Staged Investments in Entrepreneurial Financing

KOROK RAY
McDonough School of Business
Georgetown University
Washington, DC 20057
kr268@georgetown.edu

February 16, 2010

Abstract

Venture capitalists deliver investments to entrepreneurs in stages. This paper shows staged financing to be efficient. Staging lets investors abandon ventures with low early returns, and thus sorts good projects from bad. The primary implication from staging is that it is efficient to invest more in later rounds. The model yields a number of empirical implications on how the ratio of early to late round financing varies with uncertainty, the outside options of both parties, the value of the venture, the costs of investment, and project difficulty. The main results generalize in a model that includes the entrepreneur’s unknown ability.

JEL Classification code: G, D.

Keywords: Entrepreneurship, Venture capital, staged financing, optimal stopping, performance evaluation, financial contracting.
1 Introduction

“You got to know when to hold ’em, know when to fold ’em,
Know when to walk away and know when to run.”
-Kenny Rogers, The Gambler

This paper gives an efficiency-based explanation of staged financing in venture capital. Sahlman (1990), Gompers (1995), and Gompers and Lerner (1999) all document the extensive practice of venture capitalists delivering investments to new firms in stages. The current view in the venture capital literature is that staging mitigates moral hazard. Here, I argue that venture capitalists use staging as a sorting instrument. Staging investments provide the venture capitalist (VC) with the option of ending projects with low early returns. This sorts ventures into two groups: stay or quit. It is efficient to quit if the early returns are weak, and to stay otherwise. If the entrepreneur’s early output is low, it is in the interests of both the VC and the entrepreneur to discontinue work and collect their respective outside options.

The main result shows that it is efficient to assign more resources to the later stages of the project. Thus, one consequence of staging is that it is efficient for investment levels to increase in later rounds. Staging creates the possibility of termination after the early stage, and this reduces the project’s expected return. This lowers the marginal return from investment, and so the VC shades his investment downward in the early stage. Once the entrepreneur advances, the possibility of termination vanishes and the marginal return to investment rises, so the VC invests more. Those who make it to the second stage are more valuable precisely because their first stage output was sufficiently high. The VC invests more in the later stage because the new venture is “in the running” to becoming highly successful. Said differently, it is inefficient for the VC to dump too many resources into a horse that won’t finish the race.

The distinguishing feature of this paper is that the analysis operates entirely within a first-best setting. The model abstracts from conflicts of interest and agency problems between the VC and entrepreneur. While the relationship between a VC and an entrepreneur is no doubt rife with moral hazard, it is unnecessary to resort to a complex agency model to explain staged financing. A first-best analysis predicts that funding levels increase over stages, an empirical regularity that Gompers (1995) documents. This
suggests that asymmetric information is not necessary to explain all aspects of staged financing, as prior work in this literature claims. Instead, optimal decision making under uncertainty gives a clean, robust, and intuitive explanation of staged financing.

The model consists of a risk-neutral venture capitalist funding an entrepreneur over two stages. Each party invests resources (capital and labor) into each stage, and the output from the project is the total investment plus a noise term in each stage. Both parties are symmetrically uninformed on the project’s uncertainty. The project earns a positive return if the total output across both stages clears an exogenous hurdle. For example, consider a software company developing a new search engine. If the search engine is of sufficiently high quality, it has positive value and the company has potential to be taken public; otherwise, the product is worth nothing.

Both the VC and the entrepreneur have outside options in each stage. For example, the VC can fund other ventures and the entrepreneur can work on other projects. Staging investments gives the VC the option to discontinue the project, at which point both parties collect their respective outside options. The primary implication from staging investments is that it skews the efficient allocation of resources towards the later stages. The VC deliberately withholds investments in the early stages precisely because of the uncertainty from the early stage. In particular, the model shows that the VC will set a milestone after the first stage, and if the project’s output clears this milestone, the VC knows the project is sufficiently successful and therefore invests more.

In addition to this primary implication that investments increase in later rounds, the model generates several testable predictions for future empirical work. First, as the outside options of both parties increase, the VC will skew its investments even more into later rounds. Intuitively, as the parties’ outside opportunities improve, the VC has a high opportunity cost from investing, and therefore can adopt a “wait and see” approach and can postpone investments into the future. Second, as the uncertainty in the model increases, it is efficient for the VC to invest more towards the early stage. While this may seem counterintuitive, the logic follows from the option value of continuing. Because the stages are sequential, an increase in uncertainty increases the upside benefit from continuing, and this gives an extra benefit to investing in the early stage rather than the late stage. And finally, as the difficulty of project completion increases (because of market or technology factors), the VC will invest even more resources into later rounds. This oc-
curs because the VC is reluctant to invest too much money in early stage projects which are unlikely to “make it.” All of these comparative statics give concrete predictions on the ratio of early to late round financing. Future empirical work can therefore test these predictions in industries with cross-sectional heterogeneity in technological uncertainty, outside options of VCs and entrepreneurs, and market risk. None of the prior theoretical work in staged financing generates testable predictions on how the investment levels per stage vary with the economic and technological environment of the firm.

I extend the benchmark model by including an ability parameter for the entrepreneur that persists across both stages. The model considers output measures that cannot disentangle ability, investment, and uncertainty, where ability is unknown to all parties. Persistent ability induces correlation over time. If early stage output is low, both the VC and entrepreneur Bayesian update their priors on ability. They infer that low underlying ability drives low first stage output. Since ability persists into the second stage, it also drives low second stage output. They have outside options, so it is efficient to quit early on. Moreover, low ability entrepreneurs push more of their investment into later stages than high ability entrepreneurs. Thus, the main results still hold: it is still efficient to quit if the early returns are weak, and it is efficient to allocate more resources in later rounds.

Existing literature on staged financing exists exclusively within agency models of asymmetric information. A landmark paper is Neher (1999), who claims that entrepreneurs threaten to hold up VCs by reneging on investments, so VCs stage payments to reduce their bargaining power. Dividing investments into a number of stages creates inefficiencies but is necessary in overcoming the commitment problem. Landier (2002) argues that staging is one way of protecting an investor from risk when entrepreneurs have a high exit option, i.e. when bankruptcy laws are lenient and when there is little stigma associated to business failure. Bergemann and Hege (1998, 2005) study the dynamics of the optimal contract and equilibrium funding decisions in arm’s length versus relationship financing. In other work, Bergemann and Hege (2003) show that the duration of funding, though not necessarily the level of funding, increases in later stages. Cornelli and Yosha (2003) look at “window dressing,” the manipulation of information on project performance, which entrepreneurs may practice in order to continue to receive funds. Wang and Zhou (2004) finds that there are cases in which up-front financing
may be superior to staging; under staged financing, VCs will underinvest in low quality projects and potentially doom them to failure. In Yerramilli (2006), each party can hold up the other and threaten to walk away in order to press for a renegotiation of the contract. Wang and Zhou (2004) show that staged financing acts as a mechanism in contracting to control agency problems. Finally, without the ability for investors to unilaterally cancel projects, Admati and Pfleiderer (1994) argue that entrepreneurs with outside financing will be reluctant to quit unproductive ventures. All of these models take place in moral hazard and asymmetric information settings, and therefore staging is an instrument to minimize agency costs. None of the prior theoretical work explores the efficiency properties of staged financing.

The empirical literature is consistent with the primary implication of the model. For example, Sahlman (1990) and Gompers (1995) analyze the Venture Economics database and find that VCs disburse more money to firms in later stages of development. The existing theoretical papers on staged financing give mixed predictions on whether investments will increase in later rounds. In Neher (1999), investments increase over time because the VC is willing to invest more as the firm’s collateral grows. Yet Giat et al. (2009) let the VC and entrepreneur hold asymmetric beliefs about the potential value of a project, and find that staged investments can increase over time, decrease over time, or rise and then fall. Hsu (2002) analyzes staging in an options valuation framework, assuming that agents act to maximize the probability of advancing stages. Hsu (2002) finds computationally that staging tends to be more profitable to investors when ventures are in early stages and will need greater amounts of capital in the future.1 Yet, none of these papers make predictions on how the ratio of early to late stage financing changes with exogenous parameters of the environment, such as increase in uncertainty, market risk, or outside opportunities.

The paper is organized as follows. Section 2 presents the benchmark model and shows that staging investments increases total surplus. Section 3 explores the effects of staged financing on the ratio of early to late stage funding levels. Section 4 contains the

---

1 Other empirical work in venture capital documents different features of the venture capital environment. Krohmer and Lauterbach (2005) find empirically that in the final stages of a project, investment managers may be too unwilling to pull the plug on failing projects. Cuny and Talmor (2005) and Bienz and Hirsch (2009) look at the differences between the two types of staged financing that are commonly observed, staging with milestones or with rounds.
comparative statics with respect to the model parameters, and delivers secondary implications on how the funding level changes with the uncertainty in the model, technology or market risk, or the outside options of the VC or entrepreneur. Section 5 expands the model to include an ability parameter for the entrepreneur. Section 6 concludes.
2 The Model

Consider an entrepreneur working on a project (a new venture) over time. The entrepreneur seeks funding for the project from a venture capitalist (VC). Both parties are risk neutral. Production takes place across two stages, and there is no discounting. It takes time to establish a business, and the stages represent distinct phases in production. For example, the early stage involves establishing the founder’s initial business plan, while the later stage involves marketing the plan and generating advertising revenue. Let \( k_t \) be the total resources invested in the project at stage \( t \). This reflects the sum of both the entrepreneur’s and the VC’s resources (labor and capital) invested in the project. Though I call \( k_t \) investment, it includes human resources as well as financial resources. Since the focus of the analysis is on efficient resource allocation, it is not necessary to specify the entrepreneur’s and venture capitalist’s resources separately.

The total resources \( k_t \) in stage \( t = 1, 2 \) face a cost of resource function \( C(k_t) \). This is the total social cost of resources in stage \( t \). Assume \( C', C'' \) are strictly positive, so costs are separable across stages, increasing, and convex. The convexity of the cost function reflects a convex cost of investment for the venture capitalist and a convex cost of effort for the entrepreneur. A convex cost of effort is a standard assumption, while a convex cost of investment simply reflects that the VC cannot invest arbitrarily large amounts without cost.\(^2\) The convexity of the supply curve represents all the costs of raising capital to deliver funds to the entrepreneur. Output from the project is

\[
q_t = k_t + \epsilon_t.
\]

The noise terms \( \epsilon_t \) are i.i.d., and distributed symmetrically around a mean of zero and over infinite support, with cdf \( G(\cdot) \) and density function \( g(\cdot) \). Interpret \( \epsilon_t \) as a stage-specific shock unknown to anyone. The \( \epsilon_t \) captures all of the market and technological uncertainty in raising profits: novelty of the founder’s idea, viability of the business plan, viability of the business plan, viability of the business plan,

---
\(^2\)VCs draw from dedicated pools of capital that institutional investors supply. In particular, the VC raises capital in blocks called funds, usually targeted towards investments in a specific industry or technology. If the VC exhausts the fund and wants to invest more he must raise a new fund, which involves soliciting interest from limited partners (institutional investors), advertising the fund through business networks, or transferring capital from other preexisting funds. See Prowse (1998) for a full description on the capital raising process.
existence of a potential market, quality of human and physical capital, etc. Even though the noise terms are independent, later in the paper I will add an ability parameter which persists across stages. This is essentially equivalent to making the noise terms correlated over time.

A project is a pair \((V, \bar{q})\), where \(V > 0\) is the value of the project and \(\bar{q} > 0\) is the final hurdle. After stage two, the VC takes the firm public if it is of sufficiently high quality. Therefore, the final hurdle represents the minimum quality necessary for a new venture to capture a positive market price when its shares are traded on public stock markets. The value of the venture is

\[
V(q_1, q_2) = \begin{cases} 
V & \text{if } q_1 + q_2 > \bar{q} \\
0 & \text{otherwise.}
\end{cases}
\]

Output (for e.g. profits, quality, sales) has no value unless it is sufficiently high. In most new ventures, the venture is worth little unless it can eventually be taken public, or at least generate profits. Assets of firms that have either failed to go public or have not successfully obtained later round financing are usually sold at low (firesale) prices; I simply normalize these low prices to zero. Thus, \(q_t\) is the project’s internal output (prototypes, beta versions, etc) while \(V(q_1, q_2)\) measures the project’s external value based on market valuation. Throughout, call \(q_t\) the project’s output, and call \(V(q_1, q_2)\) the project’s value. Since information is symmetric in this model, both parties know the true value \(V\) but do not know whether output from the project is sufficiently high to clear the hurdle \(\bar{q}\). Observe that output levels across stages are perfect substitutes. This isolates the effects of staging on investment from the effects of technology on investment.

Suppose that both the VC and the entrepreneur have outside options in each stage. These outside options capture the value of the outside opportunities of both parties. For example, the VC has many competing investments to fund and can allocate his capital and his time elsewhere. Similarly, the entrepreneur can either work on other new ventures, or even collect a wage as an employee for another organization. Let \(\bar{u}_t\) be the sum of the outside options of the VC and entrepreneur in stage \(t\). So \(\bar{u}_t\) measures the opportunity cost of the project (time, labor, capital) to both parties.\(^3\) The venture

\(^3\)The outside options are independent of early stage output. The results of the model generalize if outside options increase linearly in output. This variation adds little insight and obscures analysis, so I omit it.
The capitalist may conduct an evaluation of the venture after stage one. In fact, the purpose of staged financing is to give the venture capitalist an intermediate reading on the new venture, with the option of ending the venture if the early returns are weak.

### 2.1 Upfront Financing

As a benchmark, suppose that the VC does not conduct an evaluation after the first stage. Importantly, there are no grounds for terminating the project after the first stage. So the VC gives all the funds for the project upfront; call this “upfront financing.” To calculate the social payoff, observe that both parties receive positive surplus only if the project is a success, i.e. that $q_1 + q_2 > \bar{q}$. The probability of success is

$$P = \Pr(q_1 + q_2 > \bar{q}) = \Pr(\epsilon_1 + \epsilon_2 > \bar{q} - k_1 - k_2) = \int_{-\infty}^{\infty} \int_{\bar{q} - k_1 - k_2 - \epsilon_1}^{\infty} g(\epsilon_1) g(\epsilon_2) \, d\epsilon_2 d\epsilon_1$$

by the independence of the errors. After integrating and using the symmetry of the errors around zero,

$$P = \int_{-\infty}^{\infty} g(\epsilon_1)[1 - G(\bar{q} - k_1 - k_2 - \epsilon_1)] \, d\epsilon_1 = \int_{-\infty}^{\infty} g(\epsilon_1) G(\epsilon_1 + k_1 + k_2 - \bar{q}) \, d\epsilon_1.$$

Therefore, the marginal effect of increasing investment on improving the probability of success is

$$\frac{\partial P}{\partial k_t} = \int_{-\infty}^{\infty} g(\epsilon_1) g(\epsilon_1 + k_1 + k_2 - \bar{q}) \, d\epsilon_1 > 0.$$ 

This expression is positive, so increasing investment makes it more likely that the project will clear the final hurdle. Moreover, observe that the right-hand side of the equality above is independent of $t$, and therefore so is the left-hand side. The VC can fund either in stage one or stage two, as they have the same effect on the project clearing the final hurdle. Thus, the probability of success increases by the same amount with investment in either stage. Since total investment is additive, stage one and stage two investment are perfect substitutes.

Since the objective of the analysis is to understand the efficient allocation of resources, it is necessary to consider the social planner’s problem, i.e., the joint payoff of the entrepreneur and the VC combined. This is the expected benefit from investments less the cost of investment in each stage. Observe that even though the production function
$V(q_1, q_2)$ is discontinuous at the point $q_1 + q_2 = \bar{q}$, the planner’s expected payoff $PV$ is continuous in $k_t$. The social planner maximizes total surplus, so the problem is

$$\max_{k_t} PV - C(k_1) - C(k_2),$$

which yields the first-order condition

$$V \frac{\partial P}{\partial k_t} \bigg|_{k_t = \hat{k}_t} = C'(\hat{k}_t),$$

where $\hat{k}_t$ denotes the optimal effort level.

The marginal cost of investment is equal to its marginal return, which is the marginal probability of success times the value of the project. Since the left-hand side is independent of $t$, the right hand side must be as well. Hence $\hat{k}_1 = \hat{k}_2 \equiv \hat{k}$; this is the efficient investment under upfront financing, and is the same in each stage. It is efficient to split investment evenly across stages since the cost of investment per stage is the same. Because the model is symmetric with respect to the VC and entrepreneur, it is possible to implement this first best solution, so the VC will split its investment evenly across stages and the entrepreneur will exert effort and deploy resources evenly across stages. For example, this is the outcome under a contracting game where the venture capitalist is the principal who proposes a contract to the agent, the entrepreneur. In this setting, since both parties are risk neutral, it is straightforward to construct a contract that implements the first-best.\(^4\)

Note that convexity of the cost function is not what guarantees that investment in both stages is the same. Investment is the same because (1) convexity of the cost function guarantees a unique solution, (2) the marginal return to investment in each period is same, and (3) the cost function is separable and identical across stages. Convexity does, however, guarantee that efficient investment increases with $V$. Collecting terms, the efficient per-stage investment level $\hat{k}$ solves

$$C'(\hat{k}) = V \int_{-\infty}^{\infty} g(\epsilon_1)g(\epsilon_1 + 2\hat{k} - \bar{q}) \, d\epsilon_1.$$  

---

\(^4\)Since contracting issues are not central to this analysis, I do not outline the details of the contracting game, such as the contract space, the bargaining power between the two parties, etc. Such a game is straightforward to construct, as the principal will pay the agent $W$ for success ($q_1 + q_2 > \bar{q}$) and $L$ for loss ($q_1 + q_2 < \bar{q}$). To guarantee full incentives to exert first best, the principal will set $W - L = V$. Further details on this contract are available from the author upon request.
The remaining constraint is a bound on the reservation utilities. The total surplus from having the entrepreneur undertake the project must be at least as large as the total outside options across both stages. So,

\[ PV - 2C(\hat{k}) \geq \bar{u}_1 + \bar{u}_2. \]

Call this the project feasibility constraint.

### 2.2 Efficiency of Staging Investments

The main reason to conduct an evaluation halfway through a project is that it provides the option to abandon the project if the early returns are low. The two parties will use first stage output to compute the expected project value \( V(q_1, q_2) \) after the second stage. This yields an expected value of continuing. Because of the outside options, it is efficient to continue only if this value exceeds these outside options.

If the VC and entrepreneur observe \( q_1 \) after the first stage and must decide whether to continue or not, their decision will depend on the observed \( q_1 \). Therefore the probability of continuing and the total surplus from continuing will also depend upon this observed \( q_1 \). The probability of clearing the final hurdle conditional on a realized value \( q_1 \) is

\[ P(q_1) \equiv \Pr(q_1 + q_2 > \bar{q} \mid q_1) = \Pr(\epsilon_2 > \bar{q} - q_1 - k_2) = G(q_1 + k_2 - \bar{q}). \]

So the total surplus conditional on a realized \( q_1 \) is

\[ S(q_1, k_2) = E_2 V(q_1, k_2 + \epsilon_2) - C(k_2) = P(q_1)V - C(k_2), \]

where \( E_t \) denotes the expectation taken over \( \epsilon_t \). Call this the continuation surplus function. For clarity, let \( S(q_1) \equiv S(q_1, k_2^*) \) be the continuation surplus evaluated at the efficient investment level \( k_2^* \). This continuation surplus function reflects the expected total surplus from continuing after a realization of first stage output \( q_1 \). The continuation decision rests entirely on this function. In particular, it is efficient to continue if and only if \( S(q_1) \geq \bar{u}_2 \). The first result below shows that the continuation surplus function is strictly increasing. This means there exists a unique cut-off output level \( q^* \) such that

---

Note that in general \( k_t^* \), which is efficient under staged financing, differs from the \( \hat{k}_t \) from the previous section, which is efficient under upfront financing.
Proposition 1  There exists a target $q^*$ such that it is efficient only for entrepreneurs with $q_1 > q^*$ to advance to the second stage.

Because $q_1 + q_2 > \bar{q}$ in order to collect positive surplus, the stages are connected; output in the early stage signals final project value. Said differently, a successful early stage (high $q_1$) means that the project will have an easier time clearing the final hurdle, and therefore a higher chance of both parties collecting surplus. Proposition 1 shows that the continuation surplus is monotonic, and this generates the cutoff target $q^*$. In practice, this $q^*$ represents the milestone in between rounds of venture financing.\(^6\) If the quality of the project clears this milestone, then the entrepreneur qualifies for the next round of funds. The assumption on outside options is key. Without outside options, it would be efficient to continue for any $q_1$ since all parties get nothing by quitting and are at least as well off continuing (recall that $V$ is always nonnegative).

Proposition 1 shows that staged financing generates more surplus than upfront financing. Under upfront financing, the VC does not collect information on early stage output, and therefore advances all projects regardless of their early performance. Under staged financing, the investor sorts projects into two groups: stay or quit. The target $q^*$ conducts the sorting, in that it allows only entrepreneurs with high output to proceed. The VC can always set the target arbitrarily low, which permits continuation for all output levels, and hence replicates upfront financing. By setting the target optimally, the VC has an additional instrument to maximize total surplus, and therefore must be weakly better off. This suggests that staging financing does more than simply minimize agency costs, as the prior literature has argued. Instead, staging is a tool to make both VCs and entrepreneurs better off.

\(^6\) For example, milestones separate early round financing (series A) from later round financing (series B). New ventures must meet certain targets, such as number of employees hired, free cash flow, research and development investments, progress on business plan, etc. These targets constitute the milestone $q^*$.
3 Effects of Staged Financing

Now that we know staged financing increases surplus, what is the efficient investment level per stage under the staged financing regime? The previous section shows that the continuation decision will take the form of a cut-off rule. Precisely, the VC sets some target (or milestone) $q^*$ after the first stage, and advances the entrepreneur only if $q_1 > q^*$. These milestones constitute the interim performance evaluations, and they separate rounds of financing. The probability of clearing the target is

$$P_1 = \Pr(q_1 > q^*) = G(k_1 - q^*).$$

As expected, this probability increases in first-stage investment since $\frac{\partial P_1}{\partial k_1} = g(k_1 - q^*) > 0$. The ex-ante probability of success is

$$P \equiv \Pr(q_1 + q_2 > \bar{q}, q_1 > q^*) = \int_{q^*}^{\infty} P(q_1)g(q_1 - k_1) dq_1,$$

where $P(q_1) = G(q_1 + k_2 - \bar{q})$ is the interim probability of clearing the final hurdle and capturing $V$ for each realization of $q_1$. Notice that

$$\frac{\partial P}{\partial k_2} = \int_{q^*-k_1}^{\infty} g(\epsilon_1)g(\epsilon_1 + k_1 + k_2 - \bar{q}) d\epsilon_1 > 0,$$

$$\frac{\partial P}{\partial k_1} = \frac{\partial P}{\partial k_2} + \frac{\partial P_1}{\partial k_1} G(q^* + k_2 - \bar{q})$$

$$> \frac{\partial P}{\partial k_2}. \quad (1)$$

The returns to investment are positive for both stages, but are higher for the first stage. Additional first stage investment increases the probability of success in two ways. First, it increases $q_1 + q_2$ and thus directly increases the probability of final success. Second, it increases first stage output ($q_1 = k_1 + \epsilon_1$) and so improves the chance of advancing to the second stage. Hence the marginal benefit of first stage investment exceeds the marginal benefit of second stage investment. It is incorrect to conclude from this, however, that it is efficient to invest more in the first stage, since this analysis ignores the cost of investment, as well as takes $q^*$ as given, whereas in fact $q^*$ is determined simultaneously with the optimal $k_t$. Increasing first stage investment increases the chances of advancing to the second stage and thus increases the probability of bearing the cost of
a second stage investment. At the optimum this cost is enough to push $k_1$ below $k_2$. To see this, it is necessary to solve the social planner’s problem.

The conditional probability of clearing the final hurdle, given that the entrepreneur has reached the target, is

$$Q = \Pr(q_1 + q_2 > \bar{q}|q_1 > q^*)$$

So the ex-ante probability of clearing the hurdle $\bar{q}$ is $P = P_1Q$. If the entrepreneur passes the intermediate target $q^*$, the planner gets $V$ if he clears $\bar{q}$ and zero otherwise, and bears cost $C(k_2)$. If he doesn’t pass the intermediate target, the planner gets only $\bar{u}_2$. So the total surplus is

$$P_1[QV - C(k_1) - C(k_2)] + (1 - P_1)[\bar{u}_2 - C(k_1)].$$

Rearranging terms gives the planner’s problem

$$\max_{k_t, q} PV - C(k_1) + (1 - P_1)\bar{u}_2 - P_1C(k_2),$$

subject to project feasibility. The last term above is the cost of advancing to the second stage. This cost is increasing in first-stage investment. As the VC invests more in stage one, he increases the expected second-stage cost, since larger first-stage investments increase the probability of making it to the second-stage. This cost forces first-stage investment downward, ultimately below even second-stage investment. More generally, it is possible to write the planner’s objective function in terms of the continuation surplus function. So the planner solves

$$\max_{k_t, q} \int_{q}^{\infty} S(q_1, k_2)g(q_1 - k_1)dq_1 + (1 - P_1)\bar{u}_2 - C(k_1),$$

where investment levels $k_t$ and the target $q$ are the planner’s choice variables, and $(k_t^*, q^*)$ denote the efficient choices. The first term is the expected value of continuing: the continuation surplus function integrated over all realizations of $q_1 > q^*$. The middle term $(1 - P_1)\bar{u}_2$ is the expected value of abandoning the project. Both parties collect their outside options if the project does not clear the target, which occurs with probability $1 - P_1$. Note that $C(k_2)$ does not appear in the objective function explicitly because it is embedded in $S(q_1, k_2)$. The planner bears the cost of $C(k_2)$ only in the event that the entrepreneur advances.
3.1 Primary Implication: Dynamic Resource Allocation

The following proposition solves the planner’s problem (2) for the efficient allocation of resources across stages, and is the main result.

**Theorem 1** It is efficient to invest more in the second stage \((k_1^* < k_2^*)\).

Since the planner sets the target optimally, the marginal return of an entrepreneur who cleared the target exceeds the marginal return of an entrepreneur in the first stage. Formally,

\[
C'(k_1^*) = E[S'(q_1)] < E[S'(q_1)|q_1 > q^*] = C'(k_2^*).
\]

The mean marginal return conditional on \(q_1 > q^*\) exceeds the unconditional mean. Since marginal costs are increasing, this implies that \(k_1^* < k_2^*\). The marginal return to investment is lower in stage one precisely because the entrepreneur may not advance to the second stage. In this case, he bears the cost \(C(k_1)\) but acquires the benefit \(V\) not with certainty but with probability less than one. This lowers the marginal return in stage one relative to stage two. At the optimum the VC selects \(k_t\) to set the marginal costs equal to the marginal returns, and so he shades investment downward in the early stages. He will allocate more resources in the later stages of the project, where the marginal return is higher. Rewriting the first order conditions in terms of the specific production function here yields

\[
C'(k_1^*) = V \frac{\partial P}{\partial k_2} < V \frac{\partial P}{\partial k_2} = V \frac{\partial Q}{\partial k_2} = C'(k_2^*).
\]

Thus, those who make it to the second stage are more valuable precisely because their first stage output was sufficiently high. The VC invests more in the later stage because the new venture is “in the running” to becoming highly successful. It is inefficient for the VC to dump too many resources into a horse that won’t finish the race.

Gompers (1995) provides data on staged financing, which fits the predictions of this model. His data comes from a sample of 794 venture capital backed firms randomly selected from the Venture Economics database. Figure 3.1 shows his summary statistics on amounts of funding by stage for all firms in the sample. Without question, it is clear that funding increases in the later stages. This is precisely \(k_1 < k_2\). Moreover, his regression results (Panel B of page 1479) show that average late stage investments
 exceed average early stage investments by $1.3 to $2.03 million, and exceed average middle stage investments by $0.7 to $1.21 million. Therefore, the predictions of this model are at least consistent at a first pass with the existing empirical work of Gompers (1995).

The premise of Gompers (1995) is to explain staged financing as an instrument to mitigate agency conflicts. While the data is consistent with such an interpretation, there is no formal model in his paper. In contrast, our formal model not only generates the same prediction that venture capitalists will stage investments and such investments rise with time, but it also generates new comparative statics on the outside options and output variance. The agency models on staged financing cited above say nothing on these last two points. We now turn to these comparative statics of the model.

### 4 Secondary Implications: Comparative Statics

While some prior theoretical work has made explicit predictions on the ratio of investment levels over stages (Neher (1999), Giat et al. (2009), Hsu (2002)), these papers have not made predictions on how this ratio $\frac{k_1}{k_2}$ varies in different environments, such as industries with different levels of market risk, projects with different levels of technological feasibility, or more prestigious venture capitalists with better outside options. The objective of this section is to explore how the endogenous variables $(k_t^*, q^*)$ vary with
the exogenous parameters \((\bar{u}_2, v, \bar{q})\). In particular, the aim is to produce a number of secondary implications that can be tested against real venture capital data. If this ratio increases, the VC skews the investment mix towards the early stage, and vice versa. To get traction on the model, parameterize the cost function as \(C(k_t) = \lambda k_t^\gamma\) for some constants \(\lambda > 0, \gamma > 1\). The next result predicts how the ratio of investment levels varies with the outside options of both parties.

**Proposition 2** The ratio \(\frac{k_1}{k_2}\) decreases in the outside options \(\bar{u}_2\).

As the outside options improve, it is efficient to invest even more money in later rounds. The outside options represent the opportunity cost of alternative investments in the second stage. As this opportunity cost increases, investors are more reluctant to fund ventures since alternative opportunities are promising. This causes the efficient investment level \(k_1\) to sink. However, once those ventures do clear the hurdle it is efficient to invest more, so \(k_2\) rises. The net effect is that the ratio \(k_1/k_2\) sinks. Ultimately good outside opportunities allow VCs to withhold early round investments relative to later round investments. Similarly, good outside opportunities for the entrepreneur make it tempting for him to abandon projects with low early returns, and therefore this will cause him to invest less effort and resources into the project in the early stage. Ultimately, the outside options capture the opportunity cost of investment, and therefore measure the tolerance for poor projects. With high outside options, this tolerance is low, and therefore both parties invest less in the first stage.

In practice, there is wide heterogeneity among venture capitalists and entrepreneurs in terms of their outside options. For example, VC firms with successful records in bringing new ventures to an IPO and generating outside profits for their limited partners will often enjoy high outside options. These VC firms are routinely flooded with capital from limited partners as well as with proposals from many different entrepreneurs.\(^7\) Similarly, entrepreneurs vary in their outside options as well. Successful managers at existing companies or entrepreneurs with a prior record of performance in new companies will no doubt enjoy multiple offers from management teams and VCs alike. Provided it is possible to measure the outside options of the VC or the entrepreneur, Proposition 2

\(^7\)The VC firms that funded the major internet companies in the late 1990s such as Kleiner Perkins or Sequoia Capital enjoy higher outside options than less successful VC firms.
gives a clear prediction on how the ratio of investment levels over stages will vary with these outside options.

To what extent does the first best analysis here prevent or hinder empirical verification of the predictions? In particular, is it necessary to specify the breakdown of the outside options between the two parties in order to test Proposition 2? Sorensen (2007) structurally estimates a two-sided matching model against Silicon Valley VCs and entrepreneurs. He finds that high quality VCs match with high quality entrepreneurs, i.e. he finds evidence of positive sorting. This suggests that the outside options of the entrepreneur and VC “move together” — when the VC has high outside options (high quality), so does the entrepreneur. This makes Proposition 2 especially relevant, as the parameter of interest is the total outside option of both parties. These options will be high among parties of high quality and low among parties of low quality.

Now consider what happens with an increase in the output variance, i.e. the variance on the error distribution $g$. For the remaining implications, let the cost of investment be quadratic ($\gamma = 2$) and the error distribution be uniform.

**Proposition 3** As the output variance increases, $k_t$ decreases while $\frac{k_1}{k_2}$ increases. If $\bar{u}_2 > \frac{V}{2}$, then $q^*$ increases.

As the variance of $g$ increases, it is clear that this will choke off investment in both stages. This is the same intuition from the Lazear and Rosen (1981) tournament model, in which increased noise reduces effort incentives. What is not obvious is whether the decrease in investment is larger in stage one versus stage two. It seems plausible that an increase in noise will cause the entrepreneur to withhold investment in the early round, and work harder in the later rounds. While this logic is compelling, it is misleading.

An increase in output variance affects later stage investment more than first stage investment. This occurs because the marginal benefit to first stage investment exceeds the marginal benefit to second stage investment (see (1)), since investment in the first stage not only affects the probability of clearing the final hurdle $\bar{q}$ but also of clearing the milestone $q^*$. Because the investor and the entrepreneur can quit the project if output does not clear $q^*$, an increase in the output variance increases the upside benefit from continuing. A larger first stage investment increases the chance of capturing this upside, and this gives an extra benefit to investing in the early stage rather than the later
stage. Therefore an increase in output variance decreases investment in both stages, but decreases late stage investment more than early stage investment.

Recall that the stage specific noise terms represent market and technological uncertainty at each stage. In reality, there is clearly a variation between different industries on the variance of this uncertainty. For example, some industries may have high uncertainty at the market level, possibly reflecting difficulty in bringing a new firm to market because of the strategic position of incumbents. On the other hand, some industries may exhibit high technological uncertainty, deriving from the production function itself; for example, the biological process of drug development may impose higher uncertainty on new biotech firms than technological uncertainty in other industries. This variation in market and technological uncertainty can be exploited to predict variation in the ratio of investment levels over stages. Finally, observe from Proposition 3 that if the investor and entrepreneur have sufficiently good outside options ($\bar{u} > \frac{V}{2}$), then they will set higher targets when the output variance increases. So in more risky industries, it is efficient to set a higher milestone to justify later round financing.

**Proposition 4** As the value of the venture $V$ increases, $q^*$ decreases while $k_1$ and $k_2$ both increase.

This comparative static is perhaps the most straightforward. As the venture becomes more valuable, it is efficient to invest more in each round. Said differently, each party is more willing to invest more and bear a higher cost of investment if the resulting benefit increases. The proof of the proposition shows that as the variance in the noise terms becomes sufficiently large, the ratio of early to late investments $k_1/k_2$ does not vary with $v$. Therefore, even though the VC invests more in each stage, the ratio of investments across stages eventually stays constant. On top of this, higher valuation ventures should exhibit lower milestones ($q^*$) between early and late stages. Therefore since $q^*$ and $k_1^*$ both increase, this increases the probability of success, since $P_1 = G(k_1^* - q^*)$. Intuitively, the venture is more valuable, and so it becomes more desirable to pass at the interim stage, as this generates surplus for both parties. Passing the interim hurdle is made easier by simultaneously increasing investments in each stage and decreasing the milestone, thereby increasing the probability of investment. It is efficient to do this precisely because the end game prize $V$ is worth more.
In practice, measuring $q^*$ directly may be difficult, as VCs may not have hard, objective criteria when deciding whether to continue funding projects or not. For example, part of the evaluation may be based on instinct for whether the project will be successful or not. Nonetheless, higher milestones are harder to clear than lower milestones, and therefore VCs that set high milestones will abandon many ventures at the interim stage. Similarly, VCs with low milestones will tell most of its entrepreneurs to continue. Therefore one such empirical proxy for $q^*$ is the number of firms abandoned at the interim stage divided by the total number of firms funded at the outset. In this sense, the milestone $q^*$ reflects the quit rate or abandonment rate of the VC and entrepreneur.

**Proposition 5** As the final hurdle $\bar{q}$ increases, $k_1$ decreases, $q^*$ increases, and $k_2$ is unchanged.

Recall that $\bar{q}$ is the final hurdle that output must clear in order for both parties to receive value from the venture, and therefore reflects the fundamental difficulty of project completion (because of market or technology factors). As the hurdle increases, it is efficient to decrease first stage investments and leave late stage investments unchanged, thus decreasing the ratio of early to late stage investments. Formally, $\bar{q}$ affects the planner’s payoffs only through the probability of success $P$. Specifically, for every $q_1$, $P(q_1) = G(q_1 + k_2 - \bar{q})$ decreases in $\bar{q}$. Therefore the expected benefit $PV$ decreases in $\bar{q}$. As the benefit sinks, the VC lowers costly investment $k_1$. In fact, a marginal increase in $\bar{q}$ has the opposite effect of a marginal increase in $V$. As $\bar{q}$ increases, the VC simultaneously decreases $k_1$ and increases $q^*$, thus lowering the probability of clearing the target, since $P_1 = G(k_1 - q^*)$. In other words, when the project’s difficulty increases, this lowers the expected benefit to the VC, so he reduces the probability of advancing the entrepreneur at the intermediate stage.

The concrete empirical prediction is that industries with higher final hurdles should observe more funding in later stages (lower $\frac{k_1}{k_2}$). Observe that this is the opposite prediction from an increase in variance, as predicted by Proposition 3. As an empirical matter, it will be important to distinguish high hurdles from high risk industries. The empirical measures of these variables may be close even though the theoretical concepts are quite different. For example, consider the market for an AIDS vaccine. It is plausible that AIDS research is both highly risky (high variance on $g$), and that it is very difficult to
discover an actual vaccine (high final hurdle $\bar{q}$). Finally, the comparative static with respect to the intermediate target is particularly elegant: $\frac{\partial q^*}{\partial \bar{q}} = 1$. For every unit increase in the final hurdle, it is efficient to increase the milestone by exactly that amount.

The last comparative static of this section involves the cost of investments $\lambda$. On the VC side, the cost of investment includes the transaction costs of deploying capital from existing funds, as well as time and labor spent attracting additional capital from institutional investors through raising a new fund. For example, the large influx of capital from public equity into private equity over the last twenty years (Prowse (1998)) has made it easier for VCs to raise new funds, and constitutes a reduction in the cost of investment $\lambda$. As such, Proposition 6 predicts that second stage investments will increase.\(^8\) For the entrepreneur, the cost of investment includes his cost of effort as well as the cost of deploying his own capital in the firm. For example, suppose in the very early stage of the venture, the entrepreneur finances the project with his own savings and a small loan from the bank. Interest rates that govern the bank loan will affect his cost of investment, and hence higher interest rates correspond to higher $\lambda$, and according to Proposition 6, this results in the VC setting a higher milestone and making a larger second round investment. Recall that the cost of investment is $C(k_t) = \frac{1}{2}(k_t)^2$, so $\lambda$ is a parameter that scales the cost of investment.

**Proposition 6** *As the cost of investment $\lambda$ increases, $k_2$ decreases and $q^*$ increases.*

Market factors which increase the cost of investment will decrease later round investments and increase the milestone $q^*$.

5 Ability

When projects fail, it is difficult to separate technology failure from employee failure. This happens because output measures cannot disentangle the entrepreneur’s ability from technological uncertainty when assessing the performance of a project. This section models optimal dynamic decision-making when ability and uncertainty are impossible

\(^8\)According to Prowse (1998), the venture capital market grew from $600 in 1980 to over $4 billion in 1984. In fact, the market grew five-fold from the years 1980 to 1984 itself. The volume of venture capital available post-1980 dwarfs the pre-1980 levels.
to observe separately. Ability is a shorthand for project-specific ability, measuring the quality of the match between the entrepreneur and the project.

The benchmark model now includes an underlying ability parameter that persists through both stages. Ability $a$ is distributed according to a prior distribution $f(\cdot)$ with support $A$. Since neither the VC nor the entrepreneur know the entrepreneur’s ability, the uncertainty is symmetric. Suppose that

$$q_t = a + k_t + \epsilon_t.$$  

Clearly, higher ability increases output. More precisely, ability and investment are substitutes, so more able (higher $a$) entrepreneurs can invest less to generate the same output as less able entrepreneurs. Including a persistent ability term induces correlation in output across stages, as the error term is now effectively at $\epsilon_t$. High ability which generates high output today will also generate high output tomorrow. This is the main intuition which drives the results of this section.

For each $a \in A$, the probability of passing the intermediate target $q^*$ is

$$\Pr(q_1 > q^*) = Pr(\epsilon_1 > q^* - a - k_1) = G(a + k_1 - q^*).$$

This probability increases in $a$, so higher ability entrepreneurs are more likely to clear the target. The ex-ante probability of passing the target, averaging over $A$, is

$$P_1 \equiv \int_A G(a + k_1 - q^*)f(a)da.$$  

Conditional on output realization $q_1$, the probability of passing the final hurdle $\bar{q}$ for each $a \in A$ is

$$\Pr(q_1 + q_2 > \bar{q}|q_1) = \Pr(\epsilon_2 > \bar{q} - q_1 - a - k_2) = G(k_2 + a + q_1 - \bar{q}).$$

Similarly, high ability entrepreneurs are more likely to clear the final hurdle as well. The ex-ante probability of clearing $\bar{q}$ averaging over all $a \in A$ is

$$P(q_1) \equiv \int_A G(k_2 + a + q_1 - \bar{q})f(a|q_1)da.$$  

Let $S(q_1, k_2)$ be the continuation surplus function, i.e. the expected surplus of continuing, given that first stage output is $q_1$. This is the same function as before, except now it is necessary to take expectations over $a$. So

$$S(q_1, k_2) = V \int_A G(k_2 + a + q_1 - \bar{q})f(a|q_1)da - C(k_2).$$
It is efficient for the entrepreneur to continue on the project if the continuation surplus exceeds the outside options, or if \( S(q_1) \equiv S(q_1, k^*_2) \geq \bar{u}_2 \). As before, a cutoff strategy is a target \( q^* \) that the planner sets, such that the entrepreneur continues to work if \( q_1 > q^* \) but not otherwise. If \( S(q_1) \) is strictly increasing, then the entrepreneur will use a cutoff strategy.

The posterior mean \( E[a|q_1] \) measures the entrepreneur’s revised estimate of his ability after the first stage. If the posterior mean increases with \( q_1 \), then high early output signals high ability. In other words, with higher realizations of \( q_1 \), the entrepreneur updates his posterior mean, giving him a revised estimate of his ability. A high \( q_1 \) is more likely to emerge from an entrepreneur with a high ability, so the entrepreneur will revise his posterior mean upward. A fairly weak condition that guarantees that \( E[a|q_1] \) increases with \( q_1 \) is the monotone likelihood ratio property of the posterior density.

**Definition 1** The posterior \( f(a|q_1) \) satisfies monotone likelihood (MLRP) if the likelihood ratio \( L(a) \equiv \frac{f_{q_1}(a|q_1)}{f(a|q_1)} \) increases in \( a \).

**Proposition 7** Under MLRP, the entrepreneur uses a cutoff strategy. There exists a \( q^* \) such that it is efficient for entrepreneurs with \( q_1 > q^* \) to continue into the second stage.

Intuitively, introducing persistent ability induces correlation across stages. This correlation makes it worthwhile to cut projects with low output. The posterior mean \( E[a|q_1] \) increases in \( q_1 \) under MLRP. So if an entrepreneur sees a high \( q_1 \), he learns (in a precise Bayesian sense) of his high ability. Because ability persists into the second stage, he will most likely get a high \( q_2 \), and is therefore more likely to clear the final hurdle. So it makes sense for him to stay after a high \( q_1 \). The entrepreneur stays not only because output is high today, but because he learns of his high ability that leads to high output tomorrow. Since the analysis solves the social planner’s problem, the interests of the entrepreneur and VC are aligned. Thus the VC would also like to keep an entrepreneur with a high \( q_1 \) because he learns of the entrepreneur’s high ability, which will affect second stage output and the probability of clearing the hurdle. This ability-induced correlation across stages is sufficient to guarantee the use of a cutoff strategy.
5.1 Efficient Dynamic Resource Allocation

This section analyzes the efficient amount of investment across stages. The main result of the paper is robust even after including ability in the model. Specifically, the sorting effect of the efficient target \( q^* \) will bias investment upwards in the second stage.

To gain intuition on the problem, suppose that errors are distributed normally, so \( \epsilon_t \sim N(0, s^2) \), and the ability parameter is distributed \( a \sim N(a_0, t^2) \). Call \( s^2 \) the error variance and \( t^2 \) the prior variance, i.e. the variance on the prior distribution of ability. Calculation shows that the posterior density is normal with moments

\[
E[a | q_1] = \frac{t^2}{t^2 + s^2}(q_1 - k_1) + \frac{s^2}{t^2 + s^2}a_0;
\]

\[
\text{Var}(a | q_1) = \frac{s^2 t^2}{t^2 + s^2}.
\]

The posterior mean is a linear function of the output realization \( q_1 - k_1 \) and the prior mean \( a_0 \), placing more weight on the term with smaller variance. For example, as the prior variance decreases to zero, the posterior mean converges to the prior mean. So as the entrepreneur’s information on his prior improves, his posterior places little weight on his output realization and more weight on the prior mean. Similarly, as the error variance decreases to zero, the posterior mean converges to the realization \( q_1 - k_1 \). With a low error variance, the output realization gives an accurate signal of ability, and so the posterior mean reflects this. Solving the social planner’s problem shows the paper’s main result is robust.

**Proposition 8** It is efficient to invest more in the second stage \( (k_1^* < k_2^*) \).

The same intuition from the main result earlier in the paper holds here: the possibility of halfway termination lowers the marginal return to investment in stage one, and so the VC shades investment downward in the first stage. Sorting guarantees that the more able entrepreneurs advance, and they invest more because they no longer face the threat of termination that they did in the first stage. The analysis here suggests that it is efficient to invest more once all parties learn of the entrepreneurs higher output and hence higher ability.

This section has shown that the benchmark model is robust to including an ability parameter for the entrepreneur, arguably a more realistic setting. The only additional assumption needed is MLRP of the posterior density.
6 Conclusion

Staged financing is a fundamental feature of the venture capital market. VCs do not fund new ventures all at once, but instead deliver the investments in stages, forcing the project to clear a sequence of milestones in order to guarantee future funding. While the relationship between a VC and an entrepreneur is no doubt plagued by agency problems and asymmetric information, this paper shows that staged financing can be explained with a more simple and robust efficiency argument. VCs stage investments not necessarily to mitigate moral hazard nor as a response to private information, but simply because doing so maximizes total surplus. A common critique of agency models is that it takes the information environment as given, and cannot explain why both parties do not bargain or trade from second-best outcomes towards first-best outcomes. The argument here is immune to that criticism.

Not only is staged financing efficient, but it skews the allocation of investment towards later stages. This backloading of investments is an empirical regularity established in the venture capital literature. Staged financing creates the possibility of termination after the early stage, and this introduces uncertainty into the early stage. This uncertainty decreases the expected surplus in stage one, and therefore, it is efficient to invest less in stage one. Once the entrepreneur has proven his first stage output to be high ($q_1 > q^*$), this uncertainty vanishes, and expected surplus rises. Because of this, it is efficient to invest more in the later stage.

The existing empirical literature on venture capital documents that investments in later rounds exceed those in earlier rounds. Therefore, the model is consistent with existing empirical work. Moreover, the model produces a number of empirical implications that have not yet been tested, and hopefully can stimulate future empirical papers. The secondary implications of the model all predict how the ratio of investment levels over stages ($k_1/k_2$) varies with the parameters of the model, such as the outside options of both parties, the variance in the error distribution, and the difficulty of project completion. To summarize, it is efficient to invest even more in the later stages if the outside options of both parties increase, the error variance decreases, or the final hurdle $\bar{q}$ increases. I provide several suggestions for empirical proxies for the exogenous parameters of the model. This hopefully provides guidance to explain some of the variation between firms and industries in terms of their outside options, market and technological volatility, dif-
ficulty of project completion, and cost of investment. Prior agency models of venture
capital have been unable to produce these testable implications. That the model here
can be tested is both its distinguishing feature and its primary strength.

Future work in this area can extend this model in a number of promising directions.
For example, the valuation of the project $V$ is known by both parties at the outset,
though in practice the firm’s valuation is highly uncertain prior to the initial public
offering. Also, it is an open question as to how syndicates of venture capitalists investing
simultaneously in a firm will change the conclusions of this paper. I assumed throughout
that the VC acts as a single entity, though in practice a lead venture capitalist provides
the majority of the financing while secondary VCs share the risk by holding a minority
share of the equity in the firm. Ultimately, the contribution of this paper is conceptual
in nature: efficiency can explain certain features of venture capital markets more cleanly
and simply than complex moral hazard arguments. Agency theory has done much to
improve our understanding of financial contracting over the last quarter century, but
there are times when a first-best explanation can do just as well, if not better.

7 Appendix

Proof of Proposition 1

Let $S(q_1) \equiv S(q_1, k_2^*)$ and $q_1^* = k_1^* + \epsilon_1$. Since $g > 0$,

$$S'(q_1) = \frac{\partial S(q_1, k_2^*)}{\partial q_1} = P'(q_1)V = Vg(q_1 + k_2^* - \bar{q}) > 0.$$ 

So continuation surplus is strictly increasing and continuous in $q_1$. Recall that $V(q_1, q_2) \to
0$ as $q_t \to -\infty$ for some $t$. Since $\bar{u}_2 > 0$, there exists an $x$ small enough such that

$$0 < S(x) < \bar{u}_2.$$ 

Now

$$PV - C(k_1^*) - C(k_2^*) = \mathbb{E}_1 S(q_1^*) - C(k_1^*)$$

$$= \mathbb{E}_1 [\mathbb{E}_2 V(q_1^*, k_2^* + \epsilon_2) - C(k_2^*)] - C(k_1^*)$$

$$= \mathbb{E}_1 V(q_1^*, q_2^*) - C(k_2^*) - C(k_1^*)$$

$$\geq \bar{u}_1 + \bar{u}_2,$$

where the inequality follows from project feasibility. Therefore $\mathbb{E}_1 S(q_1^*) > \bar{u}_2$. By
the mean value theorem there exists a $y \in \mathbb{R}$ such that $S(y) = \mathbb{E}S(q_1^*)$, and hence
$S(y) > \bar{u}_2 > S(x)$. By the intermediate value theorem there exists a $q^* \in (x, y)$ such that $S(q^*) = \bar{u}_2$. If $S(q_1) < \bar{u}_2$, it is efficient to terminate the project. Since $S(q_1)$ is monotonically increasing in $q_1$, this holds for $q_1 < q^*$. ■

**Proof of Theorem 1.** Recall that

$$S(q_1) = P(q_1)V - C(k_2) \text{ and } P(q_1) \to 0 \text{ as } q_1 \to -\infty.$$ 

The planner solves

$$\max_{k_1, q_1} \int_{q_1}^{\infty} S(q_1, k_2)g(q_1 - k_1) dq_1 + (1 - P_1)\bar{u}_2 - C(k_1).$$

The first order conditions with respect to $q, k_2, k_1$ are

$$S(q^*) = \bar{u}_2;$$

$$\int_{q^*}^{\infty} \frac{\partial S(q_1, k_2)}{\partial k_2}g(q_1 - k_1^*) dq_1 = 0;$$

$$C'(k_1^*) = -\int_{q^*}^{\infty} S(q_1)g'(q_1 - k_1^*) dq_1 - g(q^* - k_1^*)\bar{u}_2,$$

where $S(q_1) = S(q_1, k_2^*)$, and $S'(q_1) = \frac{\partial S(q_1, k_2^*)}{\partial q_1}$.

In what follows, we write $k_t$ for $k_t^*$ for visual clarity. Substituting $S(q^*) = \bar{u}_2$ into the last equation and integrating by parts gives

$$C'(k_1) = \int_{q^*}^{\infty} S'(q_1)g(q_1 - k_1) dq_1.$$ 

From the continuation surplus function $S(q_1) = P(q_1)V - C(k_2)$,

$$S'(q_1) = g(q_1 + k_2 - \bar{q})V;$$

$$\frac{\partial S}{\partial k_2} = g(q_1 + k_2 - \bar{q})V - C'(k_2).$$

Combining these gives

$$\frac{\partial S}{\partial k_2} = S'(q_1) - C'(k_2).$$

Integrating both sides and combining with the FOC for $k_2$ yields

$$0 = \int_{q^*}^{\infty} \frac{\partial S(q_1)}{\partial k_2}g(q_1 - k_1) dq_1 = \int_{q^*}^{\infty} S'(q_1)g(q_1 - k_1) dq_1 - P_1C'(k_2).$$
where \( P_1 = \Pr(q_1 > q^*) \). Now combining with FOC for \( k_1 \) gives
\[
C'(k_1) = \int_{q^*}^{\infty} S'(q_1)g(q_1 - k_1)\,dq_1 = P_1 C'(k_2) < C'(k_2).
\]
And, since marginal costs are increasing, this means \( k_1 < k_2 \).

**Proof of Proposition 2**
Consider possible values \( \bar{u}_a \) and \( \bar{u}_b \) for \( \bar{u}_2 \), with \( \bar{u}_a \) corresponding to \( q^*_a, k^*_1, k^*_2 \). Let \( S_a(q_1) = S(q_1, k^*_2) \). And let \( \bar{u}_b \) correspond to \( q^*_b, k^*_1, k^*_2 \), where \( S_b(q_1) = S(q_1, k^*_2) \).

**Lemma 1** If \( \bar{u}_a < \bar{u}_b \), then \( q^*_a - k^*_1 < q^*_b - k^*_1 \).

**Proof:** Suppose the contrary, that \( q^*_a - k^*_1 \geq q^*_b - k^*_1 \). For clarity, write \( \varepsilon \) for \( \varepsilon_1 \). Let \( F(k^*_1, k^*_2, q^* | \bar{u}_2) \) be the value function of the social planner’s objective function, so
\[
F(k^*_1, k^*_2, q^* | \bar{u}_2) = \max_{k_1,q} \int_q^{\infty} S(q_1, k_2)g(q_1 - k_1)\,dq_1 + (1 - P_1)\bar{u}_2 - C(k_1),
\]
where \( P_1 = 1 - G(q^* - k^*_1) \).
By optimality of \( (q^*_a, k^*_1, k^*_2) \) and \( (q^*_b, k^*_1, k^*_2) \),
\[
F(k^*_1, k^*_2, q^*_a | \bar{u}_a) > F(k^*_1, k^*_2, q^*_b | \bar{u}_a) \text{ and } F(k^*_1, k^*_2, q^*_b | \bar{u}_b) > F(k^*_1, k^*_2, q^*_a | \bar{u}_b).
\]
Expanding,
\[
\bar{u}_a \int_{-\infty}^{q^*_a - k^*_1} g(\varepsilon)\,d\varepsilon + \int_{q^*_a - k^*_1}^{\infty} S_b(\varepsilon + k^*_1)g(\varepsilon)\,d\varepsilon - C(k^*_1) < \bar{u}_a \int_{-\infty}^{q^*_a - k^*_1} g(\varepsilon)\,d\varepsilon + \int_{q^*_a - k^*_1}^{\infty} S_a(\varepsilon + k^*_1)g(\varepsilon)\,d\varepsilon - C(k^*_1); \quad (A1)
\]
\[
\bar{u}_b \int_{-\infty}^{q^*_b - k^*_1} g(\varepsilon)\,d\varepsilon + \int_{q^*_b - k^*_1}^{\infty} S_a(\varepsilon + k^*_1)g(\varepsilon)\,d\varepsilon - C(k^*_1) < \bar{u}_b \int_{-\infty}^{q^*_b - k^*_1} g(\varepsilon)\,d\varepsilon + \int_{q^*_b - k^*_1}^{\infty} S_b(\varepsilon + k^*_1)g(\varepsilon)\,d\varepsilon - C(k^*_1). \quad (A2)
\]
Now, (A1) implies
\[
\int_{q_b^*-k_1^b}^{q_a^*-k_1^a} (\bar{u}_a - S_b(\varepsilon + k_1^b)) g(\varepsilon) d\varepsilon \\
+ \int_{q_b^*-k_1^b}^{\infty} (S_a(\varepsilon + k_1^a) - S_b(\varepsilon + k_1^b)) g(\varepsilon) d\varepsilon - C(k_1^a) + C(k_1^b) > 0
\]

\[
\Rightarrow \int_{q_b^*-k_1^b}^{q_a^*-k_1^a} (S_b(\varepsilon + k_1^b) - \bar{u}_a) g(\varepsilon) d\varepsilon \\
+ \int_{q_a^*-k_1^a}^{\infty} (S_b(\varepsilon + k_1^b) - S_a(\varepsilon + k_1^a)) g(\varepsilon) d\varepsilon + C(k_1^a) - C(k_1^b) < 0.
\]

And, if \(q_b^* - k_1^b \leq q_a^* - k_1^a\), then it also holds that the left-hand side is negative when \(\bar{u}_a\) is replaced by \(\bar{u}_b\), since \(\bar{u}_a < \bar{u}_b\). So
\[
\int_{q_b^*-k_1^b}^{q_a^*-k_1^a} (S_b(\varepsilon + k_1^b) - \bar{u}_b) g(\varepsilon) d\varepsilon \\
+ \int_{q_a^*-k_1^a}^{\infty} (S_b(\varepsilon + k_1^b) - S_a(\varepsilon + k_1^a)) g(\varepsilon) d\varepsilon + C(k_1^a) - C(k_1^b) < 0.
\]

But a similar calculation subtracting the left-hand side of (A2) from the right shows that the term above is positive. Contradiction. Thus, \(q_b^* - k_1^b > q_a^* - k_1^a\). ■

By the lemma, if \(\bar{u}_a < \bar{u}_b\), then
\[
\frac{(q_b^* - k_1^b) - (q_a^* - k_1^a)}{\bar{u}_b - \bar{u}_a} > 0.
\]

Taking the limits gives
\[
\frac{\partial (q^* - k_1)}{\partial \bar{u}_2} \equiv \lim_{\bar{u}_b \to \bar{u}_a} \frac{(q_b^* - k_1^b) - (q_a^* - k_1^a)}{\bar{u}_b - \bar{u}_a} \geq 0.
\]

Now \(P_1 = 1 - G(q^*-k_1)\), so
\[
\frac{\partial P_1}{\partial \bar{u}_2} = -g(q^*-k_1) \frac{\partial (q^* - k_1)}{\partial \bar{u}_2} < 0.
\]

Since \(\frac{C'(k_1)}{C'(k_2)} = P_1\), this means
\[
\frac{\partial (C'(k_1)/C'(k_2))}{\partial \bar{u}_2} = \frac{\partial P_1}{\partial \bar{u}_2} < 0.
\]
For $C(k) = \lambda k^\gamma$, $\frac{C'(k_1)}{C'(k_2)} = \left(\frac{k_1}{k_2}\right)^{\gamma-1}$, so

$$\frac{\partial P_1}{\partial \bar{u}_2} = \frac{\partial (k_1/k_2)}{\partial \bar{u}_2} \left(\gamma - 1\right) \left(\frac{k_1}{k_2}\right)^{\gamma-2} < 0 \implies \frac{\partial (k_1/k_2)}{\partial \bar{u}_2} < 0,$$

since $k_t > 0$, $\gamma > 1$. ■

**Proof of Proposition 3**

Let $g$ be uniform over $[-\beta, \beta]$, and take the cost function to be quadratic, so $C(x) = \frac{\lambda x^2}{2}$.

Then,

$$S(q_1, k_2) = V \int_{\bar{q} - q_1}^{\infty} g(q_2 - k_2) dq_2 - \frac{\lambda k_2^2}{2};$$

$$S'(q_1) = V g(\bar{q} - q_1 - k_2);$$

$$\frac{\partial S(q_1)}{\partial k_2} = S'(q_1) - \lambda k_2.$$

This gives first order conditions

$$S(q^*) = \bar{u}_2 \iff V \int_{\bar{q} - q^*}^{\infty} g(q_2 - k_2) dq_2 - \frac{\lambda k_2^2}{2} = \bar{u}_2.$$

$$\int_{q^*}^{\infty} \frac{\partial S(q_1)}{\partial k_2} g(q_1 - k_1) dq_1 = 0$$

$$\iff \int_{q^*}^{\infty} V g(\bar{q} - q_1 - k_2) g(q_1 - k_1) dq_1 = \int_{q^*}^{\infty} \lambda k_2 g(q_1 - k_1) dq_1.$$

Because $g'$ is not defined, use the equivalent formulation

$$C'(k_1) = \int_{q^*}^{\infty} S'(q_1) g(q_1 - k_1) dq_1$$

$$\iff \lambda k_1 = \int_{q^*}^{\infty} V g(\bar{q} - q_1 - k_2) g(q_1 - k_1) dq_1.$$

Take candidate $\bar{q}^*$, $\bar{k}_1$, $\bar{k}_2$ values as:

$$\bar{q}^* = \frac{8\beta \bar{u}_2 - 4\beta V + 4\bar{q}V - \frac{V^2 \beta^2}{\lambda}}{4V},$$

(A3)
\[ \tilde{k}_1 = \frac{-8\beta \bar{u}_2 + 8\beta V - 4qV + \frac{V^2}{4\lambda}}{4(4\beta^2\lambda - V)}; \quad \text{(A4)} \]

\[ \tilde{k}_2 = \frac{V}{2\beta \lambda}. \quad \text{(A5)} \]

We claim that for \( \beta \) large enough, these values satisfy the three FOCs. For sequences \( x(n), y(n) \), say \( x(n) \) is asymptotically equal to \( y(n) \) (i.e. \( x(n) \sim y(n) \)) if

\[ \lim_{n \to \infty} \frac{x(n)}{y(n)} = 1. \]

Observe first that as \( \beta \to \infty \),

\[ \lim_{\beta \to \infty} \tilde{k}_1 = 0, \quad \lim_{\beta \to \infty} \tilde{k}_2 = 0, \quad \bar{q}^* \sim \left( \frac{2\bar{u}}{V} - 1 \right) \beta. \]

Because \( \bar{u} < V \), \( \frac{2\bar{u}}{V} - 1 \in (-1, 1) \). Observe that

\[ \tilde{k}_1 - \beta < \bar{q}^* \text{ for large enough } \beta. \quad \text{(A6)} \]

This holds because \( \tilde{k}_1 - \beta \sim -\beta \), but \( \bar{q}^* \sim \alpha \beta \) for \( \alpha \in (-1, 1) \). And

\[ \tilde{k}_1 < \bar{q} - \tilde{k}_2 \text{ for large enough } \beta, \quad \text{(A7)} \]

since \( \bar{q} > 0 \), and \( \tilde{k}_1 \to 0, \tilde{k}_2 \to 0 \). Moreover,

\[ q^* > \bar{q} - \tilde{k}_2 - \beta \text{ for large enough } \beta, \quad \text{(A8)} \]

since, \( q^* \sim \alpha \beta \) for \( \alpha > -1 \), \( \bar{q} - \tilde{k}_2 - \beta \sim -\beta \). Finally,

\[ \bar{q} - q^* > \bar{q}_2 - \beta \text{ for large enough } \beta, \quad \text{(A9)} \]

since \( \bar{q} - q^* \sim (-\alpha)\beta \) for \( \alpha < 1 \), and \( \bar{q}_2 - \beta \sim -\beta \).

Now observe that for sufficiently large \( \beta \)

\[ V \int_{\bar{q} - \bar{q}^*}^{\infty} g(q_2 - \bar{k}_2) dq_2 - \frac{\lambda}{2} \bar{k}_2^2 = \frac{V}{2\beta} (\bar{k}_2 + \beta - \max(\bar{k}_2 - \beta, \bar{q} - q^*)) - \frac{\lambda}{2} \bar{k}_2^2 \]

\[ = \frac{V}{2\beta} (\bar{k}_2 + \beta - \bar{q} + \bar{q}^*) - \frac{\lambda}{2} \bar{k}_2^2 \text{ by (A9)}, \]

\[ = \bar{u}_2. \]
where the last equality follows from plugging in $\tilde{k}_2$, $\tilde{q}^*$.

**Claim 1.** For large enough $\beta$,

$$\int_{\tilde{q}^*}^{\infty} V g(\tilde{q} - q_1 - \tilde{k}_2) g(q_1 - \tilde{k}_1) dq_1 = \int_{\tilde{q}^*}^{\infty} \lambda k_2 g(q_1 - k_1) dq_1.$$ 

**Proof:** This holds iff

\[ \frac{V}{4\beta^2} (\min(\tilde{q} - \tilde{k}_2 + \beta, \beta + \tilde{k}_1) - \max(\tilde{q} - \tilde{k}_2 - \beta, -\beta + \tilde{k}_1, \tilde{q}^*)) = \frac{\lambda k_2}{2\beta} (\tilde{k}_1 + \beta - \max(\tilde{q}^*, \tilde{k}_1 - \beta)), \]

which, by (A6), (A7), and (A8) holds iff

\[ \frac{V}{4\beta^2} (\tilde{k}_1 + \beta - \tilde{q}^*) = \frac{\lambda k_2}{2\beta} (\tilde{k}_1 + \beta - \tilde{q}^*), \]

which holds iff $\frac{V}{4\beta^2} = \frac{\lambda k_2}{2\beta}$, which holds iff $\tilde{k}_2 = \frac{V}{2\beta\lambda}$. ■

**Claim 2.** For large enough $\beta$,

$$\lambda \tilde{k}_1 = \int_{\tilde{q}^*}^{\infty} V g(\tilde{q} - q_1 - \tilde{k}_2) g(q_1 - \tilde{k}_1) dq_1.$$ 

**Proof:** As before in Claim 1, this holds iff

$$\frac{V}{4\beta^2} (\tilde{k}_1 + \beta - \tilde{q}^*) = \lambda \tilde{k}_1.$$ 

Plugging in $\tilde{k}_1$ and $\tilde{q}^*$ confirms that this holds. ■

So, for large enough $\beta$, $\tilde{k}_1$, $\tilde{k}_2$, $\tilde{q}^*$ satisfy the three FOCs. Observe that $\tilde{k}_2 \sim \frac{\gamma}{\beta}$, $\tilde{k}_1 \sim \frac{\delta}{\beta}$, and $\frac{\tilde{k}_1}{\tilde{k}_2}$ approaches $\frac{V - u}{V}$.

In fact,

$$\frac{\tilde{k}_1}{\tilde{k}_2} = \frac{\beta \lambda (8\beta (V - \bar{u}_2) - 4\bar{q}V + \frac{V^2}{\beta \lambda})}{2V (4\beta^2 \lambda - V)};$$

$$\frac{\partial \tilde{k}_1}{\partial \beta} = \frac{2\lambda (4\bar{u}_2 \beta - 6V \beta + \bar{q} (V + 4\lambda \beta^2))}{(V - 4\lambda \beta^2)^2} > 0.$$
As $\beta \to \infty$, this goes as $\frac{J}{\beta^2}$ where $J$ is positive.

Eventually, increasing the width of the support of $\varepsilon_t$ makes $\frac{k_1}{k_2}$ increase to some asymptote $1 - \frac{\bar{a}}{V} < 1$. ■

**Proof of Proposition 4**

The same argument from the proof of Proposition 3 shows that the candidate values $\tilde{q}^*$, $\tilde{k}_1$, and $\tilde{k}_2$ given by (A3), (A4), (A5) will satisfy the first order conditions. Straightforward computations show that

$$\frac{\partial q^*}{\partial V} \text{ and } \frac{\partial k_2^*}{\partial V} > 0.$$  

Furthermore, for large $\beta$:

$$\frac{\partial k_1}{\partial V} \sim \frac{1}{\beta \lambda} > 0 \quad \text{and} \quad \frac{\partial (k_1/k_2)}{\partial V} \to 0. ■$$

**Proof of Proposition 5**

Straightforward computations on the candidate values $\tilde{q}^*$, $\tilde{k}_1$, and $\tilde{k}_2$ defined in (A3), (A4), and (A5) in the proof of Proposition 3 shows that

$$\frac{\partial k_2}{\partial \tilde{q}} = 0, \quad \frac{\partial q^*}{\partial \tilde{q}} = 1, \quad \frac{\partial k_1}{\partial \tilde{q}} = -\frac{V}{4\beta^2 \lambda - V}$$

for sufficiently large $\beta$. ■

**Proof of Proposition 6**

Straightforward calculations on the candidate values $\tilde{q}^*$, $\tilde{k}_1$, and $\tilde{k}_2$ from (A3), (A4), and (A5) from the proof of Proposition 3 show that

$$\frac{\partial q^*}{\partial \lambda} = \frac{V^2}{\beta \lambda^2} > 0 \quad \text{and} \quad \frac{\partial k_2}{\partial \lambda} = \frac{-V}{2\beta^2 \lambda^2} < 0. ■$$
Lemma 2 Let (MLRP) hold. If $z(a)$ is strictly increasing, then

$$
\int_{A}^{\infty} z(a)f_{q_1}(a|q_1)da > 0.
$$

Proof: Let $f(\cdot | \cdot)$ denote the conditional density of $a$, given $q_1$, and let $\varphi(\cdot | \cdot)$ be the conditional density of $q_1$, given $a$. By definition, the likelihood ratio is given by

$$
L(a) = \frac{f_{q_1}(a|q_1)}{f(a|q_1)}.
$$

By Bayes Rule,

$$
f(a|q_1) = \frac{\varphi(q_1|a)f(a)}{\int_{A}^{\infty} \varphi(q_1|a')f(a')da'}.
$$

Now

$$
f_{q_1}(a|q_1) = \frac{\varphi'(q_1|a)f(a)\int_{A}^{\infty} \varphi(q_1|a')f(a')da' - \varphi(q_1|a)f(a)\int_{A}^{\infty} \varphi'(q_1|a')f(a')da'}{(\int_{A}^{\infty} \varphi(q_1|a')f(a')da')^2}.
$$

Integrating over $A$ gives

$$
\int_{A}^{\infty} f_{q_1}(a|q_1)da = 0. \quad (A10)
$$

So $f_{q_1}$ assumes positive and negative values, and therefore so does $L(a)$. Let $a, \bar{a}$ be the lower and upper limits of $A$, respectively (they may be $\pm \infty$). By (MLRP), $L(a)$ is increasing, so there exists an $a^*$ such that $\{a : L(a) < 0\} = (a, a^*)$ and $\{a : L(a) > 0\} = (a^*, \bar{a})$. By definition of $L(a)$, $f_{q_1} > 0$ if and only if $L(a) > 0$. Hence $\{a : f_{q_1}(a|q_1) > 0\} = (a^*, a)$ and $\{a : f_{q_1}(a|q_1) < 0\} = (\bar{a}, a^*)$. Rewrite (A10) as

$$
\int_{a}^{a^*} f_{q_1}(a|q_1)da = -\int_{a^*}^{a} f_{q_1}(a|q_1)da. \quad (A11)
$$

Now $f_{q_1}(a|q_1) < 0$ if $a < a^*$, so $|f_{q_1}(a|q_1)| = -f_{q_1}(a|q_1)$ for $a < a^*$. Integrate both sides over $(a, a^*)$ and combine with (A11) to get

$$
\int_{a}^{a^*} f_{q_1}(a|q_1)da = \int_{a}^{a^*} |f_{q_1}(a|q_1)da. \quad (A12)
$$

Now $z(a)$ is strictly increasing, so

$$
\int_{a}^{a^*} z(a)|f_{q_1}(a|q_1)da < \int_{a}^{a^*} z(a^*)|f_{q_1}(a|q_1)da = \int_{a}^{a^*} z(a^*)f_{q_1}(a|q_1)da < \int_{a}^{a^*} z(a)f_{q_1}(a|q_1)da.
$$
Taking the left hand side over to the right gives
\[
\int_{\mathcal{A}} z(a) f_{q_1}(a|q_1) da \equiv \int_{a^*}^\infty z(a) f_{q_1}(a|q_1) da + \int_a^{a^*} z(a) f_{q_1}(a|q_1) da
\]
\[
= \int_{a^*}^\infty z(a) f_{q_1}(a|q_1) da - \int_a^{a^*} z(a) f_{q_1}(a|q_1) da > 0.
\]

\[\blacksquare\]

**Proof of Proposition 7**

Let \( X \equiv \{ q_1 : S(q_1) > \bar{u}_2 \} \) be the continuation set and \( P_1 = \Pr(X) \). That is, the entrepreneur advances to the second stage if \( q_1 \in X \). Then the entrepreneur solves

\[
\max_{k_t, q_1^*} \mathbb{E}_a \left[ \int_{\mathcal{X}} S(q_1) g(q_1 - a - k_1) dq_1 + \bar{u}_2(1 - P_1) - C(k_1) \right],
\]

where

\[
S(q_1) = V \int_{\mathcal{A}} G(k_2 + a + q_1 - \bar{q}) f(a|q_1 - k_1) da - C(k_2).
\]

Taking the derivative with respect to \( q_1 \) and substituting in the previous expression for \( C'(k_2) \) gives \( S'(q_1) = W + Y \) where

\[
W \equiv V \int_{\mathcal{A}} G(k_2 + a + q_1 - \bar{q}) f(a|q_1 - k_1) da;
\]

\[
Y \equiv V \int_{\mathcal{A}} G(k_2 + a + q_1 - \bar{q}) f_{q_1}(a|q_1 - k_1) da.
\]

Since \( V > 0 \), \( W \) is positive. Now the function

\[
z(a) \equiv VG(k_2 + a + q_1 - \bar{q})
\]

increases in \( a \). By Lemma 2, this means

\[
Y = \int_{-\infty}^{\infty} z(a) f_{q_1}(a|q_1 - k_1) da > 0.
\]

Thus, \( S'(q_1) > 0 \). Clearly \( S \) is continuous. By the intermediate value theorem, there exists a \( q^* \) such that \( S(q^*) = \bar{u}_2 \). Hence \( X = \{ q_1 : q_1 > q^* \} \). \[\blacksquare\]

**Proof of Proposition 8**
Let us define (for $t = 1, 2$)

$$\xi_t \equiv a + \epsilon_t, \quad \Psi_t(x) \equiv \Pr(\xi_t \leq x), \quad \text{and} \quad \psi_t(x) \equiv \Psi'_t(x).$$

Let

$$P_1 \equiv \Pr(q_1 \geq q \mid q, k_1) = 1 - \Psi_1(q - k_1);$$

$$P_2 \equiv \Pr(q_1 + q_2 \geq \bar{q} \mid (q_1 \geq q), k_t, q)$$

$$= \frac{1}{P_1} \int_{q-k_1}^{\infty} [1 - \Psi_2(\bar{q} - k_1 - k_2 - \xi_1)] \psi_1(\xi_1) d\xi_1;$$

$$P \equiv P_1 P_2.$$

Then the problem is

$$\max_{k, q} PV - C(k_1) + (1 - P_1)\bar{u}_2 - P_1 C(k_2),$$

with FOCs

$$-V[1 - \Psi_2(\bar{q} - k_2 - q)] \psi_1(q - k_1) + \psi_1(q - k_1)\bar{u}_2 + \psi_1(q - k_1)C(k_2) = 0; \quad (A13)$$

$$V \frac{\partial P}{\partial k_1} - C'(k_1) - \psi_1(q - k_1)\bar{u}_2 - \psi_1(q - k_1)C(k_2) = 0; \quad (A14)$$

$$V \frac{\partial P}{\partial k_2} - P_1 C'(k_2) = 0. \quad (A15)$$

Summing up (A13) and (A14) and integrating by parts gives

$$V \int_{q-k_1}^{\infty} \psi_2(\bar{q} - k_1 - k_2 - \xi_1) \psi_2(\xi_1) d\xi_1 = C'(k_1). \quad (A16)$$

Now

$$\frac{\partial P}{\partial k_1} = \frac{\partial P}{\partial k_2} + (1 - \Psi_2(\bar{q} - k_2 - q)) \psi_1(q - k_1); \quad (A17)$$

$$\frac{\partial P}{\partial k_2} = \int_{q-k_1}^{\infty} \psi_2(\bar{q} - k_1 - k_2 - \xi_1) \psi_1(\xi_1) d\xi_1. \quad (A18)$$

Hence, by (A15) and (A16), $P_1 C'(k_2) = C'(k_1)$. So if $C'' > 0$, $k_1 < k_2$. ■
References


