Online Appendix for Financing Risk and Innovation

B1. Proof that a simpler model will not have financing risk:

Some readers may feel that financing risk could be generated with a simpler model - after all, ignoring any effects from sale negotiations, couldn't a project turn NPV negative if early investors forecast no future investors will support it? Yes, however, such a belief would be irrational and in the future investors would rationally support it. Thus, this kind of model would NOT have rational equilibrium financing risk. Although our implications and results do hold in a world with irrational financing risk, our model helps demonstrate that no irrationality is needed, which has important implications.

To demonstrate that a simpler model would not work consider the following model.

A project with NPV Π pays off V with probability p and with a probability $(1-p)$ it needs another x. Investing x buys the investor a fraction q. A rational investor will invest as long as the fraction $q \geq \frac{x}{\Pi+x}$.

If investors forecast a No-Invest state then they will invest as long as $x \leq qpV$. This is because the present value of the project is pV so they invest if their fraction q of the PV is greater than the cost, x.

If investors forecast an Invest signal next period followed by the no invest signal then an investor who invests x for fraction q expects to receive $qpV + (1-p)(1-q)qpV$, that is, if it pays off in the first attempt (prob p) they get qV and if it pays off on the second attempt (prob $(1-p)p$)they get a fraction q of what doesn't go to the second investor $(1-q)V$. Therefore, the first investors would invest as long as $x \leq qpV + (1-p)(1-q)qpV$.

And if investors forecast infinite Invest signals then the first investor would invest as long as $x \leq qpV + (1-p)(1-q)qpV + (1-p)(1-p)pq(1-q)(1-q)V + ... = qpV \sum_{i=0}^{\infty}(1-p)^i(1-q)^i$ which equals $pV\frac{q}{1-(1-p)(1-q)}$.

It might then seem that as long as $(1-p)(1-q)qpV > 0$ that some parameters would result in a world where $x \geq qpV$ but $x \leq qpV + (1-p)(1-q)qpV$ so investors would invest only if they forecasted future investment, and thus there would be financing risk. Or $x \geq qpV$ but $x \leq pV\frac{q}{1-(1-p)(1-q)}$, which would again result in financing risk.

However, this is not correct.

If $x \geq qpV$ the q endogenously increases until the first round investor invests $(x \leq q'pV)$ or q or hits 1. That is, firms that are about to fail because they cannot get funding increase the fraction they are willing to give to get funded. Thus, the first period investor will only not invest in equilibrium if $x \geq pV$. However, if this is true then $x \geq pV\frac{q}{1-(1-p)(1-q)}$ because $\frac{q}{1-(1-p)(1-q)} \leq 1 \ \forall \ 0 \leq q \leq 1$. Thus, if the first round investor will not invest in the No-Invest equilibrium then he will not invest in the Invest equilibrium, and vice versa.

Therefore, there is no financing risk in this model - investors either always invest or never do. In order to get rational equilibrium financing risk the forecast of no financing tomorrow must fundamentally alter the outcome of the project in some way. This is what the model we present in this paper does.

B2. Information Cascades:

If it is hard for VCs to find investments, then they may focus on an area to search for investments.[21] For simplicity we narrow the world to only two possible investing areas, area A or B. Each of n investors receives a signal about whether it is better to invest in A or B, and each investor chooses an action $a_i \in \{A, B\}$ which determines the area in which they search.[22] Nature moves after investor actions are chosen and reveals which area was the better one to have searched in, which leads to a bonus to the investor of Υ_j, where either $\Upsilon_A = \Upsilon > 0$ and $\Upsilon_B = 0$ or vice versa.[23] All searches yield a single investment opportunity and searches take place in the 'pre-period' – before each round of investment elucidated in the sections above. The order in which investors must decide the sector to search is common knowledge. Each investor is a rational expected value maximizer.

Prior expectations are initially diffuse such that the probability A is the better sector equals the probability B is the better sector, i.e., $Pr(\Upsilon_A = \Upsilon) = 1/2 = Pr(\Upsilon_B = \Upsilon)$. However, there are a sequence of signals Y_i, where the signal of investor i is $y_i \in \{A^A, A^B, B^A, B^B\}$.[24] 'A' signals are more likely if A is the better sector, and vice versa if

[21]Areas might be big data, or the internet of things, or the sharing economy, or robotics etc., or alternatively different time periods or different physical locations.

[22]These n investors can be thought of as the high profile flexible firms that can rapidly deploy resources to different areas and thus make a new search decision each period. We assume that there are an infinite number of other investors searching in area A and B to insure that our results are not driven by any physical capital constraints (although we note that physical constraints would magnify the results we find).

[23]This bonus could arise if at a later time limited partners desired one sector more than another, or if future jobs for VC became more plentiful in one area so experience in that area was more valuable, or some government regulation or macroeconomic factor created an added advantage to investing in one area.

[24]The signals are conditionally independent and identically distributed. The number of signals can easily be generalized to n signals but four demonstrates the effect.

B is the better sector.[25]

$$\begin{array}{lcc} & Pr(y_i = A^A \, or \, A^B \mid Z) & Pr(y_i = B^B \, or \, B^A \mid Z) \\ \hline Z = A & p & 1-p \\ Z = B & 1-p & p \end{array} \tag{B1}$$

where $1 > p > 0.5$. Furthermore, signals A^A and B^B are the stronger signals in the sense that an investor is more likely to get the A^A signal if A is the better industry and they receive an A signal, i.e.,

$$\begin{array}{lcc} & Pr(y_i = A^A \mid y_i = A^A \, or \, A^B, Z) & Pr(y_i = A^B \mid y_i = A^A \, or \, A^B, Z) \\ \hline Z = A & q & 1-q \\ Z = B & 1-q & q \end{array} \tag{B2}$$

and the opposite for B signals, i.e.,

$$\begin{array}{lcc} & Pr(y_i = B^B \mid y_i = B^B \, or \, B^A, Z) & Pr(y_i = B^A \mid y_i = B^B \, or \, B^A, Z) \\ \hline Z = A & 1-q & q \\ Z = B & q & 1-q \end{array} \tag{B3}$$

where $p > q > 0.5$.

Thus, for example, $Pr(y_i = A^A \mid A) = Pr(y_i = B^B \mid B) = pq$. Furthermore, given A is the better industry A^A is the most likely signal, A^B is the next most likely, B^A is the next most likely, and B^B is the least likely signal.[26]

We assume logically that having few or no investors searching in an area is the signal $s_j(t) = 0$, while an area with many (specifically greater than m) searching investors is the signal $s_j(t) = 1$, where $j \in A, B$ denotes the sector. Thus, signals are now created by the actions of the set of n investors everyone is watching.

Each investor takes an action a_i (search A or search B) that maximizes his expected value given the actions of other players. It is a dominant strategy to choose to search in the area with the largest expected chance of generating Υ assuming, as in Bikhchandani

[25]This requires the assumption that $p > q$ where p and q are the probabilities of different signals defined in equations B1, B2, and B3. It also ensures that our set up meets the monotone likelihood ratio ordering required in assumption 1 of Bikhchandani et al. (1992).

[26]$Pr(y_i = A^A \mid A) = pq$, $Pr(y_i = A^B \mid A) = p(1-q)$, $Pr(y_i = B^A \mid A) = (1-q)p$, $Pr(y_i = B^B \mid A) = (1-p)(1-q)$

et al. (1992), that all off equilibrium path behavior that occurs after a cascade forms is ignored by all other players. Given that the first investor has a diffuse prior, he searches in the area of his signal. Without loss of generality assume the first investor receives either A^A or A^B and consequently searches area A. Investor 2 sees action $a_1 = A$ and signal y_2 and chooses action a_2 to maximize the chance that he searches in the best area. Action $a_1 = A$ tells player 2 that player 1 had signal A^A or A^B, therefore, with signals A^A, A^B, or B^A player 2's posterior puts greater weight on area A. That is, $Pr(\Upsilon_A = \Upsilon \mid a_1 = A, y_2 \in \{A^A, A^B, B^A\}) > 1/2$. In these cases player 2 will also chose area A. However, with signal B^B player 2's posterior puts greater weight on area B, $Pr(\Upsilon_A = \Upsilon \mid a_1 = A, y_2 = B^B) < 1/2$, and player 2 chooses B. Whether or not an information cascade forms after player 2 depends on the strength of the signals. If p is larger, then A signals imply a larger probability that $\Upsilon_A = \Upsilon$. As q increases A^A signals imply a larger probability that $\Upsilon_A = \Upsilon$ relative to A^B signals. At an extreme as q approaches p, the A^B signal conveys the same information as the B^A signal, i.e. as $q \to p$, $Pr(\Upsilon_A = \Upsilon \mid y_i = A^B) \to Pr(\Upsilon_A = \Upsilon \mid y_i = B^A)$.

An information cascade can begin at player 3 if player 3 would choose the same area given any signal. This would be rational if for all signals, y_3, if $Pr(\Upsilon_A = \Upsilon \mid a_1, a_2, y_3)$ was either less than 1/2 or greater than 1/2. This would not be possible if player 2 chose $a_2 = B$. However, if $a_2 = A$ then it is possible that $Pr(\Upsilon_A = \Upsilon \mid y_1 \in \{A^A, A^B\}, y_2 \in \{A^A, A^B, B^A\}, y_3 \in \{B^B\}) > 1/2$, that is, even if player 3 gets the most informative B signal after the first two players chose area A, the posterior belief is such that area A is the better sector.[27] For example, if $q = 0.75$ then if $p > 0.9$ player 3 would chose area A regardless of his own signal if both player 1 and 2 chose area A.[28] In this case player 3's action, $a_3 = A$ conveys no information about his signal. Thus, every future investor also rationally ignores their own signal and chooses area A. Player 3 is more likely to cascade if p is larger or q is lower.[29] This is intuitive because a higher q increases the strength of the high signals, (A^A, B^B), while lowering the strength of (A^B, B^A). The stronger the high signal the more likely it is able to provide enough information to overwhelm the

[27] For a proof see below under "Proof of Cascade Statement" as well as to find the example calculations.

[28] Note that since player 2 would chose area A only if his signal $y_2 \in \{A^A, A^B, B^A\}$, therefore action $a_2 = A$ conveys the knowledge $y_2 \in \{A^A, A^B, B^A\}$, and since player 1 chose action $a_1 = A$ player three also knows that $y_1 \in \{A^A, A^B\}$.

[29] For a proof see below under "Proof of Derivative Statements".

actions of past players, particularly if their past actions might have been based on very uninformative signals, (A^B, B^A). While if p is high enough then so much information is conveyed by the first two players choosing A that the third signal is not enough to counter it.

Alternatively, if player 2 chooses B then player 3 would also choose B if $y_3 \in \{B^B, B^A, A^B\}$ (if q is large enough) but if $y_3 \in \{A^A\}$ then player 3 chooses A. In this case player four is then facing the same information set as player 2 and so player 4 can be analyzed just as player 2 and player 5 as player 3, etc. Or alternatively player 3 would also choose B if $y_3 \in \{B^B, B^A\}$ (if q is small enough) but if $y_3 \in \{A^A, A^B\}$ then player 3 chooses A. In this case player 4 uses all three previous actions to form beliefs plus his own signal. From this we can see that if player signals oscillate between A and B signals it is possible that no information cascade occurs by n players.

Thus, it is possible that one area receives the signal $s_j(t) = 1$ or both do or neither does. This depends on m and n and when and if an information cascade occurs.[30] Furthermore, the cascade will often occur on the 'wrong' area, i.e., the area where $\Upsilon_j = 0$. Moreover this cascade can be 'fragile' in the sense that when everyone's actions are based on very few actual signals mild amounts of new information can flip the equilibrium to the other area.

Proof of Cascade Statement:

To see that if p is low enough then $Pr(\Upsilon_A = \Upsilon \mid y_1 \in \{A^A, A^B\}, y_2 \in \{A^A, A^B, B^A\}, y_3 \in \{B^B\}) > 1/2$ first note that

$$Pr(y_1 \in \{A^A, A^B\}, y_2 \in \{A^A, A^B, B^A\}, y_3 \in \{B^B\} \mid \Upsilon_A = \Upsilon) =$$
$$Pr(y_1 \in \{A^A, A^B\} \mid \Upsilon_A = \Upsilon) \cdot Pr(y_2 \in \{A^A, A^B, B^A\} \mid \Upsilon_A = \Upsilon) \cdot Pr(y_3 \in \{B^B\} \mid \Upsilon_A = \Upsilon)$$
$$= p \cdot (p + (1-p)q) \cdot (1-p)(1-q) \quad \text{(B4)}$$

because all signals are conditionally independent. Furthermore,

$$Pr(y_1 \in \{A^A, A^B\}, y_2 \in \{A^A, A^B, B^A\}, y_3 \in \{B^B\} \mid \Upsilon_B = \Upsilon) =$$
$$Pr(y_1 \in \{A^A, A^B\} \mid \Upsilon_B = \Upsilon) \cdot Pr(y_2 \in \{A^A, A^B, B^A\} \mid \Upsilon_B = \Upsilon) \cdot Pr(y_3 \in \{B^B\} \mid \Upsilon_B = \Upsilon)$$

[30] For example, if $n = 10$ and $m = 6$ then either one area gets signal $s(t) = 1$ or neither does, while if $m = 4$ then potentially both areas get the signal $s(t) = 1$ and at least one will.

$$= (1-p)\cdot(1-p+p(1-q))\cdot pq \quad \text{(B5)}$$

Therefore,

$$\begin{aligned}
&Pr(y_1 \in \{A^A, A^B\}, y_2 \in \{A^A, A^B, B^A\}, y_3 \in \{B^B\}) = \\
&\quad Pr(y_1 \in \{A^A, A^B\}, y_2 \in \{A^A, A^B, B^A\}, y_3 \in \{B^B\} \mid \Upsilon_A = \Upsilon) Pr(\Upsilon_A = \Upsilon) \\
&\quad + Pr(y_1 \in \{A^A, A^B\}, y_2 \in \{A^A, A^B, B^A\}, y_3 \in \{B^B\} \mid \Upsilon_B = \Upsilon) Pr(\Upsilon_B = \Upsilon) \\
&\quad = (p\cdot(p+(1-p)q)\cdot(1-p)(1-q))\frac{1}{2} + ((1-p)\cdot(1-p+p(1-q))\cdot pq)\frac{1}{2} \quad \text{(B6)}
\end{aligned}$$

Thus,

$$\begin{aligned}
&Pr(\Upsilon_A = \Upsilon \mid y_1 \in \{A^A, A^B\}, y_2 \in \{A^A, A^B, B^A\}, y_3 \in \{B^B\}) \\
&\quad = \frac{Pr(y_1 \in \{A^A, A^B\}, y_2 \in \{A^A, A^B, B^A\}, y_3 \in \{B^B\} \mid \Upsilon_A = \Upsilon) Pr(\Upsilon_A = \Upsilon)}{Pr(y_1 \in \{A^A, A^B\}, y_2 \in \{A^A, A^B, B^A\}, y_3 \in \{B^B\})} \\
&\quad = \frac{p(p+(1-p)q)(1-p)(1-q)}{p(p+(1-p)q)(1-p)(1-q) + (1-p)(1-p+p(1-q))pq} \\
&\quad = \frac{(p+(1-p)q)(1-q)}{(p+(1-p)q)(1-q) + (1-pq)q} \quad \text{(B7)}
\end{aligned}$$

which is greater than 1/2 if $(p+(1-p)q)(1-q) > (1-pq)q$ which reduces to $p - 2pq - qq + 2pqq > 0$ if $q = 3/4$ then this becomes $p - 3p/2 - 9/16 + 9p/8 > 0$ or $p > 9/10$.

Proof of Derivative Statements:

$\frac{\partial}{\partial q} Pr(\Upsilon_A = \Upsilon \mid y_1 \in \{A^A, A^B\}, y_2 \in \{A^A, A^B, B^A\}, y_3 \in \{B^B\}) < 0$ is true if

$$\frac{\partial}{\partial q}\frac{(p+(1-p)q)(1-q)}{(p+(1-p)q)(1-q)+(1-pq)q} < 0 \quad \text{(B8)}$$

which reduces to

$$((p+(1-p)q)(1-q)+(1-pq)q)(1-2p-2(1-p)q) < (p+(1-p)q)(1-q)(2-2p-2q) \quad \text{(B9)}$$

or

$$2ppq - 2ppqq + 2pqq - qq < p \quad \text{(B10)}$$

which is true for all $1/2 < q < p < 1$.

$\frac{\partial}{\partial p} Pr(\Upsilon_A = \Upsilon \mid y_1 \in \{A^A, A^B\}, y_2 \in \{A^A, A^B, B^A\}, y_3 \in \{B^B\}) > 0$ is true if

$$\frac{\partial}{\partial p} \frac{(p + (1-p)q)(1-q)}{(p + (1-p)q)(1-q) + (1-pq)q} > 0 \tag{B11}$$

which reduces to

$$((p+(1-p)q)(1-q)+(1-pq)q)(1-q)(1-q) > (p+(1-p)q)(1-q)((1-q)(1-q)-qq) \tag{B12}$$

or $(1-q) > -qq$, which is true for all q.