COLLABORATION NETWORK FORMATION AND THE DEMAND FOR PROBLEM SOLVERS WITH HETEROGENOUS SKILLS

KATHARINE A. ANDERSON
UNIVERSITY OF MICHIGAN, DEPARTMENT OF ECONOMICS

Abstract. Collaborative problem solving is important in a wide range of contexts, including economic production, product development, academic research, and policy making. Here, I present a formal model of collaborative problem solving, in which individuals with heterogeneous skill sets collaborate to solve problems. I show that the number of problems an individual solves is a supermodular function of her set of skills, and cannot be determined by pricing her skills individually. I then look at the network formed by the collaborative links between problem solvers. An individual's position in this network reflects the demand for her skills as a problem solver, and the overall structure of the network reflects the nature of the problem-solving community. I show that the degree distribution of the network will be fat-tailed–that is, a small number of players solve the vast majority of the problems, while most players solve relatively few. This result holds, even when skills are distributed independently across the problem solvers (Bernoulli Skills Model). The degree distribution becomes more skewed when problems are difficult for the population, and when skills are arranged into disciplines (the Ladder Model).

1. Introduction

Collaborative problem solving is important in a wide range of contexts, including economic production, product development, policy making, and academic research. In all of these, individual problem solvers work together to solve problems that none of them could solve alone. For example, research groups in a pharmaceutical firm search for new and better molecules; architectural firms design new buildings; teams of programmers create more efficient algorithms; and academic collaborations answer open scientific questions. Collaboration is widely recognized as a vital part of problem solving, because it allows diverse teams of individuals to pool their skills towards a common goal.[1] As problems become more difficult, few individuals have all of the pieces required, and collaboration becomes even more important.[2]

By linking two players who work together on a problem, we create a collaboration network. An individual's position in the network reflects her prominence in the community of collaborators and her value as a problem solver. Players with more connections are presumably more important to the community because their skills

Thanks to Scott Page and Ross O'Connell. This work was completed with the help of funding from the NSF and computing resources from the Center for the Study of Complex Systems.

[1]Philips et al. (2004), Polzer et al. (2002), Thomas-Hunt et al. (2003)

[2]Hong and Page (2001) and (2004) show the importance of collaboration in problem solving. They show that under a wide range of conditions, diverse teams of problem solvers will outperform teams of experts.

are in higher demand. The overall structure of this network reflects the nature of the problem-solving community. In particular, the degree distribution of this network shows how output is distributed across the problem solvers. In networks where the degree distribution is skewed, a few individuals solve most of the problems, while the majority solve relatively few. These network structures are important because they, in turn, govern a wide range of other interactions, including the information and ideas (Jackson and Rogers (2007), Newman (2003)), individual reputation (Golub and Jackson (2007)), and opinion formation (DeGroot (1974)).

In this paper, I present a formal model of collaborative problem solving and collaboration network formation, in which individual problem solvers have heterogeneous skills and collaborate in order to solve difficult problems. In this model, skills are pieces of knowledge useful for solving problems. For example, a skill might be familiarity with a complex tool or technique, ability as a programmer, or knowledge of a particular field.[3] Each problem solver has a subset of the total set of skills, representing her human capital. Problems, in this model, are activities requiring certain sets of skills.[4] Although individual problem solvers may have some of the skills required to solve a given problem, most problems are too difficult to be solved by an individual working alone. Thus, problem solvers in my model collaborate with others to gain access to the skills they lack. The number of problems they help solve is the demand for their skills, and proxy for their value to the collaborative community.

In the first part of this paper, I look at the relationship between an individual's skills and the demand for her as a collaborator. I show that the number of problems a player solves is a supermodular function of her set of skills, and cannot be determined by pricing her skills individually. This is because in collaborative problem solving, collaborators bring all of their skills to the problem at hand. Thus *combinations of skills* important. In particular, an individual with a useful combination of skills can outperform one with many rare skills, bringing into question the utility of one-dimensional ability measures in models of problem solving.

By linking players who collaborate together on a problem, we form a collaboration network. An individual's an individual's degree on this network is the number of problems that she helps to solve and the number of collaborators she has. In the second part of the paper, I make a connection between overall structure of this collaboration network and the distribution of skills in the problem solving population. I find that even when skills are distributed independently across players, the degree distribution of the collaboration network is highly skewed–that is, a few players solve the majority of the problems, while most players solve very few. This creates a network with a distinctive, "hub and spoke" structure, similar to that observed in empirical collaboration networks. The inequality in the distribution of degree holds even when the skills are independently distributed in the population (the

[3]Note that skills (such as the ability to program in java, or familiarity with the field of combinatorics) are distinguished from information (such as an observation about local weather conditions, or the availability of employment at a firm) by the fact that skills they are non-transferable in the short run. Whereas information can be passed easily from individual to individual, and may even be aggregated, skills cannot.

[4]For example, if the problem solvers are biologists, they may face an open research question requiring experience working with a particular organism, familiarity with a difficult lab technique, C-programming skills, knowledge of an unusual statistical tool, and familiarity with a the literature in a particular sub-field.

Bernoulli Skills Model), and becomes even more pronounced as problems become more difficult. When skills are arranged into disciplines (the Ladder Model), the degree distribution becomes even more skewed.

This paper makes contributions to several distinct literatures. First, the results of this paper have important implications for labor and industrial organization, as our economy shifts away from manufacturing towards more knowledge-based production. It is widely recognized that the US economy has undergone a transition from production based in knowledge exploitation to one based in knowledge creation (Hagel et al. (2009)). This transition[5] is associated with a wide range of effects in labor and industrial organization, including an attenuation of the distribution of output Rosen (1981) and income (see Juhn et al. (1991) and Machin (2008)) and a flattening of organizational structures (Bresnahan et al. (2002) and Rajan and Wulf (2006)). In addition, because problem solving production is an intensely cooperative effort, we observe an increasing number of collaborative connections between firms (Powell et al. (1996)). Unfortunately, most current models of production are still based in existing models of manufacturing and trade. Labor productivity in these models is denoted either by a one dimensional ability measure (eg: speed) or a labor type (eg: speciality). This makes it very difficult to answer questions particular to collaborative problem solving.

The detailed treatment of skills in this model adds considerable value, when compared with these more traditional model of labor production. Players in this model have multiple skills, and their skill sets can overlap in any of a number of ways. This treatment of labor is advantageous because it much broader than the traditional treatment, encompassing both ability-based and type-based models. It also allows us to ask questions about the value of skill *combinations*, which would not be relevant in a model with individual skills or one-dimensional ability levels. Moreover, the relationships revealed by this treatment of skills are not what we would naively expect, given our understanding of more coarse-grained models. In particular, I find that the distribution of labor demand will be skewed towards a few, highly productive individuals, similar to what is observed in empirical labor markets. Moreover, this model provides a framework for creating a more general model of organization within knowledge-based firms. I am also able to answer questions about the value of a particular skill to a particular problem solver. I show that the value of a skill depends on both the supply and demand for that skill in the population, and the set of skills the problem solver already has, indicating that optimal training decisions will be highly individualized.

This paper also contributes to the network formation literature. A growing literature has demonstrated the importance of social networks in social, political, and economic interactions.[6] The structure of collaboration networks shapes a wide range of other interactions, affecting the spread of information and ideas (Jackson and Rogers (2007), Newman (2003)), individual reputation (Golub and Jackson (2007)), and opinion formation (DeGroot (1974)). The position of an individual in the collaboration network governs her access to knowledge, tools, and information (Coleman et al (1966)), and thus it may shape the kinds of questions she addresses. The structure of collaboration networks also affects job search and hiring (Granovetter (1973) and (1995)), the adoption of new technologies (Ryan and Gross

[5]which Hagel et al. call "The Big Shift"

[6]See Jackson (2008) for a good survey of the existing literature.

(1943), Hagerstrand (1967)), and the influence of individual researchers (DeMarzo et al (2003), Golub and Jackson (2007)). In the case of academic research, network structure may even affect the course of scientific inquiry. Empirically, we observe that collaboration networks have some common structural characteristics, which transcend context. In particular, the degree distribution of these networks is fat-tailed.[7] This means that a small number of individuals participate in the vast majority of the collaborations, while most individuals participate in relatively few. This skewed degree distribution has been observed in a wide range of collaboration networks, including interfirm collaboration (Powell et al (1996), Iyer et al (2006)), creative artists in broadway plays (Uzzi and Spiro (2005)), film actors (Barabasi and Albert (1999)), jazz musicians (Gleiser and Danon (2003)), and coauthorship networks in a variety of fields.[8]

Because the structure of collaboration networks affects other behaviors, there is a premium attached to understanding the origins and determinants of that structure. Statistical models of network formation, such as preferential attachment,[9] and models based on incumbency[10] do a good job of recreating the fat-tailed network structure in empirical collaboration networks. However, these models rely on stochastic processes to drive link formation. Players do not make choices about which links to make, so they cannot answer questions about the relationship between behavior and network structure.

There have been several attempts to model network formation behaviorally. In these models, player choose their links strategically, in order to maximize their payoffs. Jackson and Wolinsky (1996) present a model of coauthorship networks, in which each link represents a single paper. Players in their model must allocate effort across various projects, and thus the payoff from a paper is inversely related to the number of links the two coauthors have.[11] Goyal and Moranga-Gonzalez (2001) construct a model of collaboration among firms, rather than individuals. Firms in their model choose a set of links and an effort level to put into research and development. The firm's immediate neighbors experience perfect spillover effects from

[7]That is, there are more players with very high degree and low degree, as compared to a random network with the same average degree. Exponential and scale-free distributions are two examples of fat-tailed distributions.

[8]Newman (2001) examines coauthorship networks for several subdisciplines of physics, biomedical fields, and computer science. Moody (2004) does the same for sociology. Goyal et al (2006) looks at economists. Acedo et al (2006) present data on researchers in management and organizational studies, and while they do not directly address the degree distribution, their data includes more high-degree nodes than would be expected in a random network, suggesting a fat-tailed distribution.

[9]First introduced by Barabasi and Albert (1999). In preferential attachment models, new nodes connect to older nodes at random, but they connect to high-degree nodes with greater probability. This creates a statistical "rich get richer" phenomenon, and the resulting network has a power law degree distribution (eg: $f(k) \propto \alpha^k$). Several variations on the preferential attachment model produce degree distributions that are an even better fit to the observed distributions–see, for example, Jackson and Rodgers (2007) and Ramansco et al (2007)

[10]Guimera et al (2005) presents a model sequential team assembly based on the balance between experience and diversity. The model is statistical, and has two parameters, representing the probability that a newcomer enters the field and the probability that an incumbent player works with the same team twice.

[11]They show that in the equilibrium of this game, players form into a collection of fully-connected groups, each of a different size and the efficient configuration arranges all of the players into partnerships, indicating that collaboration networks tend to be more connected than is efficient.

the firm's efforts, whereas unconnected firms receive imperfect spillover effects. [12] Both of these models provide insights into the relationship between incentives and network structure. However, because players are homogeneous in these models, the network structure obtained is very symmetrical. Moreover, it is impossible to answer questions about the value of particular skill sets in a model with homogeneous players.

The model I present here is behavioral–players receive payoffs for solving problems, and choose a set of links that maximizes that payoff. However, it differs from existing behavioral models in it's treatment of skills and problem solving. Here, I allow the problem solvers to be heterogeneous. This creates a rich collaborative environment, in which players seek out others with complementary skills, and allows me to ask questions about how a particular individual's skill set is valued in the community. This heterogeneity in players also breaks the symmetry of the resulting collaboration network, resulting in a network with a degree distribution similar to that observed empirical networks. This allows me to look at how the degree distribution is affected by the population of problem solvers.

Note that this model combines two distinct lines of research. Existing models of collaboration network formation do not consider the impact of skills on individual degree and overall network structure, and existing models of problem solving and collaboration do not consider the network structures that result from the interactions of individuals. The model I present in this paper bridges that gap, and provides a framework which can be used to address a wide array of new questions.

The rest of the paper is organized as follows. Section 2 presents the model, and offers a brief discussion of it's characteristics. Section 3 examines the relationship between a player's degree on this network and her set of skills. Sections 5 and 6 take a step back and look at how the overall structure of the collaboration network depends on the distribution of skills in the population. Section 7 discusses the implications of these results in the labor market and industrial organization. Section 8 presents some possible extensions and Section 9 concludes.

2. A General Model of Skills, Problem Solving, and Collaboration Networks

2.1. **Inputs: Problem Solving Population and Problems.** The inputs to this model are a single problem and a population of problem solvers.

Let $I = \{1, 2, ...N\}$ be the set of problem solvers.

Let $S = \{a_1 ... a_M\}$ denote the set of all skills.

$A_i \subseteq S$ is the subset of those skills possessed by player i, which I will call her *skill set*.[13] The players' skill sets are distributed according to Ψ, a probability measure with support $\Sigma(\Psi) \subseteq 2^S$–that is, $\Psi(A)$ is the fraction of the players in I who have the skill set $A \subseteq S$.[14]

[12] They show that the complete network is the unique pairwise stable network structure. They also compare the efficiency of equilibrium networks under different levels of firm rivalry. When firm rivalry is low, the equilibrium configuration is also efficient. When firm rivalry is high, the complete network is inefficient when compared with a network with fewer links.

[13] We could think of A_i as player i's human capital.

[14] Formally, Ψ is a *frequency* distribution–that is, Ψ is a *realized* distribution of skill sets across players, rather than a statistical one. The distinction between frequency and probability distributions disappears when N is large, but using a frequency distribution allows me to also make statements about small N as well.

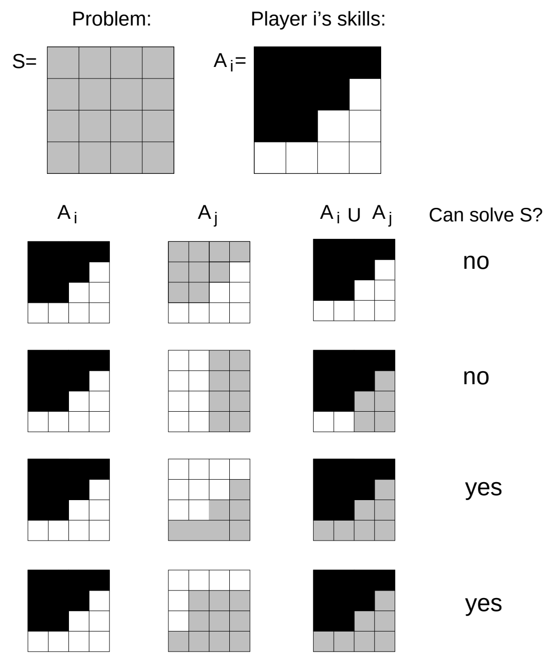

FIGURE 2.1. A graphical example of collaboration and problem solving in this model. The problem to be solved requires 16 skills, represented by the boxes. Player i has 9 of the required skills, represented by the filled boxes. Player i can solve the problem only by collaborating with someone who has the skills she lacks.

Each player is endowed with a single problem, $\omega_i \subseteq S$, which requires a subset of the skills in the population.

A *collaboration* is a subset of the players, $C \subset I$. A player and her collaborators can *solve* a problem if together hey possess all of the required skills–that is, if $\omega_i \subseteq \bigcup_{j \in C_i} A_j$ (see Figure 2.1 for an illustration).

The problem yields a payoff of 1 if solved. If the player can solve her problem alone (that is, if $A_i = \omega_i$) then she keeps the entire payoff. If she solves it with the help of other players, then she splits the payoff evenly with them, giving each a share of $\frac{1}{|C_i|}$ and retaining a similar share for herself. Each player faces a problem, and thus player i's payoff is the sum the payoff she gets from solving her own problem, plus any payoffs she gets from collaborating with others on their problems:

$$u_i = \frac{1}{|C_i|} + \sum_{j \neq i \; st \; i \in C_j} \frac{1}{|C_j|}$$

A player chooses her set of collaborators (C_i) to maximize her utility. Note that player i's payoff to solving her own problem is always positive, and thus it is always incentive compatible for her to find a solution to the problem. Since the player controls only her own collaborative decisions, a utility-maximizing player chooses C_i to minimize the number of connections she must make–in other words, she chooses a minimal subcover of the set of skills she lacks–$A_i^c = \omega_i \backslash A_i$ (see Figure 2.2 for an illustration). Let $\mathbb{C}_i$ denote the set of all minimal subcovers of A_i^c. I

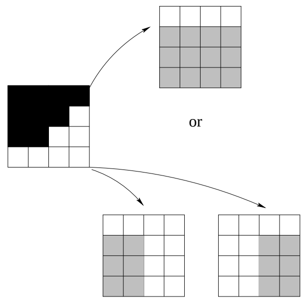

FIGURE 2.2. A graphical illustration of the player's optimization decision. Solving a problem yields a payoff of 1. Because the player splits this payoff equally with her collaborators, she optimizes by choosing the minimum number required to solve the problem.

assume that if there exist multiple minimal subcovers (ie: if $|\mathbb{C}_i| > 1$) then the player will choose a random minimal subcover, $C_i^* \in \mathbb{C}_i$.[15]

2.2. **Cost-minimizing Collaboration Networks.** For a given a set of collaborations, $C = \{C_1 ... C_N\}$, the collaboration network is represented by an adjacency matrix, $g(C)$, where $g_{ij}(C) = 1$ if $j \in C_i$. Note that the network is directed–since $j \in C_i$ does not necessarily imply $i \in C_j$, it may be that $g_{ij}(C) \neq g_{ji}(C)$. However, the links are mutual, in the sense that neither player wants to terminate a link (see Section 2.5 for further discussion). When all collaborators are chosen optimally (that is, when $C_i \in \mathbb{C}_i \, \forall i$), I will call the result a *cost-minimizing collaboration network.*

Definition. A network, $g(C)$, is a *cost minimizing network* if each player in the network chooses a minimal set of collaborators required to solve her problem–eg: if $C_i \in \mathbb{C}_i \, \forall i$.

Since the set of minimal subcovers for each player ($\mathbb{C}_i$), depends on the distribution of skills in the population, I use $\Gamma(\Psi)$ to denote the set of cost-minimizing collaboration networks for a particular distribution of skills, Ψ.

Before continuing, a brief word about network notation is in order. First, note that for ease of reading I will usually drop the argument of $g(C)$. I will denote a link from player i to player j by ij. Using a slight abuse of notation, I will use g to refer to both the adjacency matrix (as above) *and* the set of links in the network–that is, $ij \in g$ if i is connected to j in the network g. In a similar abuse of notation, I will use $g - ij$ to represent the network that results when the link ij is removed from an existing network, g, and $g + ij$ to represent the network that results when the link ij is added to the existing network, g.

2.3. **Example.** An example will help clarify the structure of this model. Suppose all of the players face the same problem requiring three skills: $\omega_i = \omega = \{a, b, c\}$ $\forall i$. Suppose the distribution of skills is such that every player has at least one skill, but no player has all of the skills required. In other words, the support of Ψ is the set $\{\{a\}, \{b\}, \{c\}, \{ab\}, \{ac\}, \{bc\}\}$. In this particular case, each player

[15]Since players are indifferent between minimal subcovers, this choice at random follows convention. The results are not sensitive to this assumption.

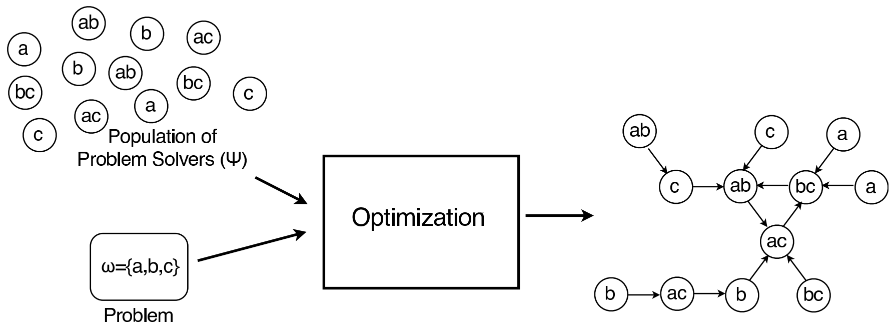

Figure 2.3. An example, illustrating the model. The inputs to the model are a problem, ω, and a population of problem solvers with a distribution of skills. In this case, there are $N = 12$ players. The players face a problem requiring three skills: $\omega = \{a, b, c\}$. Each player has either one or two skills, and they have an equal probability of having any combination. That is, $\Psi(A) = \frac{1}{6}$ for $A \in \{\{a\},\{b\},\{c\},\{ab\},\{ac\},\{bc\}\}$ and $\Psi(A) = 0$ otherwise. Players optimize their set of collaborators, and the result is a cost minimizing collaboration network. The pictured network is one example of a cost minimizing collaboration network for this problem and population.

needs to make only one link in order to solve the problem–a player with skill set $\{a\}$ must link to one of the players with the skill set $\{b, c\}$, a player with skill set $\{a, b\}$ may choose from those with skill sets $\{c\}, \{a, c\}$, and $\{b, c\}$, and so on. Figure 2.3 shows a schematic of the model. The inputs are the problem, and a particular skill distribution (in this case, $\Psi(A) = \frac{1}{6}$ for all $A \in \{\{a\},\{b\},\{c\},\{ab\},\{ac\},\{bc\}\}$ and $\Psi(A) = 0$ otherwise). The players optimize their choice of collaborators, and the result is a collaboration network. The figure shows an example network a particular set of optimal choices. The set $\Gamma(\Psi)$ is composed of many networks with the same skill distribution, but different choices of minimal subcovers.

2.4. **Discussion.** Speaking generally, there are two inputs to the model: a problem (ω) and a distribution of skills (Ψ). The output of the model is a set of cost minimizing collaboration networks, in which each player has chosen a minimal set of links in order to solve her problem.

This model correctly predicts several empirical facts about collaboration and problem solving. First, the model predicts that as problems become increasingly difficult, collaboration networks will become more densely connected. If the players have most of the skills required to solve a given problem, then the problem is a simple one for that population. In that case, most players will be able to solve the problem on their own, and the resulting collaboration network will be sparse. Conversely, if the players have few of the required skills, then the problem is a difficult one for that population and the resulting network will be more dense. This prediction is born out in data from a variety of academic fields–collaborative work has become increasingly common in mathematics,[16] physics,[17] sociology,[18] management science,[19] and economics.[20] Moreover, the literature supports a connection between increased collaboration, the difficulty of problems faced, and the increasing complexity of required methodologies.[21]

The model also predicts that problem solvers will seek out collaborators that are unlike themselves–that is, players who have complementary skills.[22] Diversity is widely recognized as contributing to the success of collaborative problem-solving groups,[23] and theoretical work indicates that diversity may even be more important to collaborative success than raw ability.[24] Empirical data backs up these assertions, indicating that collaborators are more likely to to collaborate if they have dissimilar backgrounds.[25]

2.5. **A Note on Stability and Efficiency of the Cost Minimizing Social Network.** Before considering any specific questions about the cost-minimizing collaboration network, it is worth considering the stability and efficiency of that network. Jackson and Wolinsky (1996) introduce an equilibrium concept of network stability, called *pairwise stability.* Briefly, a network is pairwise stable if no individual would prefer to terminate an existing link, and if no pair of individuals would prefer to add a link (see Appendix A for a more formal definition). Together, these two conditions ensure that links are mutual. That is, if a network is pairwise stable, then both players agree to maintain the link.

Theorem 1 states that any cost minimizing collaboration network is pairwise stable,[26] and thus all links in the network are mutual. Moreover, it states that any cost-minimizing collaboration network is strongly efficient–that is, the players extract the maximum possible value from the network.

Theorem 1. *Any cost minimizing collaboration network, $g \in \Gamma(\Psi)$, is pairwise stable and strongly efficient.*

---

[16]Grossman and Ion (2002)

[17]Barabasi et al (2002)

[18]Moody (2004)

[19]Acedo et al (2006)

[20]Laband and Tollison (2000) look at papers published in three prominant economic journals (*American Economic Review, Journal of Political Economy,* and *The Quarterly Journal of* Economics*)* from 1930-1995. The percentage of economics papers that were coauthored is around 10% in the period from 1930-1960, but rises to over 50% by 1990. The number of authors per paper also rises, from essentially 1 to 1.5 by the mid-1990s. Goyal et al (2006) notes that the average number of coauthors per individual in economics nearly doubled in the period from 1970-1999.

[21]Laband and Tollison (2000) suggest that coauthorship is more common in fields where intellectual advances are difficult or costly, and that the rise in coauthorship in economics and biology over the past 50 years could be attributed to increasingly complex methodologies, which are more costly to learn. Moody (2004)

[22]This is because in this case, players do not benefit from redundant skills. However, the result still holds for alternative production functions–we just need for the returns to a single skill to be decreasing in the number of copies of that skill obtained.

[23]Using a longitudinal study of work groups, Polzer *et al.* (2002) find that diversity improves group performance.

[24]Hong and Page (2004) shows that under a broad range of conditions, a randomly-selected group of diverse problem solvers will out-perform a group of non-diverse experts.

[25]Fafchamps et al (2006) show that economics researchers are more likely to cooperation if they have dissimilar experience and ability levels.

[26]It is actually *more* than pairwise stable, because linking players choose an optimal *set* of links from all possible sets.

Proof. See Appendix A □

This means that in any cost-minimizing collaboration network, all collaborative links are mutually beneficial, and the problem solvers choose an efficient network structure. This result is in contrast with other models of network formation–for example, Jackson and Wolinsky (1996) and Goyal and Moranga-Gonzalez (2001)–in which pairwise stability and efficiency do not coexist.[27]

3. Skills and Degree: Skill Sets and Collaborative Success

A player's *in-degree* in the network–which I will denote d_i– is the number of links that are directed towards that player and in the context of collaboration networks, and it represents the number of problems that the player helps to solve. In this section, I consider how a player's degree in the collaboration network depends on her set of skills. I show that when players can have multiple skills, a player's degree[28] in the network is a highly non-linear function of her set of skills, meaning that the value of a *combination* of skills may be greater than the sum of its parts. This result suggests that as an input to production, skills should be valued much differently than either raw materials or man hours.

3.1. Skills and Degree: An Example.

Before presenting the main results of this section, it is useful to see an example. Suppose players face a problem requiring three skills, $S = \{a, b, c\}$. Further, suppose each player has one or two of those skills, so that Ψ has support $\{\{a\}, \{b\}, \{c\}, \{ab\}, \{ac\}, \{bc\}\}$. The number of problems a player will help solve, and thus her in-degree on the network, will depend on the number of players who need her skills and the number of other players who has those same skills. For example, consider a player with the skill set $\{a\}$. She can help any player who has the complementary set of skills, $\{b, c\}$. A player with $\{b, c\}$ may ask anyone with skill a for help, including those with skill sets $\{a\}, \{a, b\}$, or $\{a, c\}$. So the expected degree of a player with skill set $\{a\}$ is

$$E[d(\{a\})] = \frac{\Psi(\{b, c\})}{\Psi(\{a\}) + \Psi(\{a, b\}) + \Psi(\{a, c\})}$$

Similarly, a player with the skill set $\{a, b\}$ can help any player who needs skill a or skill b, yielding expected degree

$$E[d(\{a, b\})] = \frac{\Psi(\{b, c\})}{\Psi(\{a\}) + \Psi(\{a, b\}) + \Psi(\{a, c\})} + \frac{\Psi(\{a, c\})}{\Psi(\{b\}) + \Psi(\{a, b\}) + \Psi(\{b, c\})} + \frac{\Psi(\{c\})}{\Psi(\{b, c\})}$$

Note that the expected degree of a player with both skills a and b is greater than $E[d(a)] + E[d(b)]$. This is because a player with both skills can help players who need skill a, players who need skill b, *and players who need both.*

[27] In both of these papers, the pairwise stable network has too many links, when compared with the efficient network.

[28] Here, and in most of the following, I will drop the modifier and refer to in-degree simply as "degree". I use in-degree because it has a clear, empirical interpretation, but the results qualitatively similar if we consider a player's degree to be the sum of his in-degree and out-degree, or use the degree of the player in a network where directed links are projected into undirected links.

3.2. **Skills and Degree: General Results.** Theorem 2 states that degree is a supermodular function of a player's set of skills. That is, regardless of the skill set required for the problem, or the distribution of skills, a player with skill set $A \cup B$ can solve at least as many problems as players with A and B put together. The sketch of the proof is similar to the above example–the set of all problems that can be solved by a player who has the skill set $A \cup B$ includes those that can be solved by a player with skill set A, and those that can be solved by a player with skill set B, and those requiring some skills from *both* sets.

Theorem 2. *For any set of skills, S, and distribution of those skills, Ψ, a player's expected degree over the networks in $\Gamma(\Psi)$ is a supermodular function of her set of skills. That is, $Ed(A \cup B) + Ed(A \cap B) \geq Ed(A) + Ed(B)$.*

Proof. Here, I will prove the result for the case where players need only one collaborator to solve their problem. The proof for the general result is similar, and can be found in Appendix B. For the sake of clarity, I consider the case where $A \cap B = \emptyset$ (again, the more general result appears in Appendix B). Since $d(A \cap B) = d(\emptyset) = 0$, we need to show that $Ed(A \cup B) \geq Ed(A) + Ed(B)$. Consider $d(A \cup B)$. Recall that a player can help anyone in the population who requires a subset of her skills. The fraction of players who need a particular subset of $A \cup B$ is the fraction who have exactly the complementary set of skills, so the fraction needing $C \subseteq A \cup B$ is $\delta(C) = \Psi(S \backslash C)$. The fraction who can supply the set C is $\sigma(C) = \sum_{D \subseteq S \backslash C} \Psi(C \cup D)$. Thus, a player with the skill set $A \cup B$ has expected degree

$$E[d(A \cup B)] = \sum_{C \subseteq A \cup B} \frac{\Psi(S \backslash C)}{\sum_{D \subseteq S \backslash C} \Psi(C \cup D)} = \sum_{C \subseteq A \cup B} \frac{\delta(C)}{\sigma(C)}$$

We can divide the subsets of $A \cup B$ into one of three categories according to the skills required:

(1) Problems requiring only skills in A
(2) Problems requiring only skills in B
(3) Problems requiring some skills from A and some from B

which gives us the following:

$$\begin{aligned} E[d(A \cup B)] &= \sum_{C \subseteq A} \frac{\delta(C)}{\sigma(C)} + \sum_{C \subseteq B} \frac{\delta(C)}{\sigma(C)} + \sum_{C \subseteq A \cup B \, and \, C \cap A, \, C \cap B \neq \emptyset} \frac{\delta(C)}{\sigma(C)} \\ &= E[d(A)] + E[d(B)] + \phi \\ &\geq E[d(A)] + E[d(B)] \end{aligned}$$

□

This theorem suggests an immediate corollary.

Corollary 3. *Adding skills to a player's skill set will never decrease her degree in a cost minimizing collaboration network.*

Proof. Suppose a player has a skill set, A, and we add a new skill that he did not have $a \notin A$. From Theorem 2, $E[d(A \cup a)] + E[d(A \cap a)] = E[d(A \cup a)] \geq E[d(A)] + E[d(a)]$, and so $E[d(A \cup a)] - E[d(A)] \geq E[d(a)] \geq 0$. □

3.3. **Bundled Skills and the Importance of Skill Combinations.** Theorem 2 indicates that an individual's importance in a community of collaborators reflects not only the supply and demand of her individual skills, but also the supply and demand of her *combination* of skills. This means that a player with a useful combination of skills might be more important to the community than a player with many skills or rare skills.

This result best illustrated using an example–consider a problem requiring five skills, $S = \{a, b, c, d, e\}$, which are distributed across $5N$ players, as shown in Table 1. Each skill is held by exactly 2 players, and thus no skill is rarer than the others. Traditionally, we might condense the information contained in this table into a single measure of ability. Player 1 and 2 have the most skills, and therefore, we would expect them to have the most value in the community. However, despite having fewer skills than players 1 and 2, player 3 will receive more links, in expectation. Note that player 3's skills are not rare, individually. However, her combination of skills *is* rare and valuable to many different people.[29] Therefore, she receives more links than a tally of her individual skills might predict.[30]

	a	b	c	d	e	in-degree
1	X	X	X			1.5
2	X	X	X			1.5
3				X	X	3
4					X	.5
5				X		.5

Table 1. An example where degree is not monotone in the size of the skill set. 5 skills are distributed across $5N$ players as shown. In this population, all skills occur with equal frequency, and therefore there are no rare skills. Players 1 and 2 have the largest skill sets. However, player 3 has more links. This demonstrates that a player with a useful combination of skills may receive more links than one with many skills.

This example also highlights another implication of Theorem 2–because a player's degree is a supermodular function of her set of skills, it is not generically possible to assign prices to individual skills in a way that captures a player's degree. This means that examining the supply and demand of single skills in isolation does not capture an individual's value to a community of problem solvers.

Corollary 4. *There need exist no vector of prices, μ, such that $\sum_{a \in A} \mu_a = d(A)$.*

To further emphasize this point, consider the skill distribution shown in Table 2. This distribution is identical to that in Table 1, except that players 4 and 5 have

[29] Note that a player's skill combination must be both rare and *useful*–that is, complementary to the skills of other players. In Table 1 players 4 and 5 both have rare combinations of skills, but the combination they have is not useful for any player, and thus they receive few links.

[30] Examples of this phenomenon are not difficult to find. For example, consider a group conducting a field experiment in a remote mountain area. This type of problem requires a wide range of skills, including the ability to pose questions, design experiments, acquire funding, collect data, and (given the remote location) mountaineering skills. In this context, player 3 represents a lab assistant who also has mountaineering experience.

both gained two skills. However, gaining these skills does not influence the degree of either player. In fact, endowing players 4 and 5 with those skills does not change any part of the degree distribution. Skills a and b have value to players 1 and 2, but not players 3 or 4–clearly, no linear weighting of the individual skills could produce that pattern.

	a	b	c	d	e	in-degree
1	X	X	X			1.5
2	X	X	X			1.5
3				X	X	3
4	X	X			X	.5
5		X	X	X		.5

Table 2. Consider the previous example, pictured in Table 1. Now, suppose players 4 and 5 were endowed with two extra skills, as shown here. Neither player's degree is affected by this change, because their combinations of skills are not useful to any of the players in the game. This example illustrates that we cannot value a player's skill set by determining the value of her skills individually.

Two characteristics of problem solving production contribute to this result. First, skills are bundled within a person. Thus, in evaluating a collaborator, that person's combination of skills must be considered as a unit. Second, players share their payoffs from solving the problem among their collaborators, and thus have an incentive to minimize the number of collaborators they work with. Together, these two factors mean that a player's as a collaborator may be more than the sum of her individual skills. This non-linearity has implications for the distribution of degree in the population as a whole. Sections 4 and 5 consider those implications in further detail.

4. Skill Distributions and the Distribution of Prominence: The Bernoulli Skills Model

Collaboration network structure governs a number of important interactions, including the spread of information, control of reputation, and even the process of finding jobs. Therefore, we would like to obtain a better understanding of the topology of these networks. In the previous section, I looked at a local measure of network topology–degree. In this section, I take a step back, and consider how the distribution of skills in the population affects the structure of the collaboration network as a whole. I use a special case where skills are independently and identically distributed to show that even when skills are distributed evenly across the population, the degree distribution is highly skewed–that is, a very small number of players help to solve the vast majority of the population's problems, while a large number of players solve very few. I then examine how the degree distribution is affected by changes in the problems faced and the distribution of skills in the population. I show that as problems become more difficult for a population, the degree distribution becomes increasingly unequal–in other words, superstars emerge.

4.1. **The Bernoulli Skills Model.** In this section, I consider a special case in which skills are distributed independently, with equal probability–that is, $Prob\left(a_i \in A | a_j \in A\right) = Prob\left(a_i \in A\right) = p \,\forall\, i \neq j \in S$. I call this the Bernoulli Skills model because each player's skill set can be thought of as the result of a set of M Bernoulli trials, each with probability p of success. This means that the distribution of skill set sizes in the population is binomial, implying that the fraction of the players who have a *particular* set of k skills is $\Psi\left(A\right) = p^k \left(1-p\right)^{M-k}$, and the fraction having *any* k skills is $\binom{M}{k} p^k \left(1-p\right)^{M-k}$.

This special case has several characteristics which make it interesting. First, because skills are completely uncorrelated and occur with the same frequency, all players with the same number of skills will have the same degree, in expectation. This construction enables a clear picture of the effects of supermodularity on the degree distribution of the network. Second, because this model has only 2 parameters–M and p–I can use this model to illustrate how the distribution of skills in the population affects the structure of the collaboration network.

4.2. **Degree Distribution of the Bernoulli Skills Model.** Let Δ denote the distribution of expected degree. That is, $\Delta\left(d\right)$ is the fraction of players who have expected degree d, where the expectation is taken over all $g \in \Gamma\left(\Psi\right)$.[31] In this particular case, I will use a convenient shorthand: $\Delta_{M,p}$ represents the distribution of expected degree when M skills are independently distributed with probability p.

Theorem 5 states the closed form expression for the degree of a player in the Bernoulli Skills Model.

Theorem 5. *Suppose players face a problem, $\omega\left(S\right)$, requiring M skills. If the skills are distributed independently with $Prob\left(a\right) = p \,\forall\, a \in S$, then the expected degree of a player with k skills is*

$$E\left[d(k)\right] = p^M \left[\left(\frac{1-p+p^2}{p^2}\right)^k - 1\right]$$

Proof. Since $\Sigma\left(\Psi\right) = 2^S$, every player needs to make only one link. Thus, we can write $E\left[d\left(A\right)\right] = \sum_{C \subseteq A} \frac{\delta(C)}{\sigma(C)}$, where $\delta\left(C\right)$ is the fraction of players who need skill set C and $\sigma\left(C\right)$ is the fraction of players who can provide skill set C. Since the skills are independent, we can separate this sum according to the size of the skill set required. If we start with the players who are lacking exactly one skill in A_i and end with players needing all of the skills, we obtain the following sum:

$$\begin{aligned} E\left[d\left(A\right)\right] &= \sum_{i \in A} \frac{p^{M-1}\left(1-p\right)}{p} + \sum_{i,j \in A} \frac{p^{M-2}\left(1-p\right)}{p^2} + ... + \frac{p^{M-k}\left(1-p\right)^k}{p^k} \\ &= \sum_{i=1}^{k} \binom{k}{i} \frac{p^{M-i}\left(1-p\right)^i}{p^i} \\ &= p^M \left[\left(\frac{1-p+p^2}{p^2}\right)^k - 1\right] \end{aligned}$$

[31] Alternatively, we might plot the distribution of degree across all networks $g \in \Gamma\left(\Psi\right)$. That is, we could set $\Delta\left(d\right) = \sum_{g \in \Gamma(\Psi)} \delta_g\left(d\right)$ where $\delta_g\left(d\right)$ is the fraction of players in network g with degree d. This choice does not affect the results.

□

Note that when skills are independent, a player's degree depends only on the size of the player's skill set, k.[32] Therefore, in this particular case, is appropriate to interpret the size of a player's skill set as her "ability"–something that we cannot do in the more general case (recall Table 1 in the previous section). This suggests a corollary to Theorem 5.

Corollary 6. *Suppose players face a problem, $\omega(S)$, requiring M skills. If the skills are distributed independently with $Prob\,(a) = p\,\forall\, a \in S$, then degree is strictly increasing in the size of the player's skill set.*

However, we still cannot price the skills individually in such a way that we capture degree, despite the fact that skills are independently distributed. Theorem 7 formalizes this statement.

Theorem 7. *Suppose the players face a problem, $\omega(S)$, requiring M skills . If the skills are distributed independently with $Prob\,(a) = p\,\forall\, a \in S$, then there exists no vector of prices, μ, such that $\sum_{a\in A}\mu_a = d\,(A)$ for all $A \subseteq S$.*

Proof. Any such vector would be required to set $\mu_a = d\,(a) = p^{M-2}(1-p)$ for all $a \in S$. But that would imply that $d\,(A) = kp^{M-2}(1-p)$ for $|A| = k$. This is clearly not true for $k > 1$. □

This also means that there is no way to price the individual skills such that a player's utility is the sum of the prices of her individual skills (Theorem 8)

Theorem 8. *Suppose the players face a problem, $\omega(S)$, requiring M skills . If the skills are distributed independently with $Prob\,(a) = p\,\forall\, a \in S$, then there exists no vector of prices, μ, such that a player can recover his utility, that is, there is exists no price vector, μ such that $\sum_{a\in A_i}\mu_a = u_i\,(A_i)$ for all $A_i \subseteq S$.*

Theorem 5 implies that despite the fact that ability is binomial, the degree distribution for the Bernoulli skills model is highly skewed–a few players have a disproportionately large number of links, while the majority of players have no links at all. As an illustration, consider a Bernoulli skills model with $M = 3$ and $p = \frac{1}{3}$. Table 3 lists the expected degree of every type of player in this case. Although the distribution of ability is binomial, the distribution of links is highly skewed towards those with more skills. For example, although the players with $\{a, b, c\}$ comprise less than 4% of the population and hold only 11% of the total skills, they help to solve nearly 50% of the problems. The vast majority of the players solve only their own problem, helping no other players at all.

[32] In the Bernoulli skills model, skills are uncorrelated and occur with equal frequency. Therefore, there is no statistical difference between two sets of skills of the same size, and degree depends only on the number of skills the player has, rather than the exact set of skills.

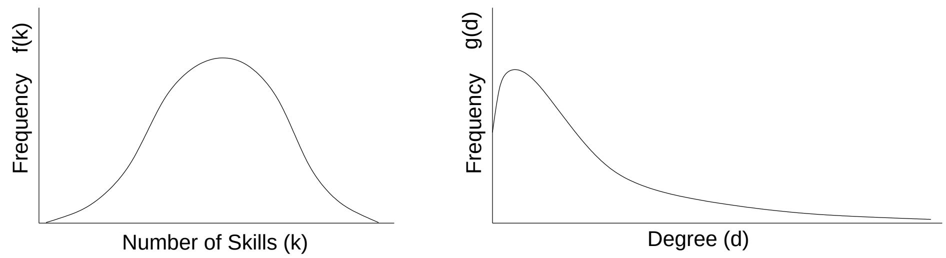

FIGURE 4.1. In the Bernoulli Skills Model, a player's degree is increasing in the size of her set of skills. The supermodularity of degree means that players with more skills receive many more links. This exaggerates any initial inequalities in the size of the skill sets. The resulting network has a very distinctive structure–a small number of players participate in a majority of the collaborations.

k	fraction of players	fraction of skills	links per player	fraction of links
3	1/27	1/9	≈ 12.7	$\approx .49$
2	2/27	4/9	≈ 1.8	$\approx .41$
1	4/27	4/9	$\approx .2$	$\approx .10$
0	8/27	0	0	0

TABLE 3. An illustration of the non-linearity of degree in a Bernoulli Skills model with $M = 3$ and $p = \frac{1}{3}$.

The fact that the degree distribution is skewed is surprising, since the distribution of skill set size (ability) is symmetrical in the Bernoulli Skills model. Figure 4.1 shows both distributions for a representative set of parameters. This example illustrates that even when the distribution of skills across players is symmetrical, the distribution of links is not–because degree is supermodular, small differences in skill set size are magnified in the degree distribution, making the distribution of links highly uneven. This structure is similar to that which we observe in empirical collaboration networks. Figure 4.2 illustrates an example: a network science coauthorship network, in which two scientists are linked if they have coauthored a paper together. The degree distribution of this network is remarkably similar to that of a cost-minimizing collaboration network such as the one pictured in Figure 4.1.

This highly centralized network structure, in which a small number of players participate in most of the collaborations, has implications for other behaviors that take place over collaborative ties. For example, suppose players use collaborative links to share information about jobs, new technologies, and open areas of study. In a network with a skewed degree distribution, most players are connected to a few, high-degree hubs. Thus, the average distance between two nodes in a collaboration network is shorter than we would find in a random network.

A player's degree in the collaboration network reflects their importance to the collaborative community, and therefore the degree distribution also reflects the distribution of influence. In networks with a highly skewed distribution, the high

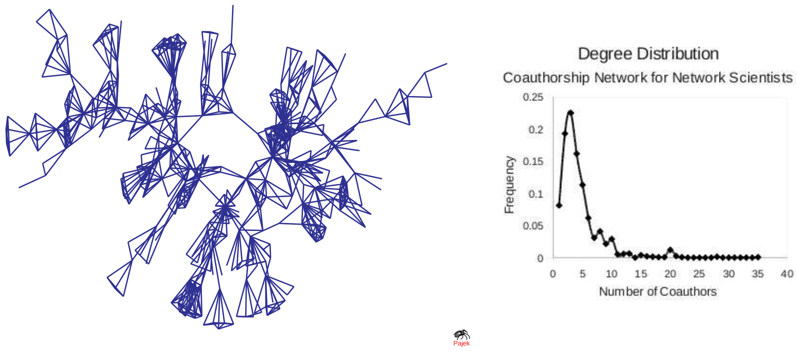

FIGURE 4.2. Empirically, most collaboration networks have a skewed degree distribution. The left panel depicts a coauthorship network for network scientists–two nodes in this network are connected if the scientists coauthored a paper together (only the largest connected component is shown here). The right panel depicts the degree distribution for this network.

degree players have a more significant impact on opinion formation, reputation, and the decisions over new technology than they would in a network with a more equitable distribution of links. In the following section, I look at the determinants of network structure more carefully–by looking at the comparative statics on the Bernoulli skills model, I examine how changing the distribution of skills in the population affects the degree distribution of the resulting collaboration network.

4.3. **Problem Difficulty and the Distribution of Degree.** In the following, will use the Gini coefficient as my measure of distributional equality. The Gini coefficient measures the area between the Lorenz curve of a distribution (in this case, the distribution of expected degree), and the line of equality. In the case of a discrete distribution with values $y_0 ... y_N$ where $y_i < y_{i+1}$, the Lorenz curve is a piecewise function connecting points (F_i, D_i) where $F_i = \sum_{k=0}^{i} \Delta(y_k)$ is the fraction of players with strictly less than y_i links, and $D_i = \frac{\sum_{k=0}^{i} \Delta(y_k) y_k}{\sum_{k=0}^{N} \Delta(y_k) y_k}$ is the fraction of the total number of links held by those players. See Figure 4.3 for an example. The gini coefficient for a discrete distribution is given by $G = 1 - \sum_{i=1}^{N} D_i (F_i - F_{i-1})$. Lower values of the gini coefficient indicate a more equal distribution of links across players, and higher values indicate a more skewed distribution of links. The coefficient is which is 0 when the distribution is perfectly equal (ie: the bottom $x\%$ of the population holds exactly $x\%$ of the links) and 1 when all of the links are held by a single player.

Theorem 9 presents a comparative static on the number of skills required to solve the problem. It shows that the degree distribution becomes more uneven when problems require more skills $(M \uparrow)$or the individual skills are less common $(p \downarrow)$.

Theorem 9. *Suppose the players face a problem, $\omega(S)$, requiring M skills. If the skills are distributed independently with $Prob(a) = p \,\forall\, a \in S$, then the gini coefficient of $\Delta_{M,p}$ (the degree distribution of the resulting network) is increasing*

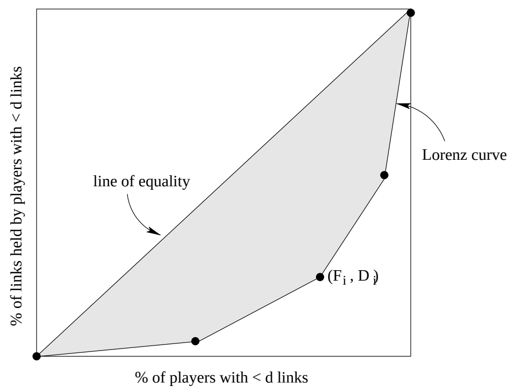

FIGURE 4.3. An example of the Gini coefficient for a discrete distribution, $\Delta(y)$. In this case, the random variable y takes on one of five values, $y_0 ... y_4$. The Gini coefficient is the area of the shaded region between the line of equality and the Lorenz curve.

in M. That is, the distribution of links in the collaboration network is more uneven when the problem being solved requires more skills.

Proof. Using the fact that $\Delta_{M,p}(d(k)) = \binom{M}{k} p^k (1-p)^{M-k}$ for $k \in \{0, 1, 2, ...M\}$, the gini coefficient of the degree distribution $\Delta_{M,p}$ is

$$G = 1 - \frac{\left(1-p^M\right)(1-p)}{p^M \left[(1-p+p^2) \left[\left(\frac{1-p+p^2}{p^2} \right)^M - 1 \right] - M(1-p) \right]}$$

It can be shown that $\frac{\delta G}{\delta M} > 0$, meaning that the degree distribution becomes more uneven as the number of skills required for the problem increases (holding the probability of having a skill constant). □

Figure 4.4 illustrates this comparative static graphically.

Together, the two parameters of the Bernoulli skills model reflect the difficulty of the problem the population faces–a problem is difficult if it requires many skills, or if the average individual has only a few of them. Theorem 9 can therefore be interpreted as a comparative static on problem difficulty–the model predicts that as problems become increasingly difficult, a few "superstars" emerge, who participate in most of the collaborations and help solve a disproportionate number of problems.

5. Skill Ladders: Specialization and Degree

In the previous section, I considered a special case in which skills are entirely uncorrelated. Although there are cases where problem-solving skills are essentially uncorrelated, we would also like to understand the impact of correlations between skills. In this section, I consider a case where skills are divided into disciplines, and the skills within a discipline build on one another, much calculus builds on algebra and algebra builds on arithmetic. I show that when skills are correlated in this way, the degree distribution of the collaboration network becomes even more unequal than when skills are uncorrelated. This suggests that as fields become

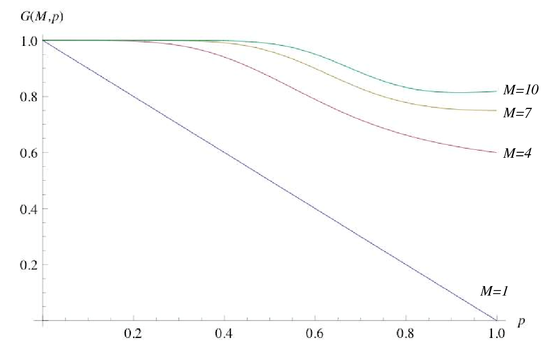

FIGURE 4.4. A graph of the gini coefficient, $G(p, M)$, for different populations. The four curves pictured represent different values of M, and the x axis represents the value of p. The difficulty of a problem rises as the number of skills required increases or the probability of having a particular skill falls.

increasingly specialized, a very small number of players will tend to dominate the collaboration network.

5.1. **Notation and Definitions.** Some additional notation is needed to formalize this concept of specialization. I will define a *ladder* to be an ordered set of skills, $L = \{a_1, a_2, a_3...a_l\} \subseteq S$, such that any player who has the i^{th} skill in the set must have all of the skills that precede it in the set.[33]

Definition 10. A *ladder* is an ordered set of skills, $L = \{a_1, a_2, a_3...a_l\} \subseteq S$, such that $Prob\,(\text{have } a_i | \text{have } a_{i+1}) = 1$. An example of a single ladder with 6 skills is shown in Figure 5.1.

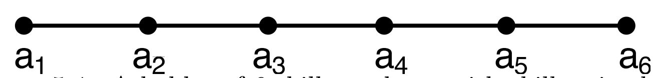

FIGURE 5.1. A ladder of 6 skills–a player with skill a_i in this set must have all of the skills that precede it: $a_1...a_{i-1}$.

Here, I consider a special case where the skills in S are partitioned into m ladders of equal length.[34] The set of all ladders is denoted $\hat{S} = \{L_1...L_m\}$. Figure 5.2 shows an example with 12 skills arranged into four ladders.

[33] Page (2007) introduces this concept of skill ladders, where each skill builds on the one before it.

[34] Obviously, since the length of a ladder is an integer, there will only be equal-length ladders if m divides M evenly. To simplify the exposition, I have written the following as if this is true. However, all of the the following results hold if the ladders are equal length up to integer constraints, which allows for cases where m does not divide M evenly.

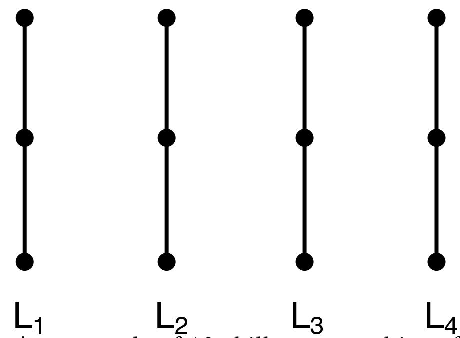

FIGURE 5.2. An example of 12 skills arranged into four ladders of equal length.

I will call a player who has all of the skills in a single ladder an "expert" in that ladder, and I will call the set of ladders that player i is an expert in $\hat{A}_i \subseteq \hat{S}$.

Definition 11. A player is an *expert* in a ladder L_k if she possesses all of the skills in that ladder.

One additional assumption will allow us to compare the results in this section to the results of the Bernoulli skills model. Assume that the conditional probability of having the next skill in a ladder is the same for all skills–that is, $Prob\,(\text{have } a_i | \text{ have } a_{i-1}) = p$ for all a_i.[35] The probability of being an expert in a ladder of l skills is then p^l.

The number of ladders, m, will be our measure of how specialized the problem-solving skills are. Thanks to the above assumption, the case where $m = M$ corresponds to the Bernoulli skills model. On the other extreme, $m = 1$, and all of the skills are arranged in a single ladder. Before considering ladders of arbitrary length, I will first look at this case, where $m = 1$.

5.2. **Example: a single ladder of skills.** Suppose the skills in the set S comprise a single ladder of length M. Because $Prob\,(\text{have } a_i | \text{have } a_{i+1}) = 1\ \forall i \in S$, a player's skill set can be represented by the number of skills she has ($|A_i| = k$ implies $A_i = \{a_1, a_2 ... a_k\}$) .

The linking behavior in this case is very simple. The only players who have skill a_M are those who also have skills $a_1 ... a_{M-1}$. All of the players who don't have all M skills link to one of the players who does. The resulting collaboration network is a set of isolated stars, each with $\frac{1-p^M}{p^M}$ links, on average.

Figure 5.3 compares the network structure in the case with one skill ladder ($m = 1$) to the network structure in the Bernoulli skills model, where skills are independent ($m = M$). The two networks have the same number of skills and players, and players have the same probability of having an additional skill. This means that the probability of having all of the skills required to solve the problem

[35]In other words, I assume that putting a skill at the end of a ladder doesn't change the essential difficulty of obtaining that skill. We could imagine cases where putting a skill at the top of a ladder would make the skill easier to obtain (eg: because it builds on previous experience). We could also imagine a case where skills at the top of the ladder are more difficult to obtain (eg: because they are more demanding than the skills that came before). These would both make interesting extensions.

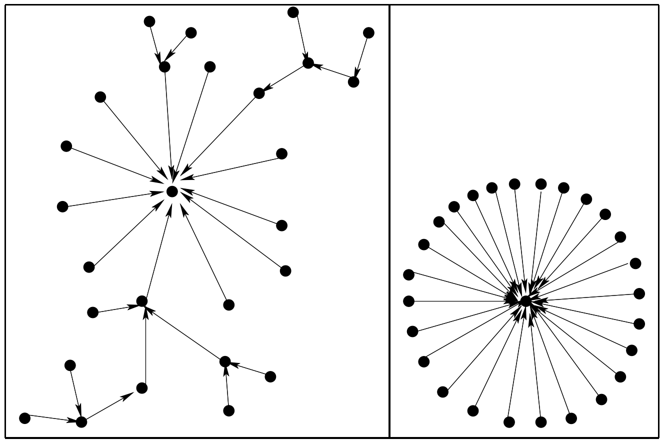

FIGURE 5.3. Two collaboration networks with 27 players. In both cases, the players are solving a problem requiring $M = 3$ skills. These networks represent the two ends of a spectrum of skill specialization, with $m = M$ ladders on the left and $m = 1$ ladders on the right.

is the same in both networks. Moreover, in both, exactly one player has all of the skills required. However, the two networks have a much different structure.

5.3. **Results for m ladders.** Using insight gained from this example, I can derive a more general result. Suppose the skills are arranged in m equal-length ladders.[36] As in the previous example, the only players with all of the skills in a ladder are those who are experts in that field. Thus, a model with m ladders reduces to a Bernoulli skills model with m independent skills. Theorem 12 presents a closed-form expression for a player's degree in the case with m skill ladders.

Theorem 12. *If Ψ is a distribution of skills such that $\hat{S} = \{L_1 ... L_m\}$ is a partition of S into m equal-length ladders with $Prob\left(have\ a_i^j |\ have\ a_{(i-1)}^j\right) = p$, then a player with the skill set A will have expected degree $Ed(A) = p^M \left[\left(\frac{1-p^{\frac{M}{m}}+p^{2\frac{M}{m}}}{p^{2\frac{M}{m}}} \right)^k - 1 \right]$, where k is the number of disciplines the player is an expert in.*

Proof. The ladders are of equal length, so the length of a single ladder is $\frac{M}{m}$, and the probability that a player is an expert in any one ladder is $p^{\frac{M}{m}}$. A player receives a link only if she is an expert in a field. Define a new set of skills that correspond to the set of ladders: $\hat{S} = \{L_1 ... L_m\}$. The player's new skill set is $\hat{A}_i$, where $L_k \in \hat{A}_i$ if she is an expert in ladder L_k. Each of these new skills has a probability equal to the probability of being an expert in that field, so define $\hat{p} = p^{M/m}$. The probability of being an expert in a particular ladder is independent of the probability of being an expert in any other ladder, so this problem reduces to one with m independent skills, with probability $p^{\frac{M}{m}}$. The result then is a simple extension of Theorem 5. □

[36]Again, the results are the same if the ladders are equal length to integer constraints.

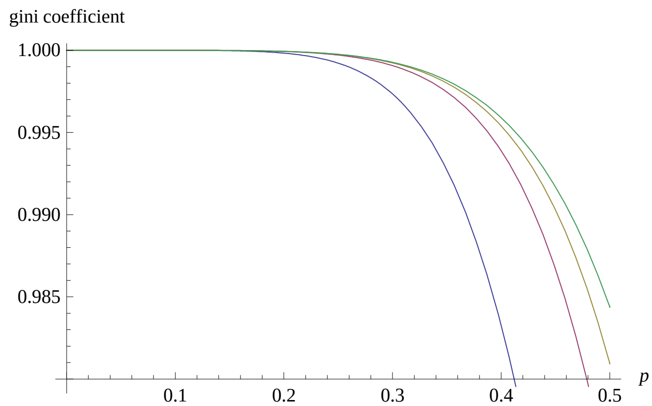

FIGURE 5.4. The gini coefficient for $M = 10$, divided into different sets of ladders. The bottom curve pictures the case where $m = 10$, the second curve pictures the case where $m = 5$, the third shows $m = 2$, and the fourth shows $m = 1$.

We can now do a comparative static on the number of skill ladders, to see how the number of "disciplines" affects the structure of the collaboration network. Theorem 13 indicates that as the skills are condensed into fewer and fewer disciplines, the collaboration network becomes increasingly skewed.

Theorem 13. *Suppose S skills are arranged in m ladders of equal length, with constant conditional probability $Prob\left(have\ a_i^j |\ have\ a_{i-1}^j\right) = p$ and $Prob\left(have\ a_1^j\right) = p\ \forall j = 1...m$. The gini coefficient of the resulting network is decreasing in the number of ladders, m. That is, when there are fewer skill ladders, the degree distribution becomes increasingly uneven.*

Proof. Recall the gini coefficient for the Bernoulli Skills model, presented in the proof for Theorem 9. A model with m skill ladders is equivalent to a Bernoulli Skills model with m skills and $p = p^{M/m}$. Substituting into the previous equation, we obtain the following:

$$G = 1 - \frac{(1-p^m)\left(1-p^{\frac{M}{m}}\right)}{p^m\left[\left(1-p^{\frac{M}{m}}+p^{2\frac{M}{m}}\right)\left[\left(\frac{1-p^{\frac{M}{m}}+p^{2\frac{M}{m}}}{p^{2\frac{M}{m}}}\right)^m-1\right]-m\left(1-p^{\frac{M}{m}}\right)\right]}$$

which is decreasing in the number of ladders, m. Figure 5.4 shows how the gini coefficient depends on the number of ladders for the case where $M = 10$.

□

Skills often build on one another because fields are specialized. The result in Theorem 13 can thus be interpreted in terms of specialization and reliance on experts–as skills become increasingly specialized, we would predict that the degree distribution would become increasingly unequal. This is because when skills are specialized, most players are not useful collaborators. As a result, most problems are solved by a few, high-degree experts. On the other hand, when skills are distributed independently, most players are capable of being useful to someone, and thus the networks will tend to have a much more even distribution of links.

A	$\Psi(A)$	$E[d(A)]$
$\emptyset$	$\frac{10}{36}$	0
a	$\frac{10}{36}$	$\frac{1}{18}$
b	$\frac{5}{36}$	$\frac{1}{6}$
c	$\frac{2}{36}$	$\frac{5}{6}$
ab	$\frac{5}{36}$	$\frac{5}{9}$
ac	$\frac{2}{36}$	$2\frac{5}{9}$
bc	$\frac{1}{36}$	6
abc	$\frac{1}{36}$	$18\frac{1}{18}$

Table 4. In this example, the problem requires 3 skills: $S = \{a, b, c\}$. The skills are distributed independently with $Prob\,(have\,a) = \frac{1}{2}$, $Prob\,(have\,b) = \frac{1}{3}$, and $Prob\,(have\,c) = \frac{1}{6}$. This table shows the frequency of each skill set, and the expected degree of an individual with those skills.

6. A Comparison to a Model with a Single Skill

One of the major contributions of the model presented above is its detailed treatment of problem-solving skills. Rather than having a single specialty or "type", players in this model have sets of skills, which may overlap. These two models allow for different levels of analysis. In a "type-based" model, a player's skills are represented by a single unit–her type. In contrast, a "skill-based" model represents a player's skills individually, allowing players' skill sets to overlap and interact in complex ways. This finer-grained approach has significant impacts on outcomes. In this section, I examine those impacts by directly comparing a type-based model of problem-solving to a skill-based model.

In the previous sections, I showed that when players have sets of skills, the relationship between a player's skills and her value as a problem solver is often highly non-linear. In particular, an individual's combination of skills may be valuable, even if her skills are not valuable individually. Since this nonlinearity of degree is driven by the desirability of combinations of skills, we would expect to regain the linear relationship between skills and degree in the special case where players' skill sets do not overlap–that is, where they have a type or specialty.

Consider the following pair of examples. First, suppose a problem requires 3 skills, $S = \{a, b, c\}$, and all three skills are distributed independently to each player–that is, $Prob\,(have\,skill\,i\,|\,have\,skill\,j) = Prob\,(have\,skill\,i)\,\forall i \neq j \in S$. This means that and the probability of having skill set A is $\Psi(A) = \prod_{i \in A} p_i$. Let $p_a = Prob\,(have\,skill\,a) = \frac{1}{2}$, $p_b = Prob\,(have\,skill\,b) = \frac{1}{3}$, and $p_c = Prob\,(have\,skill\,c) = \frac{1}{6}$. A player in this population will have, on average, exactly one skill. Table 4 shows the expected degree for all 8 possible combinations of skills. Note that we cannot price individual skills in such a way that it characterizes a player's degree in the network. To see this, suppose that such a pricing scheme existed–that is, suppose $\exists \mu = [\mu_a, \mu_b, \mu_c]$ such that $\sum_{i \in A} \mu_i = Ed\,(A)$. Then clearly $\mu\,(i) = Ed\,(i)$ for $i \in S$. But that price scheme would predict that $Ed\,(\{b, c\}) = 1$, whereas the actual degree of a player with those skills is $Ed\,(\{b, c\}) = 6$.

Now, consider a modification of this example–suppose that instead of skills being distributed independently, we assume that each player has exactly one skill, with

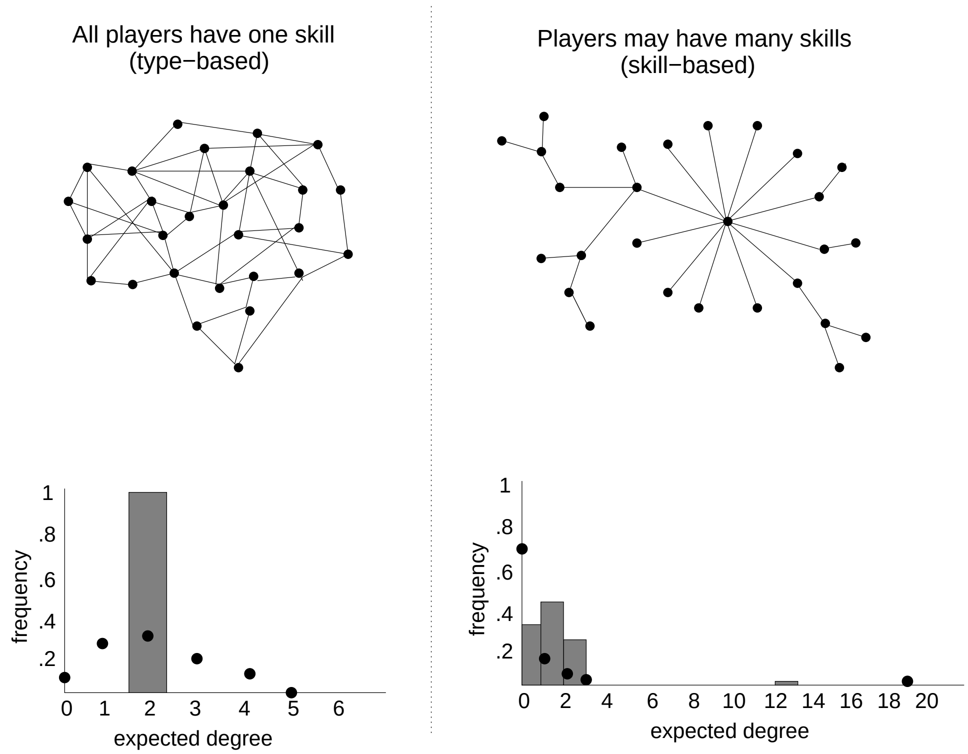

FIGURE 6.1. Contrasting the structure of collaboration networks resulting from type-based and skill-based models of collaboration. In both cases, the problem faced requires three skills. In the network on the left, each player has a single skill. In the network on the right, the three skills are distributed independently with probability $p = \frac{1}{3}$. The bottom panels show both the distribution of degree in the pictured networks (black dots) and the distribution of expected degree in the set of cost minimizing networks (grey bars).

$\Psi(a) = \frac{1}{2}, \Psi(b) = \frac{1}{3}, \Psi(c) = \frac{1}{6}$. The population in this example shares many characteristics with the previous population–the frequency of each skill in the population remain the same, and in both cases, the average player holds one skill. However, in contrast with the previous example, skills in this population can be priced individually. If a player has skill i, her expected degree in the cost minimizing collaboration network is $\frac{1-\Psi(i)}{\Psi(i)}$. If we set the prices of skills a, b, and c to be $\mu_a = 1$, $\mu_b = 2$, and $\mu_c = 5$, then each player's degree is a linear function of her endowments, weighted by the prices.

The structure of the collaboration network is also considerably different when players' skill sets can overlap. For example, consider the two networks and associated degree distributions pictured in Figure 6.1. In both cases, the players are solving a problem requiring three skills: $S = \{a, b, c\}$. In the network on the left, each player has exactly one skill, and all three skills occur with equal probability

$\left(\Psi(a) = \Psi(b) = \Psi(c) = \frac{1}{3}\right)$. In the network on the right, the skills are distributed independently, as in the Bernoulli skills model, with $Prob\,(have\,skill\,i) = \frac{1}{3}$ for $i = a, b, c$. In both of these networks, the average player has one skill. However, the degree distribution of the collaboration network is much different in the two cases. Every player in the left hand network has, on average, $\frac{1-\frac{1}{3}}{\frac{1}{3}} = 2$ links, and the distribution of links in a typical cost-minimizing collaboration network is symmetric around that value. In the left hand network, the network structure is much different. Because combinations of skills are valuable, players with more skills help to solve a disproportionate number of problems. The initial inequalities in the distribution of skills in the population are magnified, resulting in a network of interconnected stars.

This pair of examples illustrates two points. First, the type-based model is a special case of the skill-based model–in particular, it is the case where each player has exactly one skill. Second, it highlights the value of a more detailed treatment of skills in modeling problem solving. Although in some contexts it is appropriate to assume that each player has a type or specialty, this assumption is not necessarily benign. That modeling choice impacts our predictions about the value of certain individuals in the community, as well as impacting the structure of the collaboration network.

In particular, this comparison highlights the difference between modeling the value of labor in manufacturing production and modeling the value of labor in problem-solving. In manufacturing, players can plausibly be given a type or specialty, and the wages of an individual are a function of their type. As production transitions away from manufacturing towards problem-solving, it becomes more difficult to classify workers according to a type or specialty. Thus, the shift from manufacturing production to problem-solving can be modeled as a shift from a type-based model to a skill-based model. This comparison is particularly valuable because it allows us to draw conclusions and make predictions about the labor market as we transition from manufacturing to knowledge-based industries. The following section explores this in greater detail.

7. Implications: Labor Markets and Industrial Organization

The results of the previous sections have implications on employment, individual welfare, and training of workers in knowledge-based industries, as well as the internal organization of firms within those industries. This section examines some of these implications in greater detail.

7.1. **Extending the Model to Employment by Firms.**

Before going further it is worth taking a detour to look at the applicability of this model outside of collaboration. Although the results of the previous section were framed in terms of collaboration between individuals, it is relatively simple to extend this model to firm/employer labor market interactions. Suppose there are two types of agents in an economy: firms and problem solvers. Firms face problems, which have value if solved. Let δ be a probability measure on the skill sets required to solve the firms' problems–that is, $\delta(A)$ is the probability that a firm faces a problem requiring the skills $A \subseteq S$, with $\sum_{A \subseteq S} \delta(A) = 1$. The firms hire problem-solvers to work on projects, giving them each a share of the proceeds from solving the problem.[37] The expected number of projects a player contributes to can be calculated much the same way as before:

$$E\left[d\left(A\right)\right]=\sum_{C\subseteq A}\frac{\delta\left(C\right)}{\sigma\left(C\right)}$$

The sole difference between this case and that in the previous sections is that now the demand for a particular set of skills is decoupled from the supply of those skills. Since the proof of the results in this paper do not depend on such a coupling, we can obtain a supermodularity result analogous to Theorem 2 for this firm-based case as well. As a result, all of the results from previous sections apply to a case where individuals work for firms, as well as cases where individuals collaborate.

7.2. **Returns to Skills and Optimal Training Decisions.** Although individuals in this model have static skill sets, we can use it to look at the returns to skills, and thus optimal training decisions. Using the result from Theorem 2, Theorem 14 shows that in a static population, returns to acquiring new skills are non-decreasing.

Theorem 14. *If a player's utility is increasing in her degree in the collaboration network, then players experience non-decreasing returns to obtaining additional skills, holding the rest of the population fixed.*

Proof. This is a simple application of supermodularity. Theorem 2 states that $d\left(A_i\cup A_j\right)+d\left(A_i\cap A_j\right)\geq d\left(A_i\right)+d\left(A_j\right)$. Rearranging, this tells us that $d\left(A_i\cup A_j\right)-d\left(A_i\right)\geq d\left(A_j\right)-d\left(A_i\cap A_j\right)$. Both sides of this inequality represent a gain in degree from obtaining the skills in the set $A_j\backslash A_i$. The fact that this gain is greater when the player already has A_i indicates that returns are non-decreasing. □

The implication of this result is that every player will obtain the maximum number of skills possible. Thus, this model predicts that a shift in the economy from manufacturing to knowledge-based production might be accompanied by increasing returns to skills, and thus increased skill acquisition.

However, that is not the whole story–individuals also have to choose which skills to obtain in which order. In a world with many possible skills, and costs to obtaining new skills, how will a worker determine the optimal skills to acquire? I can calculate the contribution of each skill in a player's skill set to the total demand for her skills using a Shapely value decomposition. The demand for player i's skills can generically be written as $d\left(A_i\right)=\sum_{C\subseteq A_i}\frac{\delta(C)}{\sigma(C)}$ where $\delta\left(C\right)=\Psi\left(S\backslash C\right)$[38] and $\sigma\left(C\right)=\sum_{D\subseteq S\backslash C}\Psi\left(C\cup D\right)$. Using this demand as a value function,[39] we can obtain an expression for the Shapely value of a skill, a, to a player, i.

Theorem 15. *The Shapley value for a skill, a, to a player, i, is*

[37] I will assume that problem solvers don't bundle themselves together and offer their services jointly for a single share of the problem-solving proceeds. This is consistant with the previous interpretation of the model.

[38] Note that this is the expression when each player needs exactly one partner to solve their problem. The results are similar for the more general case.

[39] Note that $d\left(.\right)$ satisfies both requirements for a value function: $d\left(\emptyset\right)=0$ and according to Theorem 2, $d\left(A\cup B\right)+d\left(A\cap B\right)\geq d\left(A\right)+d\left(B\right)$, which implies superadditivity.

$$\phi_{a,i}(d) = \sum_{B \subseteq A_i \setminus \{a\}} \frac{1}{\binom{|A_i|}{|B|}} \left(\sum_{C \subseteq B} \frac{\Psi(S \setminus (C \cup a))}{\sum_{D \subseteq S \setminus (C \cup a)} \Psi((C \cup a) \cup D)} \right)$$

Theorem 15 shows that the Shapley value of a skill to an individual depends on 1) the existing skill distribution in the population and 2) the set skills the individual already has. This theorem suggests that optimal training decisions should be highly individualized, because the value of a skill depends on problem solver's existing skill set.

7.3. **Variation in Labor Demand.** The number of collaborators an individual has (her degree in the collaboration network) can be interpreted as the demand for her skills as a collaborator. Empirically, it has been observed that output is highly concentrated among a small number of people, creating an extremely skewed distribution of demand. For example, Rosen (1981) observes that in certain creative fields (eg: music, film, textbook writing), a large fraction of demand goes to an extremely small number of producers. Uzi and Spiro (2005) observe a similar patter among the directors, producers, and other creative artists on Broadway. Similarly, data on academic collaborations suggests that a small fraction of researchers are responsible for the majority of output (see Newman (2001), Moody (2004), Acedo et al (2006), and Goyal et al (2006)).

This long-tailed distribution has implications for the distribution of wages and welfare (see the next section for further discussion) and thus there has been considerable interest in understanding why such a concentration in labor demand occurs. Some existing models (for example, Rosen (1981)) can induce a long-tailed distribution when there is a high premium on quality, and production technology decouples effort from output quantity (eg: in creative industries, where a single performance or album can be enjoyed by many consumers). However, such technologies are not relevant in knowledge-based industries, where effort is not decoupled from output volume.

The model presented in this paper induces a long-tailed distribution of demand in cases where collaboration is important. As noted above, the search for complementarities and the bundling of skills within individuals exaggerates existing inequalities in skill distributions, creating a long-tailed distribution of demand. Thus, this model can explain uneven output demand when effort is not fixed, but skills are varied. Moreover, the model makes predictions about the comparative statics of the distribution of output. Specifically, the distribution of demand will become more skewed as problems become more difficult. We can observe this trend in longitudinal studies of coauthorship, such as Moody (2004) and Grossman and Ion (2002). These studies observe the number of collaborations maintained by authors in the fields of sociology and mathematics, respectively. It is widely believed that the problems faced in these fields have become more difficult over time. The authors show that the upper tail of the collaboration distribution extends over time, indicating that a small number of individuals capture an ever-increasing fraction of the collaborative demand. This model connects those two trends, attributing the attenuation of the demand distribution to the increasing difficulty of the problems being faced.

7.4. **Industrial Organization.** The results that precede also have implications of the organization of knowledge-based firms. Traditional organizational structures were hierarchical, with several layers of supervisors. The theoretical underpinnings of these hierarchical structures are explored in a wide range of models, including Rosen (1982). However, evidence indicates that organizational structure within firms is changing–hierarchical structures are flattening, and workplaces are becoming more decentralized (see, for example, Bresnahan *et al* (2002) and Rajan and Wulf (2006)).

The model presented in this paper provides a more general model of organizational structure, one that allows for these more complex interrelationships between individuals. In particular, this model allows me to answer more subtle questions about how shifting conditions affect organizational structures. In knowledge-based firms, value is created through the creation of new knowledge, rather than the exploitation of existing knowledge.

My model produces strictly hierarchical structures in one special case–that in which skills are arranged in a single ladder (see section 5.2 for the details of this case). This case corresponds to a model in which ability is measured on a one-dimensional scale. One dimensional measures of ability make sense in industries that create value by exploiting an existing bank of knowledge. However, as demand for workers shifts towards industries that create value by creating new knowledge, we expect skills to be arranged in ladders less often. My model predicts that as such a shift occurs, organizational structure should move away from hierarchies, towards more complex structures.

8. Extensions

The model I present in this paper suggests a wealth of extensions. In the following, I will briefly discuss a few of the more promising of these extensions.

8.1. **Bargaining Over Surplus.** As mentioned earlier, there are several reasons to consider an even split of payoffs between collaborators–indeed, when payoffs to problem solving are non-monetary, it may be difficult to split returns any other way.[40] However, one might want to consider what happens when individuals can bargain over the surplus from solving a problem. This bargaining process gives each player a "wage" from the collaboration,[41] which I can then use to produce an income distribution, much as we produce an output distribution in this paper.

By producing an income distribution, I can attempt to explain some empirical trends in labor markets. In particular, income inequality has grown dramatically over the past 50 years, a trend that holds even if one controls for years of education and tenure (Juhn et al (1991)). The labor literature draws a connection between this trend towards greater wage inequality and the shift towards knowledge-based production in our economy. They tell a story that is potentially consistent with

[40]It is, for example, difficult to split the authorship of a publication into arbitrarily-sized shares.

[41]It is possible to show two necessary conditions for non-zero wages: first, firms must face some friction in hiring (such as search costs) and second, problem solvers must face constraints on their time.

the model presented in this paper: knowledge-based production creates increasing returns to skills, a trend which benefits high-skilled workers over low-skilled workers, widening the wage gap.[42] More recent work notes a second-order effect of the shift– the wage distribution widens most significantly in the upper tail of the distribution. According to Machin (2008), the wage difference between the 90th percentile and the 50th percentile rose throughout the 80s, 90s, and early 2000s. On the other hand, the wage difference between the 50th percentile and wages in the 10th percentile increased significantly in the 80s, only slightly in the 90s, and actually contracted in the 2000s.

These empirical observations are not inconsistent with the predictions of this model. As noted in Section 6, we can model the shift in production as a shift from a type-based model to a skill-based model. As production becomes more problem-solving based, the upper tail of the output distribution stretches dramatically, while the lower tail contracts, increasing income inequality overall. By explicitly connecting income to the output, I will be able to determine whether these results on the output distribution can be extended to the income distribution.

8.2. **Long Run Skill Acquisition.** Thus far, I have assumed that the skill distribution in the population is fixed. However, in the long run, we would expect problem-solvers to acquire new skills, based on what will optimize their expected outcomes. I have already shown that in a static world, the individual's skill acquisition decisions will depend on the current state of the world (distribution of skills in the population) and the set of skills the individual possesses. It would be interesting to look at the long-run steady state population of problem solvers that results from this dynamic skill acquisition process.

One might ask a number of question about this steady state population. Is the equilibrium distribution of skills in the population efficient? Do individual players acquire too many skills, from a societal perspective? I can also use this model to ask questions about the equilibrium levels of specialization and generalization in the problem solving population. Under what conditions should we see individuals specializing in a set of related skills? In a companion paper, I explore whether specialists and generalists can coexist in the same equilibrium population. It would also be interesting to see whether, in the extreme long run, disciplines (groups of individuals with related, specialized skills) emerge. Finally, if the distribution of problems faced changes over time, then the steady state population may be subject to shocks in the types of problems faced. A long run model with skill acquisition can answer questions about the robustness of the populations to these shocks.

9. Conclusion

In this paper, I present a model of problem-solving and collaboration. I use this model to look at the demand for a problem-solver's skills as a collaborator. I show that because a problem solver's skills are used as a unit, the number of problem she solves is a supermodular function of her set of skills–in other words, an individual's value as a problem solver is more than the sum of her individual skills.

[42]See, for example, Juhn et al (1991), who find a trend towards increased wage inequality, and attribute that gap to increasing returns to skills (which they specifically differentiate from years of education or tenure). Other, related papers include Kruger (1993) and Berman *et al* (1998). More recently, Machin (2008) refines this notion, pointing out that technological advances in the 80s eliminated many routine tasks, accelerating the increasing returns to skills.

Each additional skill multiplies the number of skill combinations that the player can use, unlocking many potential synergies with other problem solvers. Moreover, a player who has a particularly useful combination of skills may participate in more collaborations than a player who has many rare skills, but does not fill such a hole in the organization. The fact that players collaborate with those who have complementary skills also has an effect on the structure of the collaboration network as a whole–the model predicts that the degree distribution of a collaboration network will be skewed, even if skills are distributed independently across players. In other words, the model predicts that a few players will solve the majority of the problems a population faces, even if the distribution of skills in the population is symmetrical. Finally, this model connects the nature of problems and problem solvers to the structure of the collaboration network–as problems become more difficult for a population, the model predicts that the collaboration network become more centralized around a few, high degree hubs. In sum, this framework, in which problem solvers have skill sets and collaborate to solve problems, appears to be sufficiently flexible to address many questions about the returns to skills in knowledge production.

References

[1] Acedo, Francisco Jose, Barroso, Carmen, Casanueva, Cristobal, and Jose Luis Galan (2006). "Co-authorship in Management and Organizational Studies: An Empirical and Network Analysis," *Journal of Management Studies 43 (5),* pp. 957-983.

[2] Berman , Eli, John Bound, and Stephen Machin (1998). "Implications of Skill-biased Technological Change: International Evidence," *Quarterly Journal of Economics, 133 (4),* pp. 1245-1279.

[3] Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt (2002). "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-level Evidence," *The Quarterly Journal of Economics, 117 (1),* pp. 339-376.

[4] DeGroot, M.H. 1974. "Reaching a consensus," *Journal of the American Statistical Association 69,* pp. 118-121.

[5] DeMarzo, P., D. Vayanos & J. Zwiebel. 2003. "Persuasion Bias, Social Influence, and Unidimensional Opinions," *Quarterly Journal of Economics, 118, pp. 909-968.*

[6] Calvo-Armengol, Antoni (2004). "Job Contract Networks," *Journal of Economic Theory, 115 (1),* pp. 191-206.

[7] Calvo-Armengol, Antoni and Matthew O. Jackson (2004) "Social Networks in Determining Employment: Patterns, Dynamics, and Inequality," *American Economic Review*, *94 (3)*, pp. 426-454.

[8] Coleman, J.S., E. Katz, & H. Menzel. 1966. *Medical Innovation: A Diffusion Study,* Indianapolis, Ind: Bobbs-Mirrill

[9] Fafchamps, Marcel, van der Leij, Marco J., and Sanjeev Goyal (2006). "Scientific Networks and Coauthorship," *The University of Oxford, Department of Economics Discussion Paper Series (256).*

[10] Gleiser, Pablo M. and Leon Danon (2003). "Community Structure in Jazz," *Advances in Complex Systems, 16 (4),* pp 565-573.

[11] Golub, Benjamin and Mathew Jackson (2007). "Naive Learning in Social Networks: Convergence, Influence and the Wisdom of Crowds," WP.

[12] Goyal, Sanjeev, van der Leij, Marco J., and Jose Luis Moraga-Gonzalez (2006) "Economics: An Emerging Small World," *Journal of Political Economy, 144 (2),* pp 403-412.

[13] Granovetter, M. 1973. "The Strength of Weak Ties," *American Journal of Sociology, 78,* pp. 1360-1380.

[14] Granovetter, M. 1995. *Getting a Job: A Study of Contacts and Careers,* 2nd ed, Chicago: University of Chicago Press.

[15] Guimera, Roger, Brian Uzzi, Jarrett Spiro, and Luis A. Nunes Amaral (2005), "Team Assembly Mechanisms Determine Collaboration Network Structure and Team Performance," *Science, 308,* pp 697-702.

[16] Hagel, John III, John Seely Brown, and Lang Davidson (2009). *Measuring the forces of long-term change: The 2009 Shift Index,* Deloitte Center for the Edge.

[17] Hagerstrand, T. 1967. *Innovation Diffusion as a Spatial Process,* Chicago: University of Chicago Press.

[18] Hong, Lu and Scott E. Page (2001). "Problem Solving by Heterogeneous Agents," *Journal of Economic Theory, 97*, pp 123-163.

[19] Hong, Lu and Scott E. Page (2004). "Groups of diverse problem solvers can outperform groups of high-ability problem solvers," *Proceedings of the National Academy of Sciences, 101 (46),* pp. 16385-16389.

[20] Iyer, Bala, Lee, Chi-Hyon and N. Venkatraman, (2006) "Managing in a 'Small World Ecosystem': Lessons from the Software Sector" *California Management Review, 48 (3)*, pp. 28-47.

[21] Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce (1993). "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy, 101 (3),* pp. 410-442.

[22] Jackson, Matthew O. (2004). "A Survey of Models of Network Formation: Stability and Efficiency," Chapter 1 in *Group Formation in Economics; Networks, Clubs and Coalitions,* edited by Gabrielle Demange and Myrna Wooders, Cambridge University Press, Cambridge U.K., 2004.

[23] Jackson, Matthew O. (2008). *Social and Economic Networks.* Princeton University Press: Princeton, NJ.

[24] Jackson, Matthew O. and Asher Wolinsky (1996). "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory, 71 (1)*, pp. 44–74.

[25] Jackson, Matthew O. and Brian Rogers (2007)."Relating Network Structure to Diffusion Properties through Stochastic Dominance," *Advances in Theoretical Economics, 7 (1),* pp. 1-13.

[26] Kruger, Alan B. (1993). "How Computers Have Changed the Wage Structure: Evidence from Microdata 1984-1989," *Quarterly Journal of Economics, 108 (1),* pp. 33-60.

[27] Machin, Stephen (2008). "Big Ideas: Rising wage inequality," *CentrePiece, 13 (2)*, pp. 8-10.

[28] Moody, James (2004). "The Structure of a Social Science Collaboration Network: Disciplinary Cohesion from 1963 to 1999," *American Sociological Review, 69 (2)* pp. 213-238.

[29] Newman, M.E.J., (2001). "Scientific collaboration networks. I. Network construction and fundimental results," *Physical Review E, 64:* 016131.

[30] Newman, M.E,J. 2003. "Properties of Highly Clustered Networks," *Physical Reveiw E, 68 : 026121.*

[31] Phillips, Katherine W., Elizabeth A. Mannix, Margaret A. Neale, and Deborah H. Gruenfeld (2004). "Diverse groups and information sharing: The effects of congruent ties," *Journal of Experimental Social Psychology, 40,* pp. 497-510.

[32] Polzer, Jeffrey T., Laurie P. Milton, and William B. Swann, Jr. (2002). "Capitalizing on Diversity: Interpersonal Congruence in Small Work Groups," *Administrative Science Quarterly, 47 (2)*, pp. 296-324.

[33] Powell, Walter W., Kenneth W. Koput, and Laurel Smith-Doerr (1996). "Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology," *Administrative Science Quarterly, 41 (1),* pp 116-145.

[34] Rajan, Raghuram G. and Julie Wulf (2006). "The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies," *The Review of Economics and Statistics, 88 (4),* pp. 759-773.

[35] Rosen, Sherwin (1981). "The Economics of Superstars," *The American Economic Review, 71 (5),* pp. 845-858.

[36] Rosen, Sherwin (1982). "Authority, Control, and the Distribution of Earnings," *The Bell Journal of Economics, 13 (2),* pp 311-323.

[37] Ryan, B., & N.C. Gross. 1943. "The Diffusion of Hybrid Seed Corn in Two Iowa Communities," *Rural Sociology, 8,* pp.15-24.

[38] Thomas-Hunt, Melissa C., Tonya Y. Ogden, and Margaret A. Neale (2003). "Who's Really Sharing? Effects of Social and Expert Status on Knowledge Exchange Within Groups," *Management Science, 49 (4),* pp. 464-477.

[39] Topa, Giorgio (2000). Social Interactions, Local Spillovers and Unemployment," *Review of Economic Studies, 68 (2),* pp. 261-295.

[40] Uzzi, Brian, Luis AN Amaral, and Felix Reed-Tsochas (2007), "Small-world networks and management science research: a review," *European Management Review, 4,* pp 77-91.

[41] Uzzi, Brian and Jarrett Spiro (2005), "Collaboration and Creativity: The Small World Problem," *American Journal of Sociology, 111 (2),* 447-504.

Appendix A. Pairwise Stability and Efficiency

Briefly, a network is pairwise stable if no individual would prefer to terminate an existing link, and if no pair of individuals would prefer to add a link. Although this definition is usually used in undirected networks, it works equally well in the current context. Formally, a collaboration network, g, is *pairwise stable* if

(1) for all $ij \in g$, $u_i(g) \geq u_i(g - ij)$ and $u_j(g) \geq u_j(g - ij)$
(2) for all $ij \notin g$, if $u_j(g + ij) > u_j(g)$ then $u_i(g + ij) < u_i(g)$

Together, these two conditions ensure that links are mutual. That is, if a network is pairwise stable, then both players agree to maintain the link.

Theorem. *Any cost minimizing collaboration network, $g \in \Gamma(\Psi)$, is pairwise stable. In other words, $\forall ij \in g$ $u_i(g) \geq u_i(g - ij)$ and $u_j(g) \geq u_j(g - ij)$ and for all $ij \notin g$, if $u_j(g + ij) > u_j(g)$ then $u_i(g + ij) < u_i(g)$.*

Proof. Let g be a cost-minimizing collaboration network. First, consider whether any player wishes to unilaterally remove a link, $ij \in g$. Severing this link deprives player j of his share of the payoff from solving $i's$ problem $\left(\frac{1}{|C_i|+1} \geq 0\right)$, and thus he will never choose to terminate one of his incoming links. Since player i chooses a minimal set of collaborators that allowed her to solve the problem, removing a link means that she can no longer solve the problem and no longer receives a payoff. Since her share of the payoff is greater than zero $\left(\frac{1}{|C_i|+1} \geq 0\right)$, this ensures that she will also never choose to severe a link unilaterally. Finally, note that no player will ever want to add an outgoing link to a cost-minimizing collaboration network– because every player has chosen a set of collaborators optimally, any additional link would require her to further split her prize. □

Theorem. *Any cost minimizing collaboration network, $g \in \Gamma(\Psi)$, is strongly efficient. In other words, $\sum_i u_i(g) \geq \sum_i u_i(g') \; \forall g' \in G$.*

Proof. Because all value is generated from solving problems, the maximum possible value in the network is N. Since solving the problem is incentive compatible for every player, and the payoff from solving the problem is split evenly between collaborators, with no loss, the players always extract the maximum value from the network. □

Appendix B. Proof of Theorem 2

Theorem. *For any distribution of skills, Ψ, a player's expected degree over the networks in $\Gamma(\Psi)$ is a supermodular function of her set of skills. That is, $E[d(A \cup B)] + E[d(A \cap B)] \geq E[d(A)] + E[d(B)]$.*

Proof. A player with the set $A \cup B$ will be able to help players needing any subset of those skills. Let $\delta(C)$ be the demand for a particular set of skills, C. In the general case,

$$\delta(C) = \Psi(S \backslash C) + \sum_{D \,:\, \Psi(C \cup D) = 0} \Psi(S \backslash (C \cup D))$$

. The fraction who can supply the set C is $\sigma(C)=\sum_{D\subseteq S\backslash C}\Psi(C\cup D)$. Note that $\delta(C)$ and $\sigma(C)$ depend only on the particulars of the problem (S), the distribution of skills (Ψ), and the subset of skills (C). Thus, any player with the skill set $A\cup B$ has expected degree

$$E\left[d\left(A\cup B\right)\right]=\sum_{C\subseteq A\cup B}\frac{\delta\left(C\right)}{\sigma\left(C\right)}$$

We can divide the problems that a player with $A\cup B$ can solve into three groups:

(1) Requires only skills from set A: $C\subseteq A$
(2) Requires only skills from set B, including at least one found only in B : $\{C\,|\,C\subseteq B\,and\,\exists\,b\in C\,st\,b\in B\backslash A\}$
(3) Requires at least one skill from each set that can only be found in that set: $\{C\,|\,C\subseteq A\cup B,\,where\,\exists\,a,b\in C\,s.t.\,a\in A\backslash B\,and\,b\in B\backslash A\}$

Using this partition, we can write

$$\begin{aligned}E\left[d\left(A\cup B\right)\right] &= \sum_{C\subseteq A}\frac{\delta\left(C\right)}{\sigma\left(C\right)}+\sum_{C\subseteq B\,and\,C\cap B\neq\emptyset}\frac{\delta\left(C\right)}{\sigma\left(C\right)}+\sum_{C\subseteq A\cup B\,and\,C\cap A,\,C\cap B\neq\emptyset}\frac{\delta\left(C\right)}{\sigma\left(C\right)}\\ &= E\left[d\left(A\right)\right]+\sum_{C\subseteq B\,and\,C\cap B\neq\emptyset}\frac{\delta\left(C\right)}{\sigma\left(C\right)}+\phi\end{aligned}$$

which implies that

$$\begin{aligned}E\left[d\left(A\cup B\right)\right]+E\left[d\left(A\cap B\right)\right] &= E\left[d\left(A\right)\right]+\sum_{C\subseteq B\,and\,C\cap B\neq\emptyset}\frac{\delta\left(C\right)}{\sigma\left(C\right)}+\phi+E\left[d\left(A\cap B\right)\right)\right]\\ &= E\left[d\left(A\right)\right]+\left(\sum_{C\subseteq B\,and\,C\cap B\neq\emptyset}\frac{\delta\left(C\right)}{\sigma\left(C\right)}+\sum_{C\subseteq A\cap B}\frac{\delta\left(C\right)}{\sigma\left(C\right)}\right)+\phi\\ &= E\left[d\left(A\right)\right]+E\left[d\left(B\right)\right]+\phi\\ &\geq E\left[d\left(A\right)\right]+E\left[d\left(B\right)\right]\end{aligned}$$

□

Appendix C. Shapley Value of Each Skill in a Skill Set

We can calculate the contribution of each skill in a player's skill set to the total demand for her skills using a Shapely value decomposition. The demand for player i's skills can generically be written as $d\left(A_i\right)=\sum_{C\subseteq A_i}\frac{\delta(C)}{\sigma(C)}$ where $\delta\left(C\right)=\Psi\left(S\backslash C\right)$[43] and $\sigma\left(C\right)=\sum_{D\subseteq S\backslash C}\Psi\left(C\cup D\right)$. Using this demand as a value function,[44] we can obtain an expression for the Shapely value of a skill, a, to a player, i:

$$\phi_{a,i}\left(d\right) = \sum_{B\subseteq A_i\backslash\{a\}}\frac{1}{\binom{|A_i|}{|B|}}\left(\sum_{C\subseteq B}\frac{\Psi\left(S\backslash\left(C\cup a\right)\right)}{\sum_{D\subseteq S\backslash(C\cup a)}\Psi\left(\left(C\cup a\right)\cup D\right)}\right)$$

[43]Note that this is the expression when each player needs exactly one partner to solve their problem. The results are similar for the more general case.

[44]Note that $d\left(.\right)$ satisfies both requirements for a value function: $d\left(\emptyset\right)=0$ and according to Theorem 2, $d\left(A\cup B\right)+d\left(A\cap B\right)\geq d\left(A\right)+d\left(B\right)$, which implies superadditivity.

This decomposition highlights several points that have been made earlier in the paper: first, the value of a skill to a player depends on the rest of the skills that player has, and second, the value of a skill depends on the population of problem solvers.

In the particular case of the Bernoulli Skills Model, we can make several other observations. In this case, the value of skill a to player i is

$$\phi_{a,i}\left(d\right) = p^{M}\left(1-p\right)\sum_{j=0}^{k-1}\left(1-\frac{j}{k}\right)\left(\frac{1-p+p^{2}}{p^{2}}\right)^{j}$$

Skills are symmetric in the Bernoulli Skills model, and thus each skill has the same value to a particular player–that is, $\phi_{a,i} = \phi_i \forall a \in A_i$. Also, because skills are distributed independently, the value of a skill to a player is strictly increasing in the number of skills the player already has–that is, $\phi_{a,i} > \phi_{a,j}$ iff $|A_i| > |A_j|$.