

Innovation and Strategic Network Formation

KRISHNA DASARATHA

Yale University

First version received February 2020; Editorial decision March 2022; Accepted April 2022 (Eds.)

We study a model of innovation with a large number of firms that create new technologies by combining several discrete ideas. These ideas are created via private investment and spread between firms. Firms face a choice between secrecy, which protects existing intellectual property, and openness, which facilitates learning from others. Their decisions determine interaction rates between firms, and these interaction rates enter our model as link probabilities in a learning network. Higher interaction rates impose both positive and negative externalities, as there is more learning but also more competition. We show that the equilibrium learning network is at a critical threshold between sparse and dense networks. At equilibrium, the positive externality from interaction dominates: the innovation rate and welfare would be dramatically higher if the network were denser. So there are large returns to increasing interaction rates above the critical threshold. Nevertheless, several natural types of interventions fail to move the equilibrium away from criticality. One effective policy solution is to introduce informational intermediaries, such as public innovators who do not have incentives to be secretive. These intermediaries can facilitate a high-innovation equilibrium by transmitting ideas from one private firm to another.

Key words: Innovation, Social networks, Network formation, Random networks, Diffusion

JEL Codes: D85; D83; 033

1. INTRODUCTION

A growing body of empirical research suggests that interactions between inventors are an important part of innovation.¹ New technologies are often produced by combining individual insights with learning from peers, which confers large benefits on firms and inventors engaged in such learning. When highly connected clusters of firms emerge in a location, as in the technology industry in Silicon Valley, inventors in these areas are much more productive.

But frequent collaboration and learning are not assured even when inventors in a given industry co-locate (Saxenian, 1996 gives a well-known example). Rather, interaction patterns depend on firms' decisions, such as how much to encourage their employees to interact with employees from other firms. These interactions let a firm learn from other companies and inventors. There are also downsides for the firm, as employees may share valuable information. A more secretive approach

1. The benefits from interactions between inventors and movement of inventors have been quantified empirically by Akcigit, Caicedo, Miguelez, Stantcheva and Sterzi (2018), Kerr (2008), Samila and Sorenson (2011), among others.

The editor in charge of this paper was Andrea Galeotti.

allows firms to prevent potential competition by protecting intellectual property. In making these types of decisions, firms, and inventors must choose between openness and secrecy.

We develop a theory of firms' endogenous decisions about how much to interact with other firms, and the consequences for information flows and the rate of innovation. Firms' choices determine the probabilities of links in a *learning network*; we thus contribute to a theoretical literature on network formation.² We show that at equilibrium, the learning network has a special structure that would be unlikely to arise exogenously: it is at a *critical threshold* between sparse and dense networks. This implies extreme inefficiencies arise at equilibrium, and we analyse the welfare and policy implications.

We study a framework where each firm chooses two intensities: how open to be as well as how much to invest in R&D. The choices of levels of openness determine interaction rates between firms, and the probability that one firm learns from another is equal to the interaction rate between the two firms. If a given firm is more open, that firm is more likely to learn *ideas* from other firms—but other firms are more likely to learn its ideas. These ideas are valuable because they can be combined to make new *technologies*, which are finite sets of ideas (as in Weitzman, 1998). Firms can generate profits by producing technologies, but the profits from a technology are erased by competition if another firm also knows the component ideas in that technology.

Our first contribution, which is methodological, is to develop a theory of endogenous formation of random networks in the context of our economic application. Learning opportunities are random events, and their realizations determine a learning network. We therefore consider link formation decisions with uncertainty in the matching process, while the leading approach in the literature on network-formation games focuses on deterministic models that allow agents to choose particular links (Jackson and Wolinsky, 1996 and Bala and Goyal, 2000). Since we take actions to be continuous choices that translate to interaction rates, optimal behaviour satisfies first-order conditions rather than a high-dimensional system of combinatorial inequalities.

A key feature of our model is that ideas can spread several steps through this network: when one firm learns from another, the information transferred can include ideas learned from a third firm. We refer to this as indirect learning.³ Under indirect learning, firms' incentives depend on the global structure of the network. Each firm would like to learn many ideas, since then the firm could combine these ideas to produce a large number of new technologies, and much of this learning can be indirect.

Analysing the global structure of the network leads to our second contribution, which is to establish the *criticality of equilibrium*. When there are many firms, learning outcomes depend dramatically on whether the learning network is sparsely connected or densely connected. If firms' interaction rates are below a critical threshold, the learning network consists of many small clusters of firms who learn few ideas. Above the threshold, the learning network has a giant component asymptotically: a large group of firms who learn a large number of ideas and can incorporate these ideas into many new technologies. To determine where equilibrium lies with respect to this threshold, we analyse an individual firm's decision problem in each of these two domains, i.e., when other firms form a sparse or dense network.

2. An important feature of these choices is that firms combine different ideas to create new technologies (see also Acemoglu and Azar, 2020; Chen, Elliott and Koh, 2021 for related combinatorial production functions). Strategic complementarities play a crucial role in many network games (e.g. Bramoullé, Kranton and D'amours, 2014; Ballester, Calvó-Armengol and Zenou, 2006), and in our setting complementarities arise endogenously from this process of combining ideas.

3. By contrast, existing work on strategic formation of random networks largely focuses on direct connections (see e.g. Curarini, Jackson and Pin, 2009).

The main result is that the equilibrium interaction rates are at the critical threshold between sparse and dense networks. Firms would deviate to interact more if the network were likely to be sparse and deviate to interact less if the network were likely to be dense. Intuitively, in sparse networks firms would increase interaction rates to fill central positions in the network, known in sociology as “structural holes,” which enable the firm to combine ideas learned via different interactions.⁴ As others interact more, these structural holes disappear, and indeed firms tend to learn the same ideas repeatedly from different interactions. So the incentives to be more open are weaker relative to the incentives to be secretive.

We next consider the implications of the main result for welfare and policy. Since equilibrium is at the critical threshold, a giant component would emerge if all firms shifted interaction rates slightly above equilibrium levels. Firms learn relatively few ideas at the threshold but could learn many more with a bit more interaction. More learning has benefits but also leads to more competition, which may discourage private investment. We show there are still unboundedly large improvements in innovation and welfare (as the number of firms grows large) from interventions that increase interaction rates above equilibrium levels. A consequence is that increasing interaction rates is a first-order concern in designing policy. By contrast, policies targeting decisions about private investment rather than interaction, such as subsidies to R&D, have minimal effect at equilibrium—but can be valuable if paired with interventions to increase openness.

Intervening to improve equilibrium learning networks is difficult, and policies such as broadly subsidizing interaction will be undermined by firms’ equilibrium responses. We discuss one type of policy change that does induce more productive interaction patterns, which is to introduce *public innovators* who do not have incentives to be secretive. For example, governments could fund academic researchers who are especially willing to interact with other researchers, including those in industry. The key is that public innovators can serve as informational intermediaries, transmitting ideas between private firms. They play a valuable role even after considering the equilibrium response of the profit-maximizing firms, who may adjust to be more secretive.

We next explore which features of the baseline model are needed to obtain a critical equilibrium. One assumption is that learning probabilities are symmetric across pairs of firms. We show that equilibrium remains critical even when firms have different propensities to learn from others, which allows some firms to be better at protecting ideas than others. The key feature driving the main result is that there is a margin along which firms can acquire incoming links at the cost of a higher probability of outgoing links.

The criticality of equilibrium is robust to alternate specifications of the benefits from learning from others but more sensitive to the costs of outgoing links. The baseline model assumes firms’ profits are additive across technologies, but equilibrium remains critical if there are increasing or slightly decreasing returns to producing many technologies. Indeed, the crucial property is that payoffs are convex in the number of ideas learned by a firm. Equilibrium outcomes do depend, however, on how much firms stand to lose from outgoing links. In particular, the results described above assume zero profits in competitive markets. If profits under competition are instead positive (e.g. because of a first-mover advantage or markets with collusion between firms), then incentives toward secrecy will be weaker and so equilibria will be above the critical threshold. These results give testable predictions about the relationship between market structure and outcomes such as the innovation rate.

Our final results ask how formal intellectual property rights change the incentives to interact. Consider the consequences of granting patents to a positive fraction of ideas, e.g., allowing

4. The concept of structural holes, introduced by Burt (1992), refers to network positions allowing agents to combine information from different connections or spread information between groups.

hardware but not software ideas to be patented. Patents mitigate firms' incentives to be secretive but can also discourage exchange of ideas. Firms with patents are more open but are also less desirable partners in interactions (at least when ideas are only transmitted directly). The resulting adverse selection in interaction can deter firms from collaborating with others. We show that patent rights can therefore prevent any productive interactions at equilibrium. If indirect learning is important, firms with patents will be informational intermediaries, like the public innovators above. In this case, there are benefits to allowing patents, but it turns out that the optimal policy is often to only allow patents for a very small fraction of ideas.

At a technical level, this article develops tools for studying incentives in random network settings. These tools are most applicable to analysing decisions in network models with complementarities between indirect connections. Classical results in graph theory characterize the component structure of large random networks, and thus in our context the number of ideas firms will learn (Karp, 1990; Łuczak, 1990). But due to the complementarities between ideas, firms' incentives also depend on *how* these ideas are learned, e.g., via many interactions or a few interactions. To capture these complementarities, we prove a key lemma relating a firm's equilibrium action to the extent to which technologies combine ideas from distinct interactions. An additional challenge is that to understand incentives, it is not enough to analyse "leading terms." Vanishing-probability events and lower-order terms in link probabilities could substantially affect payoffs. Our analysis therefore requires careful treatment of the graph branching process governing the number of ideas learned from each interaction.

1.1. Related literature

This article relates to research in network theory, especially network formation, and to models of innovation.

At a methodological level, we develop a theory of strategic network formation with probabilistic links. A large literature since Jackson and Wolinsky (1996) and Bala and Goyal (2000) considers endogenous network formation assuming that agents can choose their links exactly.⁵ Because equilibrium is then characterized by a large system of inequalities, these models illustrate key externalities in special cases but remain largely or entirely intractable in many others. By instead considering agents making a single choice about openness under uncertainty, we obtain a smooth model of link formation that can be solved via basic optimization techniques combined with analyses of random graphs.

Under this random-network model of network formation, incentives to form links depend on the "phase transitions" between sparse and dense networks.⁶ Economic models involving phase transitions have been recently explored in the context of diffusion processes by Campbell (2013), Akbarpour, Malladi and Saberi (2018), and Sadler (2020). These models let adoption and/or seeding decisions depend on component structure in an underlying network. We instead study equilibria of a game in which agents endogenously make decisions about how much to interact with others, and find there is a subtle interplay between strategic incentives and the global network structure.

5. The pairwise stability solution concept from Jackson and Wolinsky (1996) and variants have been applied to network formation in many settings, including innovation (Goyal and Moraga-Gonzalez, 2001, König, Battiston, Napoletano and Schweitzer, 2011).

6. Golub and Livne (2010) also study network formation with phase transitions and allow payoffs to depend on distance one and two connections. An important feature of our model is that firms' decisions depend on the global network structure rather than only local connections.

An alternate approach to smooth network formation is to consider weighted networks, so that each link has an intensity (as in [Cabral et al., 2011](#); [Griffith, 2019](#); [Baumann, 2021](#)).⁷ In some settings with unweighted links, this approach yields a deterministic approximation of a game on the underlying random network. But for discrete processes such as the diffusion of an idea, the random and deterministic models can be very different. Indeed, we show that there is a critical threshold corresponding to important discontinuities in network structure and outcomes that do not arise in deterministic models such as [Cabral et al. \(2011\)](#).⁸

In existing literature on innovation, approaches incorporating interactions between firms generally model these interactions as either mechanical spillovers or learning via imitation. A common approach is to choose a convenient functional form for spillovers, usually motivated by tractability within a macroeconomic (e.g. [Kortum, 1997](#)) or network-theory ([König, Battiston, Napoletano and Schweitzer, 2012](#)) framework.⁹ By microfounding these spillovers, which arise endogenously within the innovative process, we can study how spillovers respond to policy interventions.

In a different approach, which relies on a quality-ladders framework, interactions give firms a chance to catch up as innovation proceeds vertically through improvements in the quality of existing technologies (e.g. [Perla and Tonetti, 2014](#); [König, Lorenz and Zilibotti, 2016](#); [Akcigit et al., 2018](#)). We instead explicitly model innovation horizontally as a process of combining distinct ideas, which serve as building blocks for new technologies. Related models appear in [Weitzman \(1998\)](#) and [Acemoglu and Azar \(2020\)](#), which focus on the evolution of the total amount of innovation over time and do not involve learning or informational spillovers between firms. We find that when new technologies can be created in this way, changes in interaction patterns can have much larger consequences for the rate of innovation than in quality-ladders models.¹⁰

2. MODEL

We first describe our model formally and provide an example. We will then discuss interpretations of the model and its assumptions.

2.1. Basic setup

The model includes two key concepts: ideas and technologies. Ideas are the components in technologies, and each technology combines several ideas. Firms will obtain profits from selling new technologies and must acquire ideas to produce these technologies.

There are $n > 1$ firms $1, \dots, n$. Each firm i can potentially discover a distinct *idea*, also denoted by i . We assume each firm can discover a single idea for simplicity, and [Supplementary Appendix D](#) allows firms to potentially discover multiple ideas. We let $I \subset \{1, \dots, n\}$ be the set of ideas that are discovered.

7. Relatedly, [Erol and García-Jimeno \(2020\)](#) consider a continuum model where agents endogenously and randomly form links with a positive fraction of peers.

8. [Berlant and Fujita \(2008\)](#) study a related deterministic model of knowledge creation with individuals choosing to produce knowledge alone or with a partner. Their analysis focuses on interaction patterns in small disconnected groups, while we find social networks can be more connected in important ways.

9. In simulations, [Baum, Cowan and Jonard \(2010\)](#) study the formation of innovation networks via a mechanical process.

10. By assuming a continuum of firms, macroeconomic models of imitation often implicitly restrict to network structures without a giant component.

Each firm i chooses a probability $p_i \in [0, 1]$ of discovering this idea and pays investment cost $c(p_i)$. We will assume that c is continuously differentiable, increasing, and convex with $c(0)=0$ and $\lim_{p \rightarrow 1^-} c(p)=\infty$. The realizations of discoveries are independent.

A *technology* $t=\{i_1, \dots, i_k\}$ is a set of k ideas $i_1, \dots, i_k \in I$, where $k > 1$ is an exogenous parameter capturing the complexity of technologies.¹¹ Each idea $i \in t$ must be discovered by the corresponding firm to be included in a technology. There are therefore $\binom{n}{k}$ potential technologies, and a firm i can produce more than one technology. An equivalent interpretation, which we describe in Section 2.4, is that firms can explore many potential technologies and only a small fraction are feasible to bring to market.

Each firm i chooses a level of *openness* $q_i \in [0, 1]$. Given choices q_i and q_j , the interaction rate between i and j is $\iota(q_i, q_j)$. We will assume the multiplicative interaction rate

$$\iota(q_i, q_j) = q_i q_j$$

except when a more general interaction rate is explicitly stated (in Theorem 1'). The interaction rate determines the probability that firm i learns from firm j and vice versa, as described below.

The timing of the model is simultaneous: firms choose actions p_i and q_i and then all learning occurs. We denote the vectors of actions by (\mathbf{p}, \mathbf{q}) . When actions are symmetric, we will refer to p_i by p and q_i by q .

Given actions \mathbf{p} and \mathbf{q} , we denote the set of ideas that firm i learns from others by $I_i(\mathbf{p}, \mathbf{q}) \subset I$. This is a random set depending on realizations of learning and discoveries. We now describe how learning occurs.

With probability $\iota(q_i, q_j)$, firm i learns directly from firm j . In this case, firm i learns idea j if $j \in I$. If firm i learns directly from firm j , then with probability $\delta \in [0, 1]$, firm i also learns indirectly through firm j . In this case, firm i also learns all ideas in $I_j(\mathbf{p}, \mathbf{q})$. All realizations of direct and indirect learning are independent, and in particular, firm i can learn from firm j without j learning from i .

When $\delta=0$ there is only *direct learning*, while when $\delta>0$ *indirect learning* can also occur. When $\delta>0$ we define a directed network, which we call the indirect-learning network, with nodes $1, \dots, n$ and a link from node j to node i if firm i learns indirectly through firm j .

A firm i receives payoff 1 from each *proprietary technology* t . A technology t is proprietary for firm i if (1) $i \in t$ and (2) i is the unique firm such that $j \in \{i\} \cup I_j(\mathbf{p}, \mathbf{q})$ for all $j \in t$. In words, the technology contains firm i 's idea and firm i is the unique firm that knows all ideas in the technology, so firm i has a monopoly over the technology.

If t is not a proprietary technology for firm i , then firm i receives payoff 0 from the technology t . In Section 5, we will consider more general payoff structures in which (1) payoffs are not additive across technologies and (2) firms can receive non-zero payoffs $f(m)$ if $m>0$ other firms know all ideas contained in a technology.

2.2. Payoffs

Given actions (\mathbf{p}, \mathbf{q}) and a firm i , we define $T_i(\mathbf{p}, \mathbf{q})$ to be the set of technologies t such that (1) $i \in t$ and (2) firm i knows all ideas $j \in t$. The *proprietary technologies* $PT_i(\mathbf{p}, \mathbf{q}) \subset T_i(\mathbf{p}, \mathbf{q})$ for i are then the subset of technologies $t \in T_i(\mathbf{p}, \mathbf{q})$ such that no other firm knows all ideas in t . Note that these sets are random objects depending on link realizations. The expected payoff to firm i is

$$U_i(\mathbf{p}, \mathbf{q}) = \mathbb{E}[|PT_i(\mathbf{p}, \mathbf{q})|] - c(p_i).$$

11. In the baseline model, the parameter k is the same for all firms.

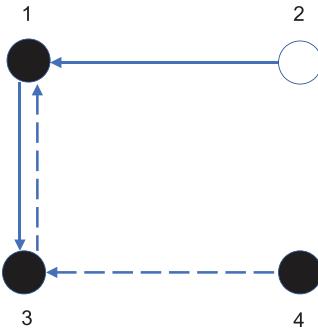


FIGURE 1

Network with four firms and $k = 3$. Black circles are firms that discover ideas while white circles do not discover ideas. Arrows are in the direction of information flow, and dashed arrows indicate direct learning while solid arrows indicate indirect learning. The only technology produced is $t = \{1, 3, 4\}$ and Firm 3 receives monopoly profit.

To further illustrate payoffs, we write the cardinality of $PT_i(\mathbf{p}, \mathbf{q})$ explicitly when $\delta = 1$. Recall that $I_i(\mathbf{p}, \mathbf{q})$ is the set of ideas learned by firm i given actions (\mathbf{p}, \mathbf{q}) . Like $PT_i(\mathbf{p}, \mathbf{q})$, this is also a random object.

When $\delta = 1$, the expected payoffs to firm i are

$$U_i(\mathbf{p}, \mathbf{q}) = p_i \cdot \mathbb{E} \left[\binom{|I_i(\mathbf{p}, \mathbf{q})|}{k-1} \right] \cdot \prod_{j \neq i} (1 - \iota(q_i, q_j)) - c(p_i).$$

A technology t that i profits from consists of i 's private idea, which is developed with probability p_i , and a choice of $(k-1)$ other ideas known to i . The firm j faces competition if and only if some firm learns all of i 's ideas, and the probability that this does not occur is $\prod_{j \neq i} (1 - \iota(q_i, q_j))$. Finally, the private investment cost is $c(p_i)$.

In general, a firm can face competition for a technology t in two ways. First, a firm j can learn all of firm i 's ideas via indirect learning. Second, a firm j can learn i 's private idea directly and then the other ideas in the technology t from links with firms other than i . The probability of the second possibility is more difficult to express in closed form, and in general depends on the technology t . We will show that when there is not too much interaction, most competition comes via the first channel.

2.3. Example

To illustrate the mechanics of the model, we describe a simple example with $n = 4$ firms and complexity $k = 3$. Suppose that firms choose some actions (\mathbf{p}, \mathbf{q}) . As an example, we consider the particular realizations such that (1) ideas are discovered by firms in $I = \{1, 3, 4\}$ and (2) Firm 1 learns indirectly through Firm 2 and directly from Firm 3, Firm 3 learns indirectly through Firm 1, and firm 3 learns directly from Firm 4.

The network and ideas are shown in Figure 1. Black circles correspond to firms with ideas $i \in I$, i.e., firms that discover ideas, while white circles correspond to firms with ideas $i \notin I$, i.e., firms that do not discover ideas. Solid arrows denote indirect learning links, while dashed arrows indicate only direct learning occurred.

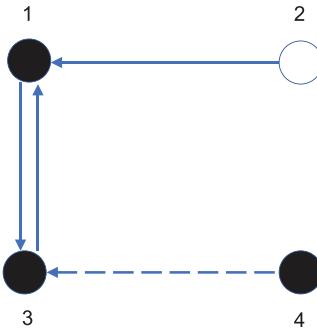


FIGURE 2

Network with four firms and $k = 3$. Black circles are firms that discover ideas while white circles do not discover ideas. Arrows are in the direction of information flow, and dashed arrows indicate direct learning while solid arrows indicate indirect learning. Firm 1 now learns indirectly from Firm 3, unlike in Figure 1. The only technology produced is $t = \{1, 3, 4\}$, and there are no profits because Firms 1 and 3 both produce t .

Since $k = 3$, the unique technology t consisting of ideas in I is $t = \{1, 3, 4\}$. The realizations of the sets $I_i(\mathbf{p}, \mathbf{q})$ of ideas learned from others are

$$I_1(\mathbf{p}, \mathbf{q}) = \{3\}, I_2(\mathbf{p}, \mathbf{q}) = \emptyset, I_3(\mathbf{p}, \mathbf{q}) = \{1, 4\}, I_4(\mathbf{p}, \mathbf{q}) = \emptyset.$$

Because Firm 3 is the unique firm such that $t \subset I_i(\mathbf{p}, \mathbf{q}) \cup \{i\}$, and we have $3 \in t$, Firm 3 produces the technology t and receives monopoly profit of 1 for that technology. There are no profits from any other technologies.

Suppose instead that Firm 1 also learns indirectly through Firm 3, as shown in Figure 2. Then, we have

$$I_1(\mathbf{p}, \mathbf{q}) = \{3, 4\}, I_2(\mathbf{p}, \mathbf{q}) = \emptyset, I_3(\mathbf{p}, \mathbf{q}) = \{1, 4\}, I_4(\mathbf{p}, \mathbf{q}) = \emptyset.$$

The only potential technology remains $t = \{1, 3, 4\}$. We now have $t \subset I_i(\mathbf{p}, \mathbf{q}) \cup \{i\}$ for both firm 1 and firm 3, so both receive the competitive profit of zero for that technology. There are also no profits from other technologies.

2.4. Interpretation and discussion

Before expressing expected payoffs of firms and defining equilibrium, we discuss interpretation and assumptions in the model.

2.4.1. Actions. Firm actions are choices (p_i, q_i) . The first component p_i corresponds to a level of investment in R&D. A small probability of a discovery is cheap, while probabilities close to one are very expensive.

The second component q_i corresponds to a level of openness or secrecy in interactions with other firms. As one example, consider a technology company's decision about whether to locate in an area with many other technology firms. Locating near other firms will lead to more casual interactions between the employees of the firm making the choice and employees of other firms (e.g. at bars and restaurants), and information can be shared in either direction in these interactions.¹² In addition to a firm's choice of location, the action q_i could include decisions

12. A large body of work describes the role of geographical proximity in driving innovation (e.g. Storper and Venables, 2004). Kelly (2009) considers the impact of geography on information-sharing and describes

such as whether to send employees to conferences and how much disclose ongoing R&D to employees.

An important feature of the model is that increasing q_i increases the probability that firm i learns from other firms but also increases the probability that other firms learn from i .¹³ The baseline model assumes that learning probabilities are symmetric: firm i learns from firm j with the same probability that firm j learns from firm i . In Section 4, we allow firms to have heterogeneous propensities to learn across firms and find this symmetric structure does not drive results.

The downside to interaction for a firm i is the increased probability of outgoing links, not an exogenous link-formation cost. Because the costs of links are an endogenous feature of the model, our equilibrium characterization does not depend on functional forms of costs, as it would with exogenous link costs separate from the innovation process.¹⁴

2.4.2. Formal and informal interactions. The model is meant to primarily describe informal interactions between employees or firms, rather than more formal arrangements such as licensing agreements or joint R&D ventures. As such, our results are most applicable to industries where formal property rights are imperfectly enforced (Section 6 discusses the interplay between informal interactions and more formal property rights). Because information transmitted via informal interactions can often spread several steps, an analysis considering global network structure is particularly relevant.

In Supplementary Appendix D, we compare the payoffs to firms with different numbers of private ideas. This analysis can also be interpreted as measuring the value of formal contracting arrangements allowing multiple firms to share ideas frictionlessly. We find that as the number of firms grows large, the benefits to such an arrangement are a vanishing fraction of a firm's expected profits.

2.4.3. Interaction rate. The multiplicative interaction rate $\iota(q_i, q_j) = q_i q_j$ has the feature that firm i 's probability of learning from another firm and that firm's probability of learning from i are both proportional to q_i . Thus, this is the (unique up to rescaling) interaction rate that arises from a random matching process in which all agents choose a search intensity and the probability of learning in each direction is proportional to that intensity.

We will show in Section 3 that under a symmetry assumption, the main result does not rely on the multiplicative functional form and extends to any $\tilde{\iota}: [0, 1] \times [0, 1] \rightarrow [0, 1]$ satisfying several mild properties.

2.4.4. Learning network. A useful assumption is that if firm i learns indirectly through firm j , then firm i learns *all* ideas known to j . This ensures that there is a well-defined learning network, and this network is a central object in our analysis. If indirect learning were not perfectly correlated across ideas, there would be a separate learning network for each idea.

a phase transition in exogenously determined networks; our approach differs in that we model firm decisions including factors such as location as an endogenous choice.

13. Stein (2008) gives a microfoundation for bilateral communication in the context of innovation.

14. Acemoglu, Makhroumi, Malekian and Ozdaglar (2017) consider a similar link cost in a setting where the benefits depend only on direct connections and the network-formation game is deterministic. Outgoing links are undesirable in their model because of a primitive preference for privacy.

2.4.5. Firm profits. The positive payoffs from producing proprietary technologies correspond to monopoly payoffs, which we normalize to 1. Formally, all technologies give the same monopoly profits and these profits are deterministic. It would be equivalent to take monopoly profits to be randomly drawn from any distribution with finite mean, as long as firms have no information about the realizations *a priori*. For example, only a small constant fraction of technologies could actually be profitable enough to produce.

If multiple firms know all ideas contained in t , then there is a competitive market and firms receive zero profits. This baseline payoff structure, which we generalize in Section 5.2, corresponds to Bertrand competition.

Our setup requires that monopolist firms must have privately developed one of the ideas in a technology to produce that technology, but competitors need not. To start a new market, some expertise and/or confidence in the quality of the relevant idea is needed. Once a market exists, however, entrants do not require this expertise, perhaps because relevant details can be obtained from the competitor's technology.

3. EQUILIBRIUM

In this section, we define and then characterize equilibrium in our model. The characterization first briefly describes investment equilibria under direct learning ($\delta=0$). The remainder of the section shows that equilibria are at a critical threshold under indirect learning ($\delta>0$) and discusses implications for innovation and policy.

Two assumptions that simplify the initial analysis are that firms are homogeneous (which we will relax in several ways, including Section 4) and that profits are equal to the number of proprietary technologies (which we relax in Section 5).

3.1. Solution concept

We begin by defining our solution concept:

Definition 1. An equilibrium $(\mathbf{p}^*, \mathbf{q}^*)$ is a pure-strategy Nash equilibrium. An equilibrium $(\mathbf{p}^*, \mathbf{q}^*)$ is an investment equilibrium if $p_i^* > 0$ for all i .

Because all choices p_i and q_i are probabilities of discoveries or interactions, we restrict to pure strategies.

If $p_i=0$ for all i , then any \mathbf{q} will give an equilibrium: if no other firms are investing, there is no reason to invest and so payoffs are zero. It is easy to see these trivial equilibria always exist, and we will focus on investment equilibria.

For some of our results, it will also be useful to make the stronger assumption that private investment is non-vanishing asymptotically. We consider a sequence of equilibria as the number of firms $n \rightarrow \infty$.

Definition 2. A sequence of equilibria $(\mathbf{p}^*, \mathbf{q}^*)$ has non-vanishing investment if

$$\liminf_n \min_i p_i^* > 0.$$

Depending on $c(\cdot)$, there may be equilibria at which all firms choose very low levels of private investment because others are investing very little. The definition excludes these partial coordination failures as well.

3.2. Direct learning

We briefly summarize results with $\delta=0$ here, and give a full analysis in Supplementary Appendix C. In this case, ideas can spread at most one step.

There exists a symmetric investment equilibrium for n large, and at any sequence of symmetric investment equilibria the interaction rate is

$$\iota(q^*, q^*) \approx \left(\frac{k-1}{n} \right)^{\frac{1}{k}}.$$

Since the interaction rate is of order $n^{-\frac{1}{k}}$, the probability that a generic firm knows all the ideas in a given technology is of order $\frac{1}{n}$. It follows that the probability that there exists competition on a given technology is constant.

For n large, each firm learns from a large number of other firms with high probability. We will see that interaction rates are much lower in the indirect-learning case. With only direct learning much more interaction is needed to generate a substantial risk of competition, so the interaction rate must be higher for potential competition to meaningfully deter openness.

3.3. Main result

Our main focus is the indirect learning case ($\delta > 0$) in which ideas can spread multiple steps. We now show that when $\delta > 0$, equilibrium networks are at the critical threshold asymptotically.

We begin by defining this critical threshold. Let the number of firms $n \rightarrow \infty$ and consider outcomes under a sequence of symmetric actions. We say that an event occurs asymptotically almost surely (a.a.s.) if the probability of this event converges to 1 as $n \rightarrow \infty$. To simplify notation, we often omit the index n (e.g. from the actions (p_i, q_i) .)

Definition 3. A sequence of symmetric actions with openness \mathbf{q} is:

- *Subcritical* if $\limsup_n \iota(q, q) \delta n < 1$
- *Critical* if $\lim_n \iota(q, q) \delta n = 1$
- *Supercritical* if $\liminf_n \iota(q, q) \delta n > 1$

The expected number of firms with links to i in the indirect-learning network is

$$\iota(q, q) \delta(n-1),$$

so the three cases distinguish networks where each firm learns indirectly less than once, approximately once, and more than once in expectation. In the subcritical case, it follows that the expected number of firms that learn a given idea is a finite constant. In the supercritical case, there is a positive probability that a given idea is learned by a large number of firms (i.e. a number growing linearly in n).

This intuition is formalized by results from the theory of random directed graphs (Karp, 1990; Luczak, 1990). Adapting their results to this setting, we have the following result. A *component* of a directed network is a strongly connected component, i.e., a maximal set of nodes such that there is a path from any node in the set to any other.

Lemma 1 (Theorem 1 of Luczak (1990)) Suppose \mathbf{q} is symmetric.

(i) If the indirect-learning network is subcritical, then a.a.s. every component has size $O(\log n)$.

(ii) If the indirect-learning network is supercritical, then a.a.s. there is a unique component of size at least $\tilde{\alpha}n$ for a constant $\tilde{\alpha} \in (0, 1)$ depending on $\lim_{n \rightarrow \infty} \iota(q, q)\delta n$, and all other components have size $O(\log n)$.

These asymptotic results each imply that large finite graphs have the component structures described with high probability. It follows from the lemma that in a subcritical sequence of equilibria, all firms learn at most $O(\log n)$ ideas a.a.s. In a supercritical sequence of equilibria, there is a positive fraction of firms learning a constant fraction of all ideas a.a.s. At a critical equilibrium, the number of ideas learned lies between the subcritical and supercritical cases.

To discuss asymmetric strategies and later heterogeneity in firms, we now generalize the notion of criticality to arbitrary strategies. Consider the matrix $(\iota(q_i, q_j)\delta)_{ij}$. The entry (i, j) is equal to the probability that firm i learns indirectly from firm j . Let λ be the spectral radius of this matrix, i.e., the largest eigenvalue.

Definition 4. A sequence of actions with openness \mathbf{q} is:

- Subcritical if $\limsup_n \lambda < 1$
- Critical if $\lim_n \lambda = 1$
- Supercritical if $\liminf_n \lambda > 1$

We will see that, as in Lemma 1, the critical threshold corresponds to the emergence of a giant component. To show this, we will combine the results of Bloznelis, Götze and Jaworski (2012) with analysis of multi-type branching processes.

Our existence result establishes that there are equilibria with non-zero investment and communication. Our characterization result shows that asymptotically, equilibrium is on the threshold between sparse and dense networks:

Theorem 1. For n sufficiently large, there exists a symmetric investment equilibrium. Any sequence of investment equilibria is critical.

Theorem 1 makes a sharp prediction about equilibrium. The theorem assumes symmetric firms and payoffs that are linear in the number of monopoly technologies produced by a firm. We will show that equilibrium remains critical with heterogeneity in firms (Section 4) and non-linear payoffs (Section 5.1) but does depend on the structure of competition (Section 5.2).

At a sequence of symmetric investment equilibria, the theorem implies that $\iota(q^*, q^*) \rightarrow \frac{1}{\delta n}$, and in particular symmetric investment equilibria are asymptotically unique. While it is easy to see there must exist a symmetric equilibrium, *a priori* there need not be an equilibrium with non-zero interaction and investment. In fact, we show that there exists a sequence of symmetric equilibria with non-vanishing investment. The existence result relies on analysis of firms' best responses in each region to show a fixed point theorem applies.

We are able to drop the assumption of symmetric strategies, which is standard in settings involving random networks (e.g. Currarini *et al.*, 2009; Golub and Livne, 2010; Sadler, 2020) and show any equilibrium is at the critical threshold. Asymmetric equilibria could feature firms with $\iota(q_i^*, q_i^*)$ above and below $\frac{1}{\delta n}$.

The proof of Theorem 1 builds on existing mathematical results on large random graphs, and generalizes them to allow complementarities between ideas and endogenous link probabilities. The first obstacle to applying existing results is that the combinatorial structure of technologies generates complementarities between ideas, so payoffs and incentives do not simply depend on the expected number of ideas learned. A second issue is that link probabilities are endogenous, so

lower-order terms in link probabilities and vanishing-probability events can matter asymptotically. We now discuss the key ideas in the proof, including how we address these challenges.

Proof. intuition We describe the basic idea of the proof in the case $\delta = 1$, and the general argument is similar. We also begin by discussing symmetric strategies.

The first-order condition for the action q_i says that at any best response, the expected cost to firm i of allowing a firm j to learn from i is equal to the expected benefit from learning from an additional firm j . A key feature is that the cost and benefit both depend on the distribution of the number of ideas learned from a given link. The proof exploits this symmetry between costs and benefits to solve for $q^* n$. We are able to do so because of the endogenous downside to outgoing links, which depends on the number of ideas that firm i learns.

We use the first-order condition at a symmetric equilibrium to obtain an expression for $q^* n$ in terms of the number of incoming links used to learn the ideas in an average proprietary technology. Consider a technology t such that i produces t and gets monopoly profits. This technology is a combination of ideas learned from different links. For example, if $k=4$, an example technology could consist of i 's private idea, two ideas learned indirectly from firm j , and one idea learned directly from firm j'' . In this example, the technology would combine ideas from three different links.

More generally behaviour will depend on the number of links utilized in learning the ideas in a technology t . We refer to this number of links as $\tau(t)$, so that $\tau(t)=3$ in the example in the previous paragraph. The key tool, which we state in the subcritical region, is:

Lemma 2. *Along any sequence of symmetric investment equilibria with $\limsup \delta \iota(q^*, q^*)n < 1$,*

$$\delta \iota(q^*, q^*)n \sim \mathbb{E}_{t \in PT_i(\mathbf{p}^*, \mathbf{q}^*)}[\tau(t)]$$

for all i .¹⁵

Lemma 2 says that the expected number of other firms from whom i learns is equal to the expected value of $\tau(t)$ for a random proprietary technology t . We give a brief intuition for the lemma. If $\tau(t)$ is higher, then there are stronger complementarities between links, because produced technologies combine ideas from more links. In this case, if a firm has a few existing links, an additional link will be more valuable than an existing link due to these complementarities. Since additional links are relatively more valuable, firms are willing to interact more.

Since $\tau(t)$ is always at least one, Lemma 2 implies that $\lim_n \iota(q^*, q^*)n \geq 1$, so there cannot be a subcritical equilibrium.

In the supercritical region, almost all proprietary technologies $t \in PT_i(\mathbf{p}^*, \mathbf{q}^*)$ are created by combining a private idea with $(k-1)$ ideas learned from observing the giant component. In particular, payoffs are determined up to lower order terms by whether firm i has a link that provides a connection to the giant component. Given such a link, additional links add little value. Thus there are not complementarities between links; indeed, links are substitutes due to the potential redundancies.

But because firms have more to lose from an outgoing link in the supercritical region, complementarities between links are needed to sustain high interaction rates. Since these complementarities are not present, there is not a supercritical equilibrium either. We check this intuition formally by a computation.

Extending results to asymmetric equilibria presents several additional technical obstacles. One is that existing mathematical results, e.g., Bloznelis *et al.* (2012), prove the component

15. The distribution over technologies t is defined in Appendix A.

structure has certain properties asymptotically almost surely. But this does not remove the possibility that vanishing-probability events distort incentives in an unknown direction. To rule this out, we show that an arbitrary subcritical sequence of equilibria, the asymptotic probability $\lim_{n \rightarrow \infty} \mathbb{P}[|I_i(\mathbf{p}, \mathbf{q})| = y]$ that firm i learns y ideas decays exponentially in y . The proof bounds $|I_i(\mathbf{p}, \mathbf{q})|$ above with the number of nodes in a multi-type Poisson branching process and then analyses this branching process. ||

The analysis of the symmetric equilibria extends to more general interaction rates. The proof of Theorem 1 also shows:

Theorem 1'. *Suppose the interaction rate is a strictly increasing and continuously differentiable function $\tilde{\iota}: [0, 1] \times [0, 1] \rightarrow [0, 1]$ satisfying $\tilde{\iota}(q, q') = \tilde{\iota}(q', q)$ for all q and q' and $\tilde{\iota}(q, 0) = 0$ for all q .¹⁶ Then for n sufficiently large, there exists a symmetric investment equilibrium. Any sequence of symmetric investment equilibria is critical.*

3.4. Discovery rate and policy implications

We next discuss consequences of Theorem 1 for innovation and welfare. There are large gains to exogenously increasing interaction, but designing policies to realize these gains is subtle.

We first define a measure of the innovation rate, which is the fraction of possible technologies that are produced. Recall that $T_i(\mathbf{p}, \mathbf{q})$ is the set of technologies known to firm i , so the set of potential technologies that are produced is the union $\bigcup_i T_i(\mathbf{p}, \mathbf{q})$. This union is a subset of the $\binom{n}{k}$ possible technologies.

Definition 5. *Given actions (\mathbf{p}, \mathbf{q}) , the discovery rate*

$$D(\mathbf{p}, \mathbf{q}) = \frac{\mathbb{E}[|\bigcup_i T_i(\mathbf{p}, \mathbf{q})|]}{\binom{n}{k}}$$

is the expected fraction of potential technologies that are produced.

Given any subcritical or critical sequence of actions, each firm learns $o(n)$ ideas asymptotically almost surely, so the discovery rate converges to zero. Along any sequence of supercritical actions, there are a positive fraction of firms learning some fraction α of ideas, so the discovery rate is non-vanishing.

Theorem 1 therefore lets us characterize the discovery rate at and near equilibrium:

Corollary 1. *Let $(\mathbf{p}^*, \mathbf{q}^*)$ be a sequence of equilibria with non-vanishing investment. Then, the discovery rate vanishes along this sequence: $\lim_n D(\mathbf{p}^*, \mathbf{q}^*) = 0$. For any $\epsilon > 0$, the discovery rate with openness $(\mathbf{p}^*, (1 + \epsilon)\mathbf{q}^*)$ is non-vanishing: $\liminf_n D(\mathbf{p}^*, (1 + \epsilon)\mathbf{q}^*) > 0$.*

The discovery rate is vanishing under equilibrium interaction patterns but would be non-vanishing with slightly more interaction. An immediate consequence is that increasing openness by any multiplicative factor has a very large effect on payoffs asymptotically:

$$\lim_n \frac{D(\mathbf{p}^*, (1 + \epsilon)\mathbf{q}^*)}{D(\mathbf{p}^*, \mathbf{q}^*)} = \infty.$$

16. The same result holds, with minor modifications to the proof, for the additive interaction rate $\tilde{\iota}(q_i, q_j) = q_i + q_j$. We assume $\tilde{\iota}(q, 0) = 0$ to avoid a technicality: for the additive interaction rate, the first-order condition for q_i may not have a solution if other firms choose sufficiently large q_j .

TABLE 1

Best responses and policy implications when firms choose symmetric strategies (\mathbf{p}, \mathbf{q}) in the subcritical, critical, and supercritical regions

	Subcritical	Critical	Supercritical
Best response q_i	High	Intermediate	Low
Discovery rate	Vanishing	Vanishing	Non-vanishing
Increasing q	Large benefit	Large benefit	Ambiguous
Increasing p	Small benefit	Intermediate	Large benefit

Corollary 1 relates to [Saxenian \(1996\)](#)'s study of the Route 128 and Silicon Valley technology industries, which found that Silicon Valley had much more open firms and grew faster. In the terminology of our model, Route 128's secrecy corresponds to equilibrium behaviour. But institutional features of Silicon Valley (including non-enforcement of non-compete clauses and common ownership of firms by venture capital firms) may have constrained firms' actions to prevent high levels of secrecy (subcritical or critical choices of q_i). Such constraints would imply a much higher innovation rate.

A natural question is whether these gains can be realized via policy interventions other than directly restricting firms' strategy spaces. One common policy approach is to subsidize private research and development. We find that increasing R&D spending has relatively little effect on its own but can be effective given a collaborative culture:

Corollary 2. *Let $(\mathbf{p}^*, \mathbf{q}^*)$ be a sequence of equilibria with non-vanishing investment. Then the derivative of the equilibrium discovery rate in private investment vanishes along this sequence: $\lim_n \frac{\partial D(\mathbf{p}^* + x\mathbf{1}, \mathbf{q}^*)}{\partial x}(0) = 0$. For any $\epsilon > 0$, the derivative of the discovery rate with openness $(1+\epsilon)\mathbf{q}^*$ in private investment is non-vanishing: $\liminf_n \frac{\partial D(\mathbf{p}^* + x\mathbf{1}, (1+\epsilon)\mathbf{q}^*)}{\partial x}(0) > 0$.*

Under equilibrium interaction patterns, a higher level of R&D does not increase the discovery rate. But if openness is already above equilibrium levels, then increasing R&D will have a much larger impact on the number of ideas discovered.

To summarize the implications of Corollaries 1 and 2 (see Table 1), at or near equilibrium outcomes, there are large gains to policies (e.g. non-enforcement of non-compete clauses, establishing innovation clusters) that encourage or require more interaction between firms and thus shift outcomes to the supercritical region. Policies to increase private investment (e.g. subsidies for R&D), however, will not shift outcomes to the supercritical region, and thus have much smaller benefits at equilibrium. But once outcomes are in the supercritical region, policies to increase private investment will have large benefits.

3.5. Public innovators

Corollary 1 showed there are large gains to increasing interaction rates above equilibrium levels. This section shows that these gains can be realized via targeted interventions that change interaction patterns.

We now show that introducing *public innovators* who are not concerned with secrecy leads to learning and innovation at the same rate as in the supercritical region. In particular, there exists a giant component of the learning network containing these public innovators. Public innovators could correspond to academics, government researchers, open-source software developers, or other researchers with incentives or motivations other than profiting from producing and selling technologies.

A public innovator i pays investment cost $c(p_i)$ and receives a payoff of one for each technology t such that: (1) $i \in t$ and (2) $j \in \{i\} \cup I_i(\mathbf{p}, \mathbf{q})$ for all $j \in t$. We will rely on the fact that for public innovators there is no downside to interactions, but not on the exact incentive structure.

All firms have the same incentives as in the baseline model, and public innovators and firms interact as in the baseline model. We now call an equilibrium symmetric if all public innovators choose the same action and the same holds for all private firms.

Proposition 1. *Suppose a non-vanishing share of agents are public innovators. Then, there exists a sequence of symmetric equilibria with non-vanishing investment, and at any sequence of equilibria with non-vanishing investment the discovery rate is non-vanishing: $\liminf_n D(\mathbf{p}^*, \mathbf{q}^*) > 0$.*

The proposition says that at equilibrium, a positive fraction of possible technologies are discovered. The ratio between the equilibrium discovery rates with and without public innovators grows unboundedly large as $n \rightarrow \infty$. The proof shows that a giant component forms around the public innovators. This holds for any positive share of public innovators and indeed could be extended to a slowly vanishing share of public innovators.

Proposition 1 assumes that firms cannot direct interactions toward public innovators or private firms. In [Supplementary Appendix E](#), we show the same result holds when interactions can be directed toward public innovators or private firms. Because public innovators are more likely to be in the giant component, private firms are willing to interact with them.

Public innovators are valuable primarily as informational intermediaries rather than for their private ideas. Because public innovators do not face costs to interaction, they will choose $q_i = 1$ at equilibrium. Therefore, public innovators can learn many ideas via interactions and transmit these ideas to other public innovators or to private firms (e.g. academics learning ideas from conferences and collaborations and then consulting for private industry). Conversely, the proposition would remain unchanged if all public innovators instead choose $p_i = 0$ and $q_i = 1$.

Empirical research on collaboration between academia and industry supports the value of academic researchers as informational intermediaries between firms. [Azoulay, Graff Zivin and Sampat \(2012\)](#) study movement of star academics and find that moves increase patent-to-patent and patent-to-article citations locally. Federal funding for universities is also linked to higher entrepreneurship locally ([Tartari and Stern, 2021](#)).

3.6. Welfare

The preceding analysis focused on the discovery rate as an outcome measure. We now extend the results to a more general measure of welfare that allows for consumer and producer surplus. We take a reduced-form approach, assuming fixed consumer surplus from each technology sold by a monopolist and each technology sold in a competitive market.

For each i , we define the *competitive technologies* $CT_i(\mathbf{p}, \mathbf{q}) \subset T_i(\mathbf{p}, \mathbf{q})$ for i to be the set of technologies $t \in T_i(\mathbf{p}, \mathbf{q})$ such that at least one other firm knows all ideas in t . Thus, each technology $t \in CT_i(\mathbf{p}, \mathbf{q})$ is either proprietary or competitive.

We define

$$W(\mathbf{p}, \mathbf{q}) = (w_{PT} + 1) \cdot \left| \bigcup_i PT_i(\mathbf{p}, \mathbf{q}) \right| + w_{CT} \cdot \left| \bigcup_i CT_i(\mathbf{p}, \mathbf{q}) \right| - \sum_{i=1}^n c(p_i),$$

where the weights w_{PT} and w_{CT} are positive. Welfare is the sum of producer surplus (monopoly profits minus R&D costs) and consumer surplus (w_{PT} from each proprietary technology and w_{CT} from each competitive technology).

Proposition 2. Fix a sequence of actions (\mathbf{p}, \mathbf{q}) with $\liminf_n \min_i p_i > 0$ and non-negative expected profits for each firm i . The following are equivalent:

- (1) There exists $\tilde{\alpha} \in (0, 1)$ such that there is a unique component of the indirect-learning network of size at least $\tilde{\alpha}n$ a.a.s.,
- (2) The discovery rate is non-vanishing: $\liminf_n D(\mathbf{p}, \mathbf{q}) > 0$, and
- (3) There exists $C > 0$ such that $\mathbb{E}[W(\mathbf{p}, \mathbf{q})] > Cn^{k-1}$ for n sufficiently large.

When there is a giant component, a positive fraction of ideas are discovered and welfare is within a constant factor of the social optimum. In the subcritical and critical regions, a vanishing share of ideas are discovered and welfare is only a vanishing share of the socially optimal level. The basic idea behind the proof is that consumer surplus grows at the same rate as the number of technologies discovered.

An immediate consequence is that Corollaries 1 and 2 and Proposition 1 imply similar statements about any welfare measure $W(\mathbf{p}, \mathbf{q})$. For example, equilibrium welfare $W(\mathbf{p}^*, \mathbf{q}^*)$ grows at rate $o(n^{k-1})$ in the baseline model but grows at a rate proportional to n^{k-1} with a positive share of public innovators. Therefore, the ratio between equilibrium welfare with and without public innovators grows unboundedly large as $n \rightarrow \infty$.

4. ASYMMETRIC LEARNING PROBABILITIES

The baseline model assumes that information flows are symmetric across pairs of firms. In practice, firms may have heterogeneous probabilities of learning from others, even given a fixed interaction rate. We next show that equilibrium remains critical with heterogeneous propensities to learn.

Suppose that firms have propensities to learn $\beta_i \in (0, 1)$, where there are finitely many propensities to learn and the fraction of firms with each propensity β_i converges as the number of firms grows large. Firm i now directly learns from firm j with probability

$$\beta_i \iota(q_i, q_j).$$

Learning otherwise occurs as in the baseline model, including indirect learning.

It is straightforward to extend Definition 4 to allow heterogeneous secrecy. We now let λ be the spectral radius of the matrix $(\beta_i \iota(q_i, q_j) \delta)_{ij}$. As before entry (i, j) is equal to the probability that firm i learns indirectly from firm j . Given the modified definition of λ , Definition 4 continues to define a critical threshold and this threshold again corresponds to the emergence of a giant component.

Theorem 2. Suppose firms have propensities to learn β . There exists an investment equilibrium for n large, and any sequence of investment equilibria is critical.

Equilibria remain critical even when the directed link probabilities are asymmetric across pairs. The characterization result extends immediately to the case in which β_i are chosen endogenously at a cost $\tilde{c}_i(\beta_i)$, which can vary across firms.¹⁷ In this case, firms can now control the likelihood of learning along two dimensions. First, higher interaction rates allow a firm to learn more from others at the expense of a higher probability of its ideas leaking. Second, firms can pay an exogenous cost to increase the probability of learning from others at a given interaction rate, and some firms may be able to do so more cheaply than others.

17. This choice can be made simultaneously with or prior to the choice of q_i .

The proof of Theorem 2 shows that decisions with asymmetric learning probabilities are similar to decisions in the baseline model. Recall that at equilibrium, the first-order condition for the openness q_i relates the value of the ideas already known to firm i with the value of increasing the interaction rate. Fixing q_i , a higher β_i increases both sides of this first-order condition because firms with higher propensities to learn have already learned more ideas but also will learn more from an additional interaction.

At potential equilibria in the subcritical region, these two forces cancel out and a firm's optimal choice of openness q_i^* is approximately independent of that firm's propensity to learn β_i . The proof of Theorem 1 used the symmetry between the value of existing links and an additional link to solve for q^*n . Even if each of these links only realizes with probability β_i , the symmetry persists and so q^*n is unchanged.

The two opposing effects would not entirely cancel in the supercritical region, because there are potential redundancies between multiple links to the giant component. These redundancies matter more for firms with higher β_i . Nevertheless, we can bound the average interaction rate when all firms choose q_i to respond optimally to the giant component size.

5. BENEFITS AND COSTS OF LINKS

In Sections 2 and 3, we studied equilibrium when expected payoffs were

$$U_i(\mathbf{p}, \mathbf{q}) = \mathbb{E}[|PT_i(\mathbf{p}, \mathbf{q})|] - c(p_i).$$

Firms' utility functions had two properties:

1. Payoffs are linear in the number of proprietary technologies, and
2. Payoffs do not depend on technologies for which the firm faces competition.

The first assumption determines the benefits from incoming links, while the second determines the costs of outgoing links.

We now relax each of these assumptions. We find that equilibrium remains critical when the returns to producing more technologies are increasing. More generally, we show that equilibrium is critical in a setting where profits are a convex function of the number of ideas learned. Changing the profit structure in competitive markets, however, leads to supercritical or subcritical equilibria. So outcomes depend on the specification of the costs of outgoing links, but are less sensitive to the specification of the gains from learning.

5.1. Concavity of profits

The baseline model assumed that a firm's profits are linear in the number of proprietary technologies. We now extend the main result to allow more general firm profits.

In practice, there may be increasing or decreasing returns to producing more technologies. Suppose that the payoffs to firm i are instead

$$|PT_i(\mathbf{p}, \mathbf{q})|^\rho - c(p_i),$$

where $\rho > 0$.

The baseline model is the case $\rho = 1$. When $\rho > 1$, there are increasing returns to controlling more monopolies. When $\rho < 1$, there are decreasing returns to controlling more monopolies. Note that these increasing or decreasing returns to scale are not determined by the innovative process, but rather by production costs or other market conditions.

Proposition 3. *There exists $\underline{\rho} \leq 1$ such that for any $\rho \geq \underline{\rho}$, any sequence of symmetric investment equilibria is critical. When $k > 2$, we have $\underline{\rho} < 1$.*

The proposition shows that the prediction of critical equilibria is not knife-edge with respect to ρ . In particular, increasing returns to scale cannot move interactions above the critical threshold. As long as $k \neq 2$, slightly decreasing returns to scale will not move interactions below the critical threshold either.

Consider a firm i that does not face competition. We show that under the conditions of the proposition, the firm's profits are convex in $|I_i(\mathbf{p}, \mathbf{q})|$. As a result, learning additional ideas is more appealing relative to protecting existing ideas, so openness will not decrease below the critical region. Checking convexity is delicate when $\rho < 1$, because in this case firm profits are the composition of the binomial coefficient $\binom{|I_i(\mathbf{p}, \mathbf{q})|}{k-1}$, which is convex, and the polynomial, $|PT_i(\mathbf{p}, \mathbf{q})|^{\rho}$, which is concave.

We also show that for any ρ , openness will not increase enough to push equilibrium into the supercritical region either. At a potential supercritical sequence of symmetric investment equilibria, profits are driven by the event that firm i learns αn ideas from the giant component and produces approximately $(p^*)^k \binom{\alpha n}{k-1}$ proprietary technologies. Firm i chooses q_i to maximize the probability of this event. Asymptotically the optimal q_i is independent of the payoffs from this event since these payoffs are very large, and therefore the optimal q_i is independent of ρ . Given this, the calculation is the same as in the case $\rho = 1$ (Theorem 1), where there is no supercritical sequence of investment equilibria.

More generally, the proof shows that our criticality result relies on two features of the payoff function. First, payoffs for a firm i that does not face competition are convex in the number of ideas $|I_i(\mathbf{p}, \mathbf{q})|$ learned by i . Second, payoffs grow at a polynomial rate in the number of ideas learned by i .

We can state this formally when $\delta = 1$, so that when firm i learns from j it will learn all ideas known to firm i . In this case, we let the profits for firm i be $\phi(|I_j(\mathbf{p}, \mathbf{q})|)$ when firm i discovers its private idea ($i \in I$) and no firm learns from i , and 0 otherwise. We will assume that $\phi(\cdot)$ is strictly increasing and continuously differentiable.

Proposition 4. *Suppose $\delta = 1$ and payoffs when no firm learns from i and $i \in I$ are equal to $\phi(|I_i(\mathbf{p}, \mathbf{q})|)$, where $\phi(x)$ is convex and $\frac{\phi(x_j)}{\phi(x'_j)} \rightarrow 1$ along any sequence of (x_j, x'_j) such that $\frac{x_j}{x'_j} \rightarrow 1$. Then any sequence of symmetric equilibria with non-vanishing investment is critical.*

The assumption that $\frac{\phi(x_j)}{\phi(x'_j)} \rightarrow 1$ along any sequence of (x_j, x'_j) such that $\frac{x_j}{x'_j} \rightarrow 1$ bounds the rate of growth of $\phi(\cdot)$. In particular, this assumption holds if $\phi(x) = Cx^d + O(x^{d-1})$ for any $C > 0$ and any real $d \geq 1$. If payoffs instead grow at an exponential rate in the number of ideas, then a supercritical equilibrium is possible because an additional idea may be very valuable (see [Acemoglu and Azar, 2020](#) for a related effect).

A special case is that firms can produce technologies of multiple complexities k , perhaps with different payoffs for technologies of different complexities. The growth condition will hold as long as the allowed complexities are bounded (independent of n). A consequence of the proposition in this case is that equilibrium profits are driven by technologies of the highest feasible complexity for n large. At a critical equilibrium, most profits come from rare events where a firm learns many ideas, and when this occurs the firm can produce many more technologies of higher complexities. This suggests that if firms can choose the complexity k of their products, firms will produce more complex technologies (at least in large markets).

5.2. Profits under competition

We found in Theorem 1 that equilibrium lies on the critical threshold. This result is robust to different payoffs structures for monopolist firms. We now show that Theorem 1 does depend on the structure of competition and altering the payoffs from competitive markets can lead to supercritical or subcritical outcomes.

To generalize the payoff from technologies, we will now assume that firm i receives payoffs $f(m)$ from a technology t such that $i \in t$, firm i learns all other ideas in t , and m other firms learn all ideas in t . We assume $f(\cdot)$ is weakly decreasing and maintain the normalization $f(0)=1$.

A simple case is $f(m)=a < 1$ for all $m > 0$. The parameter a represents the profits in competitive markets for firms i such that the idea i is included in the technology. These profits could correspond to a first-mover advantage or a higher quality product due to the firm's expertise. For example, firms whose idea is included in a technology could bring the product to market first and obtain positive profits before competitors arrive (see [Boldrin and Levine, 2008](#) for details). Alternately, these firms could produce higher quality technologies and make positive profits due to this product differentiation. The analysis in previous sections corresponded to the case $a=0$.

We can also allow $f(m) < 0$, which could correspond to a fixed cost of production that must be paid before competition is known. We assume that firms make a single decision about whether to produce the technologies that they learn.¹⁸

Proposition 5. (i) If $0 < f(1) < 1$ and $f(m) \geq 0$ for all m , then any sequence of symmetric investment equilibria is supercritical.

(ii) If $f(m) < 0$ for all $m > 0$, then any sequence of symmetric investment equilibria is subcritical.

We can restate the proposition in terms of the discovery rate, and find that there is more innovation when there are positive profits under competition than in the Bertrand case:

Corollary 3. (i) If $0 < f(1) < 1$ and $f(m) \geq 0$ for all m , then the discovery rate is non-vanishing along any sequence of symmetric equilibria with non-vanishing investment: $\liminf_n D(\mathbf{p}^*, \mathbf{q}^*) > 0$.

(ii) If $f(m) < 0$ for all $m > 0$, then the discovery rate vanishes along any sequence of symmetric equilibria with non-vanishing investment: $\lim_n D(\mathbf{p}^*, \mathbf{q}^*) = 0$.

Part (i) of the proposition says that if the potential downside to enabling competitors is not as large, then firms will be more willing to interact. This pushes the equilibrium from the critical threshold into the supercritical region. Proposition 5(i) introduces an additional force to classic debates on whether firms are more innovative in more competitive markets (see [Cohen and Levin, 1989](#) for a survey). While much of this literature considers how competition changes firms' private incentives to conduct R&D, the proposition considers its effect on interaction and learning between firms.

Part (ii) of the proposition says that increasing the costs of competition discourages interaction, and pushes the equilibrium to the subcritical region. There need not be an investment equilibrium if payoffs under competition are sufficiently negative.

The proof of (ii) is more involved, as we must characterize payoffs at a potential critical sequence of equilibria. The key lemma shows that at the critical threshold, we have $\mathbb{E}_{t \in PT_i(\mathbf{p}, \mathbf{q})}[\tau(t)] \rightarrow 1$, i.e., most of a firm i 's proprietary technologies only include ideas learned

18. If firms can condition their production decision on the flow of ideas, the analysis becomes more complicated.

from one other firm. To prove this lemma, we use a pair of coupling arguments to show that most profits come from rare events in which a single link (indirectly) lets a firm learn many ideas.

Proposition 5 describes how the equilibrium learning network depends on competition qualitatively, but we can also relate equilibrium outcomes to the innovation rate within the supercritical region. We now show that increasing the payoffs from competitive outcomes will increase the innovation rate even within the supercritical region.

Consider a function $g(m)$ such that $g(m) \geq 0$ for all m , $g(m)$ is weakly decreasing in m , and $g(m) \rightarrow 0$ as $m \rightarrow \infty$. Set $f_x(0) = 1$ and $f_x(m) = xg(m)$ for $m \geq 1$, so that monopoly payoffs are constant and competitive profits are weakly increasing in x . For example, higher values of x can correspond to a larger first-mover advantage.

Proposition 6. *There exists $x^* > 0$ such that for $x \in [0, x^*]$, with payoffs $f_x(m)$:*

- (i) *There exists an investment equilibrium for n large;*
- (ii) *The limit of the discovery rate $\lim_n D(\mathbf{p}^*, \mathbf{q}^*)$ along any sequence of symmetric equilibria with non-vanishing investment $(\mathbf{p}^*, \mathbf{q}^*)$ exists and does not depend on the sequence of equilibria;*
- (iii) *The limit $\lim_n D(\mathbf{p}^*, \mathbf{q}^*)$ is weakly increasing in x .*

When payoffs are higher in competitive markets, the upside to learning from others is potentially higher and the downside of information leakage is lower. So individuals are willing to interact more, and thus the discovery rate is higher, as x increases. The proposition shows this comparative static locally, because for x small we can leverage our analysis of equilibrium under Bertrand payoffs to show that an investment equilibrium exists and is well-behaved.

6. PATENT RIGHTS

In the baseline model, technologies could only be protected via secrecy. We now consider the possibility that a positive fraction of firms receive patents on their ideas. As motivation, suppose that firms discover different types of ideas and patent law determines which types are patentable. For example, [Bessen and Hunt \(2007\)](#) discuss the boundaries of patent law in the software industry and how those boundaries have changed over time.

6.1. Equilibrium

Suppose a fraction $b \in (0, 1)$, of firms receive a *patent* on their private ideas. In this case, other firms cannot use this private idea, either as monopolists or competitors. Formally, a firm i receives payoff 1 from each technology t such that (1) idea $i \in t$; (2) firm i knows all $j \in t$ and no other $j \in t$ receive patents; and (3) either i receives a patent or i is the unique firm that knows all $j \in t$. Else the firm receives payoff 0 from the technology t .

To focus on how patents relate to informal interactions, we analyse model patents in a very simple way. In particular, the model will not include imperfect patent rights and/or licensing of patents, and licensing can increase the innovation rate.

A firm now chooses either a level of openness $q_i(0)$, which is the action without a patent, or $q_i(1)$, which is the action with a patent. We will refer to the choices at symmetric equilibria as $q^*(0)$ and $q^*(1)$.

For the first part of the following result ($\delta = 0$), we will also assume that each firm i pays cost $\epsilon > 0$ for each realized link. The purpose of this cost is to break near-indifferences in favour of lower interaction rates. We observe in [Supplementary Appendix C](#) that without patents, a small link cost ϵ has little effect on the equilibrium.

Proposition 7. Suppose a fraction $b \in (0, 1)$ of firms receive patents. If $\delta = 0$, then with $k = 2$ and any link cost $\epsilon > 0$, there does not exist an investment equilibrium for n large. If $\delta > 0$, then $\iota(q^*(0), q^*(0))$ is $o(1/n)$ along any sequence of symmetric investment equilibria.

With only direct learning ($\delta = 0$), the proposition says that positive investment cannot be sustained at equilibrium for n large. This is because of an adverse-selection effect that discourages social interactions.

Because firms receiving patents have no need for secrecy, firms with patents choose very high interaction rates $q_i(1)$. Thus, most interactions are with firms with patents. On the other hand, firms with patents are undesirable to interact with because their ideas cannot be used by others. Because of this adverse selection in the matching process, firms without patents will have much lower expected profits than in the model without patents. When $k = 2$ and there is an arbitrarily small cost to links, this has the effect of shutting down all interaction and investment.

This contrasts with our results on direct learning with no patent rights (Section 3.2 and [Supplementary Appendix C](#)), where there is an investment equilibrium with substantial interaction. With $k > 2$ and patent rights, the adverse selection effect persists but no longer prevents any equilibrium investment. In this case, the interaction rate between firms without patents is much lower asymptotically than in the result with no patent rights ([Supplementary Appendix C.1](#)).

The direct learning result also suggests a more general adverse-selection effect in strategic network formation. Suppose that agents with lower link formation costs are also less valuable partners for connections. If agents cannot discriminate in their link formation decisions, the composition of the pool of potential partners will discourage connections.

With indirect learning, firms with patents still do not provide private ideas to others, but can now serve as informational intermediaries (like the public innovators in Section 3.5). A giant component forms around the firms with patents. Firms without patents now have some interactions with firms without patents, who can transmit ideas from other firms without patents. Interactions between pairs of firms without patents, however, are rare.

Proposition 7 relates to literature on firms' choices between formal and informal intellectual property protections, particularly patents vs. secrecy (e.g. [Anton and Yao, 2004](#); [Kultti, Takalo and Toikka, 2006](#)). We focus not on the choice between formal and intellectual property rights but on the interplay between the two.¹⁹

The proposition assumes that firms cannot direct interaction toward or away from firms with patents, but we could also allow firms to choose separate interaction rates for peers with and without patents. Under direct learning ($\delta = 0$), firms without patents would only interact with each other and the adverse selection effect would disappear.²⁰ Under indirect learning, firms with patents can still play a similar role as informational intermediaries when directed interaction is permitted. The analysis is very similar to [Supplementary Appendix E](#), so we omit details here.

6.2. Discovery rate

We can use Theorem 1 and Proposition 7 to ask when patent rights produce more innovation and what the optimal share b^* of patentable ideas would be. In the direct-learning case, the proposition gives conditions under which patents decrease innovation and welfare.

19. A second contrast is to theoretical findings on patents and follow-up innovation (e.g. [Scotchmer and Green, 1990](#); [Scotchmer, 1991](#); [Bessen and Maskin, 2009](#)). This literature investigates when granting patent rights for an idea decreases follow-up innovations involving that idea. In our random-interactions setting, patent rights can not only decrease follow-up innovations involving patented ideas but also decrease follow-up innovations involving *other* unprotected ideas.

20. The analysis of equilibrium interaction patterns among these firms would be the same as in [Supplementary Appendix C](#).

In the indirect-learning case, the equilibrium discovery rate and social welfare are higher (for n large) with interior patent rights $b \in (0, 1)$ than without patent rights because firms with patents are valuable as intermediaries. Under indirect learning, we can ask what value of b maximizes the discovery rate. Any positive b provides the benefits of information intermediaries, and so there is a tradeoff between the higher private profits obtained by firms with patents and the social benefits provided by firms without patents, whose ideas can be used by others. For high k , the optimal value of b converges to zero as n grows large.

This is easiest to see when $\delta = 1$. In this case, we can compute that the share of firms without patents whose ideas are learned by the giant component is $e^{-bq^*(0)n} \approx \frac{1}{2}$. The number of technologies produced is

$$(p^*)^k \binom{\frac{1}{2}(1-b)n}{k-1} \left(b + \frac{1}{2} \cdot (1-b)n \right) + o(n^{k-1}), \quad (1)$$

because approximately $b + \frac{1}{2} \cdot (1-b)$ firms learn from the giant component and each learns approximately $\frac{1}{2}(1-b)n$ unpatented ideas. The leading term in expression (1) is maximized by $b=0$, so the patent share b^* maximizing the discovery rate converges to zero.

7. CONCLUSION

We have studied strategic network formation in large random graphs in the context of an economic application to innovation and social learning. The model is particularly suited to analysis of informal interactions, e.g. between employees of firms, which cannot be fully governed by formal contracts. We find that in these settings, if there are many firms and ideas can travel multiple steps, the global structure of the learning network has stark consequences for incentives and payoffs. In particular, expected payoffs and welfare are much higher when there is enough interaction to support a giant component.

The techniques developed in this article could be extended or applied in several directions. For example, we have studied a static model, but in a dynamic version of the model past events would create new incentives toward openness or secrecy. More broadly, while we have focused on a network-formation game with a tradeoff between secrecy and learning, we have developed more widely applicable tools for questions in network economics, particularly those with complementarities between connections. Outside of network formation, the same techniques can also be applied to problems such as optimal seeding for diffusion processes or search and matching models where indirect connections are important.

Supplementary Data

[Supplementary data](#) are available at *Review of Economic Studies* online.

Acknowledgments. I am deeply grateful to Drew Fudenberg, Benjamin Golub, Matthew Rabin, and Tomasz Strzalecki for their guidance and support. I would also like to thank Andrea Galeotti and three anonymous referees for their very helpful suggestions and Daron Acemoglu, Mohammad Akbarpour, Leonie Baumann, Yochai Benkler, Laura Doval, Edward Glaeser, Kevin He, Nir Hak, Scott Kominers, Shengwu Li, Jonathan Libgober, George Mailath, Michael Rose, Stefanie Stantcheva, Eduard Talamàs, Omer Tamuz, Rakesh Vohra, and Alexander Wolitzky for valuable comments and discussion.

A. PROOF OF THEOREM 1

We will first prove Theorem 1. We begin with two lemmas, which bound equilibrium actions and the number of ideas learned at equilibrium. We then prove the theorem for symmetric equilibria. The symmetric case illustrates the main

ideas of the proof. We complete the proof by extending the analysis of the symmetric case to show that any sequence of investment equilibria, which need not be symmetric, is critical.

In all appendices, we say that $f(n) \sim g(n)$ if $f(n)/g(n) \rightarrow 1$ as $n \rightarrow \infty$.

A.1. Preliminary lemmas

Our first lemma, which bounds agents' actions uniformly, will be useful for the subcritical and supercritical cases. Given \mathbf{q} , we let $\bar{q} = \max_i q_i$ and $\underline{q} = \min_i q_i$ be the maximum and minimum choices of openness.

Lemma 3. *Consider any sequence of investment equilibria $(\mathbf{p}^*, \mathbf{q}^*)$. There exists a constant \bar{C} such that $\iota(\bar{q}, \underline{q})n \leq \bar{C}$ for all n .*

Proof of Lemma 3 Suppose not. Relabelling firms, we can assume that $q_1^* = \bar{q}$ for each n . Passing to a subsequence if necessary, we can take $q_1^* \sqrt{n} \rightarrow \infty$ as $n \rightarrow \infty$.

We first assume for the sake of contradiction that the expected number of times firm 1 learns directly is unbounded. Passing to a subsequence we can assume that the expected number of times firm 1 learns directly converges to infinity, i.e., $\sum_{j \neq 1} \iota(q_1^*, q_j^*) \rightarrow \infty$.

We now fix n large and consider the payoffs to firm 1 after deviating to choose q_1 . By the Chernoff bound, the probability that no firm learns indirectly from firm 1 decays exponentially in q_1 .

We claim that $\binom{|I_1(\mathbf{p}^*, (q_1, q_{-1}^*))|}{k-1}$ grows at most at a polynomial rate in q_1 . Let X be a random variable equal to the number of ideas learned by learning from a random firm j , each chosen with probability proportional to q_j . For \mathbf{q} such that $\sum_{j \neq 1} \iota(q_1, q_j^*)$ is an integer m , the random variable $\binom{|I_1(\mathbf{p}^*, (q_1, q_{-1}^*))|}{k-1}$ is first-order stochastically dominated by $\binom{\sum_{j=1}^m X_j}{k-1}$, where X_1, \dots, X_m are i.i.d. random variables distributed as X . This in turn within a constant multiple of $\left(\sum_{j=1}^m X_j\right)^{k-1}$. By Rosenthal's inequality (Rosenthal, 1970), the growth rate of the expectation of this sum of moments in m is at most polynomial in m .

Therefore, the payoffs to firm 1 conditional on no firm learning indirectly from 1 are at most polynomial in q_1 . The probability of this event decays exponentially, so for n sufficiently large, firm 1 could profitably deviate to a smaller choice of q_1 . This gives a contradiction.

The remaining case is that $q_1^* \sqrt{n} \rightarrow \infty$ but $\sum_{j \neq 1} \iota(q_1, q_j^*)$ is bounded. Then, there must exist a sequence of i such that $q_i^*/q_1^* \rightarrow 0$, and for some such sequence of i we have $\sum_{j \neq i} \iota(q_i^*, q_j^*) \rightarrow 0$. This gives a contradiction since then for n large, firm i could profitably deviate to choose $q_i = (\sum_{j \neq i} q_j^*)^{-1}$. This complete the proof of the lemma. \parallel

Lemma 4. *Consider a subcritical sequence of actions such that $\iota(q_i, q_i)n$ is bounded above uniformly. Then*

$$\lim_{n \rightarrow \infty} \mathbb{P}[|I_i(\mathbf{p}, \mathbf{q})| = y]$$

decreases at an exponential rate in y .

Proof of Lemma 4 Because the sequence of actions is subcritical, we can assume that the matrix $(\iota(q_i, q_j)\delta)_{ij}$ has spectral radius at most $\bar{\lambda} < 1$ for all n sufficiently large. Increasing q_i for some i and therefore also $\bar{\lambda}$, we can assume without loss of generality that there are at most K choices of q_i for each n . Here the number of distinct actions K can depend on the initial upper bound $\bar{\lambda}$ and distribution of actions. We denote the number of firms choosing q_i by $n(q_i)$.

We will bound the number of firms j with a path from j to i in the indirect learning network above by the number of nodes in a multi-type branching process with the number of successors distributed as Poisson random variables. The types will correspond to the (at most K) choices of q_i .

For each firm i , the number of firms choosing q_j that firm i learns from indirectly is a binomial random variable with success probability $\delta_i(q_i, q_j)$ and at most $n(q_j)$ trials. We claim that for any $C > 1$, for n sufficiently large such a random variable is first-order stochastically dominated by a Poisson random variable with parameter $C\delta_i(q_i, q_j)n(q_j)$. By Theorem 1(f) of Klenke and Mattner (2010), this holds if

$$(1 - \delta_i(q_i, q_j))^n \geq e^{-C\delta_i(q_i, q_j)\delta n}.$$

The inequality is satisfied for n sufficiently large since $C > 1$.

Choose $1 < C < 1/\bar{\lambda}$ and assume n is large enough for the previous claim to hold. Then the number of firms that firms j with a path from j to i in the indirect learning network is first-order stochastically dominated by the number of nodes in the multi-type branching process such that the number of successors of a node of type q_i of each type q_j is distributed as a Poisson random variable with mean $C\delta_i(q_i, q_j)n(q_j)$. Call this number of nodes y' .

We want to show that $y' < \infty$ with probability one and the probability that $y' = y$ decays exponentially in y . The (at most $K \times K$) matrix $C\delta\iota(q_i, q_j)n(q_j)$ has spectral radius at most $C\bar{\lambda} < 1$ because $(C\iota(q_i, q_j)\delta)_{ij}$ does. Therefore, by Theorem 2 of Section V.3 of Athreya and Ney (1972), the probability that $y' = \infty$ is zero.

Let Z_j be the number of nodes in the j th generation of the branching process. By Theorem 1 of Section V.3 of Athreya and Ney (1972), the probability that $Z_j > 0$ is of order at most $(C\bar{\lambda})^j$. So the probability that $Z_T > 0$ decays exponentially in T .

Dropping nodes with zero probability of interaction if necessary, we can assume that all $q_i > 0$. By the Perron–Frobenius theorem, there exists an eigenvector of $(\delta\iota(q_i, q_j)n(q_j))_{ij}$ with positive real entries and eigenvalue equal to the spectral radius of this matrix. We call this eigenvector \mathbf{v} .

We claim that the probability that there are more than Tv_i nodes of some type q_i in one of the generations $1, \dots, T$ decays exponentially in T . There is one node in generation zero. Suppose that there are at most Tv_i nodes of each type q_i in generation j . Then by our construction of \mathbf{v} , the number of nodes of each type q_i in generation $j+1$ is Poisson with mean at most $Tv_i C\bar{\lambda}$. A Poisson random variable of mean $Tv_i C\bar{\lambda}$ is the sum of T Poisson random variables of mean $v_i C\bar{\lambda}$. So by the central limit theorem, the probability that there are at least Tv_i such nodes decays exponentially in T , independent of j . This implies the claim.

Therefore, the probability that there are more than Tv_i nodes decays exponentially in T . We have completed the proof that $y' < \infty$ with probability one and the probability that $y' = y$ decays exponentially in y . Finally, $|I_i(\mathbf{p}, \mathbf{q})|$ is first-order stochastically dominated by the sum of y' Bernoulli random variables with n trials and success probability $\max_i \iota(q_i, q_i)$. The statement of the lemma now follows by the central limit theorem. \parallel

A.2. Symmetric investment equilibria

We begin by describing the structure of this section.

1. We first show that there does not exist a subcritical sequence of symmetric investment equilibria. The proof assumes such a sequence exists for the sake of contradiction and characterizes equilibrium learning and behaviour via two lemmas. The first lemma gives a first-order condition for the choice of q_i . This first-order condition is used to prove a second lemma, which shows that $\delta\iota(q^*, q^*)n$ is approximately equal to an expectation $\mathbb{E}[\tau(t)]$, where $\tau(t)$ is the number of links needed to learn the ideas in the technology t . Because $\tau(t) \geq 1$ for all t , this implies that $\liminf_n \delta\iota(q^*, q^*)n$ is at least one. But, this contradicts the assumption that the sequence of equilibria is subcritical.
2. We then show that there does not exist a supercritical sequence of symmetric investment equilibria. We show that given a supercritical sequence of symmetric actions, the payoffs for firm i can be computed (to first order) based on whether firm i learns the ideas learned from the giant component and which firms learn from i . This is because the number of ideas that firm i learns from outside the giant component is very likely to be small, so such ideas have a small effect on payoffs. We compute the highest-order term in the firm i 's payoffs explicitly and show that each firm i would prefer to deviate to a lower choice of q_i for n large.
3. We finally show that there exists a symmetric investment equilibrium for n large. The argument shows that there is a unique best response near the critical region. Given the preceding analysis, the intermediate value theorem then implies the best response function must have a fixed point.

After passing to a subsequence if necessary, we can assume that any sequence of equilibria is either subcritical, critical, or supercritical.

Subcritical case: We begin with a lemma expressing the first-order condition for q_i at a subcritical sequence of investment equilibria.²¹

Lemma 5. *Along any sequence of subcritical investment equilibria,*

$$\delta \sum_{j \neq i} \frac{\partial \iota(q_i, q_j^*)}{\partial q_i}(q_i^*) \cdot \mathbb{E} \left[\binom{|I_i(\mathbf{p}^*, \mathbf{q}^*)|}{k-1} \right] \sim \mathbb{E} \left[\frac{\partial \binom{|I_i(\mathbf{p}^*, (q_i, q_{-i}^*))|}{k-1}}{\partial q_i}(q_i^*) \right]$$

for each i .

Proof of Lemma 5 We first claim that because the sequence of equilibria is subcritical, competition that is not based on learning all of firm i 's ideas indirectly is lower order. More formally, let $T_i(\mathbf{p}, \mathbf{q})$ be the set of technologies t such that $i \in t$ and firm i learns all other ideas $j \in t$.²² The claim is that if there does not exist a link from firm i to another firm j in

21. Note that we do not require a symmetry assumption, so we can also apply this lemma in the asymmetric case.
22. Recall that $PT_i(q_i, q_{-i}) \subset T_i(\mathbf{p}, \mathbf{q})$ is the subset of technologies t such that no other firm learns all ideas in t .

the indirect-learning network, the conditional probability $\mathbb{E}_{t \in T_i(\mathbf{p}^*, \mathbf{q}^*)}[\mathbf{1}_{t \in PT_i(\mathbf{p}^*, \mathbf{q}^*)}]$ that $t \in PT_i(\mathbf{p}^*, \mathbf{q}^*)$ for a technology $t \in T_i(\mathbf{p}^*, \mathbf{q}^*)$ chosen at random converges to one. Here, each technology t with $i \in t$ is chosen with probability proportional to the probability that firm i knows all ideas in t , i.e., $t \in T_i(\mathbf{p}^*, \mathbf{q}^*)$.

Suppose $t \in T_i(\mathbf{p}^*, \mathbf{q}^*)$ and no firm learns indirectly from i . Choose some $j \in t$ distinct from i . By Lemma 4, the probability that a given firm j' learns $y \geq y$ ideas decays exponentially in y at a rate independent of n . By independence, the probability that firm j' learns ideas i and j is at most $o(\frac{1}{n})$. Therefore, the probability that any firm j' learns ideas i and j is at most $o(1)$. The technology $t \in PT_i(\mathbf{p}, \mathbf{q})$ if there is no such j' for any $j \in t$ distinct from i , so this proves the claim.

Thus, we can express the expected utility of player i choosing $p_i \in [0, 1)$ and $q_i \geq 0$ as

$$p_i \mathbb{E} \left[\binom{|I_i(\mathbf{p}^*, (q_i, q_{-i}^*))|}{k-1} \right] \prod_{j \neq i} (1 - \delta \cdot \iota(q_i, q_j^*)) - c(p_i) + o(1).$$

We first note that the optimal q_i does not depend on p_i , but instead is chosen to maximize

$$\mathbb{E} \left[\binom{|I_i(\mathbf{p}^*, (q_i, q_{-i}^*))|}{k-1} \right] \prod_{j \neq i} (1 - \delta \cdot \iota(q_i, q_j^*)) + o(1).^{23}$$

The first-order condition gives

$$\delta \mathbb{E} \left[\binom{|I_i(\mathbf{p}^*, \mathbf{q}^*)|}{k-1} \right] \sum_{j \neq i} \frac{\partial \iota(q_i, q_j^*)}{\partial q_i} (q_i^*) (1 - \delta \cdot \iota(q_i, q_j^*))^{-1} \sim \delta \mathbb{E} \left[\frac{\binom{|I_i(\mathbf{p}, (q_i, q_{-i}^*))|}{k-1}}{\partial q_i} (q_i^*) \right].$$

Finally, we have

$$\sum_{j \neq i} \frac{\partial \iota(q_i, q_j^*)}{\partial q_i} (q_i^*) (1 - \delta \cdot \iota(q_i, q_j^*))^{-1} \rightarrow \sum_{j \neq i} \frac{\partial \iota(q_i, q_j^*)}{\partial q_i} (q_i^*)$$

because $\limsup \delta \iota(\bar{q}, \bar{q}) n$ is bounded by Lemma 3. \parallel

Given $t \in PT_i(\mathbf{p}^*, \mathbf{q}^*)$, let $\tau(t)$ be the smallest number of (direct or indirect) links such that firm i would still know all ideas $j \in t$ with only $\tau(t)$ of its links.

We next prove Lemma 2. In the statement, each technology $t \in PT_i(\mathbf{p}^*, \mathbf{q}^*)$ under a given realization of all random variables is chosen with probability proportional to the probability of that realization.

Lemma 2. *Along any sequence of symmetric investment equilibria with $\limsup \delta \iota(q^*, q^*) n < 1$,*

$$\delta \iota(q^*, q^*) n \sim \mathbb{E}_{t \in PT_i(\mathbf{p}^*, \mathbf{q}^*)} [\tau(t)]$$

for all i .

Proof of Lemma 2 We will apply Lemma 5, which gives

$$\delta \cdot \frac{\partial \iota(q_i, q^*)}{\partial q_i} (q^*) \cdot \mathbb{E} \left[\binom{|I_i(\mathbf{p}^*, \mathbf{q}^*)|}{k-1} \right] \sim \frac{1}{n} \frac{\partial \mathbb{E} \left[\binom{|I_i(\mathbf{p}^*, (q_i, q_{-i}^*))|}{k-1} \right]}{\partial q_i} (q^*)$$

at a symmetric equilibrium.

Let Γ be the set of weakly increasing tuples $\gamma = (\gamma_1, \dots, \gamma_{l(\gamma)})$ of integers such that $\sum_{j=1}^{l(\gamma)} \gamma_j = k-1$. We write $l(\gamma)$ for the length of the tuple γ .

Let X_j be i.i.d. random variables with distribution given by the number of ideas that firm i would learn from a firm j' conditional on learning directly from j' . That is, X_j is distributed as the sum of a Bernoulli random variable with success probability p^* (corresponding to direct learning) and a random variable distributed as $|I_j(\mathbf{p}^*, \mathbf{q}^*)|$ with probability δ and zero otherwise.

Then we claim that

$$\mathbb{E} \left[\binom{|I_i(\mathbf{p}^*, \mathbf{q}^*)|}{k-1} \right] = \sum_{\gamma \in \Gamma} \binom{n-1}{l(\gamma)} \iota(q^*, q^*)^{l(\gamma)} \mathbb{E} \left[\prod_{j=1}^{l(\gamma)} \binom{X_j}{\gamma_j} \right] + o(1). \quad (\text{A.1})$$

The right-hand side counts the expected number of choices of $k-1$ ideas learned (directly or indirectly) via different neighbours j , allowing for the same idea to be chosen multiple times via distinct neighbours. To show the claim, we must argue that the contribution from choices of $k-1$ ideas including such repetitions is $o(1)$.

23. Note that $I_i(\mathbf{p}, \mathbf{q})$ does not depend on p_i .

By Lemma 3, the probability that there are links from j to i in the indirect learning network for l firms j decays exponentially in l . On the other hand, the probability of learning the ideas in a component via multiple direct connections converges to zero because each component is $o(n)$ a.a.s. By Lemma 4, the contribution from the this vanishing probability event vanishes asymptotically. This gives the claim in equation (A.1).

We can express the right-hand side of Lemma 5 similarly. Recall the right-hand side counts the number of additional sets of $k-1$ distinct ideas that would be known to i if i added a direct link to an additional random agent. We then have

$$\frac{1}{n} \mathbb{E} \left[\frac{\partial \binom{|I_i(\mathbf{p}^*, (q_i, q_{-i}^*))|}{k-1}}{\partial q_i} (q^*) \right] = \sum_{\gamma \in \Gamma} \binom{n-2}{l(\gamma)-1} \iota(q^*, q^*)^{l(\gamma)-1} \mathbb{E} \left[\prod_{j=1}^{l(\gamma)} \binom{X_j}{\gamma_j} \right] + o(1). \quad (\text{A.2})$$

The same argument shows that the contribution from choices of $k-1$ ideas at least one of which is learned via multiple links is $o(1)$.

Substituting equations (A.1) and (A.2) into Lemma 5, we find that

$$\iota(q^*, q^*) \delta \sim \mathbb{E} \left[\frac{\binom{n-2}{l(\gamma)-1}}{\binom{n-1}{l(\gamma)}} \right],$$

where the expectation is taken over all $(k-1)$ -element sets of ideas in $I_i(\mathbf{p}^*, \mathbf{q}^*)$, and for each such set, $l(\gamma)$ is the number of direct links on a path to at least one idea in the set. Thus,

$$\lim_{n \rightarrow \infty} \iota(q^*, q^*) \delta n = \lim_{n \rightarrow \infty} \mathbb{E}[l(\gamma)] = \lim_{n \rightarrow \infty} \mathbb{E}_{t \in PT_i(\mathbf{p}^*, \mathbf{q}^*)} [\tau(t)].$$

||

We must have $\tau(t) \geq 1$ for all t . So by Lemma 2, we have $\liminf_n \delta \iota(q^*, q^*) n \geq 1$ at any subcritical sequence of symmetric investment equilibria. This contradicts the definition of a subcritical sequence of equilibria.

Supercritical case: Suppose there exists a supercritical sequence of symmetric investment equilibria. We have $\liminf \iota(q^*, q^*) \delta n > 1$ along this sequence, and we can pass to a convergent subsequence under which $\lim \iota(q^*, q^*) n$ exists or is infinite.

Theorem 1 of Karp (1990) shows that a.a.s. the number of firms that all firms in the giant component learn from is $\alpha n + o(n)$ for a constant α increasing in $\lim \iota(q^*, q^*) n$ and that the number of agents outside the giant component observed by any agent is $o(n)$.

If firm i chooses q_i , the probability of i learning all ideas known to the giant component is $(1 - (1 - \delta \iota(q_i, q^*))^{\alpha(n-1)+o(n)})$. Conditional on this event, firm i learns $p^*(\alpha n + o(n))$ ideas. Therefore, we have

$$\mathbb{E} \left[\binom{|I_i(\mathbf{p}^*, (q_i, q_{-i}^*))|}{k-1} \right] = (p^*)^{k-1} (1 - (1 - \delta \iota(q_i, q^*))^{\alpha n + o(n)}) ((\alpha(n-1))^{k-1} + o(n^{k-1})).$$

In particular, to solve for firm i 's choice of q_i to first order, we need only consider technologies consisting of i 's private idea and $(k-1)$ ideas learned by the giant component. The probability that such a technology faces competition is

$$(1 - \delta \cdot \iota(q_i, q^*) - (1 - \delta) \cdot \iota(q_i, q^*) \cdot h(\alpha))^{n-1} + o(1),$$

where $h(\alpha)$ is the fraction of firms j such that all of j 's ideas are learned by some firm that learns all ideas known to the giant component. The term $\delta \cdot \iota(q_i, q^*)$ corresponds to the possibility of a firm j indirectly learning all of firm i 's ideas. The term $(1 - \delta) \cdot \iota(q_i, q^*) \cdot h(\alpha)$ corresponds to the possibility of a firm learning firm i 's idea via a firm j directly learning (but not indirectly learning) from i and indirectly learning the ideas learned by the giant component.

Thus, we are looking for q_i maximizing

$$\mathbb{E} \left[\binom{|I_i(\mathbf{p}^*, (q_i, q_{-i}^*))|}{k-1} \right] (1 - \delta \cdot \iota(q_i, q^*) - (1 - \delta) \cdot \iota(q_i, q^*) \cdot h(\alpha))^{n-1} \quad (\text{A.3})$$

asymptotically. We claim the derivative of expression (A.3) in q_i is equal to zero at q^* only if $\iota(q^*, q^*) \delta n \leq 1$. *A fortiori*, we can instead show this for

$$\mathbb{E} \left[\binom{|I_i(\mathbf{p}^*, (q_i, q_{-i}^*))|}{k-1} \right] (1 - \delta \cdot \iota(q_i, q^*))^{n-1}.$$

This is because for n large, the derivative of this expression in q_i is positive if the derivative of expression (A.3) is positive.

The first-order condition for the latter expression at $q_i = q^*$ implies that

$$(p^*)^k n (\alpha n)^{k-1} \delta (1 - \delta \iota(q^*, q^*))^{n-1} (\delta \alpha (1 - \delta \iota(q^*, q^*)) (1 - \delta \iota(q^*, q^*))^{\alpha(n-1)-1} - (1 - (1 - \delta \iota(q^*, q^*))^{\alpha(n-1)})$$

is $o(n^{k-1})$. Solving for $\iota(q^*, q^*)$ such that this holds, we obtain

$$\lim_n (1 - \delta \iota(q^*, q^*))^{\alpha(n-1)-1} = \frac{1}{1 + \delta \alpha}.$$

Thus,

$$\lim_n e^{-\delta\alpha\iota(q^*, q^*)n} = \frac{1}{1+\delta\alpha}. \quad (\text{A.4})$$

The left-hand side is the asymptotic probability that a firm does not indirectly learn from a firm which learns all ideas known to all firms in the giant component, and this is $1-\alpha$. So

$$1-\alpha = \frac{1}{1+\delta\alpha},$$

or equivalently $\alpha(\alpha\delta+1-\delta)=0$. Since $\delta \leq 1$, our sequence of solutions to equation (A.4) must have $\alpha=0$, and therefore cannot be supercritical.

We have now shown that any sequence of symmetric investment equilibria is critical. We next prove that there exists a symmetric investment equilibrium for n large.

Existence: We first show that given private investment $p > 0$, there exists a level of openness q that is a best response when all other firms choose (p, q) . We then show that when all firms choose openness q , there exists an optimal p for all firms.

Let $BR(q)$ be the set of best responses q when all other firms choose $q_j=q$ and $p_j=p>0$. Note that the set $BR(q)$ does not depend on the value of $p>0$.

Lemma 6. *There exists $\kappa > 0$ such that for n sufficiently large, if $\delta\iota(q, q)n \in (1-\kappa, 1+\kappa)$ then $BR(q)$ is a single-valued correspondence.*

Proof of Lemma 6 Let $\epsilon > 0$. Taking κ sufficiently small, we first claim that we can choose $M > 0$ such that

$$\lim_n \mathbb{P}[|I_i(\mathbf{p}, \mathbf{q})| > M] < \epsilon$$

for all q such that $\delta\iota(q, q)n \in (1-\kappa, 1+\kappa)$. For any $q < q'$, the random variable $|I_i(\mathbf{p}, \mathbf{q})|$ is first-order stochastically dominated by $|I_i(\mathbf{p}, \mathbf{q}')|$. So we can assume $q = 1+\kappa$. For M and n sufficiently large, the probability of learning at least M ideas is bounded above by twice the probability of learning all ideas learned by the giant component, which is less than ϵ for κ sufficiently small.

Now let $M' > 0$. We next claim that we can choose $\kappa > 0$ such that

$$\liminf_n \mathbb{E}[|I_i(\mathbf{p}, \mathbf{q})|] > M'$$

for all q such that $\delta\iota(q, q)n \in (1-\kappa, 1+\kappa)$.

Let $x > 0$ and define $q(x)$ by $\iota(q(x), q(x)) = \frac{1-x}{\delta n}$. For any $q' < q$, the random variable $|I_i(\mathbf{p}, \mathbf{q}')|$ is first-order stochastically dominated by $|I_i(\mathbf{p}, \mathbf{q})|$. So it is sufficient to show that

$$\lim_{\lambda \rightarrow 0} \lim_{n \rightarrow \infty} \mathbb{E}[|I_i(\mathbf{p}, \mathbf{q}(x))|] \rightarrow \infty.$$

We can bound $|I_i(\mathbf{p}, \mathbf{q}(x))|$ below by the expected number of firms j with a path from j to i in the indirect learning network. By Theorem 11.6.1 of Alon and Spencer (2004), the limit of this quantity as $n \rightarrow \infty$ is equal to the number of nodes in a Poisson branching process with parameter $1-x$. As $x \rightarrow 0$, this number of nodes converges to infinity. This proves the claim.

Suppose all other firms choose $q_{-i}=q$ for some q such that $\delta\iota(q, q)n \in (1-\kappa, 1+\kappa)$. It is sufficient to show that we can choose $\kappa > 0$ such that for n large, there is a unique q_i such that the derivative $\frac{\partial U_i(\mathbf{p}, (q_i, q_{-i}))}{\partial q_i}$ is equal to zero. By the argument in Lemma 3, we can restrict to q_i such that $\iota(q_i, q) \leq \bar{C}/\delta$ for some constant \bar{C} .

We can again check this first-order condition in the case that no firm has learned indirectly from i , as else payoffs are zero. Fix $\epsilon > 0$. Let v be the expected value of a single incoming link. We claim that for n sufficiently large, we can take κ so that the derivative in q_i of the expected marginal cost of an outgoing link is greater than $\delta v/2$ for q_i such that $\iota(q_i, q) \leq \bar{C}/\delta$.

The marginal cost of an outgoing link is equal to $\delta(|I_i(\mathbf{p}, (q_i, q_{-i}))|) + o(\mathbb{E}(|I_i(\mathbf{p}, (q_i, q_{-i}))|))$. Fixing $\epsilon > 0$ and M sufficiently large, we can assume the probability that firm i has already learned at least M ideas is at most ϵ . For any $M' > 0$, we have

$$\liminf_n \mathbb{E}[|I_i(\mathbf{p}, \mathbf{q})|] > M'$$

for κ sufficiently small. Taking M' sufficiently large relative to M , we can assume that the change in $\binom{|I_i(\mathbf{p}, (q_i, q_{-i}))|}{k-1}$ from an additional outgoing link is greater than $v/2$. This proves the claim.

We also claim that for κ small enough and n sufficiently large the derivative in q_i of the expected marginal value of an additional incoming link (conditioning on no firm learning indirectly from i) is less than $\delta v/4$.

The probability that firm i learns from d firms decays exponentially in d . By Rosenthal's inequality (Rosenthal, 1970), the payoffs to learning from d firms grow at most at a polynomial rate in d . Thus we can choose \bar{d} such that an

arbitrarily small share of firm i 's expected payoffs are from the event that firm i learns from more than \bar{d} other firms for n large.

We can condition on the event that firm i learns from at most \bar{d} other firms. Fixing $\epsilon > 0$ and M sufficiently large, we can assume the probability that firm i has already learned at least M ideas is at most ϵ . Taking M' sufficiently large relative to M , we can assume the expected number of ideas learned by each other firm is at least M' .

We want to evaluate the change in the marginal value of an additional incoming link when i gains an incoming link. We consider two cases depending on whether firm i learns at least M ideas from each of two or more incoming links.

The probability that i learns at least M ideas from each of two or more incoming links is at most $\epsilon^2 \binom{\bar{d}+2}{2}$ and the expected payoffs conditioning on this event are within a constant multiple of the expected payoffs conditioning on learning at least M ideas. So the payoffs from this event are a vanishing fraction of v asymptotically.

Suppose firm i learns at least M ideas from at most one link. The change in the marginal value of an additional incoming link when i gains an incoming link is equal to the number of choices of $k-1$ ideas learned by firm i , at least one of which comes from each of these two new incoming links. We have assumed firm i does not learn more than M ideas from both of these links, and therefore the number of such technologies is at most $M \binom{l_i(\mathbf{p}, q_i, q_{-i})}{k-2}$. Taking M' sufficiently large relative to M , we can take this quantity to be an arbitrarily small fraction of v for n large. Combining these arguments shows the claim.

We have shown that conditioning on no firm learning indirectly from i , the derivative in q_i of the expected marginal cost of an additional outgoing link is at least $\delta v/2$ while the derivative in q_i of the expected marginal value of an additional incoming link is less than $\delta v/4$. So the derivative in $\iota(q_i, q_{-i})$ of expected payoffs conditional on no firm learning indirectly from i is strictly decreasing, and therefore the first-order condition is satisfied uniquely, for n sufficiently large. \parallel

By the lemma, we can choose κ such that $BR(q)$ induces a function on the set of q such that $\delta\iota(q, q)n \in (1-\kappa, 1+\kappa)$ for n large. Fix such a κ . Because payoffs are continuous in \mathbf{q} , this function $BR(q)$ is continuous.

We have shown in our analysis of the subcritical region that when $\iota(q, q) = \frac{1-\kappa}{\delta n}$, any element q' of $BR(q)$ has $\delta\iota(q', q)n$ approximately equal to the expectation of $\tau(t)$ over proprietary technologies.²⁴ Because $\tau(t) \geq 1$ for all t , we have

$$\liminf_n \iota(BR(q), q)n \geq \frac{1}{\delta}.$$

We have shown in our analysis of that supercritical region that if $\iota(q, q) = \frac{1+\kappa}{\delta n}$, then

$$\limsup_n \iota(BR(q), q)n < \frac{1}{\delta}.$$

So for n large, we have $BR(q) > q$ when $\iota(q, q) = \frac{1-\kappa}{\delta n}$ while $BR(q) < q$ when $\iota(q, q) = \frac{1+\kappa}{\delta n}$. By the intermediate value theorem, there is a fixed point of the function $BR(q)$ with $\delta\iota(q, q)n \in (1-\kappa, 1+\kappa)$ for n large.

Choose any such fixed point q^* of $BR(q)$. Fix any firm i and let the potential proprietary technologies $PPT_i(\mathbf{q})$ be the set of technologies t such that firm i will receive monopoly profits for t if all ideas in the technology t are discovered. This is a random object depending on the realizations of interactions but not on the realizations of private investment, and $PPT_i(\mathbf{q}) \cap I = PT_i(\mathbf{p}, \mathbf{q})$. We have shown that \mathbf{q}^* is critical, so $\mathbb{E}[|PPT_i(\mathbf{q})|] \rightarrow \infty$.

A symmetric equilibrium corresponds to p^* satisfying

$$p^* = \operatorname{argmax}_p p(p^*)^{k-1} \mathbb{E}[|PPT_i(\mathbf{q})|] - c(p).$$

Taking the first-order condition, a symmetric equilibrium corresponds to p^* satisfying

$$c'(p^*) = (p^*)^{k-1} \mathbb{E}[|PPT_i(\mathbf{q})|].$$

Because $c(\cdot)$ is continuously differentiable and convex with $c'(0) \geq 0$ and $c(p) \rightarrow \infty$ as $p \rightarrow 1$, while $\mathbb{E}[|PPT_i(\mathbf{q})|] \rightarrow \infty$, there exists a solution for n sufficiently large. So there exists a symmetric investment equilibrium for n sufficiently large.

A.3. Arbitrary investment equilibria

To complete the proof of Theorem 1, it remains to extend our characterization from symmetric equilibria to arbitrary equilibria. It is again sufficient to show that we cannot have a supercritical sequence of investment equilibria or a subcritical sequence of investment equilibria, and we treat each case separately.

We first consider the supercritical case, and show that there exists a giant component with the same relevant properties as in our analysis of symmetric equilibria. We then consider the subcritical case, which follows the same basic outline as in our analysis of symmetric equilibria.

24. We stated this result above at equilibrium, but only used that the firm was choosing a best response.

Supercritical case: Because the sequence of actions is supercritical, we can assume that the matrix $(\iota(q_i^*, q_j^*)\delta)_{ij}$ has spectral radius at least $\underline{\lambda} > 1$ for all n sufficiently large.

We first claim that there exists $\underline{\alpha} > 0$ such that for all n , there is a component of the learning network containing at least $\underline{\alpha}n$ firms a.a.s. It is sufficient to show this after decreasing q_i for some i , and therefore also $\underline{\lambda}$. By Lemma 3, we have $q_i^* \leq \bar{C}/\sqrt{n}$ for each i . We can therefore assume without loss of generality that there are at most K choices of q_i for each n . Here the number of distinct actions K can depend on the initial upper bound $\underline{\lambda}$. We denote the number of firms choosing q_i by $n(q_i)$.

By Theorem 1 of Bloznelis, Götze and Jaworski (2012), the largest component has at least $\underline{\alpha}n + o(n)$ nodes a.a.s., where $1 - \underline{\alpha}$ is the extinction probability of the multi-type branching process with types corresponding to choices of q_i and the number of successors of type q_j of a node of type q_i distributed as a Poisson random variable with mean $\delta\iota(q_i, q_j)n(q_j)$. By Theorem 2 of Section V.3 of Athreya and Ney (1972), this extinction probability satisfies $1 - \underline{\alpha} > 0$ for n large since $\lambda > \underline{\lambda} > 1$. This proves the claim, and we now return to studying the original actions \mathbf{q} .

Because there is a component of the learning network containing at least $\underline{\alpha}n$ firms with probability at least ϵ , we can also choose \underline{C} such that \underline{q} is at most $\frac{\underline{C}}{\sqrt{n}}$ for all n . To see this, note that the payoffs to choosing $q_i = \frac{1}{\sqrt{n}}$ are of order n^{k-1} . On the other hand, the payoffs to choosing $q_i = \frac{\underline{C}}{\sqrt{n}}$ are bounded above by $2(1 - e^{-\underline{C}\bar{C}})n^{k-1}$. So for \underline{C} sufficiently small, the expected payoffs to choosing $q_i = \frac{\underline{C}}{\sqrt{n}}$ conditional on any realization of all links between firms other than i are less than the expected payoffs to choosing $q_i = \frac{1}{\sqrt{n}}$ at equilibrium.

As n grows large, for all i and j distinct the probability that $j \in t$ for a uniformly chosen $t \in PT_i(\mathbf{q}^*)$ approaches zero. We will show that for any sequence of best responses q_i for firm i , the expected number of interactions $\sum_{j \neq i} \iota(q_i^*, q_j^*)$ has a unique limit which is independent of i .

To show this, we next claim that there is at most one component of linear size a.a.s. To do so, we will use equation (5) of Bloznelis, Götze and Jaworski (2012). In the notation of Bloznelis, Götze and Jaworski (2012), the type space S will be $S = [\underline{C}, \bar{C}]$, and the kernel $\kappa(s, s') = ss'$. We will identify the type of an agent i with q_i^*/\sqrt{n} .

The space of distributions $\Delta(S)$ over types is compact. Fix such a distribution. Relabelling so that q_i^* are increasing in i , we can generate a random network for each n by taking the action q_i^* of agent i to be s/\sqrt{n} , where s is the $(i/n)^{th}$ quantile of the distribution. As $n \rightarrow \infty$, by equation (5) of Bloznelis, Götze and Jaworski (2012), the largest component of the learning network learns $\alpha n + o(n)$ ideas for some $\alpha \in [0, 1]$. It follows from Theorem 1 and the same approximation techniques used in that paper that the second largest component learns $o(n)$ ideas. Because the space of distributions $\Delta(S)$ is compact, this convergence of component sizes is uniform. So passing to a convergent subsequence if necessary, we can assume that there is a unique giant component learning $\alpha n + o(n)$ ideas a.a.s., where $\alpha > 0$.

The payoffs to choosing q_i are then equal to $(p^*)^k \binom{\alpha n}{k-1}$ times the probability that firm i learns all ideas known to the giant component and no firm j learns i 's idea and all ideas known to the giant component, plus a term of order $o(n^{k-1})$. Formally, if G_1 is the set of firms whose ideas are learned by some firm that learns all ideas known to the giant component, the action q_i is chosen to maximize

$$\binom{\alpha n}{k-1} \left(1 - \prod_{j \in G_1} (1 - \iota(q_i, q_j^*)) \right) \prod_{j \in G_1} (1 - \iota(q_i, q_j^*)) \prod_{j \notin G_1} (1 - \delta\iota(q_i, q_j^*)) + o(n^{k-1}).$$

Taking the first-order condition, we must have

$$\delta \sum_{j \in G_1} q_j^* \sim \left(1 - \prod_{j \in G_1} (1 - \iota(q_i, q_j^*)) \right) \left(\sum_{j \in G_1} q_j^* + \delta \sum_{j \notin G_1} q_j^* \right).$$

Since the right-hand side is increasing in q_i , the solution has a unique limit $\lim_n \sum_{j \neq i} \iota(q_j^*, q_i^*)$. This limit does not depend on i . Passing to a subsequence if necessary, we can assume the limit $\lim_{n \rightarrow \infty} \sum_{j \neq i} \iota(q_j^*, q_i^*)$ exists and is independent of i . Moreover, this limit must be greater than $\frac{1}{\delta}$ for equilibrium to be supercritical. But then the same calculation as in the symmetric case shows that the best response q_i for all firms is at most $\frac{1}{\sqrt{n}}$ asymptotically, which gives a contradiction.

Subcritical case: The largest component of the learning network has at most $o(n)$ nodes a.a.s. By Lemma 3, we can choose \bar{C} such that \bar{q} is at most $\frac{\bar{C}}{\sqrt{n}}$ for all n . We now proceed to derive a characterization of equilibrium as in Lemma 2. We will then use this characterization to show the result.

We now let X_j be i.i.d. random variables with distribution given by the number of ideas that firm i would learn from a firm j conditional on learning directly from j . That is, X_j is distributed as the sum of a Bernoulli random variable with success probability p_j^* (corresponding to direct learning) and a random variable with the distribution of $|I_j(\mathbf{p}^*, \mathbf{q}^*)|$ with probability δ and equal to zero otherwise.

Therefore, we can express the expected number of technologies firm i can make if it discovers its own idea as

$$\mathbb{E} \left[\binom{|I_i(\mathbf{p}^*, \mathbf{q}^*)|}{k-1} \right] = \sum_{\gamma \in \Gamma} \sum_{j_1, \dots, j_{l(\gamma)} \neq i} \prod_{r=1}^{l(\gamma)} q_i^* q_{j_r}^* \mathbb{E} \left[\prod_{r=1}^{l(\gamma)} \binom{X_{j_r}}{\gamma_r} \right] + o \left(\mathbb{E} \left[\binom{|I_i(\mathbf{p}^*, \mathbf{q}^*)|}{k-1} \right] \right).$$

The second summation is over choices of $l(\gamma)$ distinct firms other than i .

Taking the first-order condition in q_i ,

$$\frac{1}{n} \mathbb{E} \left[\frac{\partial (I_i(\mathbf{p}^*, (q_i, q_{-i}^*)))}{\partial q_i} (q^*) \right] = \frac{1}{n} \sum_{\gamma \in \Gamma} \sum_{j_1, \dots, j_{l(\gamma)} \neq i} \left(\sum_{r=1}^{l(\gamma)} q_{j_r}^* \prod_{r' \neq r} q_{j_{r'}}^* \right) \mathbb{E} \left[\prod_{r=1}^{l(\gamma)} \binom{X_{j_r}}{\gamma_r} \right] + o \left(\mathbb{E} \left[\binom{|I_i(\mathbf{p}^*, \mathbf{q}^*)|}{k-1} \right] \right).$$

By Lemma 5, we have

$$\delta \left(\sum_{j \neq i} q_j^* \right) \sum_{\gamma \in \Gamma} \sum_{j_1, \dots, j_{l(\gamma)} \neq i} \prod_{r=1}^{l(\gamma)} q_{j_r}^* q_{j_{r'}}^* \mathbb{E} \left[\prod_{r=1}^{l(\gamma)} \binom{X_{j_r}}{\gamma_r} \right] \sim \sum_{\gamma \in \Gamma} \sum_{j_1, \dots, j_{l(\gamma)} \neq i} \left(\sum_{r=1}^{l(\gamma)} q_{j_r}^* \prod_{r' \neq r} q_{j_{r'}}^* \right) \mathbb{E} \left[\prod_{r=1}^{l(\gamma)} \binom{X_{j_r}}{\gamma_r} \right].$$

Rearranging,

$$\sum_{\gamma \in \Gamma} \sum_{j_1, \dots, j_{l(\gamma)} \neq i} \prod_{r=1}^{l(\gamma)} q_{j_r}^* q_{j_{r'}}^* \mathbb{E} \left[\prod_{r=1}^{l(\gamma)} \binom{X_{j_r}}{\gamma_r} \right] \left(\delta q_i^* \left(\sum_{j \neq i} q_j^* \right) - l(\gamma) \right) \sim 0.$$

In particular, we have

$$\delta q_i^* \left(\sum_{j \neq i} q_j^* \right) \sim \mathbb{E}_{t \sim G_i(\mathbf{p}^*, \mathbf{q}^*)} [\tau(t)] \quad (\text{A.5})$$

where the expectation is taken with respect to the appropriate distribution $G_i(\mathbf{p}^*, \mathbf{q}^*)$ over technologies.

As in the symmetric case above, this implies that $\liminf_{n \rightarrow \infty} \delta q_i^* \left(\sum_{j \neq i} q_j^* \right) \geq 1$ for each i . So the limit inferior of the row sums of $(\delta I_i(q_i^*, q_j^*))_{i,j}$ is at least one. Thus, the spectral radius of this matrix also satisfies $\limsup_n \lambda \geq 1$, which contradicts our assumption that the sequence of equilibria is subcritical.

We conclude that any sequence of investment equilibria must be critical, which proves Theorem 1.

REFERENCES

- ACEMOGLU, D. and AZAR, P. D. (2020), "Endogenous Production Networks", *Econometrica*, **88**, 33–82.
- ACEMOGLU, D., MAKHDOUMI, A., MALEKIAN, A. and OZDAGLAR, A. (2017), "Privacy-Constrained Network Formation", *Games and Economic Behavior*, **105**, 255–275.
- AKBARPOUR, M., MALLADI, S. and SABERI, A. (2018), "Diffusion, Seeding, and the Value of Network Information" (Working Paper).
- AKCIGIT, U., CAICEDO, S., MIGUELEZ, E., STANTCHEVA, S. and STERZI, V. (2018), "Dancing with the Stars: Innovation through Interactions" Available at SSRN 2647049.
- ALON, N. and SPENCER, J. H. (2004), *The Probabilistic Method* (Hoboken, NJ: Wiley).
- ANTON, J. J. and YAO, D. A. (2004), "Little Patents and Big Secrets: Managing Intellectual Property", *RAND Journal of Economics*, **35**, 1–22.
- ATHREYA, K. and NEY, P. (1972), *Branching Processes* (Berlin: Springer).
- AZOULAY, P., GRAFF ZIVIN, J. S. and SAMPAT, B. N. (2012), in "The Diffusion of Scientific Knowledge across Time and Space", in *The Rate and Direction of Inventive Activity Revisited*, (Chicago, IL: University of Chicago Press) 107–155.
- BALA, V. and GOYAL, S. (2000), "A Noncooperative Model of Network Formation", *Econometrica*, **68**, 1181–1229.
- BALLESTER, C., CALVÓ-ARMENGOL, A. and ZENOU, Y. (2006), "Who's Who in Networks. Wanted: The Key Player", *Econometrica*, **74**, 1403–1417.
- BAUM, J. A., COWAN, R. and JONARD, N. (2010), "Network-Independent Partner Selection and the Evolution of Innovation Networks", *Management Science*, **56**, 2094–2110.
- BAUMANN, L. (2021), "A Model of Weighted Network Formation", *Theoretical Economics*, **16**, 1–23.
- BERLIANT, M. and FUJITA, M. (2008), "Knowledge Creation as a Square Dance on the Hilbert Cube", *International Economic Review*, **49**, 1251–1295.
- BESSEN, J. and HUNT, R. M. (2007), "An Empirical Look at Software Patents", *Journal of Economics & Management Strategy*, **16**, 157–189.
- BESSEN, J. and MASKIN, E. (2009), "Sequential Innovation, Patents, and Imitation", *The RAND Journal of Economics*, **40**, 611–635.
- BLOZNELIS, M., GÖTZE, F. and JAWORSKI, J. (2012), "Birth of a Strongly Connected Giant in an Inhomogeneous Random Digraph", *Journal of Applied Probability*, **49**, 601–611.
- BOLDRIN, M. and LEVINE, D. K. (2008), "Perfectly Competitive Innovation" *Journal of Monetary Economics*, **55**, 435–453.
- BRAMOULLÉ, Y., KRANTON, R. and D'AMOURS, M. (2014), "Strategic Interaction and Networks", *American Economic Review*, **104**, 898–930.
- BURT, R. S. (1992), *Structural Holes: The Social Structure of Competition* (Cambridge, MA: Harvard University Press).
- CABRALES, A., CALVÓ-ARMENGOL, A. and ZENOU, Y. (2011), "Social Interactions and Spillovers", *Games and Economic Behavior*, **72**, 339–360.

- CAMPBELL, A. (2013), "Word-of-Mouth Communication and Percolation in Social Networks", *American Economic Review*, **103**, 2466–2498.
- CHEN, J., ELLIOTT, M. and KOH, A. (2021), "Capability Accumulation and Conglomeratization in the Information Age" Available at SSRN 2753566.
- COHEN, W. M. and LEVIN, R. C. (1989), "Empirical Studies of Innovation and Market Structure", *Handbook of Industrial Organization*, **2**, 1059–1107.
- CURRARINI, S., JACKSON, M. O. and PIN, P. (2009), "An Economic Model of Friendship: Homophily, Minorities, and Segregation", *Econometrica*, **77**, 1003–1045.
- EROL, S. and GARCÍA-JIMENO, C. (2020), "Civil Liberties and Social Structure" Available at SSRN.
- GOLUB, B. and LIVNE, Y. (2010), "Strategic Random Networks" Available at SSRN 1694310.
- GOYAL, S. and MORAGA-GONZALEZ, J. L. (2001), "R&D Networks", *Rand Journal of Economics*, **32**, 686–707.
- GRIFFITH, A. (2019), "A Continuous Model of Strong and Weak Ties" (Working Paper).
- JACKSON, M. O. and WOLINSKY, A. (1996), "A Strategic Model of Social and Economic networks", *Journal of Economic Theory*, **71**, 44–74.
- KARP, R. M. (1990), "The Transitive Closure of a Random Digraph", *Random Structures & Algorithms*, **1**, 73–93.
- KELLY, M. (2009), "Technological Progress under Learning by Imitation", *International Economic Review*, **50**, 397–414.
- KERR, W. R. (2008), "Ethnic Scientific Communities and International Technology Diffusion", *The Review of Economics and Statistics*, **90**, 518–537.
- KLENKE, A. and MATTNER, L. (2010), "Stochastic Ordering of Classical Discrete Distributions", *Advances in Applied Probability*, **42**, 392–410.
- KÖNIG, M. D., BATTISTON, S., NAPOLETANO, M. and SCHWEITZER, F. (2011), "Recombinant Knowledge and the Evolution of Innovation Networks", *Journal of Economic Behavior & Organization*, **79**, 145–164.
- (2012), "The Efficiency and Stability of R&D Networks", *Games and Economic Behavior*, **75**, 694–713.
- KÖNIG, M. D., LORENZ, J. and ZILIBOTTI, F. (2016), "Innovation vs. Imitation and the Evolution of Productivity Distributions", *Theoretical Economics*, **11**, 1053–1102.
- KORTUM, S. S. (1997), "Research, Patenting, and Technological Change", *Econometrica*, **65**, 1389–1419.
- KULTTI, K., TAKALO, T. and TOIKKA, J. (2006), "Simultaneous Model of Innovation, Secrecy, and Patent Policy", *American Economic Review*, **96**, 82–86.
- ŁUCZAK, T. (1990), "The Phase Transition in the Evolution of Random Digraphs", *Journal of Graph Theory*, **14**, 217–223.
- PERLA, J. and TONETTI, C. (2014), "Equilibrium Imitation and Growth", *Journal of Political Economy*, **122**, 52–76.
- ROSENTHAL, H. P. (1970), "On the Subspaces of L^p ($p > 2$) Spanned by Sequences of Independent Random Variables", *Israel Journal of Mathematics*, **8**, 273–303.
- SADLER, E. (2020), "Diffusion Games", *American Economic Review*, **110**, 225–70.
- SAMILA, S. and SORENSEN, O. (2011), "Venture Capital, Entrepreneurship, and Economic Growth", *The Review of Economics and Statistics*, **93**, 338–349.
- SAXENIAN, A. (1996), *Regional Advantage* (Cambridge, MA: Harvard University Press).
- SCOTCHMER, S. (1991), "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law", *Journal of Economic Perspectives*, **5**, 29–41.
- SCOTCHMER, S. and GREEN, J. (1990), "Novelty and Disclosure in Patent Law", *The RAND Journal of Economics*, **21**, 131–146.
- STEIN, J. C. (2008), "Conversations among Competitors", *American Economic Review*, **98**, 2150–62.
- STORPER, M. and VENABLES, A. J. (2004), "Buzz: Face-to-Face Contact and the Urban Economy", *Journal of Economic Geography*, **4**, 351–370.
- TARTARI, V. and STERN, S. (2021), "More than an Ivory Tower: The Impact of Research Institutions on the Quantity and Quality of Entrepreneurship" (Working Paper).
- WEITZMAN, M. L. (1998), "Recombinant Growth", *The Quarterly Journal of Economics*, **113**, 331–360.