.

Picking Winners in Rounds of Elimination

Suzanne Scotchmer[1]

October 5, 2005

Abstract: A defect of intellectual property as an incentive mechanism is that researchers must pay the costs of research long before receiving a reward. Self-finance is often not an option. For potential funders such as venture capitalists and public sponsors, the problem is to find the most promising researchers or projects, and to do so before the research is complete or the commercial prospects certain. Funding involves rounds of elimination, which serve as a screening mechanism. In this paper I study elimination mechanisms with and without memory, and show how the structure of elimination rounds involves a tradeoff between screening and costs. Public sponsors and venture capitalists may deviate in different ways from what is optimal.

[1]Department of Economics and GSPP, University of California, 2607 Hearst Ave, Berkeley, CA 94720-7320. scotch@berkeley.edu. This paper was funded by the Toulouse Network for Information Technology and by the Kaufmann Foundation to the Lester Center of IMIO at Haas Business School. Please cite as "preliminary draft" with date. Before citing, please check the author's website for updates: http://socrates.berkeley.edu/~scotch.

1 Introduction

Among the defects of intellectual property as an incentive system is that an innovator must fund the research up front. Since researchers are often liquidity constrained, this is a real impediment. As early as the 19th century, Thomas Edison established an "invention factory" to draw venture capital and contract funding.

In the modern world, there are two major ways to overcome liquidity constraints and to minimize the need for self-finance. On the private side, the solution is venture capital and contract research. On the nonprofit side, a solution is grant funding from the government and foundations.

In order to make money, venture capitalists must spend their funds on productive researchers. Public funding agencies such as NSF and NIH face a similar burden to separate the wheat from the chaff. Both sets of institutions are trying to support useful research that is conceived and carried out by someone else. They thus face a common problem, which is to figure out which researchers or projects to fund. The funding must be given before the research is accomplished, the outcome is known, or, at least in the beginning, before the researcher has proved his mettle. How does the venture capitalist or grantor know that he or she is funding worthwhile projects or researchers?

A feature shared by venture funding, grant funding and the academic tenure process is that support is given in stages. In the venture-capital process, funding is typically in "rounds." If a previous round of funding did not show progress, later rounds will not be forthcoming. In the grant process, if a grantee does not perform, future grants will not be forthcoming. And in the tenure process, an assistant professor who does not perform will not be promoted. Most venture-capital-backed projects vanish after initial rounds of funding.

This paper is a first pass at trying to understand how the rounds of funding should be structured in order to optimize screening. I assume that ideas for R&D are "scarce" (Scotchmer 2004, ch. 2) in the sense that different researchers have different research programs, to which they are tethered, and that researchers are differentiated by their prospects for success, which may be unknown to both them and their funders. The prospects of a given researcher or project cannot be distinguished without some investment.

In the next two sections I compare the screening properties of two funding schemes, a double-elimination scheme in which a company or project will be brought to fruition if it demonstrates sufficient progress in each of two preliminary rounds, and a single-elimination scheme with a single demonstration period of twice the length. In both schemes, the standard for success may be a choice variable which affects both the number and quality of researchers that survive. I consider double-elimination schemes with and without "memory," where memory means that the standard for success in round 2 can take account of the level of success in round 1. Without memory, the standard for success in round 2 must be the same for all survivors of round 1. With memory, the standard in round 2 can be weakened to account for a particularly strong showing in round 1. The question is whether there is any sense in which the double-elimination scheme is more or less effective than the single-elimination scheme with a longer demonstration period.

In section 2, I focus on the lumpiness of discoveries. Researchers are differentiated by their Poisson hit rates of discoveries, and the objective is to find the researchers with the higher hit rates. Assuming that observable discoveries are relatively rare, so that any positive number is taken as success, I show that the double-elimination scheme generates a higher-quality population of survivors than the single-elimination sheme with longer duration, but also has fewer survivors. Thus, the double-elimination scheme will be appropriate if the objective is to weed out a lot of researchers or projects, while also finding the better ones.[2]

However, the sponsor may want to increase the number of survivors with as little erosion of quality as possible. With lumpy research outputs, it is difficult to adjust the number of survivors. In section 3, I assume that there is a more finely grained observation of research productivity, normally distributed, so that the same number of survivors can be assured in both elimination schemes. I show that, with equal numbers of survivors, their expected quality is greater in the single-elimination scheme of longer duration than in the double elimination scheme. Similarly, if the standards are adjusted to equalize the average quality of survivors in both schemes, the single-elimination scheme has more survivors. Thus, the main result is to show that, in these senses, the single-elimination scheme is better at distinguishing high-quality researchers or projects from low-quality researchers or projects.

However, this does not mean that funding in rounds is suboptimal. Rounds of funding

[2]If discovery is according to a Poisson process, rounds of elimination involve some loss relative to a mehcanism that takes account of the dates of discovery. Nevertheless, many screening mechanisms have rounds of elimination with fixed duration.

eliminate unpromising researchers earlier, and thus reduce costs. In section 4, I discuss how the theorems on survivorship apply to VC financing and to incubator programs such as SBIR grants and DARPA funding.

In Section 5 I contrast the stage funding considered in this paper with grant funding that requires no self-finance at all, such as the grants administered by the NIH and NSF.

2 Screening with Lumpy Discoveries

I refer to the two elimination schemes as a *double-elimination scheme* and a *single-elimination scheme.* The double-elimination scheme can be with or without memory. In both double-elimination schemes, there are two rounds of funding. The researcher receives funding to found the company if he is successful in both rounds. In the single-elimination scheme, the intermediate round of funding is skipped. A researcher receives funds to found a company if he had at least two successes in the combined two periods.

If the double-elimination scheme is with memory, survival in round two is easier if the firm had a good showing in round one. Without memory, the funder only knows that the firm survived the first round, but not whether it survived easily.

I show that the expected ability of the survivors is higher in the double-elimination scheme, with or without memory, than in the single-elimination scheme. This shows that screening can be relatively accurate even if it saves costs by eliminating unproductive researchers early. On the other hand, in the double-elimination scheme, the sponsor ends up with fewer funded researchers.

To model lumpy discoveries, I suppose that each researcher is endowed with an ability parameter λ which determines the Poisson rate at which he has observable successes. Suppose the Poisson parameter λ is distributed among researchers according to a distribution function with density h on $[0, \infty]$.

The probability that the researcher has at least one success in a single period is $1 - F(0, \lambda) = 1 - e^{-\lambda}$. The probability that a researcher receives at least one success in each of two successive periods is $(1 - F(0, \lambda))^2$. Now double the time period and the hit rate. Then $1 - F(1, 2\lambda) = 1 - e^{-2\lambda}(1 + 2\lambda)$ is the probability that there are 2 or more successes in a combined period of twice the duration. The probability that the researcher survives the double-elimination without memory is the probability that he has at least one success in each of two periods, namely, $\left[1 - e^{-\lambda}\right]^2$. The probability that the researcher survives the double-elimination scheme with memory is the probability that he receives 2 or more successes in the first period plus the probability that he receives 1 success in the first period and at least 1 success in the second period, that is, $\left(1-e^{-\lambda}\left(1+\lambda\right)\right)+\lambda e^{-\lambda}\left(1-e^{-\lambda}\right)=1-e^{-\lambda}-\lambda e^{-2\lambda}$.

Let D and $\bar{D}$ be the respective probabilities that a random researcher survives the double-elimination schemes with and without memory, and let S be the probability of surviving the single elimination in a longer period.

$$
\begin{aligned}
D &= \int_0^{\infty}\left[1-e^{-\lambda}\right]^2\ h(\lambda)\ d\lambda \\
\bar{D} &= \int_0^{\infty}\left[1-e^{-\lambda}-\lambda e^{-2\lambda}\right]\ h(\lambda)\ d\lambda \\
S &= \int_0^{\infty}\left[1-e^{-2\lambda}\left(1+2\lambda\right)\right]\ h(\lambda)\ d\lambda
\end{aligned}
$$

It can easily be checked that $D<\bar{D}<S$. Intuitively, the first inequality follows because the researcher may survive in the double-elimination scheme with memory even if he does not survive in the double-elimination scheme without memory. His first-round performance may compensate for bad second-round performance. The second inquality follows because the researcher may throw himself out of the pool in the first round of elimination, when he could have recovered if given a chance in a second round.

The expected abilities of survivors under the double-elimination scheme without memory and with memory are

$$
\begin{aligned}
E_D(\lambda) &= \frac{1}{D}\int_0^{\infty}\lambda\ \left[1-e^{-\lambda}\right]^2\ h(\lambda)\ d\lambda \\
E_{\bar{D}}(\lambda) &= \frac{1}{\bar{D}}\int_0^{\infty}\lambda\ \left[1-e^{-\lambda}-\lambda e^{-2\lambda}\right]\ h(\lambda)\ d\lambda
\end{aligned}
$$

The expected ability of survivors under the single-elimination scheme is

$$E_S(\lambda)=\frac{1}{S}\int_0^{\infty}\lambda\ \left[1-e^{-2\lambda}\left(1+2\lambda\right)\right]\ h(\lambda)\ d\lambda$$

We will now show that the double-elimination scheme results in higher expected ability for survivors than the single-elimination scheme, whether or not the double-elimination scheme has memory. However, fewer researchers survive.

Proposition 1 *Suppose that each researcher makes discoveries according to a Poisson process with an indiosyncratic hit rate λ, and the researchers' hit rates have a random distribution with density h. The expected ability of researchers who survive the double-elimination* *scheme, with or without memory, is higher than of those who survive the single-elimination scheme, but there are fewer survivors in the double-elimination scheme.*

$$E_D(\lambda) > E_S(\lambda), \quad E_{\bar{D}}(\lambda) > E_S(\lambda), \quad \text{and} \quad D < \bar{D} < S$$

Proof: We have already argued that $D < \bar{D} < S$.

We first show that $E_D(\lambda) > E_S(\lambda)$. That is, the double-elimination scheme without memory leads to higher-ability survivors than the single-elimination scheme. Define a function G by

$$G(\lambda) = \frac{\left[1 - e^{-\lambda}\right]^2}{D} - \frac{\left[1 - e^{-2\lambda}(1 + 2\lambda)\right]}{S}$$

Then G is a continuous function. Since each of the terms in G integrates to 1 when multipled by $h(\lambda)$, there are domains of λ with positive measure where G is positive and negative respectively. We will show that G is negative for low values of λ and positive for high values of λ. That is, there exists m such that $G(\lambda) < 0$ for $\lambda \in (0, m)$ and $G(\lambda) > 0$ for $\lambda \in [m, \infty)$. It then follows that $E_D(\lambda) > E_S(\lambda)$, since the low values of λ are being weighted more heavily with the density $\frac{\left[1-e^{-2\lambda}(1+2\lambda)\right]}{S} h(\lambda)$ and the high values of λ are being weighted more heavily with the density $\frac{\left[1-e^{-\lambda}\right]^2}{D} h(\lambda)$.

Write G as

$$G(\lambda) = \frac{\left[1 - e^{-2\lambda}(1 + 2\lambda)\right]}{D} \left[\frac{\left[1 - e^{-\lambda}\right]^2}{\left[1 - e^{-2\lambda}(1 + 2\lambda)\right]} - \frac{D}{S}\right]$$

Then, since $\left[1 - e^{-2\lambda}(1 + 2\lambda)\right] / D > 0$, it is enough to show that the bracketed term is increasing with λ or indeed that $\left[1 - e^{-\lambda}\right]^2 / \left[1 - e^{-2\lambda}(1 + 2\lambda)\right]$ is increasing. The derivative of the logarithm is

$$\frac{d}{d\lambda} \log \frac{\left[1 - e^{-\lambda}\right]^2}{\left[1 - e^{-2\lambda}(1 + 2\lambda)\right]} = \frac{2e^{-\lambda}}{\left[1 - e^{-\lambda}\right]} - \frac{4\lambda e^{-2\lambda}}{\left[1 - e^{-2\lambda}(1 + 2\lambda)\right]}$$

which is positive for $\lambda > 0$.

We now show that $E_{\bar{D}}(\lambda) > E_S(\lambda)$. Reasoning as above, it is enough to show that

$$\frac{\left[1 - e^{-\lambda} - \lambda e^{-2\lambda}\right]}{\left[1 - e^{-2\lambda} - 2\lambda e^{-2\lambda}\right]}$$

is increasing. The derivative of the logarithm satisfies

$$e^{\lambda} \frac{d}{d\lambda} \log[\cdot] = \frac{1 + 2\lambda e^{-\lambda} - e^{-\lambda}}{\left[1 - e^{-\lambda} - \lambda e^{-2\lambda}\right]} - \frac{4\lambda e^{-\lambda}}{\left[1 - e^{-2\lambda} - 2\lambda e^{-2\lambda}\right]}$$

which is positive for all $\lambda > 0$. $\square$

Thus, if the objective is only to maximize the ability of funded researchers while reducing cost, it is better to have two rounds of elimination, dumping some of the competitors after the first round. But this also reduces the total number of survivors.

Proposition 1 seems inconsistent with the intuition that "more information is better than less." If it is costless to gather information, as in allowing all researchers to continue both periods in a single round of elimination, then one would suspect that better decisions could be made as to which researchers ultimately survive. There has to be some sense in which that intuition is correct. It is not entirely obvious because, in the double-elimination scheme, it is only some of the agents who are not allowed to continue. The relative merits of the two schemes depend on who is eliminated. The next section makes precise the sense in which more information is better than less.

3 Screening with Finely Grained Research Outputs

The assumption in the previous section is that discoveries are lumpy. The discovery is the minimum observable research output, perhaps a patent, and it happens relatively rarely. As a consequence, the mechanism is restricted to eliminations after each period or elimination after the combined periods, but this choice affects the overall number of survivors (funded projects) as well as their expected quality. If there is a large budget to spend, the best option is a single elimination after two periods, in order to have enough survivors on which to spend the funds. However this has the disadvantage that the expected quality is lower than with two rounds of screening.

The above model does not sort out whether the survivors' higher expected ability in the double elimination scheme is because screening is more effective with two elimination rounds, or simply because there are fewer survivors. We now sort this out with more finely grained observations. We first consider screening without memory, and then consider screening with memory.

With two rounds of screening and no memory, let k_1, k_2 be the minimums in the two rounds that are required for success, and let the researcher's success rates be x_1, x_2. The researcher survives two rounds of elimination without memory if $x_1 > k_1, x_2 \geq k_2$. Let K be the number of successes required in a single round of twice the length. The same researcher survives the single elimination round if $x_1 + x_2 \geq K$.

The standards for success that mimic the situation of the previous section satisfy $k_1 = k_2$ and $k_1 + k_2 = K$. This means that the same total number of successes are required in both schemes, but to survive the double-elimination scheme, there is a restriction on how they are distributed between two periods. As in the case of lumpy research outputs, the double-elimination scheme results in higher-quality survivors, but fewer of them (Proposition 2(a)).

We will now consider the normal approximation to the Poisson distribution, so that the mean and variance of a researcher's productivity are the same. Let $\Phi(\cdot)$ be the standard normal distribution function. Let H be a distribution of means (and variances) μ with density h and nondegenerate support in $(-\infty, \infty)$.

3.1 Without memory

Conditional on ability μ and the standards for success k_1, k_2 in each round, the probability of surviving two rounds is

$$\left(1 - \Phi(\frac{k_1 - \mu}{\sqrt{\mu}})\right)\left(1 - \Phi(\frac{k_2 - \mu}{\sqrt{\mu}})\right) = \Phi(\frac{\mu - k_1}{\sqrt{\mu}})\Phi(\frac{\mu - k_2}{\sqrt{\mu}})$$

The probability that the average number of successes is greater than k in a single, longer round of elimination (e.g., $k = \frac{K}{2}$, using the above notation) is

$$1 - \Phi\left(\sqrt{2}\left(\frac{k - \mu}{\sqrt{\mu}}\right)\right) = \Phi\left(\sqrt{2}\left(\frac{\mu - k}{\sqrt{\mu}}\right)\right)$$

Denote the survival probabilities by

$$\begin{aligned} D(k_1, k_2) &\equiv \int \Phi(\frac{\mu - k_1}{\sqrt{\mu}})\Phi(\frac{\mu - k_2}{\sqrt{\mu}})\ h(\mu)\, d\mu \\ S(2k) &\equiv \int \Phi\left(\sqrt{2}\left(\frac{\mu - k}{\sqrt{\mu}}\right)\right)\ h(\mu)\, d\mu \end{aligned}$$

The expected abilities of survivors in the two regimes are

$$\begin{aligned} E_D(\mu | k_1, k_2) &= \int_{-\infty}^{\infty} \mu\ \frac{\Phi(\mu - k_1)\Phi(\mu - k_2)}{D(k_1, k_2)}\ h(\mu)\, d\mu \\ E_S(\mu | 2k) &= \int_{-\infty}^{\infty} \mu\ \frac{\Phi\left(\sqrt{2}(\mu - k)\right)}{S(2k)}\ h(\mu)\, d\mu. \end{aligned}$$

If $D(k_1, k_2) = S(K)$ then $K > k_1 + k_2$. If $K \le k_1 + k_2$, the successes (x_1, x_2) might satisfy $x_1 < k_1, x_2 > k_2$ but $x_1 + x_2 \ge K$. In that case the researcher would survive in the single-elimination scheme even though he would fail one of the rounds in the double-elimination scheme.

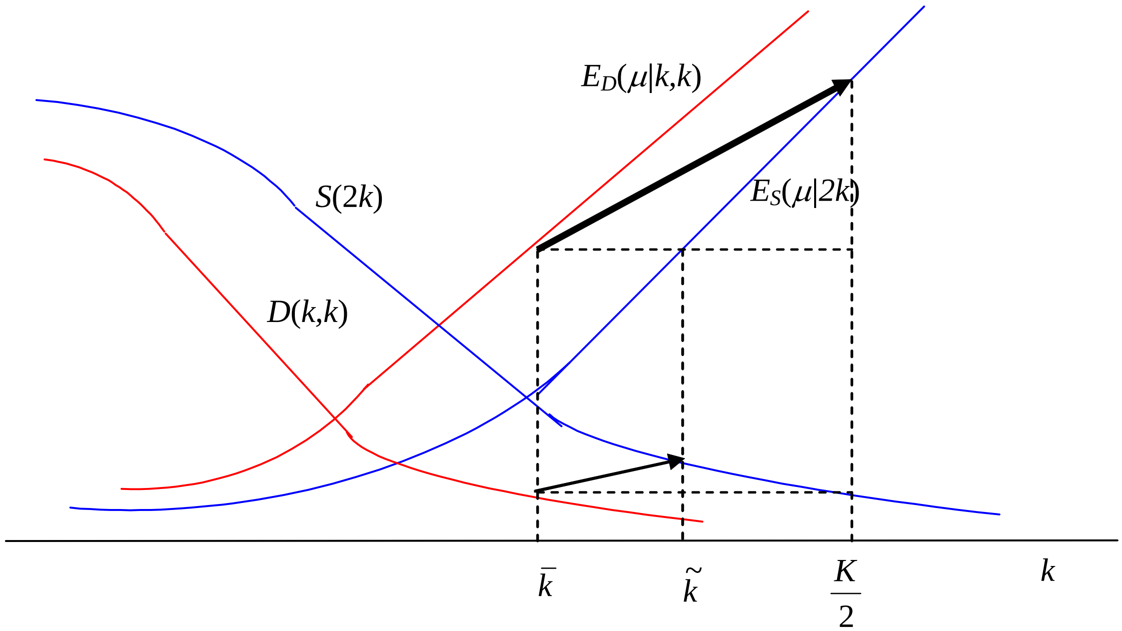

Figure 1: Single and Double Eliminations

It will also be useful for our discussion in section 4 to define a variable for the probability of surviving the first round in the double-elimination scheme: We will call this $D_1(k_1)$:

$$D_1(k_1) \equiv \int \Phi(\frac{\mu - k_1}{\sqrt{\mu}})\ h(\mu)\, d\mu$$

Proposition 2(a) below says that, as in the Poisson case, if the number of successes required to survive two rounds of elimination are the same as required to survive a single, longer round, the average ability of survivors is higher in the double-elimination. However, Proposition 2(b) says that if the standards for success are adjusted to equalize the expected numbers of survivors, this is reversed. In that case, the single-elimination scheme then has higher-ability survivors. Proposition 2(c) says that if the expected ability of survivors is the same, the double-elimination scheme has fewer survivors.

Proposition 2 is illustrated in figure 1 for the case that $k_1 = k_2$. In figure 1 it is straightforward that $D(k,k)$, $S(2k)$ are downward sloping, and that $E_S(\mu|2k)$, $E_D(\mu|k,k)$ are upward sloping. It is also straightforward that $D(k,k) < S(2k)$ for each k, and perhaps unsurprising, but worth proving, that $E_S(\mu|2k) < E_D(\mu|k,k)$ for each k (Proposition 2(a)). What is less obvious is how the difference between $D(k,k)$ are $S(2k)$ relates to the difference between $E_S(\mu|2k)$ and $E_D(\mu|k,k)$. Proposition 2(b) says hat the higher dark arrow in figure 1 points upward. Proposition 2(c) says the lower dark arrow points upward. Further, this would be true regardless of whether $k, \tilde{k}, \frac{K}{2}$ are shifted left or right. That the arrows point upward does not follow from the monotonicity properties of the functions.

Proposition 2 is proved in section 6.

Proposition 2 *Suppose that each researcher's discoveries are normally distributed and that the means of the researchers' distributions are distributed according to some H with density h. Suppose that the double-elimination scheme does not have memory.*

(a) *If survival in the double-elimination and single-elimination schemes require the same total number of successes, the expected ability of survivors in the double-elimination scheme is larger than in the single-elimination scheme, but the number of survivors is smaller.*

$$\text{Given } k_1, k_2,\ E_D(\mu|k_1, k_2) > E_S(\mu|k_1 + k_2) \text{ and } D(k_1, k_2) < S(k_1 + k_2)$$

(b) *If the standards for success in the double-elimination and single-elimination schemes ensure the same numbers of survivors, the expected ability of survivors in the single-elimination scheme is higher than in the double-elimination scheme.*

$$\text{Given } k_1, k_2, K \text{ such that } D(k_1, k_2) = S(K),\ E_D(\mu|k_1, k_2) < E_S(\mu|K)$$

(c) *If the standards of success ensure equal abilities of survivors in both schemes, the probability of surviving the single-elimination scheme is larger than the probability of surviving the double-elimination scheme, provided the latter is less than one. That is,*

$$\text{Given } k_1, k_2, K \text{ such that } E_D(\mu|k_1, k_2) = E_S(\mu|K),\ D(k_1, k_2) < S(K).$$

3.2 With memory

The same conclusions also hold if the double-elimination scheme has memory, and figure 1 also describes that case.

Let $f(k_1, k_2, \mu)$ be the probability that an agent with mean μ survives the double-elimination scheme when the first round requires $x_1 \geq k_1$ and the second round requires $x_1 + x_2 \geq k_1 + k_2$. With memory, survival at round 2 requires only that $x_2 \geq k_1 + k_2 - x_1$, which is a weaker requirement than the requirement without memory, namely, $x_2 \geq k_2$.

We again assume that the researcher's successes in each period are normally distributed with mean μ, and, for easier notation, variance 1.

$$\begin{aligned}
f(k_1,k_2,\mu) &= \frac{1}{2\pi\mu}\int_{k_1}^{\infty} e^{-\frac{1}{2}\frac{(y-\mu)^2}{\mu}} \int_{(k_1+k_2)-y}^{\infty} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\mu}} dx\ dy \\
&= \frac{1}{2\pi\sqrt{\mu}}\int_{k_1}^{\infty} e^{-\frac{1}{2}\frac{(y-\mu)^2}{\mu}} \int_{\frac{-(y-(k_1+k_2)+\mu)}{\sqrt{\mu}}}^{\infty} e^{-\frac{1}{2}z^2} dz\ dy \\
&= \frac{1}{2\pi}\int_{\frac{k_1-\mu}{\sqrt{\mu}}}^{\infty} e^{-\frac{1}{2}w^2} \int_{-w+\frac{(k_1+k_2)-2\mu}{\sqrt{\mu}}}^{\infty} e^{-\frac{1}{2}z^2} dz\ dw \\
&= \frac{1}{\sqrt{2\pi}}\int_{\frac{k_1-\mu}{\sqrt{\mu}}}^{\infty} e^{-\frac{1}{2}w^2}\Phi\left(w+\frac{2\mu-(k_1+k_2)}{\sqrt{\mu}}\right)\ dw
\end{aligned}$$

Let $\hat{f}(K,\mu)$ be the probability that the researcher survives the single-elimination scheme when survival requires $x_1+x_2>K$. Surviving the single elimination can be thought of as surviving two rounds, with $x_1>-\infty$ in the first round, and $x_1+x_2>K$ in the second round. We shall use the following as the definition of $\hat{f}$, analogously to f, although it also holds that $\hat{f}(K,\mu)=1-\Phi(\sqrt{2}\frac{(\frac{K}{2}-\mu)}{\sqrt{\mu}})$.

$$\hat{f}(K,\mu)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2}\Phi\left(y+\frac{2\mu-K}{\sqrt{\mu}}\right)\ dy$$

Let $\bar{D}(k_1,k_2)$ and $\hat{S}(K)$ $(=S(K))$ be, respectively, the expected numbers of survivors in the double-elimination and single-elimination schemes with memory, and let the expected abilities of survivors be, respectively, $E_{\bar{D}}(\mu|k_1,k_2)$ and $E_{\hat{S}}(\mu|K)$.

$$\begin{aligned}
E_{\bar{D}}(\mu|k_1,k_2) &= \int \mu\frac{f(k_1,k_2,\mu)}{\bar{D}(k_1,k_2)}h(\mu)\,d\mu \\
\bar{D}(k_1,k_2) &= \int f(k_1,k_2,\mu)\,h(\mu)\,d\mu \\
E_{\hat{S}}(\mu|K) &= \int \mu\frac{\hat{f}(K,\mu)}{\hat{S}(K)}h(\mu)\,d\mu \\
\hat{S}(K) &= \int \hat{f}(K,\mu)h(\mu)\,d\mu
\end{aligned}$$

Proposition 3 is proved in section 7.

Proposition 3 *Suppose that each researcher's discoveries are normally distributed and that the means of the researchers' distributions are distributed according to some H with density h on support $(-\infty,\infty)$. Suppose that the double-elimination scheme has memory.*

(a) *If survival in the double-elimination and single-elimination schemes require the same total number of successes, the expected ability of survivors in the double-elimination scheme* *is higher, but the number of survivors is lower. More particularly, given* k_1, k_2

$$\begin{aligned} \Phi\left(\frac{k_2-k_1}{\sqrt{2}}\right)\hat{S}\left(k_1+k_2\right) &< \bar{D}\left(k_1,k_2\right)<\hat{S}\left(k_1+k_2\right) \\ \text{and } E_{\bar{D}}\left(\mu|k_1,k_2\right) &> E_{\hat{S}}\left(\mu|k_1+k_2\right) \end{aligned}$$

(b) *If the expected number of survivors is the same in the double-elimination scheme and single-elimination scheme, the expected ability of survivors in the double-elimination scheme is smaller than in the single-elimination scheme. That is,*

$$\text{Given } k_1,k_2,K \text{ such that } \bar{D}\left(k_1,k_2\right)=\hat{S}\left(K\right),\ E_{\bar{D}}\left(\mu|k_1,k_2\right)<E_{\hat{S}}\left(\mu|K\right).$$

(c) *If the standards of success ensure equal abilities of survivors in both games, the probability of surviving the single-elimination scheme is larger than the probability of surviving the double-elimination scheme, provided the latter is less than one. That is,*

$$\text{Given } k_1,k_2,K \text{ such that } E_{\bar{D}}\left(\mu|k_1,k_2\right)=E_{\hat{S}}\left(\mu|K\right),\ \bar{D}(k_1,k_2)<\hat{S}(K).$$

Finally, the following is proved in section 8.

Proposition 4 *Suppose that each researcher's discoveries are normally distributed and that the means of the researchers' distributions are distributed according to some* H *with density* h *on support* $(-\infty,\infty)$. *Then, in expectation, there will be more survivors if the double-elimination scheme has memory than if not, but the survivors of the double-elimination scheme with memory will have lower expected ability:*

$$\text{Given } k_1,k_2,\quad E_D\left(k_1,k_2\right)<E_{\bar{D}}\left(k_1,k_2\right) \text{ and } E_D\left(\mu|k_1,k_2\right)>E_{\bar{D}}\left(\mu|k_1,k_2\right)$$

4 Applications

In a sense, Propositions 2(b)(c) and 3(b)(c) express the intuitive notion that more information is better than less. This is why a single elimination after two rounds is better than eliminations after each round. If the unlucky researchers are not eliminated after period one, they might overturn their bad luck in the second period. This is especially useful to good researchers, those with high values of μ or λ. However, in arguing that the single elimination after two periods favors the better researchers, we also have to contend with the fact that the standards for success also affect the *number* of survivors.

It should be noticed that the propositions contain no assumptions about the distributions of λ and μ except that they have nondegenerate support. (With degenerate support, all survivors would have to have the same ability.) The weak assumptions on the support will matter when the propositions about the abilities of survivors are embedded in an analysis of games where researchers may self-select to participate. This will be the next stage of research.

We now return to one of the motivating questions, which is how to fund a startup. The most immediate application of the propositions above is to publicly funded programs, where the sponsor begins by fertilizing many flowers, and then picks the best ones. SBIR grants and DARPA funding have some of that quality.

To be concrete, consider the case that the funding scheme has no memory, but there are finely grained research benchmarks – the regime of Proposition 2. Assume for simplicity that the value of the resulting technology, if the company is funded, is proportional to μ. Then, assuming that the double-elimination scheme has the same standard for success k_D in both periods, the social value generated with the single- and double-eliminations are, respectively,

$$\begin{aligned} V_S\left(2k_S\right) &= S\left(2k_S\right) E_S\left(\mu|2k_S\right) - 2c \\ V_D\left(k_D\right) &= D\left(k_D\right) E_S\left(\mu|k_D\right) - \left(1 + D_1\left(k_D\right)\right)c \end{aligned}$$

where c is the cost per period of the research.

Suppose that there is a fixed amount of funding to allocate, say M. Then the budget constraint in the single-elimination scheme is $M = 2c$, which immediately determines k_S, the number of firms that can be supported. The budget constraint in the double-elimination scheme is $M = (1 + D_1\left(k_D\right))c$, which determines k_D, the number of firms that can be supported in the double-elimination scheme. Because the sponsor funds a smaller proportion of researchers' time for each funded company in the double-elimination scheme, more such companies will be funded. That is, $D\left(k_D\right) > S\left(2k_S\right)$. This follows only from the budget constraint. However, since the double elimination scheme is a less effective screen, it is not obvious whether $D\left(k_D\right) E_D\left(\mu|k_D\right)$ is greater than or smaller than $S\left(2k_S\right) E_S\left(\mu|2k_S\right)$. Nevertheless,

Remark 1 *For sufficiently large c, the double-elimination scheme is optimal in the sense that it maximizes the social value of research outputs, net of costs, given a budget constraint.*

Venture capitalists face a different problem than public sponsors. In their research environment, the main reward to R&D is intellectual property, and the researcher is the gatekeeper to that reward. The task of the venture capitalist is to exploit the researcher's lack of liquidity to get equity in the intellectual property and resulting firm. Due to the researchers' illiquidity, the venture capitalist may be forced to make choices early. Thus, there is a natural reason that venture capitalists will "err" on the side of early eliminations, creating a funding scheme that is more akin to the double-elimination scheme described above rather than the single-elimination scheme.

I now give some preliminary remarks on what determines venture capital finance, and when it is given.

Self-finance followed by VC funding puts the researcher in a vulnerable position. By the time the researcher has shown promise, e.g., after round one, he has sunk considerable money into the project, and may have no way forward except venture capital. If there were only one venture capitalist making a take-it-or-leave-it offer, he would be in jeopardy of not covering his sunk costs. Anticipating that outcome, he would not embark on the project in the first place. The system would collapse.

Competition among venture capitalists will diminish this threat, since competition on the VC side of the market transfers bargaining power to the researcher.[3] Thus, when the researcher sells part of his company in the first round of funding, the value taken away by the VC will be commensurate with his funding contribution, rather than, for example, transferring the expected value of the company.

Researchers who can self-finance create selection problems for venture capitalists, especially if the capacity for self-finance is not observable. Researchers with high liquidity will typically not sell equity, since there is no advantage to trading ownership for money, at least if the researcher is risk neutral and provided the researcher is reasonably confident of eventual success. Thus a researcher who seeks venture capital when he can self-finance should be inherently suspect.

This line of reasoning leads to the paradoxical conclusion that venture capitalists should be more generous in their terms the more they are needed. If venture capital is not really

[3] The notion here is that "ideas are scarce" in the sense discussed in chapter 2 of Scotchmer (2004). Two agents with the same commercial prospects, captured in μ, may nevertheless have completely different projects. This is a convenient modeling assumption for my present purposes, although it would also be useful to investigate bargaining power on both sides of the market, as will occur when researchers have closely competing projects.

needed, then the researchers seeking finance are more likely to be those with unpromising projects. Venture capitalists should naturally be wary, which may undermine the venture capital industry entirely.

The easier case is when the researcher's liquidity, or lack of liquidity, is observable. Then if the researcher can sell equity to the highest bidder in a round of finance, the price of equity will depend on the statistical evidence up to that point, e.g., in what I have called the first round. Although the supposition of two rounds is obviously stylized, the model generates the insight that liquidity constraints will force the venture capitalist to make early decisions, eliminating some of the researchers at round one even when that is suboptimal.

5 Up Front Grants: A Comparison

The funding schemes discussed so far are aimed at projects that involve a series of steps, where there are intermediate indicators of a project's prospects. Because there are indicators along the way, funding can be given in rounds, and unpromising researchers or projects can be eliminated.

If there are no observable intermediate steps, or if the transactions costs of eliminating researchers in midproject are too great, projects may be funded as all-or-nothing propositions. That, of course, exacerbates the moral hazard problems. Potential grantees may overstate what they can deliver in order to get funding, and then redirect the one-time windfall to another use. This may be attractive even if it burns bridges with the funding agencies. The feature that limits moral hazard in stage funding is that the main payoff comes at the end, for example, in the form of a profitable company. The company will not materialize unless the researcher performs.

For comparison, I now review the grants model of Maurer and Scotchmer (2004), which does not rely on any self-finance. Like the rounds of funding described above, a grant system must try to avoid funding low-ability researchers. That is accomplished by self-selection of high-ability researchers to stay in the system, while low-ability researchers cheat the grantor once and are then eliminated. I will recapitulate this argument, and then point out that, when ideas are large but infrequent, the grant system is very costly. Ideas may be infrequent either because of the nature of the technology or because transactions costs makes it optimal to define a fundable idea in that way. Thus, for the large-scale projects that are typically the object of venture capital finance, a grant system would not be very effective.

Following Maurer and Scotchmer, assume for simplicity that all ideas proposed for funding have the same value and cost, (v, c), but that researchers receive ideas at different rates $\lambda \geq 0$ per period. The objective of the grant agency is to reward high-λ researchers.

When the researcher thinks of a fundable idea, she can file a grant proposal with the sponsor, and the sponsor will decide whether or not to fund it. If the funding body is a VC, the VC would then take possession of the resulting knowledge. For a researcher with research ability λ, the expected present discounted value of investing in all the ideas conceived in a given period at date t is the following, when r is the discount rate and the size of the grant per idea is ρ :

$$\frac{\lambda}{(1+r)^t}(\rho - c)$$

If the researcher proposes a research program she cannot perform, or simply chooses to shirk, she will pocket the grant without spending the cost c. But this can only be done once, since further grant funding will not be forthcoming. High-ability researchers will not cheat in this way, as they have too much to lose in future grants.

If the researcher cheats, her net loss from lost future grants would be

$$\sum_{t=1}^{\infty} \frac{\lambda}{(1+r)^t}(\rho - c) = \lambda(\rho - c) \sum_{t=1}^{\infty} \frac{1}{(1+r)^t} = \frac{\lambda}{r}(\rho - c)$$

The researcher will perform instead of pocketing the money if the cost c that could be saved by cheating is smaller than the profit on future grants that would be lost:

$$\begin{aligned} c &\leq \frac{\lambda}{r}(\rho - c) \qquad (1) \\ \text{or} \quad c &\leq \frac{\lambda}{\lambda + r}\rho \end{aligned}$$

If the inequality (1) holds for any λ, it also holds for any researcher with a higher value of λ. High-ability researchers will self-select to perform research in this system, while low-ability researchers cheat once and drop out.

For a fixed rate of idea formation λ, researchers will only perform if the reward-to-cost ratio ρ/c is high enough. And, of course, they will never perform if $\rho < c$. As compared to the amount of funding that would be required with full information, which is $\rho = c$, the grant system will be expensive. Researchers must be overrewarded in order to keep them coming back, instead of taking the one-time profit of cheating.

The grant system will be expensive if research ideas are infrequent, that is, if λ is low. To keep honest researchers in the system, the ratio ρ/c, hence $\rho - c$, must be relatively high. Researchers must be rewarded considerably more than their cost. Otherwise, they will have an incentive to propose an idea they can't deliver, accept the grant without investing, and withdraw from the system.

As compared to the high cost of grant funding, the rounds of funding described above, which start with self-finance and end with shared ownership, look very attractive.

6 Proof of Proposition 2

Instead of writing the standards for success in the double-elimination scheme as k_1, k_2, we will write them as

$$\left\{ \begin{array}{ll} k_1 = & k - a \\ k_2 = & k + a \end{array} \quad | \quad a \geq 0 \right\} \tag{2}$$

Since the probabilities of success in the double elimination game are the same if k_1 and k_2 are reversed, there is no loss of generality in restricting to $a \geq 0$.

Abusing notation, write $D(a,k)$ for $D(k-a, k+a)$ and $E_D(\mu|k,a)$ for $E_D(\mu|k-a, k+a)$. Define a function $\Delta(\cdot; k, a, K)$on $(-\infty, \infty)$ by

$$\begin{aligned} \Delta(\mu; k, a, K) &= \frac{(1-\Phi(k-a-\mu))(1-\Phi(k+a-\mu))}{D(k,a)} - \frac{1-\Phi(\sqrt{2}\left(\frac{K}{2}-\mu\right))}{S(K)} \\ &= \frac{\Phi(\mu-k+a)\Phi(\mu-k-a)}{D(k,a)} - \frac{\Phi(\sqrt{2}\left(\mu-\frac{K}{2}\right))}{S(K)} \end{aligned} \tag{3}$$

Then

$$E_D(\mu|k,a) - E_S(\mu|K) = \int_{-\infty}^{\infty} \mu \quad \Delta(\mu; k, a, K)\, h(\mu)\, d\mu$$

Since

$$0 = \int_{-\infty}^{\infty} \Delta(\mu; k, a, K)\, h(\mu)\, d\mu$$

there are domains where Δ is positive and domains where Δ is negative, and the sign of $E_D(\mu|k,a) - E_S(\mu|K)$ depends on where they are. The proofs of Proposition 2(a) and 2(b) consist in characterizing these domains, as shown in figures 2 and 3.

We will need the first and second derivatives:

$$\Delta'(\mu;k,a,K) = \tfrac{1}{D(k,a)\sqrt{2\pi}} e^{-\frac{1}{2}(\mu-k-a)^2-\frac{1}{2}(\mu-k+a)^2} \times$$
$$\left[\Phi(\mu-k-a)\, e^{\frac{1}{2}(\mu-k-a)^2} + \Phi(\mu-k+a)\, e^{\frac{1}{2}(\mu-k+a)^2}\right]$$
$$-\frac{1}{S(K)\sqrt{\pi}} e^{-\left(\mu-\frac{K}{2}\right)^2} \tag{4}$$

$$\Delta''(\mu;k,a,K) = \tfrac{1}{D(k,a)\sqrt{2\pi}} e^{-\frac{1}{2}(\mu-k-a)^2-\frac{1}{2}(\mu-k+a)^2} \times$$
$$\left\{\frac{2}{\sqrt{2\pi}} - \left[(\mu-k-a)\,\Phi(\mu-k+a)e^{\frac{1}{2}(\mu-k+a)^2} + (\mu-k+a)\,\Phi(\mu-k-a)e^{\frac{1}{2}(\mu-k-a)^2}\right]\right\}$$
$$+\frac{2\left(\mu-\frac{K}{2}\right)}{\sqrt{\pi}S(K)} e^{-\left(\mu-\frac{K}{2}\right)^2} \tag{5}$$

It will be useful in several places to use the identity

$$\begin{aligned} e^{(\mu-k)^2} e^{-\frac{1}{2}(\mu-k+a)^2} e^{-\frac{1}{2}(\mu-k-a)^2} &= e^{-a^2} \\ e^{\left(\mu-\frac{(k_1+k_2)}{2}\right)^2-\frac{1}{2}(\mu-k_1)^2-\frac{1}{2}(\mu-k_2)^2} &= e^{-\frac{1}{4}(k_2-k_1)^2} \end{aligned} \tag{6}$$

Part 2(a): First we argue that $D(k,a) < S(2k)$. For every pair of successes (x_1, x_2) such that the firm survives the double elimination, it also survives the single elimination, but not vice versa. Even if $(x_1 + x_1) \geq 2k$, it happens with positive probability that $x_1 < k - a$, where x_1 is the realization that comes first.

We now show that $E_D(\mu|k,a) - E_S(\mu|2k) > 0$. It is enough to show that for some m

$$\begin{aligned} \Delta(\mu;k,a,2k) &< 0 \text{ for } \quad \mu < m \\ \Delta(\mu;k,a,2k) &> 0 \text{ for } \quad \mu > m \end{aligned} \tag{7}$$

To show (7) we use the derivative (4), which is the following under the hypotheses of part 3(a).

$$\Delta'(\mu;k,a,2k) =$$
$$\frac{1}{D(k_1,a)\sqrt{2\pi}}\left[\Phi(\mu-k+a)\, e^{-\frac{1}{2}(\mu-k-a)^2} + \Phi(\mu-k-a)\, e^{-\frac{1}{2}(\mu-k+a)^2}\right]$$
$$-\frac{1}{S(2k)\sqrt{\pi}} e^{-(\mu-k)^2} \tag{8}$$

$\Delta'(\mu; k, a, 2k) = \frac{e^{-(\mu-k)^2}}{D(k,a)\sqrt{2\pi}} \times$

$$\left[\Phi(\mu - k + a)\, e^{-\frac{1}{2}(\mu-k-a)^2+(\mu-k)^2} + \Phi(\mu - k - a)\, e^{-\frac{1}{2}(\mu-k+a)^2+(\mu-k)^2} - \frac{\sqrt{2}D(k_1, a)}{S(2k)}\right] \quad (9)$$

First, there exists d such that

$$\Delta(\mu; k, a, 2k) < 0, \qquad \Delta'(\mu; k, a, 2k) < 0 \qquad \text{for } \mu \in (-\infty, d) \quad (10)$$

It follows from (3) that $\Delta(\mu; k, a, 2k) \to 0$ and from (8) that $\Delta'(\mu; k, a, 2k) \to 0$,as $\mu \to -\infty$. The latter follows because the first and second terms in the brackets of (9) go to zero as $\mu \to -\infty$. To see this, write the first term in brackets (similarly, the second term) as

$$\begin{aligned} \Phi(\mu - k + a)\, e^{-\frac{1}{2}(\mu-k-a)^2+(\mu-k)^2} &= \Phi(\mu - k + a)\, e^{\frac{1}{2}(\mu-k+a)^2} e^{-\frac{1}{2}(\mu-k+a)^2-\frac{1}{2}(\mu-k-a)^2+(\mu-k)^2} \\ &= \Phi(\mu - k + a)\, e^{\frac{1}{2}(\mu-k+a)^2} e^{-a^2} \end{aligned}$$

where the second line uses (6). Changing variables and assuming $\mu - k + a < 0$,

$$\Phi(\mu - k + a)\, e^{\frac{1}{2}(\mu-k+a)^2} e^{-a^2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} \sqrt{\frac{z^2}{z^2 + (\mu - k + a)^2}}\, e^{-\frac{1}{2}z^2} dz$$

But the righthand side goes to zero pointwise as $\mu \to -\infty$. This shows (10).

Due to (10), it must hold that $\Delta''(\mu_1; k, a, 2k) \geq 0$ at the smallest extreme point, say μ_1, where $\Delta'\left(\mu_{1;}; k, a, 2k\right) = 0$. We now show that there is no other extreme point.

Define a function $\tilde{\Delta}''$ as the second derivative, $\tilde{\Delta}''(\mu; k, a, 2k) = \Delta''(\mu; k, a, 2k)$ on a domain $\{\mu \in (-\infty, \infty) \,|\Delta'(\mu; k, a, 2k) = 0\}$, using (8). Then, using (5), (8), and (6), $\tilde{\Delta}''(\mu; k, a, 2k)$ satisfies the following where it is defined:

$$\begin{aligned} &\tilde{\Delta}''(\mu; k, a, 2k)\sqrt{\pi}D(k, a)\, e^{(\mu-k)^2} e^{a^2} = (\mu - k)\, \tfrac{e^{a^2} D(k,a)}{S(2k)} \\ &\quad + \frac{1}{\sqrt{\pi}} + \frac{a}{\sqrt{2}}\left[\Phi(\mu - k + a)\, e^{\frac{1}{2}(\mu-k+a)^2} - \Phi(\mu - k - a)\, e^{\frac{1}{2}(\mu-k-a)^2}\right] \end{aligned} \quad (11)$$

Since the function defined by $e^{\frac{1}{2}x^2}\Phi(x)$ is increasing with x, the last two terms in (11) are positive.

Suppose that μ_1 is not the only extreme point. Let μ_2 be the next larger extreme point, so that $\Delta'(\mu_2; k, a, 2k) = 0$. Since the derivative is continuous, it would have to hold that

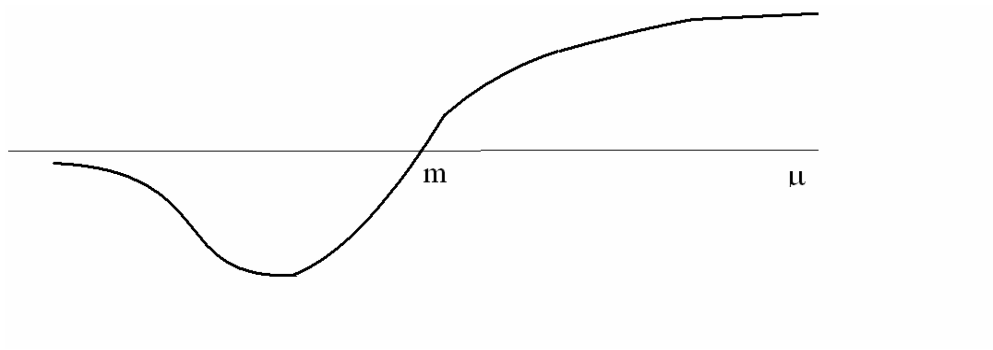

Figure 2: $\Delta(\mu; k_1, k_2, k_1 + k_2)$, $\delta(\mu; k_1, k_2, k_1 + k_2)$

$\Delta''(\mu_2; k, a, 2k) \leq 0$. But this is a contradiction, since, using (11), if $\Delta''(\mu_1; k, a, 2k) \geq 0$, then $\Delta''(\mu_2; k, a, 2k) > 0$. Thus, there is a single extreme point μ_1. In addition, $\Delta(\mu; k, a, 2k) < 0$ for $\mu \in (-\infty, \mu_1)$ and $\Delta'(\mu; k, a, 2k) > 0$ for $\mu > \mu_1$. Since $\int \Delta(\mu; k, a, 2k)\, h(\mu)\, d\mu = 0$, there are domains where $\Delta(\mu; k, a, 2k) > 0$. These features together imply (7), as shown in figure 2.

Part 2(b): Let k, a, K be such that $D(k, a) = S(K)$. Then $(k-a)+(k+a) = 2k < K$. If $K \leq (k-a)+(k+a)$, then every pair of successes (x_1, x_2) that permits success in the double elimination, $x_1 > (k-a)$, $x_2 > (k+a)$, also permits success in the single elimination, but not vice versa. With positive probability the successes (x_1, x_2) would survive the single elimination but not the double elimination. To achieve the same probabilities of success, it must therefore hold that $k < \frac{K}{2}$.

It suffices for $E_D(\mu|k, a) - E_S(\mu|K) < 0$ that (12) holds.

$$\begin{aligned} \Delta(\mu; k, a, K) &> 0 \text{ for } \quad \mu < m \\ \Delta(\mu; k, a, K) &< 0 \text{ for } \quad \mu > m \end{aligned} \tag{12}$$

In fact, we will show that Δ has the shape in figure 3.

Using $D(k, a) = S(K)$, it follows from (3) and (4) that

$$\begin{aligned} \lim_{\mu \to -\infty} \Delta(\mu; k, a, K) &= \left[\frac{0}{D(k, a)} - \frac{0}{S(K)}\right] = 0 \\ \lim_{\mu \to \infty} \Delta(\mu; k, a, K) &= \left[\frac{1}{D(k, a)} - \frac{1}{S(K)}\right] = 0 \\ \lim_{\mu \to -\infty} \Delta'(\mu; k, a, K) &= \lim_{\mu \to \infty} \Delta'(\mu) = 0 \end{aligned} \tag{13}$$

I now argue that there exist d, b such that

$$\begin{array}{ll} \Delta'(\mu;k,a,K)>0,\ \Delta(\mu;k,a,K)>0 & \text{for } \mu\in(-\infty,d) \\ \Delta'(\mu;k,a,K)>0,\ \Delta(\mu;k,a,K)<0 & \text{for } \mu\in(b,\infty) \end{array} \tag{14}$$

Using $D(k,a)=S(K)$, the derivative (4) can be written

$$\Delta'(\mu;k,a,K)\sqrt{2\pi}D(k,a)\,e^{\left(\mu-\frac{K}{2}\right)^2}=-\sqrt{2} \tag{15}$$
$$+\left[\Phi(\mu-k-a)e^{-\frac{1}{2}(\mu-k+a)^2+\left(\mu-\frac{K}{2}\right)^2}+\Phi(\mu-k+a)e^{-\frac{1}{2}(\mu-k-a)^2+\left(\mu-\frac{K}{2}\right)^2}\right]$$

Then Δ' becomes positive as either $\mu\to\infty$ or $\mu\to-\infty$, since each term in the brackets of (15) blows up. Changing variables, write the first term (similarly the second term) as

$$\begin{aligned} &\Phi(\mu-k-a)e^{\frac{1}{2}(\mu-k-a)^2}\ \times\ e^{-\frac{1}{2}(\mu-k-a)^2-\frac{1}{2}(\mu-k+a)^2+\left(\mu-\frac{K}{2}\right)^2} \\ =\ &\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0}\sqrt{\frac{z^2}{z^2+(\mu-k-a)^2}}e^{-\frac{1}{2}z^2}dz\ \times\ e^{-\mu(K-2k)}e^{-a^2+\left(\frac{K}{2}\right)^2-k^2} \end{aligned}$$

which becomes large as $\mu\to-\infty$. Then (14) holds because $\Delta\to0$ and Δ' is positive as $\mu\to-\infty$ or $\mu\to\infty$.

To complete the proof that Δ is shaped as in figure 3, let the function $\tilde{\Delta}''$ be defined as the second derivative, $\tilde{\Delta}''(\mu;k,a,K)=\Delta''(\mu;k,a,K)$, on a domain $\{\mu\in(-\infty,\infty)\,|\Delta'(\mu;k,a,K)=0\}$, that is, using (5) and (6) and setting (15) equal to zero. Then $\tilde{\Delta}''(\mu;k,a,K)$ satisfies the following where it is defined:

$$\tilde{\Delta}''(\mu;k,a,K)\sqrt{\pi}D(k,a)\,e^{(\mu-k)^2+a^2}=$$

$$(\mu-K+k)\ e^{(\mu-k)^2-\left(\mu-\frac{K}{2}\right)^2}\ e^{a^2} \tag{16}$$
$$+\frac{1}{\sqrt{\pi}}+\frac{a}{\sqrt{2}}\left[\Phi(\mu-k+a)\,e^{\frac{1}{2}(\mu-k+a)^2}-\Phi(\mu-k-a)\,e^{\frac{1}{2}(\mu-k-a)^2}\right]$$

Due to (13) and (14), there are values μ where $\Delta'=0$. Let μ_1 be the lowest value where $\Delta'(\mu_1;k,a,K)=0$. Then $\Delta''(\mu_1;k,a,K)\le0$. Let μ_2 be the next larger value where $\Delta'(\mu_2;k,a,K)=0$. Then it must hold that $\Delta''(\mu_2;k,a,K)\ge0$. Our objective is to show that there is no larger value $\mu_3>\mu_2$ such that $\Delta'(\mu_3;k,a,K)=0$. If there were such a value, it would have to hold that $\Delta''(\mu_3;k,a,K)\le0$. Since the function defined by $e^{\frac{1}{2}x^2}\Phi(x)$ is increasing with x, the last term in (16) is positive. Thus, since $\Delta''(\mu_2;k,a,K)\ge0$ and $\mu_3>\mu_2$, $\Delta''(\mu_3;k,a,K)>0$, a contradiction. We conclude that there are only two extreme values, μ_1 and μ_2, and the function Δ is shaped as in figure 3.

Proof of 2(c). It is easy to show that $\frac{d}{d\bar{K}}E_S\left(\mu|\bar{K}\right)>0$ and $\frac{d}{d\bar{K}}S\left(\bar{K}\right)<0$. Let $k,\bar{K},a$ satisfy $D(k,a)=S\left(\bar{K}\right)$. By the intermediate value theorem, since $E_D(\mu|k,a)-E_S(\mu|2k)<$

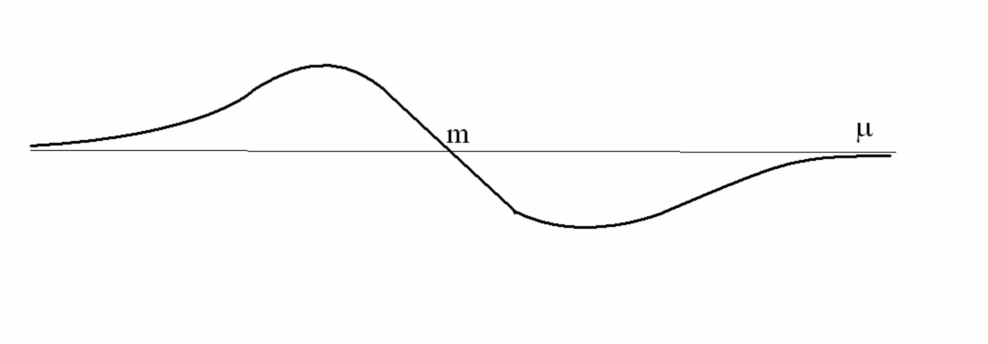

Figure 3: $\Delta(\mu; k_1, k_2, K), \quad \delta(\mu; k_1, k_2, K)$

$0 < E_D(\mu|k, a) - E_S(\mu|\bar{K})$, there exists $K \in (2k, \bar{K})$ such that $0 = E_D(\mu|k-a, k+a) - E_S(\mu|K)$. But then $S(K) > S(\bar{K}) = D(k, a)$. □

7 Proof of Proposition 3

$$
\begin{aligned}
f(k_1, k_2, \mu) &= \frac{1}{\sqrt{2\pi}} \int_{\frac{k_1-\mu}{\sqrt{\mu}}}^{\infty} e^{-\frac{1}{2}y^2} \Phi\left(y + \frac{2\mu - (k_1+k_2)}{\sqrt{\mu}}\right) dy \\
&= \frac{1}{2\pi} \int_{\frac{k_1-\mu}{\sqrt{\mu}}}^{\infty} e^{-\frac{1}{2}w^2} \int_{-w+\frac{(k_1+k_2)-2\mu}{\sqrt{\mu}}}^{\infty} e^{-\frac{1}{2}z^2} dz\, dw
\end{aligned}
$$

Let $\nu_1 = \frac{k_1-\mu}{\sqrt{\mu}}, \nu_2 = \frac{k_2-\mu}{\sqrt{\mu}}$

$$
\begin{aligned}
2\pi \frac{d}{d\nu_1} f &= \frac{d}{d\nu_1} \int_{\nu_1}^{\infty} e^{-\frac{1}{2}w^2} \int_{-\infty}^{w-\nu_1-\nu_2} e^{-\frac{1}{2}z^2} dz\, dw \\
&= -\int_{\nu_1}^{\infty} e^{-\frac{1}{2}w^2} e^{-\frac{1}{2}(w-\nu_1-\nu_2)^2} dw - e^{-\frac{1}{2}\nu_1^2} \int_{-\infty}^{-\nu_2} e^{-\frac{1}{2}z^2} dz \\
&= -e^{-\frac{1}{4}(\nu_1+\nu_2)^2} \frac{\sqrt{\pi}}{\sqrt{2\pi}} \int_{\frac{1}{\sqrt{2}}(v_1-v_2)}^{\infty} e^{-\frac{1}{2}z^2} dz - \sqrt{2\pi} e^{-\frac{1}{2}\nu_1^2} \Phi(-v_2) \\
&= -e^{-\frac{1}{4}(\nu_1+\nu_2)^2} \sqrt{\pi} \Phi\left(\frac{1}{\sqrt{2}}(v_2 - v_1)\right) - \sqrt{2\pi} e^{-\frac{1}{2}\nu_1^2} \Phi(-v_2)
\end{aligned}
$$

$$
\frac{dv_1}{d\mu} = \frac{d}{d\mu} \frac{k_1 - \mu}{\sqrt{\mu}} = \frac{-\sqrt{\mu} - \frac{1}{2}\mu^{-\frac{1}{2}}(k_1 - \mu)}{\mu} = \frac{-1}{2\sqrt{\mu}}\left(1 + \frac{k_1}{\mu}\right)
$$

$$
\begin{aligned}
2\pi \frac{d}{d\nu_2} f &= -\int_{\nu_1}^{\infty} e^{-\frac{1}{2}w^2} e^{-\frac{1}{2}(w-\nu_1-\nu_2)^2} dw \\
&= -e^{-\frac{1}{4}(\nu_1+\nu_2)^2} \frac{\sqrt{\pi}}{\sqrt{2\pi}} \int_{\frac{1}{\sqrt{2}}(v_1-v_2)}^{\infty} e^{-\frac{1}{2}z^2} dz \\
&= -e^{-\frac{1}{4}(\nu_1+\nu_2)^2} \sqrt{\pi} \Phi\left(\frac{1}{\sqrt{2}}(v_2 - v_1)\right)
\end{aligned}
$$

$$
\begin{aligned}
\hat{f} &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{1}{2}w^2}\int_{-w+\frac{K-2\mu}{\sqrt{\mu}}}^{\infty} e^{-\frac{1}{2}z^2}dz\ dw \\
v &= 2\frac{\frac{K}{2}-\mu}{\sqrt{\mu}}, \frac{dv}{d\mu} = \frac{-1}{2\sqrt{\mu}}\left(\frac{K+2\mu}{\mu}\right)
\end{aligned}
$$

$$
\begin{aligned}
2\pi\frac{d}{dv}\hat{f} &= -\int_{-\infty}^{\infty} e^{-\frac{1}{2}w^2}e^{-\frac{1}{2}(w-v)^2}dw \\
&= -e^{-\frac{1}{4}\nu^2}\sqrt{\pi}
\end{aligned}
$$

Given k_1, k_2, K, define a function $\delta(\cdot; k_1, k_2, K)$ on $\mu \in (-\infty, \infty)$ by

$$
\delta(\mu; k_1, k_2, K) = \left[\frac{f(k_1, k_2, \mu)}{\bar{D}(k_1, k_2)} - \frac{\hat{f}(K, \mu)}{\hat{S}(K)}\right] \tag{17}
$$

which is the difference in probabilities of survival under the double-elimination scheme and the single elimination scheme. We prove the proposition by proving that δ has the shape in figure 2 and figure 3, respectively, under the hypotheses of Proposition 3(a) and 3(b). We do this in each case by showing that the second derivative of δ is monotone when restricted to its extreme points (values μ where $\delta' = 0$).

The derivative of δ is the following, where Φ is the cumulative standard normal distribution.

$$
\begin{aligned}
&4\sqrt{\pi}\delta'(\mu; k_1, k_2, K) = \\
&\frac{\mu^{-3/2}}{\bar{D}(k_1, k_2)}\left[\sqrt{2}(\mu + k_1)e^{-\frac{1}{2}\nu_1^2}\Phi(-v_2) + \Phi\left(\frac{1}{\sqrt{2}}(v_2 - v_1)\right)e^{-\frac{1}{4}(\nu_1+\nu_2)^2}(2\mu + k_1 + k_2)\right] \\
&-\frac{\mu^{-3/2}}{\hat{S}(K)}e^{-\frac{1}{4}\nu^2}(2\mu + K)
\end{aligned}
$$

The second derivative is

Now impose $\bar{D} = \hat{S}$ ane $\delta' = 0$ (ignore multiplier $\mu^{-3/2}$)

$$
\begin{aligned}
&4\sqrt{\pi}\mu^{3/2}\delta' = \\
&\left[\sqrt{2}(\mu + k_1)e^{-\frac{1}{2}\nu_1^2}\Phi(-v_2) + \Phi\left(\frac{1}{\sqrt{2}}(v_2 - v_1)\right)e^{-\frac{1}{4}(\nu_1+\nu_2)^2}(2\mu + k_1 + k_2)\right] \\
&-e^{-\frac{1}{4}\nu^2}(2\mu + K)
\end{aligned}
$$

$$
\begin{aligned}
&(d/d\mu)\, 4\sqrt{\pi}\mu^{3/2}\delta' \\
= \quad & \frac{1}{\sqrt{2}\mu^2} e^{-\frac{1}{2}\nu_1^2}\Phi\left(-v_2\right)\left[2\mu^2 + \left(k_1-\mu\right)\left(k_1+\mu\right)^2\right] \\
& + \frac{1}{\mu^2\sqrt{\pi}} e^{-\frac{1}{2}\left(v_2^2+v_1^2\right)}\left[\left(k_2-\mu\right)\left(k_2+\mu\right)\left(\mu+k_1\right) - \frac{\left(k_2-k_1\right)}{4\sqrt{\pi}}\mu^{1/2}(2\mu+k_1+k_2)\right] \\
& + \Phi\left(\frac{1}{\sqrt{2}}\left(v_2-v_1\right)\right) e^{-\frac{1}{4}\left(\nu_1+\nu_2\right)^2}\left[2 + \frac{1}{4\mu^2}\left(k_1+k_2-2\mu\right)\left(k_1+k_2+2\mu\right)^2\right] \\
& - \frac{1}{4\mu^2}\left(K-2\mu\right)\left(K+2\mu\right)^2 e^{-\frac{1}{4}\nu^2} - 2e^{-\frac{1}{4}\nu^2}
\end{aligned}
$$

$$
\begin{aligned}
&(d/d\mu)\, 4\sqrt{\pi}\mu^{5/2}\delta' \\
= \quad & \frac{1}{\sqrt{2}} e^{-\frac{1}{2}\nu_1^2}\Phi\left(-v_2\right)\left[2\mu^2 + \left(k_1-\mu\right)\left(k_1+\mu\right)^2\right] \\
& + \frac{1}{\sqrt{\pi}} e^{-\frac{1}{2}\left(v_2^2+v_1^2\right)}\left[\left(k_2-\mu\right)\left(k_2+\mu\right)\left(\mu+k_1\right) - \frac{\left(k_2-k_1\right)}{4\sqrt{\pi}}\mu^{1/2}(2\mu+k_1+k_2)\right] \\
& + \Phi\left(\frac{1}{\sqrt{2}}\left(v_2-v_1\right)\right) e^{-\frac{1}{4}\left(\nu_1+\nu_2\right)^2}\left[2\mu^2 + \frac{1}{4}\left(k_1+k_2-2\mu\right)\left(k_1+k_2+2\mu\right)^2\right] \\
& - e^{-\frac{1}{4}\nu^2}\left[2\mu^2 + \frac{1}{4}\left(K-2\mu\right)\left(K+2\mu\right)^2\right]
\end{aligned}
$$

Let $k_1 = k_2$

$$
\begin{aligned}
& 4e^{\nu_1^2}\sqrt{\pi}\mu^{3/2}\delta' = \\
& \left(\mu+k_1\right)\left[\sqrt{2}e^{\frac{1}{2}\nu_1^2}\Phi\left(-v_2\right)+1\right] - e^{\nu_1^2}e^{-\frac{1}{4}\nu^2}\left(2\mu+K\right)
\end{aligned}
$$

$$
\begin{aligned}
& \frac{d}{d\mu} 4e^{\nu_1^2}\sqrt{\pi}\mu^{5/2}\delta' \\
= \quad & \left[\sqrt{2}e^{\frac{1}{2}\nu_1^2}\Phi\left(-v_2\right)\mu^2 + \mu^2\right] + \left(\mu+k_1\right)\left(k_1^2-\mu^2\right)e^{\frac{1}{2}\nu_1^2}\Phi\left(-v_2\right)\frac{1}{\sqrt{2}} \\
& + (\mu+k_1)^2\mu^{\frac{1}{2}}e^{\frac{1}{2}v_1^2}\frac{1}{2\sqrt{\pi}} - 2e^{\nu_1^2-\frac{1}{4}\nu^2} + 2\left(2\mu+K\right)\left(k_1^2-\left(\frac{K}{2}\right)^2\right)e^{\nu_1^2-\frac{1}{4}\nu^2}
\end{aligned}
$$

$$
\begin{aligned}
= \quad & \mu^2\left[\sqrt{2}e^{\frac{1}{2}\nu_1^2}\Phi\left(-v_2\right)+1\right] + \left(\mu+k_1\right)\left(k_1^2-\mu^2\right)e^{\frac{1}{2}\nu_1^2}\Phi\left(-v_2\right)\frac{1}{\sqrt{2}} \\
& + (\mu+k_1)^2\mu^{\frac{1}{2}}e^{\frac{1}{2}v_1^2}\frac{1}{2\sqrt{\pi}} - 2e^{\nu_1^2-\frac{1}{4}\nu^2} + 2(\mu+k_1)\left[\sqrt{2}e^{\frac{1}{2}\nu_1^2}\Phi\left(-v_2\right)+1\right]\left(k_1^2-\left(\frac{K}{2}\right)^2\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \left[\sqrt{2}e^{\frac{1}{2}\nu_1^2}\Phi\left(-v_2\right)+1\right]\left[\mu^2+2(\mu+k_1)\left(k_1^2-\left(\frac{K}{2}\right)^2\right)\right] \\
&\quad +(\mu+k_1)e^{\frac{1}{2}\nu_1^2}\left[\Phi\left(-v_2\right)\frac{1}{\sqrt{2}}\left(k_1^2-\mu^2\right)+\frac{\mu^{\frac{1}{2}}}{2\sqrt{\pi}}\right] \\
&\quad -2e^{\nu_1^2-\frac{1}{4}\nu^2} \\
&= \mu^2\left[e^{\frac{1}{2}\nu_1^2}\left(\sqrt{2}\Phi\left(-v_2\right)-(\mu+k_1)\right)+1\right] \\
&\quad -\left[\sqrt{2}e^{\frac{1}{2}\nu_1^2}\Phi\left(-v_2\right)+1\right]\left[2(\mu+k_1)\left(\left(\frac{K}{2}\right)^2-k_1^2\right)\right] \\
&\quad +\frac{1}{2}(\mu+k_1)e^{\frac{1}{2}\nu_1^2}\left[\sqrt{2}\Phi\left(-v_2\right)k_1^2+\frac{\mu^{\frac{1}{2}}}{\sqrt{\pi}}\right] \\
&\quad -2e^{\nu_1^2-\frac{1}{4}\nu^2}
\end{aligned}
$$

Part 3(a) Now impose $k_1+k_2=K$. We first show that $\Phi\left(\frac{k_2-k_1}{\sqrt{2}}\right)\hat{S}\left(k_1+k_2\right)<\bar{D}\left(k_1,k_2\right)<\hat{S}\left(k_1+k_2\right)$. For every pair of successes (x_1,x_2) such that the firm survives the double elimination, it also survives the single elimination, but not vice versa. Conditional on $x_1+x_2\geq K=k_1+k_2$, it holds with positive probability that $x_1<k_1$, where x_1 is the realization that comes first. Thus, $\bar{D}\left(k_1,k_2\right)<\hat{S}\left(k_1+k_2\right)$. The researcher survives the double elimination if the following two inequalities hold, which imply that $x_1+x_2\geq k_1+k_2$ and $x_1\geq k_1$.

$$
\begin{aligned}
x_2-x_1 &\leq k_2-k_1 \qquad (18)\\
x_1+x_2 &\geq k_1+k_2
\end{aligned}
$$

Since $x_1-\mu$ and $x_2-\mu$ are independent standard normal variables, the random variables defined by $\frac{x_2-x_1}{\sqrt{2}}$ and $\frac{x_1+x_2-2\mu}{\sqrt{2}}$ are independent standard normal variables for each μ (Mood, Graybill and Boes 1974, p. 191). Using independence, a lower bound on the probability of surviving the double elimination is the probability of the first line in (18), which is $\Phi\left(\frac{k_2-k_1}{\sqrt{2}}\right)$, times the probability of the second line, which is $\hat{S}\left(k_1+k_2\right)$. Hence $\Phi\left(\frac{k_2-k_1}{\sqrt{2}}\right)\leq\frac{\bar{D}(k_1,k_2)}{\hat{S}(k_1+k_2)}$. In fact the inequality must be strict, since the inequalities (18) are sufficient but not necessary. With positive probability there are realizations (x_1,x_2) which allow survival in both elimination schemes, but do not satisfy (18). (Given $x_1>k_1$, let $x_2>k_2-(k_1-x_1)$.)

The difference in expected abilities of survivors under the two schemes can be written

$$
E_{\bar{D}}\left(\mu|k_1,k_2\right)-E_{\hat{S}}\left(\mu|k_1+k_2\right)=\int\mu\;\delta\left(\mu\right)\;h\left(\mu\right)d\mu
$$

Since

$$0 = \int \delta(\mu)\ h(\mu)\, d\mu$$

it suffices for $E_{\bar{D}}(\mu|k_1, k_2) - E_{\hat{S}}(\mu|k_1 + k_2) > 0$ that, as in figure 2, there exists m such that

$$\begin{aligned} \delta(\mu; k_1, k_2, k_1 + k_2) &< 0 \quad \text{for} \quad \mu \in (-\infty, m) \\ \delta(\mu; k_1, k_2, k_1 + k_2) &> 0 \quad \text{for} \quad \mu \in (m, \infty) \end{aligned}$$

Using (17) and $\bar{D}(k_1, k_2) < \hat{S}(k_1 + k_2)$,

$$\begin{aligned} \lim_{\mu \to -\infty} \delta(\mu; k_1, k_2, k_1 + k_2) &= \left[\frac{0}{\bar{D}(k_1, k_2)} - \frac{0}{\hat{S}(k_1 + k_2)}\right] = 0 \\ \lim_{\mu \to \infty} \delta(\mu; k_1, k_2, k_1 + k_2) &= \left[\frac{1}{\bar{D}(k_1, k_2)} - \frac{1}{\hat{S}(k_1 + k_2)}\right] > 0 \end{aligned} \tag{19}$$

The first derivative (**??**) can be written as

$\delta'(\mu; k_1, k_2, k_1 + k_2) =$

$$\begin{aligned} &\frac{1}{\sqrt{\pi}\bar{D}(k_1, k_2)} e^{-\left(\mu - \frac{(k_1 + k_2)}{2}\right)^2} \times \\ &\left[\frac{1}{\sqrt{2}} e^{-\frac{1}{2}(\mu - k_1)^2} e^{\left(\mu - \frac{(k_1 + k_2)}{2}\right)^2} \Phi(\mu - k_2) + \Phi\left(\frac{k_2 - k_1}{\sqrt{2}}\right) - \frac{\bar{D}(k_1, k_2)}{\hat{S}(k_1 + k_2)}\right] \end{aligned} \tag{20}$$

We now show that there exists d such that

$$\delta(\mu; k_1, k_2, K) < 0 \quad \text{if} \quad \mu \in (-\infty, d) \tag{21}$$

It follows from (20) that $\delta' \to 0$ as $\mu \to -\infty$. The first term in the brackets goes to zero as $\mu \to -\infty$, since

$$\begin{aligned} & e^{-\frac{1}{2}(\mu - k_1)^2} e^{\left(\mu - \frac{(k_1 + k_2)}{2}\right)^2} \Phi(\mu - k_2) \\ = \ & e^{\left(\mu - \frac{(k_1 + k_2)}{2}\right)^2 - \frac{1}{2}(\mu - k_1)^2 - \frac{1}{2}(\mu - k_2)^2} \times e^{\frac{1}{2}(\mu - k_2)^2} \Phi(\mu - k_2) \\ = \ & e^{-\frac{1}{4}(k_2 - k_1)^2} e^{\frac{1}{2}(\mu - k_2)^2} \Phi(\mu - k_2) \to 0 \end{aligned}$$

Since $\Phi\left(\frac{k_2 - k_1}{\sqrt{2}}\right) - \frac{\bar{D}(k_1, k_2)}{\hat{S}(k_1 + k_2)} < 0$, δ' becomes negative for small μ.

We will complete the argument by showing that there is only one value μ where $\delta' = 0$, and in a neighborhood of that value, δ' is increasing. Then, since δ is positive on some of its domain, it must have the shape in figure 2.

Write the second derivative (**??**) as

$$\delta''(\mu; k_1, k_2, k_1+k_2) \;\; \bar{D}(k_1,k_2)\, e^{\left(\mu - \frac{(k_1+k_2)}{2}\right)^2} \sqrt{2\pi} =$$
$$-(\mu - k_1)\, e^{-\frac{1}{2}(\mu-k_1)^2} e^{\left(\mu - \frac{(k_1+k_2)}{2}\right)^2} \Phi(\mu - k_2) + \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\mu-k_1)^2} e^{-\frac{1}{2}(\mu-k_2)^2} e^{\left(\mu - \frac{(k_1+k_2)}{2}\right)^2}$$
$$+2\sqrt{2}\left(\mu - \frac{(k_1+k_2)}{2}\right)\left[-\Phi\left(\frac{k_2-k_1}{\sqrt{2}}\right) + \frac{\bar{D}(k_1,k_2)}{\hat{S}(k_1+k_2)}\right]$$

Using (20), at μ such that $\delta'(\mu; k_1, k_2, k_1+k_2) = 0$,

$$\frac{1}{\sqrt{2}} e^{-\frac{1}{2}(\mu-k_1)^2} e^{\left(\mu - \frac{(k_1+k_2)}{2}\right)^2} \Phi(\mu - k_2) = -\Phi\left(\frac{k_2-k_1}{\sqrt{2}}\right) + \frac{\bar{D}(k_1,k_2)}{\hat{S}(k_1+k_2)} \tag{22}$$

Let the function $\tilde{\delta}''$ be defined as the second derivative, $\tilde{\delta}''(\mu) = \delta''(\mu)$, on a domain $\{\mu \in (-\infty,\infty) \,|\, \delta'(\mu) = 0\}$. Then, using (22) and (6), $\tilde{\delta}''(\mu)$ satisfies the following where it is defined:

$$\tilde{\delta}''(\mu; k_1, k_2, k_1+k_2) \;\; \sqrt{2\pi}\bar{D}(k_1,k_2)\, e^{\left(\mu - \frac{(k_1+k_2)}{2}\right)^2} =$$
$$-(\mu - k_1)\, e^{-\frac{1}{2}(\mu-k_1)^2} e^{\left(\mu - \frac{(k_1+k_2)}{2}\right)^2} \Phi(\mu - k_2) + \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{4}(k_2-k_1)^2}$$
$$+2\left(\mu - \frac{(k_1+k_2)}{2}\right) e^{-\frac{1}{2}(\mu-k_1)^2} e^{\left(\mu - \frac{(k_1+k_2)}{2}\right)^2} \Phi(\mu - k_2)$$
$$= e^{-\frac{1}{4}(k_2-k_1)^2} \frac{1}{\sqrt{2}}\left[e^{\frac{1}{2}(\mu-k_2)^2} \Phi(\mu-k_2)(\mu-k_2) + \frac{1}{\sqrt{2\pi}}\right] \tag{23}$$

Due to (21) and continuity, and because $\delta > 0$ somewhere on its domain, δ must have at least one extreme point. Let μ_1 be the smallest value where $\delta = 0$. It must hold that $\delta''(\mu_1) \geq 0$. But then, since the bracketed term in (23) is increasing with μ, $\delta''(\mu) > 0$ at any larger μ where $\delta' = 0$. Thus, there is only a single value μ where $\delta' = 0$, proving that δ has the shape given in figure 2.

Part 3(b): Choosing k_1, k_2, K such that $\bar{D}(k_1,k_2) = \hat{S}(K)$, we will show that δ has the shape given in figure 3, which implies that $E_{\bar{D}}(\mu|k_1,k_2) - E_{\hat{S}}(\mu|K) > 0$. We will show, as in figure 3, that there is a dividing value, say m, such that

$$\begin{aligned} \delta(\mu; k_1, k_2, K) &> 0 \quad \text{for} \quad \mu < m \\ \delta(\mu; k_1, k_2, K) &< 0 \quad \text{for} \quad \mu > m \end{aligned}$$

Using (17) and $\bar{D}(k_1, k_2) = \hat{S}(K)$,

$$\lim_{\mu\to-\infty} \delta(\mu; k_1, k_2, K) = \left[\frac{0}{\bar{D}(k_1,k_2)} - \frac{0}{\hat{S}(K)}\right] = 0$$

$$\lim_{\mu\to\infty} \delta(\mu; k_1, k_2, K) = \left[\frac{1}{\bar{D}(k_1,k_2)} - \frac{1}{\hat{S}(K)}\right] = 0$$

We will first argue that there exist d, b such that

$$\begin{array}{lll} \delta(\mu; k_1, k_2, K) > 0 & \text{if} & \mu \in (-\infty, d) \\ \delta(\mu; k_1, k_2, K) < 0 & \text{if} & \mu \in (b, \infty) \end{array} \tag{24}$$

Given that $\delta \to \infty$ as $\mu \to \infty$ or $\mu \to -\infty$, it is enough to show that $\delta' > 0$ for μ sufficiently small and for μ sufficiently large. It follows from (**??**) that as $\mu \to \infty$ or $\mu \to -\infty$, $\delta' \to 0$. Using $\bar{D}(k_1, k_2) = \hat{S}(K)$, (6), and (**??**),

$$\delta'(\mu; k_1, k_2, K)\sqrt{\pi}\bar{D}(k_1,k_2)\, e^{\left(\mu - \frac{(k_1+k_2)}{2}\right)^2} =$$

$$\frac{1}{\sqrt{2}} e^{\left(\mu - \frac{(k_1+k_2)}{2}\right)^2} e^{-\frac{1}{2}(\mu-k_1)^2}\Phi(\mu - k_2) + \Phi\left(\frac{k_2-k_1}{\sqrt{2}}\right) - e^{-\left(\mu-\frac{K}{2}\right)^2} e^{\left(\mu - \frac{(k_1+k_2)}{2}\right)^2} \tag{25}$$

and

$$\delta'(\mu)\sqrt{\pi}\bar{D}(k_1,k_2)\, e^{\frac{1}{2}\left(\mu - \frac{(k_1+k_2)}{2}\right)^2} =$$

$$\begin{aligned} &\frac{1}{\sqrt{2}} e^{\frac{1}{2}\left(\mu - \frac{(k_1+k_2)}{2}\right)^2} e^{-\frac{1}{2}(k_2-k_1)^2}\Phi(\mu - k_2) \\ &+\Phi\left(\frac{k_2-k_1}{\sqrt{2}}\right) e^{-\frac{1}{2}\left(\mu - \frac{(k_1+k_2)}{2}\right)^2} - e^{-\left(\mu-\frac{K}{2}\right)^2} e^{\frac{1}{2}\left(\mu - \frac{(k_1+k_2)}{2}\right)^2} \end{aligned} \tag{26}$$

As $\mu \to -\infty$ (and using $K > k_1 + k_2$), the second line of (25) becomes positive, hence $\delta' > 0$. As $\mu \to \infty$, the second and third terms of (26) go to zero while the first term blows up, hence $\delta' > 0$. Since $\delta \to 0$ as $\mu \to \infty$ or $\mu \to -\infty$, these imply (24).

We now complete the proof that δ is shaped as in figure 3. There are clearly values μ where $\delta' = 0$, since δ is positive for low values of μ and negative for high values of μ. Let μ_1 be the smallest value where $\delta' = 0$. Using (24), it holds that $\delta''(\mu_1) \leq 0$. Let μ_2 be the next larger value where $\delta' = 0$. Using continuity of the derivative, $\delta''(\mu_2) \geq 0$. Our objective is to show that there is no larger value $\mu_3 > \mu_2$ such that $\delta'(\mu_3) = 0$.

Using (22) and (6), at any μ where $\delta'(\mu) = 0$, it holds that

$$\frac{1}{\sqrt{2}} e^{-\frac{1}{4}(k_2-k_1)^2} e^{\frac{1}{2}(\mu-k_2)^2}\Phi(\mu - k_2) = e^{-\left(\mu-\frac{K}{2}\right)^2} e^{\left(\mu - \frac{(k_1+k_2)}{2}\right)^2} - \Phi\left(\frac{k_2-k_1}{\sqrt{2}}\right) \tag{27}$$

Using (**??**) and (6) and $\bar{D}(k_1,k_2)=\hat{S}(K)$ it holds that

$$\delta''(\mu)\sqrt{\pi}\bar{D}(k_1,k_2)\,e^{\left(\mu-\frac{(k_1+k_2)}{2}\right)^2} =$$

$$-\frac{1}{\sqrt{2}}(\mu-k_1)\,e^{-\frac{1}{4}(k_2-k_1)^2}e^{\frac{1}{2}(\mu-k_2)^2}\Phi(\mu-k_2)+\frac{1}{2\sqrt{\pi}}e^{-\frac{1}{4}(k_2-k_1)^2} \tag{28}$$
$$-2\Phi\left(\frac{k_2-k_1}{\sqrt{2}}\right)\left(\mu-\frac{(k_1+k_2)}{2}\right)+2\left(\mu-\frac{K}{2}\right)e^{-\left(\mu-\frac{K}{2}\right)^2}e^{\left(\mu-\frac{(k_1+k_2)}{2}\right)^2}$$

Let $\tilde{\delta}''$ be a function on $\{\mu\in(-\infty,\infty)\,|\delta'(\mu)=0\}$ defined by $\tilde{\delta}''(\mu)=\delta''(\mu)$. Substituting (27) into (28), so that δ'' is restricted to μ where $\delta'(\mu)=0$,

$$\tilde{\delta}''(\mu)\sqrt{\pi}\bar{D}(k_1,k_2)\,e^{\left(\mu-\frac{(k_1+k_2)}{2}\right)^2}=-(K-k_2-k_1)\,\Phi\left(\tfrac{k_2-k_1}{\sqrt{2}}\right)$$
$$+e^{-\frac{1}{4}(k_2-k_1)^2}\frac{1}{\sqrt{2}}\left[e^{\frac{1}{2}(\mu-k_2)^2}\Phi(\mu-k_2)(\mu+k_1-K)+\frac{1}{\sqrt{2\pi}}\right] \tag{29}$$

If μ_3 is the smallest value larger than μ_2 where $\delta'=0$, it would have to hold that $\delta''(\mu_3)\leq 0$. However, it follows from (29) that if $\delta(\mu_2)=0$ and $\delta''(\mu_2)\geq 0$, and if $\delta'(\mu_3)=0$, then it also holds that $\delta''(\mu_3)>0$. This is because the term in large brackets of (29) is increasing in μ. Since this is a contradiction, there is no such μ_3, and δ is shaped as in figure 3.

Part 3(c). It is easy to show that $\frac{d}{d\bar{K}}E_{\hat{S}}\left(\mu|\bar{K}\right)>0$ and $\frac{d}{d\bar{K}}\hat{S}\left(\bar{K}\right)<0$. Let $k_1,k_2,\bar{K}$ satisfy $\bar{D}(k_1,k_2)=\hat{S}\left(\bar{K}\right)$. By the intermediate value theorem, since $E_{\bar{D}}(\mu|k_1,k_2)-E_{\hat{S}}(\mu|2k)<0<E_{\bar{D}}(\mu|k_1,k_2)-E_{\hat{S}}\left(\mu|\bar{K}\right)$, there exists $K\in(k_1+k_2,\bar{K})$ such that $0=E_{\bar{D}}(\mu|k_1,k_2)-E_{\hat{S}}(\mu|K)$. But then $\hat{S}(K)>\hat{S}\left(\bar{K}\right)=\bar{D}(k_1,k_2)$. □

8 Proof of Proposition 4

Given (k_1,k_2) let $\Gamma(\cdot;k_1,k_2)$ be a function on $(-\infty,\infty)$ defined by

$$\Gamma(\mu;k_1,k_2)=\frac{1}{D(k_1,k_2)}\Phi(\mu-k_1)\,\Phi(\mu-k_2)-\frac{1}{\bar{D}(k_1,k_2)}f(\mu;k_1,k_{2w})$$

Then

$$E_D(\mu|k_1,k_2)-E_{\bar{D}}(\mu|k_1,k_2)=\int\Gamma(\mu;k_1,k_2)\;h(\mu)\,d\mu$$

Since

$$\int\Gamma(\mu;k_1,k_2)\;h(\mu)\,d\mu=0$$

it is enough to show that Γ has the shape in figure 2, and in particular, that there exists m such that

$$\begin{array}{lll} \Gamma(\mu;k_1,k_2)<0 & \text{for} & \mu<m \\ \Gamma(\mu;k_1,k_2)>0 & \text{for} & \mu>m \end{array} \tag{30}$$

Write (k_1,k_2) as in (2) and D for $D(k-a,k+a)$, $\bar{D}$ for $\bar{D}(k-a,k+a)$. Then $\Gamma'(\mu|k-a,k+a)=$

$$\frac{1}{D\sqrt{2\pi}}\left[\Phi(\mu-k-a)\,e^{-\frac{1}{2}(\mu-k+a)^2}+\Phi(\mu-k+a)\,e^{-\frac{1}{2}(\mu-k-a)^2}\right]$$
$$-\frac{1}{\sqrt{2\pi}\bar{D}}\left[\Phi(\mu-k-a)\,e^{-\frac{1}{2}(\mu-k+a)^2}+\sqrt{2}\Phi\left(\frac{2a}{\sqrt{2}}\right)e^{-(\mu-k)^2}\right]$$

$\Gamma'(\mu|k-a,k+a)\sqrt{2\pi}=$

$$\begin{aligned} =\ & \left[\frac{1}{D}-\frac{1}{\bar{D}}\right]e^{-\frac{1}{2}(\mu-k+a)^2}e^{-\frac{1}{2}(\mu-k-a)^2}e^{\frac{1}{2}(\mu-k-a)^2}\Phi(\mu-k-a) \\ & +\frac{1}{D}\Phi(\mu-k+a)^{-\frac{1}{2}(\mu-k+a)^2}\,e^{-\frac{1}{2}(\mu-k-a)^2}e^{\frac{1}{2}(\mu-k+a)^2}-\frac{1}{\bar{D}}e^{-(\mu-k)^2}\sqrt{2}\Phi\left(\sqrt{2}a\right) \end{aligned}$$

$$\begin{aligned} =\ & \left[\frac{1}{D}-\frac{1}{\bar{D}}\right]e^{-(\mu-k)^2}e^{-a^2}e^{\frac{1}{2}(\mu-k-a)^2}\Phi(\mu-k-a) \\ & +\frac{1}{D}\Phi(\mu-k+a)\,e^{\frac{1}{2}(\mu-k+a)^2}e^{-(\mu-k)^2}e^{-a^2}-\frac{1}{\bar{D}}e^{-(\mu-k)^2}\sqrt{2}\Phi\left(\sqrt{2}a\right) \end{aligned}$$

$\Gamma'(\mu|k-a,k+a)\sqrt{2\pi}e^{(\mu-k)^2}e^{a^2}=$

$$\begin{aligned} =\ & \left[\frac{1}{D}-\frac{1}{\bar{D}}\right]e^{\frac{1}{2}(\mu-k-a)^2}\Phi(\mu-k-a) \\ & +\frac{1}{D}\Phi(\mu-k+a)\,e^{\frac{1}{2}(\mu-k+a)^2}-\frac{1}{\bar{D}}e^{a^2}\sqrt{2}\Phi\left(\sqrt{2}a\right) \end{aligned} \tag{31}$$

Since the latter is increasing with μ, if Γ is increasing at any μ, it is increasing at any larger μ. Since $\Gamma\to 0$ as $\mu\to-\infty$ and (31) becomes negative as $\mu\to-\infty$, and since Γ must be negative and positive on different parts of its domain, we can conclude that (30) holds.

References

[1] Gompers, P. and J. Lerner. 1999. *The Venture Capital Cycle.* Cambridge, MA: MIT Press.

[2] Maurer, S., and S. Scotchmer. 2004. "Procuring Knowledge." In G. Libecap, ed., *Advances in the Study of Entrepreneurship, Innovation and Growth,* 1-31. Amsterdam: JAI Press.

[3] Mood, A., F. Graybill and D. Boes. 1974. *Introduction to the Theory of Statistics.* New York: McGraw-Hill.

[4] Scotchmer, S. 2004. *Innovation and Incentives.* Cambridge, MA: MIT Press.