). Then clearly, the owner will choose to take his firm public if and only if

$$V_{ipo}(k) > V_{pri}(k). \tag{82}$$

Figure 25 below plots these two value functions, V_{pub} and V_{pri} as well as the value of simply using their initial wealth to buy an annuity, $u((1-\beta)w)/(1-\beta)$. Though it is slightly hard to see, the value of doing an IPO uniformly dominates the value of running the firm as a private company. The reason is that the owner of a private firm, while undertaking a *privately optimal* dividend and investment policy, is nevertheless

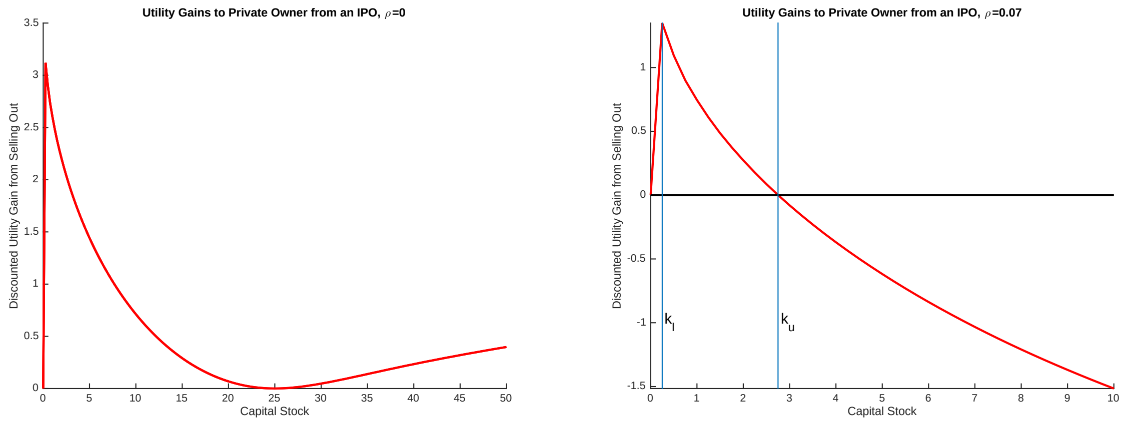


Figure 25: Effect of transactions costs ρ on the value of going public

adopting a *suboptimal policy* from the standpoint of maximizing the market value of the company. The distorted dividend and investment policies that we illustrated above, plus the slower trajectory of capital accumlation due to a private owner's incentive to pay dividends in every period are costly in terms of lowering the present value of what the owner could consume if he/she sold the market to a professional manager whose objective is to maximize the market value of the firm. In essence, it is better for the owner to use the *annuity market* to smooth their consumption, than to attempt to do this on their own by distorting their investment and dividend policy. By doing an IPO, the owner allows the new management to adopt value maximizing investment and dividend policies and the owner is free to take these proceeds and smooth their consumption stream in the annuity market. This is another example of what is known as a *separation theorem* in the finance literature.

The left hand panel of figure 25 illustrates the gains to going public for a private owner with utility function $u(c) = c^{.7}$ where $r_m = r_p = 0.05$ and the proportional transactions costs for doing an IPO is $\rho = 0$. We see that consistent with Theorem 11, the gains from an IPO are positive for all values of k except $k = 0$ and $k = k^* = 25$ where they are zero. The right hand panel of figure 25 illustrates the gains to going public when $\rho = 0.07$ and the fixed fee $F = 0.5$. The 7 percent commission rate is typical for IPOs underwritten by investment banks in the US and other countries. We see with the higher commission rate, the gains from doing an IPO are negative for $k > k_u = 2.75$, and also an owner who has no access to credit markets is unable to afford the fixed cost of the IPO when $k < k_l = .25$ since the cash flow $f(k) = \sqrt{k}$ is insufficient to enable the owner to pay F in this case. So the interval $(k_l, k_u) = (.25, 2.75)$ represents the interval of capital where it is optimal for the owner to take his firm public with an IPO and then sell out.

We see that in the presence of transactions costs, it is no longer better to go public regardless of the initial capital stock of the firm. It is only optimal for firms that are in a "sweet spot" with a minimal amount of capital to be able to afford the fixed costs of an IPO but not too much capital where the commission constitutes an unacceptably large tax on the owner. In addition, as the private firm's capital gets close to k^*, the incentive to do an IPO falls as we showed in Theorem 11, since the firm is close to a point where it can reach k^* quickly anyway using its own retained earnings. Thus, we have succeeded already in providing an answer to both "why" and "when" private firms go public.

Now we consider an additional motivation for a firm to go public: access to credit markets. It is often claimed that due to agency reasons and information costs, public firms have superior access to credit markets than private firms. We will start by considering the most extreme case where the owner has no ability to borrow as a private firm but if the firm were public, it would face no borrowing constraints. Clearly, when this is the case the incentive to go public is enhanced, since we have already shown how the ability to borrow increases the value of a public company.

Corollary 11.1 (Access to credit markets enhances the value of going public) *Suppose the assumptions of Theorem 11 hold except that public firms have unlimited access to credit markets and face no borrowing constraints at an interest rate $r_b = r_m = r_p$. Then the incentives for going public are higher, and when there are both fixed and proportional transactions costs to undertaking an IPO, the interval of capital $[k_l, k_u]$ where it is optimal for the owner to do an IPO is larger than in the case where public firms have no access to capital markets.*

Figure 26 illustrates Corollary 11.1 by plotting the gains to going public, $v_{ipo}(k) - v_{priv}(k)$, in the case where the post-IPO public firm has full access to credit markets and can borrow unlimited amounts at $r_b = 0.05$, the same as the owner's personal interest rate and the rate at which the public firm's dividends are discounted. The biggest gain to going public now occurs at $k = 0$: an owner with no initial capital cannot borrow as a private firm and thus cannot get his firm off the ground. But once the firm goes public it obtains access to credit markets, which enables the firm to borrow enough to attain the optimal steady state capital stock k^* with only a single period delay. Thus, the owner of a private firm is able to unlock tremendous value by the act of going public, and in effect, gain access to credit in an indirect manner that would not be possible for him to do if the firm remainded private.

We can see from figure 26 that due to the increased value that access to credit markets confers on the post-IPO public firm, the interval of capital (k_l, k_u) for which going public is optimal widens from $(.25, 2.75)$ to $(.25, 5)$. We can see that the gains to doing an IPO decline very rapidly as k increases, and private firms with the least capital are the ones that gain the most from going public. Underwriters of IPOs might take note of this and waive any up-front fixed fees for doing an IPO for the smallest private firms in exchange for a larger commission ρ. For example if the underwriter waived the fixed fee, $F = 0$ and raised the commission from $\rho = 0.07$ to $\rho = 0.08$, it would capture new business from the smallest private firms and the incremental commission revenue (since the valuation of a new public firm is $V_{pub}(0) = 50$), makes up for lost fixed feea However fewer medium-sized private firms would do IPOs, since k_u falls from

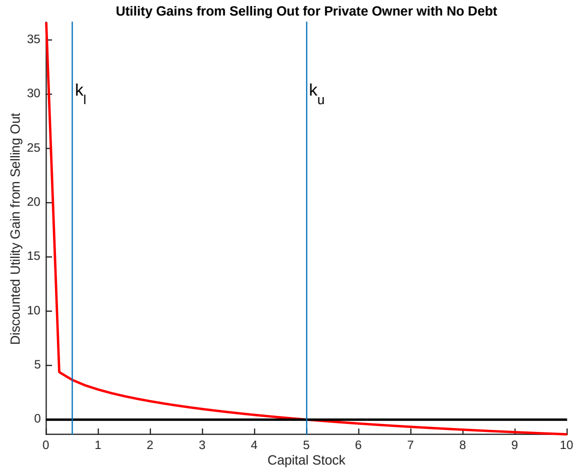


Figure 26: Gains to going public with $\rho = 0.07$ when post-IPO public firm has access to credit markets

5 when $\rho = .07$ to 4.25 when $\rho = 0.08$.

Finally, suppose both private publics firms have equal (unlimited) access to credit markets at an interest rate $r_b = r_m = r_p$. Then the gain to doing an IPO drops precipitously, and unless the transactions costs of doing an IPO are very close to zero, private firms will not choose to go public. This is illustrated in figure 27 which plots the discounted utility gains to having an IPO and selling out for an owner who can borrow up to a limit of $\underline{B} = -30$ for varying levels of initial debt b.

The left hand panel of the figure plots the gains for the case where $b = 0$, and we can see that the utility gains are very small, and largest for larger values of k above the steady state capital stock. This is due to the inefficiency that exists for an owner of a private firm that we discussed in the previous sectioin: *if the firm were somehow endowed with initial capital stock k that is greater than the optimal steady state value k^*, the owner's desire to consumption can cause the owner to undertake investment to slow the depreciation in the capital stock back to k^*, whereas a public firm never does investment when $k > k^*$ so that capital depreciates back to k^* as rapidly as possible.* By undertaking an IPO the private owner can avoid this investment inefficiency caused by his own desire to consumption smooth.

However for values of $k < k^*$, the private owner's access to borrowing results in rapid convergence of the capital stock to k^*, though not always immediately in period $t = 1$ (even though the firm can borrow enough to achieve k^* in period $t = 1$). Thus there are small inefficiencies in the private owner's investment policy created by the desire to consumption smooth, and these small inefficiencies translate into the small gains we calculate for going public in the left panel of figure 27.

The right panel of figure 27 shows how initial debt increases the private owner's incentive to undertake an IPO, but by in large the effect is not big. These gains were calculated under the assumption that $\rho = 0$, and if there are even very small costs to undertaking an IPO (say $\rho = 0.01$) then the gains to an IPO are

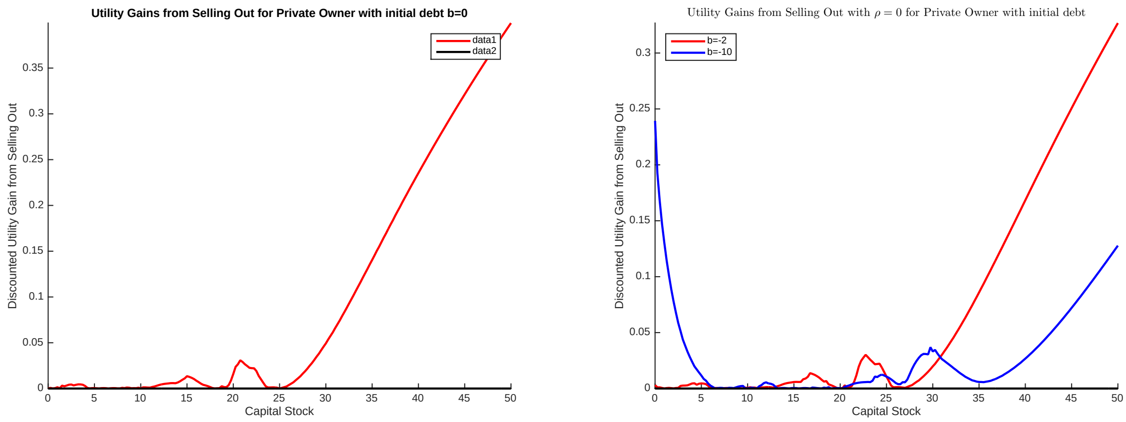


Figure 27: Effect of private firm access to credit on the gains to going public

negative for all values of k. The reason initial debt creates larger gains to doing an IPO is clear: *the initial debt serves as an encumbrance that slows the owner's ability to reach the optimal steady state capital stock* k^*, *and combined with the additional inefficiencies due to the owner's desire to consumption-smooth, this creates large investment inefficiencies that translate into large gains from doing an IPO.* We summarize this discussion as

Corollary 11.2 (Access to credit markets by a private held firm reduces the value of going public) *If the owner of the privately held firm has access to credit markets, the incentives for undertaking an IPO are dramatically reduced, and unless the transactions costs associated with undertaking an IPO are near zero, the private owner will have no incentive to go public. The incentives for going public are stronger if: a) the borrowing constraints the private owner faces are tighter, and b) the owner of the private firm is encumbered with initial debt.*

5 IPOs with partial cash-outs

Most IPOs do not entail a 100% sell-off of the original owner's stake in the company. Intead, the original owner retains a partial ownership stake in the firm, and only takes part of the IPO proceeds in cash to finance consumption or other investment projects. The other important role of a partial cash-out is that when the original owner continues to own a significant share of the post-IPO company, the share of the IPO proceeds that the owner does not "cash out" are re-invested in the company, thereby providing a new infusion of capital to the firm after the IPO that is not reflected in our analysis of an IPO with a 100% cash out by the original owner. Thus, an IPO can have two effects that can boost the value of the firm: 1) the IPO can switch the objective function of the firm from utility maximization of the private owner to one of value maximization in the market (an effect we describe as a "moral hazard effect"), and 2) to the extent that the original owner reinvests some of the proceeds of the IPO back into the company, it represents a new source of capital to the firm (an effect we describe as the "financing" or "leverage" effect of the IPO).

Suppose the firm is originally a privately owned firm by a sole owner, and the owner chooses to take the firm public via an IPO and retain only a fraction $\alpha \in (0,1)$ of his/her original 100% ownership stake in the firm. Thus, after the IPO the original owner will own a fraction α of the firm (i.e. α is fraction of shareholdings still owned by the founder of the firm) and the outside investors who bought shares in the new firm will own the remaining fraction $1-\alpha$ of the firm's shares.

The "IPO proceeds" equal the total amount the founder receives from selling shares in the newly public firm to the new "outside investors" and the founder can either reinvest these funds to increase the capital stock (and hence future profit/dividend stream of the firm), or take some or all of the proceeds as a "cash out" for private consumption purposes (e.g. to buy an annuity). Or the founder might want to reinvest some of the IPO proceeds in other nascent investment projects such as to found some other new firm. We will let the symbol $\omega \in [0,1]$ represent the fraction of the IPO proceeds that the owner chooses to take out for consumption or other investment purposes, and thus the fraction $1-\omega$ is reinvested in the firm.

Let $P(k,\alpha,\omega)$ represent the IPO proceeds received by a founder/owner of a private firm who decides to take the firm public when it has initial capital k, and the owner chooses to retain an ownership share α after the IPO, and to "cash out" a fraction ω of the IPO proceeds and reinvest the remaining fraction $1-\omega$. We assume that the fractions α and ω are publicly observable, as a newly public firm must meet various accounting standards that are designed to protect outside investors from fraud such as "take the money and run" schemes that are patent ripoffs of unsuspecting investors. It is one function of intermediaries such as investment banks to do the *due diligence* to investigate a private firm that wishes to go public with an IPO and verify that the company really does exist and the founder will not "take the money and run" after an IPO. Thus, the reputation of the investment bank intermediary, in addition to market regulation (such as is done by government agencies such as the Securities and Exchange Commission) helps to convince outsider investors that an IPO is legitimate and is not a thinly disguised take the money and run scheme.

We assume that an investment bank intermediary incurs costs of doing the due diligence and insuring that a private firm that wants to go public via an IPO is legitimate. The investment bank recovers the costs of providing these services by charging a proportional fee $\rho \in (0,1)$ plus, possibly, a fixed fee F. Thus if the gross proceeds of the IPO are $P(k,\alpha,\omega)$, the net proceeds received by the founder from the investment bank (after it deducts its fees) are $(1-\rho)P(k,\alpha,\omega)-F$. Initially we will study the IPO in a "frictionless market" setting where the costs of doing due diligence are zero, and hence we initially assume that $\rho = F = 0$. In this case, the gross and net proceeds of the IPO coincide.

In a market where k, α and ω are public information, and where the operations of a public firm are sufficiently regulated by both government regulators and the discipline of market competition, the public will also have a rational expectation that the newly public firm operates to maximize the discounted stream of dividend payments to its shareholders. In this case we can write an equation for the new proceeds of the IPO as

$$P(k,\alpha,\omega) = (1-\alpha)V(k+(1-\omega)[P(k,\alpha,\omega)(1-\rho)-F]). \tag{83}$$

where $V(k)$ is the value of a public company with capital stock k as defined in the Bellman equation (2) in section 2 above, which is the value of the company after an IPO with $\alpha = 1$. In equation (2) we assume that the net proceeds $P(k,\alpha,\omega)(1-\rho)-F \geq 0$, otherwise it is not clear that the founder would see any benefit to doing the IPO. Further we assume that $k > (1-\omega)F$. This implies that the function $V(P) = (1-\alpha)V(k,(1-\omega)[P-F])$ satisfies $V(0) > 0$, and together with the strict concavity of V implies that there is a unique solution $P(k,\alpha,\omega)$ to equation (83). Furthermore, it is easy to see from the strict concavity, that at this solution we have $1 > (1-\alpha)(1-\omega)(1-\rho)V'(k+(1-\omega)[P(k,\alpha,\omega)(1-\rho)-F])$. This implies that, in equilibrium, if an additional dollar were raised in the IPO, the amount of this extra dollar, net of IPO costs and the fraction of proceeds taken out by the founder, will raise the market value of the fraction of the shares held by the outside investors, $1-\alpha$, by less than 1 dollar.

Equation (83) tells us that the IPO proceeds will equal the value of the outside shareholders' share of the firm *after* the original founder has reinvested the fraction $1-\omega$ of the net proceeds $P(k,\alpha,\omega)(1-\rho)-F$ received from the investment bank as new capital for the newly public firm. The IPO proceeds is implicitly defined as the solution to equation (83) above. Due to the strict concavity of $V(k)$, there is a unique solution to (83) for each $k \geq 0$ and each $\alpha \in (0,1)$, and $\omega \in [0,1]$. The Implicit Function Theorem guarantees that $P(k,\alpha,\omega)$ is continuously differentiable in its arguments k, α and ω for almost all values of k, α and ω with derivatives

$$\begin{aligned}
\frac{\partial}{\partial k}P(k,\alpha,\omega) &= \frac{(1-\alpha)V'(k+(1-\omega)[P(k,\alpha,\omega)(1-\rho)-F])}{1-(1-\alpha)(1-\omega)(1-\rho)V'(k+(1-\omega)[P(k,\alpha,\omega)(1-\rho)-F])} \\
\frac{\partial}{\partial \alpha}P(k,\alpha,\omega) &= \frac{-V'(k+(1-\omega)[P(k,\alpha,\omega)(1-\rho)-F])}{1-(1-\alpha)(1-\omega)(1-\rho)V'(k+(1-\omega)[P(k,\alpha,\omega)(1-\rho)-F])} \\
\frac{\partial}{\partial \omega}P(k,\alpha,\omega) &= \frac{-V'(k+(1-\omega)[P(k,\alpha,\omega)(1-\rho)-F])}{1-(1-\alpha)(1-\omega)(1-\rho)V'(k+(1-\omega)[P(k,\alpha,\omega)(1-\rho)-F])}
\end{aligned} \tag{84}$$

It follows that $P(k,\alpha,\omega)$ is increasing in k and decreasing in α and ω, as we would naturally expect.

We can view $P(k,\alpha,\omega)/(1-\alpha)$ as the market's rational expectation of the total value of the firm following an IPO where it has full knowledge of the fraction of the firm owned by the founder after the IPO, and the fraction of the IPO proceeds that the founder cashed out for consumption or other purposes, leaving only the fraction $1-\omega$ of the net proceeds as the amount of new investment the firm actually undertakes as a result of the IPO. It also is contingent on the assumption that after the IPO the firm will be run in a discounted profit maximizing manner, even if the owner retains a majority stake in the company after the IPO. Thus, our theory of rational market valuation following an IPO and partial cash out encompasses both the moral hazard and financing/leverage effects that an IPO can have on the valuation of a company that we discussed above.

Figure 28 illustrates how an IPO can be used as leverage, substantially increasing a firm's value by reinvesting a fraction of the IPO proceeds to acquire more capital, which further increases the value of the firm. We focus on a small private firm that has an initial capital stock of $k = 3$ when it decides to go public via an IPO. We plot the value of the firm as a fraction of α, the fraction of the firm that the founder chooses to own after the IPO. We assume that $\omega = 0$, so that the owner does not divert any of the IPO proceeds to any other purposes except reinvestment in the firm. The left panel of figure 28 plots the total value of the firm, the amount reinvested, and the value of the share of the firm owned by the founder as a function of α.

Notice that the value of the founder's interest in the firm is zero when $\alpha = 0$. Clearly it would make no sense for the founder to sell off his/her entire ownership interest and then reinvest all proceeds back into a firm he/she no longer owns: this would be a nice gift to the new shareholders but not something that the founder would want to do absent a peculiar sense of altruism to outside investors. The case where $\alpha = 1$ corresponds to a situation where the founder decides to take the company public but without raising any new capital from outside investors. There is no new investment resulting from the IPO in this case, and the value of the firm is equal to the value we already calculated in section 4 under 100% sell off option, namely \$52.47.

The black line in the left hand panel of figure 28 shows that amount of the IPO that is reinvested in the firm as a function of α. If the owner was to be so nice to sell off his/her entire ownership interest to outside invstors and reinvest the entire IPO proceeds in the firm, the firm would attain its maximum value of \$217.58, which equals the amount of new capital the original owner reinvests in the firm. However if the owner were to retain 50% ownership, the value of the firm is \$110.41, which is double the amount the original owner reinvests in the firm when $\alpha = .5$. Thus, the owner obtains a 100% return from doing an IPO and reinvesting half of their ownership stake in the firm, even though the outside investors will also benefit from this investment made by the founder. The founder's net worth after this deal is \$55.21, which exceeds the founder's net worth from the option of selling off his entire ownership stake and not reinvesting any of the IPO proceeds back in the company, \$52.47. Thus, some degree of apparent "altruism" towards the outside shareholders by the founder is actually in the founder's self-interest.

If we consider which value of α maximizes the founder's net worth after the IPO (assuming ω is fixed at 0), we find that $\alpha = .64$ and the founder's net worth (i.e. the value of his/her ownership stake in the post IPO firm) is \$57.47. The founder invests \$32.08, and the total value of the firm is \$89.13 after this investment. Thus, the return on this investment is equal to $(57.47 - 32.08)/32.08 = .7915$. This represents a very high return even though the founder is not able to capture all of the benefit from this investment: the outside investors reap 36% of the increase in the firm value resulting from the founder's reinvestment of the \$32.08 in IPO proceeds back into the capital stock of the firm.

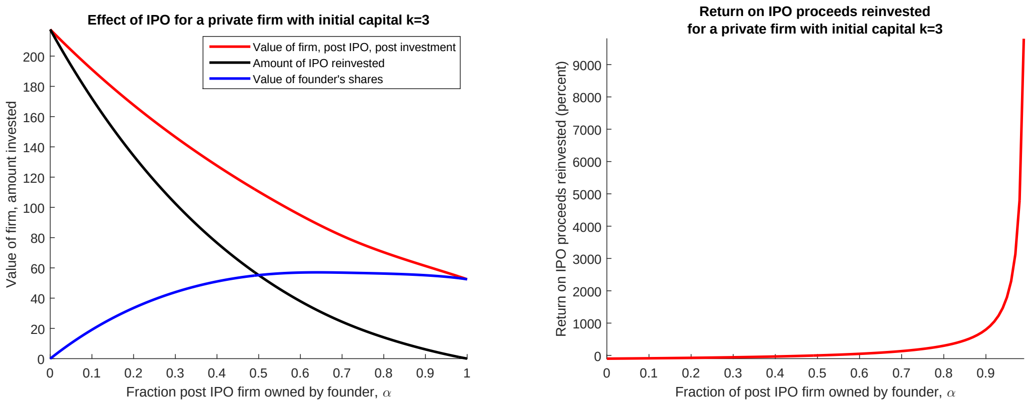


Figure 28: Value of a firm after IPO, and return to investment from IPO proceeds

The right hand panel of figure 28 plots the rate of return on the marginal dollar the owner reinvests in the firm, as a function of α. The first dollar reinvested has an exceptionally high rate of return in this example. Naturally there are diminishing returns to investment and so the return falls as α (the fraction the owner cashes out) decreases quick to zero as α tends to zero. However even when $\alpha = 0.64$, the founder obtains a 79% return on their investment as we noted above.

To complete the model, we now discuss the founder's choice of α and ω. It is simplest to consider the case where the only motive for a cash out is to buy an annuity to smooth consumption. If the owner retains ownership of a fraction α of the company following the IPO, the owner could initially invest this fraction of the IPO proceeds in the company (to benefit from the effective leverage or financing effect of the IPO) and then immediately sell off this residual stake after the IPO and the investment in new capital is completed. Then the owner could purchase an annuity with the total proceeds. Under this formulation of the owner's problem we have that the optimal values of α^* and ω^* solves

$$(\alpha^*(k), \omega^*(k)) = \underset{\alpha\in[0,1]}{argmax}\, \underset{\omega\in[0,1]}{argmax}\, u((1-\beta)(\omega + \alpha/(1-\alpha)P(k,\alpha,\omega))/(1-\beta). \tag{85}$$

Notice that the optimal fraction to cash out depends on the size of owner's initial capital stock k when the firm is privately held, just prior to doing the IPO. Since u is monotonically increasing, the founder's problem reduces to simply maximizing the value of his/her net worth following the IPO, where the net worth is a combination of the cash taken out of the IPO proceeds, $\omega P(k,\alpha,\omega)$, plus the value of the founder's sharedholdings in the post-IPO company, $\alpha V(k + (1-\omega)[P(k,\alpha,\omega)(1-\rho) - F]) = \alpha P(k,\alpha,\omega)/(1-\alpha)$. Thus, the founder's problem reduces to

$$(\alpha^*(k), \omega^*(k)) = \underset{\alpha\in[0,1]}{argmax}\, \underset{\omega\in[0,1]}{argmax} (\omega + \alpha/(1-\alpha))P(k,\alpha,\omega) \tag{86}$$

Figure 29 plots the net worth of the founder, $(\omega + \alpha/(1-\alpha))P(k,\alpha,\omega)$ as a function of (α,ω). It turns out that this function is symmetric as a function of (α,ω) about the diagonal line $\alpha = \omega$. As a result we find

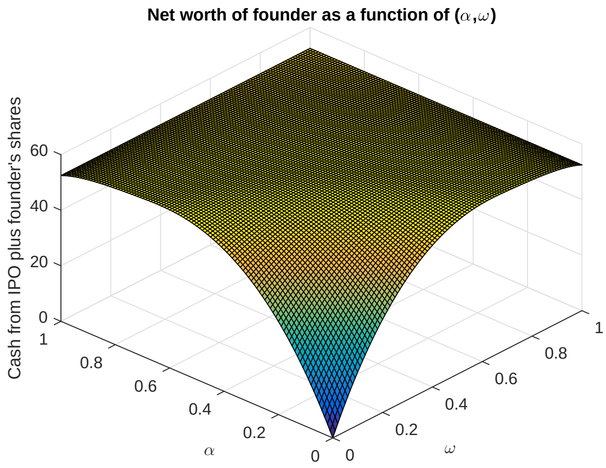


Figure 29: Net worth of founder, as a function of (α, ω)

two symmetrically located optimal solutions, $(\alpha^*(k), \omega^*(k)) = (.42, .38)$ and $(\alpha^*(k), \omega^*(k)) = (.38, .48)$, and both yield the optimal level of net worth for the founder equal to \$57.04. We see that when we fix α, if α is sufficiently small, the founder's net worth is initially increasing in ω and then decreasing, so there is an optimal value of the cash out fraction $\omega^*(k, \alpha)$ for any fixed α. By symmetry, there is also an optimal value of the fraction of ownership $\alpha^*(k, \omega)$ that the founder should retain for any fixed cash out fraction ω provided ω is not too close to 1.

However if we fix a value for α that is sufficiently large, say $\alpha > .7$, then the net worth of the founder is monotonically decreasing in ω and thus the optimal value $\omega^*(k, \alpha) = 0$ when α is sufficiently large. That is, the founder does not want to cash out if he/she decides to retain a sufficiently large stake in the firm: the return to reinvesting in the firm is higher. By symmetry this is also true for α when ω is fixed as a value that is sufficiently large: the optimal ownership stake is zero $\alpha^*(k, \omega) = 0$. Thus, if the founder precommits to cashing out a sufficiently large share of the IPO proceeds, the founder will also will not find it optimal to retain any ownership interest in the firm.

The optimal combination $(\alpha^*(k), \omega^*(k))$ represents the tradeoff between the founder's desire to reinest in the firm, but tempered by the disincentive effect of the fact that the larger the amount the founder sells to outside investors, the less the founder benefits from reinvesting the IPO proceeds back into the firm.

6 Conclusions

This paper introduced a simple theory of why and when firms go public. These are our main results:

Why go public? Because risk aversion and the desire of a private owner to consumption smooth causes him/her to distort their investment and financial policy, resulting in slower growth and less ambitious operation of the company. Though the behavior of the owner is *privately optimal* it is sub-optimal with respect to the criterion of maximization of market value of the firm. Provided the fixed and variable transaction costs of undertaking an IPO are not to high, the owner can gain by taking the firm public so it can be run by a risk-neutral manager in a way that maximizes the present value of dividends, and thus the equity valuation of the company.

When to go public? We showed that especially for firms with a non-concave production function (which we argued is relevant for firms that undergo multiple stages of growth as they evolve from a small scale firm with little capital to a large scale firm with a large capital stock that can exploit significant global returns to scale from expansion), the optimal time to go public is when the firm is at the "Goldilocks size" — not too small and not too big. Firms that are too small and which face borrowing constraints that prevent them from investing to quickly reach the efficient scale of operations will not experience large returns from going public, and will not be able to afford the significant fixed and proportional transactions costs of an IPO. Firms that have already acquired a large capital stock can afford to do an IPO but have already have nearly the level of capital they need to to achieve the efficient scale of operation, and so can easily use a combination of debt financing and retained earnings to acquire the remaining capital necessary to achieve their optimal scale.

Our theory is only a beginning. There are many ways to extend our simple model to make it more realistic such as adding uncertainty and learning as we discussed in section 2.2. Our eventual goal is to use this theory as a basis for *structural estimation* of the IPO decision using panel data on Indian firms over the period 1990 to 2005. Such an exercise can help us determine whether our theory really is capable of providing a good approximation to the behavior of actual firms, and explaining why some firms go public and others don't. A successful theory may be useful for policy making to better understand the "declining dynamism" in the US economy by Decker, Haltiwanger, Jaarmin and Miranda (2016), "The post-2000 period in the U.S. has exhibited declines both in various indicators of business dynamism and in aggregate productivity growth. The decline in both has been most dramatic in the High Tech sectors of the economy." The declining dynamism is reflected in a secular decline in the number of IPOs, as Ritter (2013) notes "From 1980–2000, an annual average of 310 operating companies went public in the United States. During 2001–12, on average, only 99 operationg companies went public. This decline occurred in spite of the doubling of real gross domestic product (GDP) during this 33-year period. The decline is even more severe for small-company initial public offerings (IPOs), for which the average volume dropped 83 percent, from 165 IPOs a year during 1980–2000 to only 28 a year during 2001–12." This secular decline occured despite policies such as the JOBS Act of 2012 which was "intended to encourage the funding of small businesses, primarily by easing various securities regulations."

Our theory can be extended to incorporate the impact of various types of regulations and policies affecting IPOs, including regulations such as the Sarbanes-Oxley Act that was passed in the wake of the Enron debacle and was designed to improve accounting, financial disclosures, company oversight and corporate responsibility. It remains an open question whether well intended regulations designed to prevent corporate abuses have had unintended effects that have contributed to the decline in IPOs. As Ritter argues, "It is possible that, by making it easier to raise money privately, creating some liquidity without being public, restricting the information that stockholders have access to, restricting the ability of public market shareholders to constrain managers after investors contribute capital, and driving out independent research, the net effect of the JOBS Act might be to reduce the flow of capital into young high-technology companies or the number of IPOs of small emerging growth companies."

However our theory suggests it is not just regulatory burdens that could contribute to the trend: we find that the costs of undertaking IPOs, with the large 7% underwriting commission and associated fixed costs, can also be a significant deterrent of using IPOs to finance growth relative to debt financing or retained earnings. But most importantly it is the owner's *expectations* and beliefs about the potential gains to significant capital investments that is a decisive driver of the decision to do an IPO. If competition in the product market is tougher making it harder for newly ascending companies to break the "glass ceiling" of competition from larger well established rivals, then even drastic reductions or even outright elimination of regulatory burdens and reductions in the investment banking costs of doing IPOs (perhaps via policies that encourage freer entry into the underwriting sector) may not have a significant effect on the number of IPOs, especially with the expansion of "private capital" and "venture capital" as alternative potentially more efficient alternative to IPOs. As Decker *et. al.* note, "Theo open question is why firms with high realizations of productivity, especially those in the High Tech sector, do not experience the same high growth as before."

Appendix: Solving the Model by Discrete Policy Iteration

This appendix describes the numerical solution of the models in this paper using the Howard (1960) *policy iteration* algorithm. This algorithm was originally developed to solve infinite horizon stationary Markovian dynamic programming problems (often abbreviated as MDPs for Markovian Decision Problems) on a *finite* state spaces. The optimal investment and dividend problem is superficially not a finite state MDP in the following senses: 1) the state space is continuous (the entire positive real line, $k \geq 0$), and 2) the problem is deterministic, rather than stochastic. Despite these differences, we show that policy iteration can still be applied, but to solve the problem on a finite subset or *grid* of points in the state space and then to apply *linear interpolation* to construct an approximate value function and decision rule essentially by "connecting the dots" where the "dots" are the calculated value function and optimal investment/divident policy at values of k on a pre-defined grid of points $\{k_1,\ldots,k_n\}$ where $k_1 = 0$ and $k_j < k_{j+1}$, $1 = 1,\ldots,n-1$. There is quite a bit of flexibility in how one chooses a grid, but we will show that even for relatively small n and a "naive" choice of equally spaced grid points, it is possible to obtain a very accurate approximation of $V(k)$, $I(k)$ and $D(k)$. The most important choice is the value k_n which constitutes an effective "upper bound" on the capital stock. It is important to "guess" a value for k_n that is large enough so that $I(k_n) = 0$. Otherwise if the guess of the upper bound is too small and $I(k_n) > 0$, this poor initial choice of upper bound can lead to substantial errors in the calculated V, I and D functions.

The basic ideal of how policy iteration works is explained well for the case of finite state spaces in Howard (1960) or Bertsekas (1987), but for the case where the state space has uncountably many states, policy iteration can also be defined but it takes somewhat more advanced functional analysis, see e.g. Puterman (1978). We will describe policy iteration first in the case where the state space is continuous, but it is important to consider a "truncated" version of the problem on a finite interval $[0,K]$ for some $K > 0$ sufficiently large. The reason for truncating the problem is that much of the standard functional analysis machinery is based on use of the *sup norm* $\|V\| = \sup_k |V(k)|$ but this will equal ∞ if the state space is the entire positive real line $[0,\infty)$ if the function V is not bounded.

However once we consider a bounded interval, we can define the Banach space $C(K)$ of all bounded, continuous functions on the interval $[0,K]$, and for this function space, the sup-norm is well defined. In particular if we define the *Bellman operator* $\Gamma : C(K) \rightarrow C(K)$ by

$$\Gamma(V)(k) = \max_{0 \leq I \leq k} [f(k) - I + \beta V(k(1-\delta) + I)] \tag{87}$$

we can show that Γ is a *contraction mapping,* i.e. it satisfies

$$\|\Gamma(V) - \Gamma(W)\| \leq \beta \|V - W\| \tag{88}$$

and via the well-know Banach Fixed Point Theorem (also known as the Contraction Mapping Theorem), Γ has a unique fixed point $V = \Gamma(V)$. This unique fixed point is the value function for the truncated problem given by the Bellman equation (2).

Policy iteration is an iterative method for finding the solution to the Bellman equation which is equivalent to finding the fixed point to the Bellman operator Γ. The standard method for finding a fixed point is the method of *successive approximation* and it is based on any initial guess V_0 and an updated estimate V_1 is produced by evaluating Γ on the initial guess V_0, or $V_1 = \Gamma(V_0)$. Then we use V_1 to produce another estimate $V_2 = \Gamma(V_1)$ and we continue this iteration in general as

$$V_j = \Gamma(V_{j-1}) \quad j = 1, 2, \ldots \tag{89}$$

until we find that the changes in the successive iterates are less than a specified convergence tolerance ε, i.e. until some iteration j such that $\|V_j - V_{j-1}\| < \varepsilon$. The Contraction Mapping Theorem guarantees that for any $V_0 \in C(K)$ we have

$$\lim_{j\to\infty} V_j = \Gamma(V_{j-1}) = V = \Gamma(V) \tag{90}$$

so the method of successive approximations is guaranteed to converge from any initial guess V_0. A drawback of successive approximations is that is converges only *geometrically,* that is, we have

$$\|V_j - V\| \le \beta^j \|V_0 - V\| \tag{91}$$

so that for β close to 1, the rate of convergence of the estimated value function V_j to the true value function V is very very slow.

However policy iteration is a much faster algorithm that usually converges to the *exact solution V in a finite number of iterations, regardless of how close* β *is 1.* This is technically true only in finite state MDPs, but for continuous state MDPs there is a close analog of this result, namely that *policy iteration is equivalent to Newton's method and will converge at a quadratic rate.* This implies that the error in approximating the fixed point, $\|V_j - \Gamma(V_j)\|$, where V_j is the j^{th} iterate produced by the policy iteration algorithm will be very small after only a "small" number of iterations j even for β very close to 1.

Policy iteration is a combination of two "sub-iterations": 1) policy improvement and 2) policy valuation. We explain policy valuation first. Policy valuation is a method to find the value function V_I corresponding to any given investment policy I. Given that we are considering only truncated investment problems, we will initially consider only a subclass of decision rules that satisfy the constraints 0) $I(k)$ is a continuous function of $k \in [0, K]$, i.e. $I \in C(K)$, 1) $0 \le I(k) \le f(k)$ (feasibility), and 2) $k(1-\delta) + I(k) \le K$ for all $k \in [0, K]$. The latter constraint ensures that the mapping Γ_I defined by

$$\Gamma_I(V)(k) = f(k) - I(k) + \beta V(k(1-\delta) + I(k)) \tag{92}$$

makes Γ_I a well defined operator on $C(K)$, i.e. for any $W \in C(K)$ we have $\Gamma_I(W) \in C(K)$. Further Γ_I can be shown to be a constraction mapping, and thus it has a unique fixed point $V_I = \Gamma_I(V_I)$. We now show that Γ_I is an *affine operator,* that is it is a "shifted linear operator" given by

$$\Gamma_I(W)(k) = D_I(k) + \beta E_I(W)(k) \tag{93}$$

where $W \in C(K)$, $D_I(k) = f(k) - I(k)$, and E_I is a linear operator on $C(K)$ defined by

$$E_I(W)(k) = W(k(1-\delta) + I(k)) \tag{94}$$

The constraint on the set of allowable investment rules I implies that if $W \in C(K)$ then $E_I(W) \in C(K)$ so that E_I is an operator on $C(K)$ and further it is a *linear operator* since we have

$$E_I(V+W)(k) = V(k(1-\delta)k + I(k)) + W(k(1-\delta) + I(k)) = E_I(V)(k) + E_I(W)(k) \tag{95}$$

Now define the *norm* of the linear operator E_I by $\|E_I\|$ as follows

$$\|E_I\| = \sup_{V \neq 0} \frac{\|E_I(V)\|}{\|V\|} \tag{96}$$

It is not hard to show that for any $V \in C(K)$ we have $\|E_I(V)\| \leq \|V\|$ which implies that $\|E_I\| \leq 1$, and further, using the example of a function $W(k) = 1$ for $k \in [0,K]$, it is trivially true that $\|E_I(W)\| = 1$, which implies that $\|E_I\| = 1$.

Since we have established that E_I is a linear operator, we can write the equatino for V_I, the fixed point of the operator Γ_I as

$$V_I(k) = D_I(k) + \beta E_I(V)(k) \tag{97}$$

or

$$[\mathscr{I} - \beta E_I](V)(k) = D_I(k) \tag{98}$$

where $\mathscr{I}$ is the *identity operator* on $C(K)$, i.e. $\mathscr{I}(W) = W$ for all $W \in C(K)$ (we use the funny scripted version of capital letter "I" here, $\mathscr{I}$, to distinguish the identity operator from the investment function I). It is easy to show that $\mathscr{I}$ is a linear operator, and thus $[\mathscr{I} - \beta E_I]$ is also a linear operator. Suppose that this linear operator is *invertible.* Then we have the solution

$$V_I = [\mathscr{I} - \beta E_I]^{-1} D_I \tag{99}$$

where $[\mathscr{I} - \beta E_I]^{-1}$ is the inverse operator of $[\mathscr{I} - \beta E_I]$, which is itself also a linear operator. We can show that the inverse operator exists by a geometric series argument. We conjecture that

$$[\mathscr{I} - \beta E_I]^{-1} = \sum_{j=0}^{\infty} \beta^j E_I^j \tag{100}$$

where E_I^j is the linear operator formed as the $j-$fold composition of the operator E_I, i.e. $E_I^2 = E_I(E_I)$ and recursively, $E_I^j = E_I(E_I^{j-1})$. Since $\beta \in (0,1)$ and $\|E_I\| = 1$, it follows that the norm of the right hand side of the *Neumann series expansion* of the inverse operator $[\mathscr{I} - \beta E_I]^{-1}$ in equation (100) is finite (and equals $1/(1-\beta)$) and this establishes that the inverse operator exists.

So in summary policy valuation enables us to obtain a "closed form" expression for $V_I = [\mathscr{I} - \beta E_I]^{-1}D$, which is the value of the firm implied by a given feasible investment policy I. This is why we call this *policy valuation* because V_I represents the *value of the investment policy I*.

Now consider the next sub-iteration of the policy iteration algorihm: *policy improvement.* Using the value function V_I we now seek an *improved policy* I' given by

$$I'(k) = \underset{\substack{0 \le \iota \le f(k) \\ k(1-\delta)+\iota \le K}}{argmax} \; [f(k) - \iota + \beta V_I(k(1-\delta)+\iota)] \tag{101}$$

where we use the notation ι to denote a candidate investment value that we are optimizing over in order to find a new better policy $I'(k)$. Given I' we can now return to the policy valuation step to find the value $V_{I'}$ of this new, improved policy I'

$$V_{I'} = [\mathscr{I} - \beta E_{I'}]^{-1} D_{I'} \tag{102}$$

We can show that $V_{I'} \ge V_I$, i.e. I' really is an *improved* policy that results in a higher value for the firm. However if $V_{I'} = V_I$, then the new policy is not a *strict improvement* over the previous policy I at any $k \in [0,K]$, and at that point *policy iteration has converged.* It is not hard to show that at convergence, $V_I = \Gamma(V_I)$, i.e. V_I is a solution to the Bellman equation, and since this is unique by the Contraction Mapping Theorem, the policy iteration algorithm has succeeded to find the fixed point to Bellman's equation.

The formulas above seem "theoretical" since they involve inversion of linear operators on $C(K)$ which are infinite-dimensional objects. However we can approximate these "infinite dimensional" operators with finite-dimensional operators on a large but finite dimensional Euclidean space R^n. We achieve this via the device of *discretization* and solving the problem on a finite grid of n points $\{k_1,\dots,k_n\} \subset [0,K]$.

When we have a finite grid, we can produce a continuous piecewise-linear approximation by *linear interpolation.* For example suppose we have a given function $W(k)$ but suppose that we only have access to values of this function at n grid points $\{k_1,\dots,k_n\} \subset [0,K]$. That is we know the n values $\{w_1,\dots,w_n\}$ where $w_j = W(k_j)$, $j = 1,\dots,n$. How can we approximate the true value $W(k)$ at some $k \in [0,K]$ that is not one of these grid points? This is quite easy: k must lie between two successive grid points, i.e. $k \in (k_{j-1},k_j)$ and so we can represent k as a convex combination of these grid points using a weight (or it could be interpreted as a "probability") $p(k)$ given by

$$p(k) = \frac{k - k_{j-1}}{k_j - k_{j-1}} \tag{103}$$

so we can write $k = p(k)k_j + (1-p(k))k_{j-1}$. Then using the weight $p(k)$ we can produce the following approximate value $\hat{W}(k)$

$$\hat{W}(k) = p(k)w_j + (1-p(k))w_{j-1} \tag{104}$$

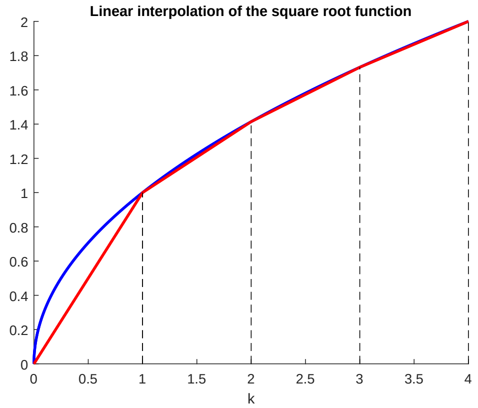


Figure 30: Interpolated square root function $f(k) = \sqrt{k}$

Figure 30 below shows the square root function on the interval $[0,4]$ and its linear interpolation using a grid of five equally spaced points $\{k_1,k_2,k_3,k_4,k_5\} = \{0,1,2,3,4\}$.

Using interpolation, we can carry out policy iteration over a grid of n points on the interval $[0,K]$ and nearly all of the operations become *finite* because the set of piecewise linear functions with nodes (or "knot points") at a grid of n points $\{k_1,\dots,k_n\}$ is an *n-dimensional subspace of* $C(K)$. Our goal is to try to approximate the true value function $V \in C(K)$ with an approximate value function V_n that lives in the n-dimensional subspace of $C(K)$ that consists of all continuous functions whose values over the entire interval $[0,K]$ are linearly interpolated from their values at the n points $\{k_1,\dots,k_n\}$.

So suppose we are given an investment policy I whose values are known at each of the grid points $\{k_1,\dots,k_n\}$ and are determined by linearly interpolation of the known values $\{I(k_1),\dots,I(k_n)\}$ for other values of $k \in [0,K]$. Recall our general equation for the policy valuation step

$$V_I = D_I + \beta E_I V_I \tag{105}$$

where E_I is the (infinite-dimensional) linear operator that "implements" the evaluation of V at a given point $k(1-\delta)+I(k) \in [0,K]$, i.e.

$$E_I(V)(k) \equiv V(k(1-\delta)k + I(k)) \tag{106}$$

Now consider restricting the domain of allowable values of k to just the n grid points $\{k_1,\dots,k_n\}$. Then the infinite-dimensional version of the policy-evaluation equation (105) above become a *system of n linear equations in n unknowns in R^n*

$$V_I = D_I + \beta E_I V_I \tag{107}$$

where now we have V_I and D_I are *vectors* in R^n given by

$$V_I = \begin{pmatrix} V_I(k_1) \\ V_I(k_2) \\ \\ \cdots \\ \\ V_I(k_{n-1}) \\ V_I(k_n) \end{pmatrix} \quad D_I = \begin{pmatrix} D_I(k_1) \\ D_I(k_2) \\ \\ \cdots \\ \\ D_I(k_{n-1}) \\ D_I(k_n) \end{pmatrix} \tag{108}$$

and E_I is an $n \times n$ *transition probability matrix* which implements the interpolation operation. That is, consider the first row of E_I. It will be all zeros except for at most two adjacent non-zero elements with values $1 - p_I(k_1)$ and $p_I(k_1)$, repectively. Recall that we can interpolate $V_I(k_1(1-\delta) + I(k_1))$ using its known values $(V_I(k_1), \ldots, V_I(k_n))$ on the grid $(k_1, \ldots, k_n)$ as follows

$$V_I(k_1(1-\delta) + I(k_1)) = p_I(k_1)V_I(k_j) + (1 - p_I(k_1))V_I(k_{j-1}) \tag{109}$$

where j indexes the grid point k_j such that $k_1(1-\delta) + I(k_1) \in [k_{j-1}, k_j]$ and $p_I(k_1)$ is given by

$$p_I(k_1) = \frac{k_1(1-\delta) + I(k_1) - k_{j-1}}{k_j - k_{j-1}}. \tag{110}$$

Thus the first row of E_I will have $p_I(k_1)$ in its j^{th} column and $1 - p_I(k_1)$ in its $(j-1)^{\text{st}}$ column and all other columns equal zero. It follows that the first row will sum to 1 by construction. This same idea applies to all other rows of E_I so we conclude it as the form of a Markov *transition probability matrix* i.e. all elements are between 0 and 1 and each row sums to 1.

Using E_I and D_I it is now a matter of linear algebra to solve for $V_I \in R^n$

$$V_I = [\mathscr{I} - \beta E_I]^{-1} D_I \tag{111}$$

except that now $\mathscr{I}$ is the $n \times n$ *identity matrix* (which is also the "identity operator" on R^n). Using $V_I \in R^n$ we can extend it to a continuous function of k over all of the interval $[0, K]$ via linear interpolation, so we can also interpret V_I as an element of the n-dimensional subspace of $C(K)$ of functions which are linearly interpolated from their values at the n grid points $\{k_1, \ldots, k_n\}$.

Given V_I we can now do the *policy improvement step* to see if we can find a better investment policy $I'(k)$ by optimizing over investment at each of the n grid points k_j, $j = 1, \ldots, n$.

$$\begin{aligned} I'(k_j) &= \underset{\substack{0 \le \iota \le f(k_j) \\ k_j(1-\delta)+\iota \le K}}{\mathit{argmax}} \left[f(k_j) - \iota + \beta \hat{V}_I(k_j(1-\delta) + \iota) \right] \\ &= \underset{\substack{0 \le \iota \le f(k_j) \\ k_j(1-\delta)+\iota \le K}}{\mathit{argmax}} \left[f(k_j) - \iota + \right. \\ &\qquad \beta \left[p(k_j(1-\delta) + \iota) V_I(k_l) + (1 - p(k_j(1-\delta) + \iota) V_I(k_{l-1})) \right]] \end{aligned}$$

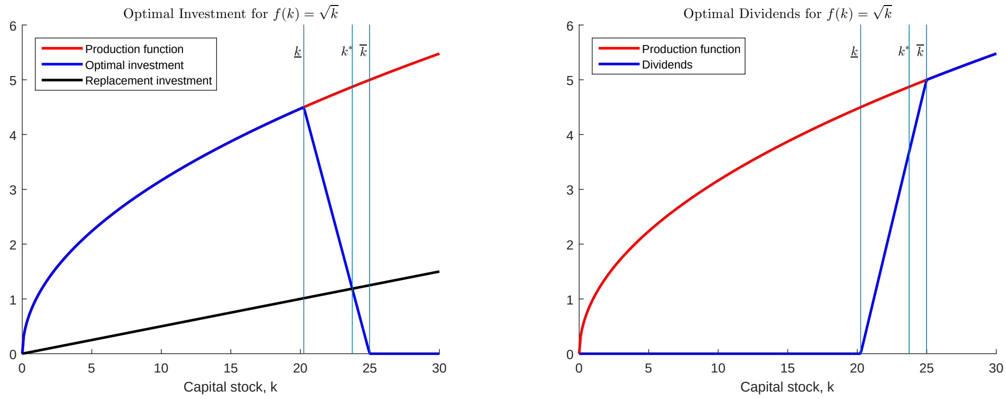


Figure 31: Approximate optimal investment and dividend policy for $f(k) = \sqrt{k}$

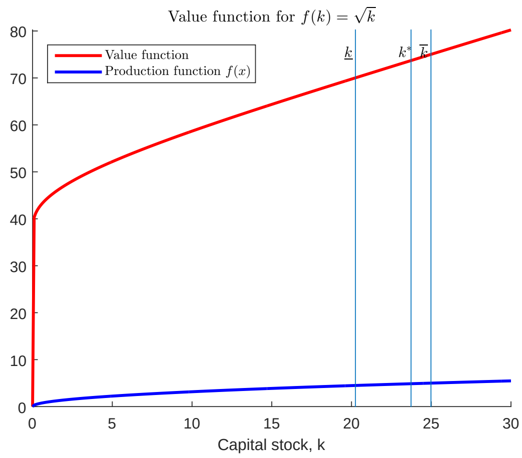


Figure 32: Value function for $f(k) = \sqrt{k}$

where $k_j(1-\delta)+\iota \in [k_{l-1}, k_l]$ for some index $l \in \{1,\ldots,n\}$. If $I'(k) = I(k)$ for all grid points $k \in \{k_1,\ldots,k_n\}$ (or equivalently if $V_{I'} = V_I$), then **stop:** policy iteration as converged to a V that solves the Bellman equation (though restricted to the finite dimensional subspace of $C(K)$ of functions defined by linear interpolation at the n grid points $\{k_1,\ldots,k_n\}$). If not, then using the improved policy I' we return to the policy valuation step (107) to calculate $V_{I'}$ and continue until the policy iteration process converges.

Figure 31 presents the approximate decision rules for investment and dividends computed by poicy iteration with $n = 301$ grid points, equally spaced from $k_1 = 0$ to $k_{301} = 30$, a spacing of 0.1 apart. Policy iteration converged after 20 iterations, resulting in a (sup norm) change in value functions of 5.97×10^{-13}. We see that the computed solutions look virtually identical to the true solutions plotted in figure 1 in section 2. Figure 32 also plots the interpolated value function from policy iteration and it also looks virtually identical to the true value function in figure 1.

There are approximation errors but they are small. Figure 33 plots the approximation errors at the grid points $\{k_1,\ldots,k_n\}$ for two different solutions, one using policy iteration with $n = 150$ grid points,

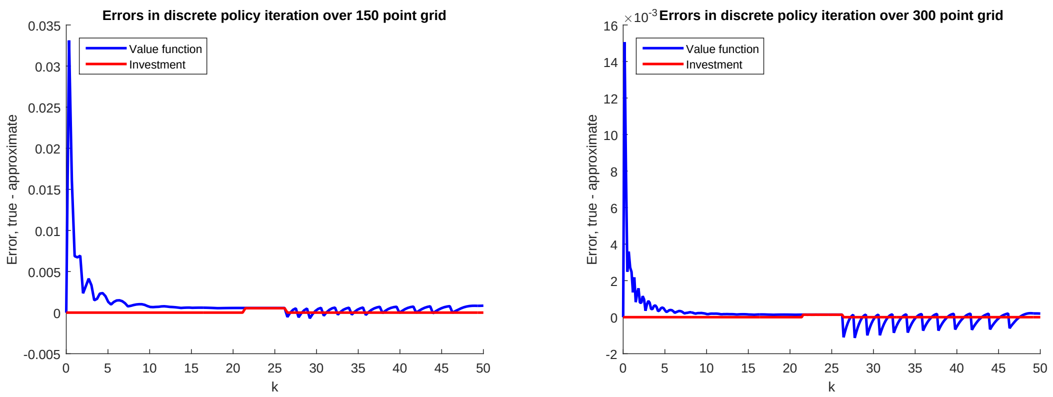


Figure 33: Approximation errors in value function and investment rule with $n = 150$ and $n = 300$ grid points

and the other using $n = 300$ grid points, in both cases equally spaced over the interval $[0,30]$. Generally we would expect that using a "finer grid" i.e. a larger number of grid points n, should result in a better approximation. This is the case here, though it required a more sophisticated version of interpolation than simple linear interpolation of the ordinates $\{V(x_1),\dots,V(x_n)\}$ in the policy improvement step. We used *piecewise cubic hermit polynomial interpolation* as implemented in the `pchip` function of Matlab. The `pchip` function interpolates the ordinates in a way that gaurantees continuous differentiability of the interpolated function at the grid points (unlike what happens with simple linear interpolation, where the derivatives are generally discontinuous at the grid points) and the interpolated function is *shape preserving* which is particularly important in this case to preserve the concavity of the value function over the entire domain.

In each policy improvement step we used the Matlab `fminbnd` function to numerically search for the optimal value of investment $\iota \in [0, f(k_j)]$ at each grid point k_j, $j = 1,\dots,n$. When we used simple linear interpolation, the interpolated value function has more and more discontinuities in its derivative as n gets large. This appears to create problems for the Matlab optimizer, and when n gets sufficiently large, the approximation actually starts to degrade. This is not the case when the `pchip` interpolater was used. The approximation error reduces as the number of grid points increases, though there is diminshing returns to increasing n. Further accuracy can be achieved by using the strict concavity of the value functions and using a Newton or bisection algorithm to find optimal investment as a solution to the first order condition $1 = \beta V'(k(1-\delta)+I)$, using the fact that the `pchip` interpolated results in a piecewise quadratic expression for V' that makes it easy to employ Newton's method to solve for the value of I that satisfies the first order condition.

Overall, we have demonstrated that the DPI algorithm seems to be capable of finding a good approximation to the true value function and decision rules.