Three Essays on Interconnectedness Between Forms of Organizational Learning

by

John S. Chen

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Business Administration)
in The University of Michigan
2014

Doctoral Committee:

Professor Gautam Ahuja, Co-chair
Associate Professor Hart Posen, Co-chair
Professor Scott Page
Assistant Professor Maxim V. Sytch
Assistant Professor Minyuan Zhao

DEDICATION

To my wife Lisa and our daughters Justina and Amelia

ACKNOWLEDGEMENTS

Before embarking on a Ph.D., I had little idea of what an exhilarating, intellectually rewarding, challenging, and at times difficult process it would be. In helping me navigate the process, I owe much to my advisor Hart Posen. His endless patience and obvious care that he put into mentoring me through the Ph.D. process were foundational to my experience at Michigan. A number of other faculty with whom I had the privilege to interact also had a profound effect on my scholarly perspectives. Gautam Ahuja and his incomparable eye for a contribution showed me a great deal about what good research should be and why creating it can be so elusive but rewarding. Sendil Ethiraj taught me the value of looking at research through a carefully critical lens. Arvids Ziedonis, Rosemarie Ziedonis, Scott Masten and Minyuan Zhao served as wonderful role models for being not only excellent scholars but kind and humble people. Jerry Davis is a true master in the art of subtle, entertaining, but insightful critique. Jerry's witty replies to my ideas would often lead me to discover, only much later, that I had committed some important oversight. Francine LaFontaine was a great source of encouragement in the remarkably numerous times I bumped into her taking walks with her husband on campus. Scott Page was a generous source of ideas, and he gave me the unique challenge of walking and thinking fast at the same time. Maxim Sytch was a big help with his sagely advice on transitioning to faculty life.

I also am grateful for the emotional and intellectual support provided by my fellow graduate students at Michigan. Thanks to David Zhu, David Benson, and Vivek Tandon for sagely advice on how to navigate dissertation and job market challenges.

I could not have asked for better cohort group in the program than Sun Park and Na-Eun Cho. Sun, you are officially forgiven for getting my car towed the day before starting the Ph.D. program. A special mentioni goes to Joon Oh, Marek Zapletal, Charles Zhang, Sarah Stith, and Na-Eun Cho, my officemates who helped maintain levity during my last and most stressful years at Michigan. I also want to thank Shinwoo Kang, who never took things too seriously and put me at ease. I will always remember Maria Farkas, the prim and proper New Englander who, despite being a frequent dispenser of good advice, often singed my ears with hilarious remarks that I cannot repeat. Finally, I had the privilege getting to know several of the newer Ph.D. students across different departments: Heeyon Kim, Gigi Giustiziero, Heewon Chae, Sara Ryoo, Guy Shani, Suntae Kim, Yong Hyun Kim, Beatriz Pereira, Laura Rees, and Jason Stornelli. I am confident we will see great things from them in the near future.

I thank my friends at Huron Hills Church for their encouragement and unwavering support. They definitely gave more than they received from me. I am grateful to God for both the joys that put a smile on my face and the trials that have refined me not just as scholar but, more importantly, as a person.

I owe much to my parents for their love and belief in me. I especially appreciate the fact that they never pushed me into anything. I am their son first and foremost, and I have the full assurance that they love me equally irrespective of who I am or what I do.

I reserve my last acknowledgements for the ones who have trekked most closely with me on this scholarly journey: Lisa, my wife and best friend, and our daughters Justina and Amelia. Lisa walked by my side every step of the way and made so many sacrifices as I gave up a secure career in technology to go on this crazy adventure. If that is not love, then I do not know what is. Justina and Amelia provided me my most indelible memories of my doctoral studies. After a long night of deep and often

stressful immersion in my research, I would always go to my girls' rooms, see their blissful sleeping faces, and walk away with a wonderful peace in my heart.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Three Essays on Interconnectedness Between Forms of Organizational Learning

by

John S. Chen

Co-Chairs: Gautam Ahuja and Hart Posen

This dissertation builds from the observation that organizational learning takes on multiple forms or dimensions, many of which are interconnected. It comprises three essays that explore the implications of this observation for (1) the dynamics of entry within an industry, (2) cooperation, often among rivals, in an industry, and (3) the riskiness of decision-making within a firm, respectively.

The first essay theorizes a process whereby external knowledge is used in the process of experiential learning. This represents a departure from most prior work, which typically treats vicarious and experiential learning dichotomously. The key implication of this mechanism is that it represents a means by which entrants hold a learning advantage over incumbents, as this essay empirically demonstrates using a full census of U.S. banks.

The second essay proposes a linkage between knowledge generated by a group of collaborating firms and knowledge internal to the firm: firms contribute their internal knowledge to a collaboration as a vehicle for interacting with other firms and absorb knowledge on other firms' contributions. By contributing, firms therefore build a collaboration-specific absorptive capacity. In contrast to prior conceptions

of absorptive capacity, the knowledge that facilitates knowledge absorption is visible to all firms, and as such, high levels of knowledge contributions potentially reveal valuable strategic information to partnering firms. These hypotheses are tested using detailed meeting level data from the wireless industry.

In contrast to the focus on learning sources in the first two essays, the third essay examines inter-temporal interconnections in a firm's learning objectives. In particular, the study examines the linkage between exploration prior to the emergence of a problem and the need to manage the riskiness of adaptation afterwards. While exploration indeed provides knowledge that guides post-problem adaptation, which reduces the risk of searching over alternatives of differing and otherwise unknown quality, we show that exploration also skews firms' preferences towards solutions that in and of themselves have higher performance variation. As such, exploration can in fact increase the risk of problem-solving efforts. The arguments and findings for this essay are formalized through a multi-armed bandit computational model.

CHAPTER I

Introduction

This dissertation investigates interconnectedness among distinct forms of organizational learning across three studies that examine the implications of this interconnectedness for (1) the dynamics of entry within an industry, (2) cooperation, often among rivals, in an industry, and (3) the riskiness of decision-making within a firm. While interconnectedness among learning modes is somewhat abstract in an organizational setting, it is perhaps easier to see at an individual level, such as in the game of chess. Learning in chess takes on multiple, interdependent forms. For instance, learning through experience cannot be considered in isolation from learning through others because the scenarios a player learns to play depend on situations partially dictated by her opponents and, conversely, what the player learns from her opponents depends on her own understanding of chess strategy from prior experience. As another example, vicarious learning may also come through teachers, but the teacher's knowledge may be challenging to internalize absent the student encountering real game situations that trigger a need to learn responses. Finally, learning to manage risk in chess play depends partially on the sources of vicarious learning. A player may wish to learn low-risk chess strategies, but her playing environment – that is, the playing style of the opponents she tends to face – may dictate that the optimal strategies for her to learn are instead risky ones (i.e., strategies leading to many wins

and losses, but few draws).

In the first study, Hart Posen and I examine interconnectedness between experiential and vicarious learning. In particular, we argue a learning mechanism by which external knowledge facilitates the process of experiential learning, which stands in contrast to the separable treatment of vicarious and experiential learning mechanisms typically seen in prior research. While a firm's ability to absorb (Cohen and Levinthal, 1990) – that is, comprehend and assimilate – external knowledge is central to prior literature's explanations for the extent of vicarious learning, it is not a firm's ability but rather the impetus to gain external knowledge that is central under the mechanism proposed in this study.

As such, we propose that entrants benefit more than incumbents from learning at the intersection of own experience and external knowledge. An important implication of this proposition is that it helps explain the preponderance of successful entry in most industries (Geroski, 1995), despite entrants suffering a "liability of newness" (Stinchcombe, 1965) in that they rarely have the knowledge and capabilities at founding to survive and have limited ability to absorb knowledge from others. While radical innovation by entrants offers a compelling explanation (Henderson, 1993; Reinganum, 1983; Schumpeter, 1934), Geroski's study also finds significant successful entry in the numerous industries marked almost exclusively by incremental innovations. It is in these contexts where the proposed mechanism of learning at the intersection of own experience and external knowledge perhaps holds the most explanatory power for how entrants "catch up" to incumbents and become viable within an industry.

In the second study, I examine interconnectedness between internal learning – that is, learning facilitated by internal experience or prior knowledge – and learning through multi-firm collaborations, such as research consortia or standardization bodies. Internal and collaborative learning processes are rarely considered interdependently in the multi-firm collaboration (and dyadic collaboration) literature. For

example, the research consortia literature considers internal and collaborative learning as distinct and largely independent modes (Cassiman and Veugelers, 2002; Katz, 1986; Sakakibara, 2003). In contrast, I theorize a process by which knowledge sharing in a collaboration enhances a firm's (collaboration-specific) absorptive capacity, which in turn facilitates firm-level learning.

This process implies that knowledge contributions to a collaboration enhance the performance of a firm's innovations that build on knowledge generated by the collaboration. However, we further theorize that this beneficial effect of knowledge contributions is limited: higher levels of contribution may signal to other firms information, such as technology focus or expectations of the collaboration's benefits, that intensifies the competition for post-collaborative innovations.

In contrast to the first two studies examining interconnectedness among knowledge sources for learning, the third study, which I jointly conducted with Hart Posen, examines the inter-temporal interconnectedness among a firm's objectives for learning. In particular, we examine the linkage between exploration prior to the emergence of a problem and the need to manage the riskiness of adaptation afterwards. We challenge the seemingly obvious prescription that increased exploration of contingent actions prior to the emergence of a problem reduces the risk associated with solving the problem. Using a computational bandit model (Gittins and Jones, 1979), we show that while this exploration indeed provides knowledge that guides an otherwise blind and hence risky search for solutions to the problem, it also biases firms towards riskier solutions.

This study extends prior work showing how endogenous risk aversion emerges from risk-neutral search (March, 1996) and implications thereof (Denrell and March, 2001). In contrast to this prior research, wherein risk aversion emerges from a search for the best alternative in a stable, problem-free environment, we show how risk seeking in problem-solving emerges as a byproduct of pre-problem search. In particular, risk

seeking emerges from alternatives tried and discarded prior to the occurrence of a problem (i.e., in the pre-problem environment) but ultimately adopted as a solution to the problem.

CHAPTER II

An Advantage of Newness: Vicarious Learning Despite Limited Absorptive Capacity

2.1 Introduction

New entrants are endowed with the pre-entry knowledge of their founders (Agarwal et al., 2004; Dencker et al., 2009; Helfat and Lieberman, 2002; Huber, 1991). However, this knowledge alone is often insufficient to generate the performance levels necessary to survive. Indeed, entrants appear to suffer a "liability of newness" (Stinchcombe, 1965), which has been attributed to factors including a lack of resources, limited organizational knowledge and capabilities, and underdeveloped organizational practices. Entrants can bridge the gap between their initial endowments and the capabilities necessary to compete effectively by learning vicariously from the knowledge of incumbents. But given the important role of absorptive capacity (Cohen and Levinthal, 1990), how can entrants learn vicariously when they lack stocks of prior related knowledge at founding? In answering this question, we theorize that the process of experiential learning facilitates vicarious learning, particularly for entrants.

Organizations, both entrants and incumbents, seek to learn directly from their own experience (Argote, 1999; Lieberman, 1984; March and Olsen, 1975) and vicariously from the knowledge and experience of rivals (Levitt and March, 1988; Miner

and Haunschild, 1995; Posen et al., 2013). Progress in understanding organizational learning has been facilitated by the assumption that experiential learning and vicarious learning are independent processes. Studies have typically focused on one mode of learning alone, whereas those that examine both modes simultaneously treat them as theoretically additive(Argote et al., 1990; Herriott et al., 1985).

Building on recent research that relaxes the assumption of independence (Baum and Dahlin, 2007; Schwab, 2007; Simon and Lieberman, 2010), our central thesis is that, particularly for entrants, experiential learning and vicarious learning are more closely related than prior research suggests. Conditional on the size of the pool of external knowledge available, the amount of vicarious learning is a function of both the capacity to absorb external knowledge and the impetus to obtain it (Greve, 2005). This implies two mechanisms through which experiential learning may facilitate vicarious learning. The first mechanism derives from the observation that the outcome of experiential learning augments a firm's knowledge stock. An absorptive capacity logic (Cohen and Levinthal, 1990) suggests that the ability to learn vicariously is enhanced by accumulated experience, which facilitates the identification and utilization of external knowledge.

The second mechanism, which is the main focus of our attention, derives from the observation that experiential learning is a process as well as an outcome. This suggests a far richer interaction between experiential and vicarious learning. Central to our theory is a "problem" (Nickerson and Zenger, 2004; Nickerson et al., 2012). Experiential learning occurs when, in the course of engaging in productive activity, problems are identified, experiments are performed as solutions are sought, and solutions are implemented (Arrow, 1962; March and Olsen, 1975). These problem-solving attempts do not happen in a vacuum, isolated from the knowledge and experience of others. Many solutions to problems are stimulated by learning about what other firms do. Thus, problems that arise in the process of experiential learning trigger a

search for external knowledge, such that vicarious learning occurs as a by-product of the ongoing process of experiential learning. The strength of this mechanism is a decreasing function of a firm's existing stock of problem-solving knowledge. As such, the impact of experiential learning as a facilitator of vicarious learning is relatively small for incumbents, but it is a significant driver of entrants' learning. In this sense, entrants enjoy an advantage of newness.

To illustrate the intersection of experiential and vicarious learning, consider the example of an entrant engaged in the manufacture of mobile phones.[1] After commencing initial production, the entrant observes higher defect rates than anticipated. Testing reveals signal strength varies widely across units. The firm must solve the problem rapidly, but there are many potential causes of the excess defects, such as management practices, employee skills, production planning, production processes, and product design. The entrant initially attempts to draw from its own production knowledge for guidance in identifying a solution. When this internal knowledge is unavailable to solve the problem, the entrant seeks knowledge outside the firm. It begins by disassembling competitors' phones to compare signal quality at various internal design points, and in doing so, it discovers simple differences in the production process that account for the problem. Here, a problem encountered in the experiential learning process directs vicarious learning since there are too many potential problems with the phone (or any other complex product) for a firm to solve proactively (that is, absent the problems actually occurring).

The remainder of this paper is organized as follows: In the next section, we review the literatures on experiential and vicarious learning. Then we more fully develop our theory at their intersection and implications for differences between entrants and incumbents. Next, we describe our empirical setting: the U.S. commercial banking

---

[1]This illustration reflects a real problem encountered during a prior career, though some minor details have been modified to protect the identity of the firm and more clearly convey the theoretical mechanisms.

industry from 1984 to 1998. In this context, we are able to construct learning curves for a complete census of firms. We then present our empirical analysis, in which we examine problems that arise in the experiential learning process and identify the extent to which they trigger vicarious learning by entrants and incumbents. Finally, we offer concluding remarks, including contributions to the broader literature on organizational learning.

2.2 Theory and Hypotheses

In this study, we focus attention on the intersection of two types of organizational activities oriented toward the accumulation of knowledge: experiential learning and vicarious learning. These mechanisms of learning share a common foundation in that they attempt to replicate one's own, or others', successful choices and to avoid those choices proven unsuccessful (March, 2010). We begin with a brief discussion of the theoretical basis of the two types of learning and then present our argument for learning at their intersection.

2.2.1 Experiential Learning

An organization learns from its own experience. Experiential learning can be conceptualized as a history-dependent process in which adjustments to organizational practices are based, in part, on knowledge accumulated during an organization's past experience (Herriott et al., 1985; March and Olsen, 1975). Consider an organization engaged in day-to-day productive activity. The management of the organization can be conceptualized as choosing among alternative ways of performing a task (Herriott et al., 1985; March, 2010). An alternative may represent a single organizational choice, but more generally, it represents a configuration of choices governing how the organization engages in productive activity. March (2010, p. 16) described organizations that "(1) act by choosing an alternative from among those available, (2) record

a result and evaluate it in terms of its success, (3) replicate the choice of alternatives associated with successes ..."

Thus, the locus of attention in experiential learning is a problem that emerges in the process of engaging in productive activity and the subsequent search for solutions (Arrow, 1962; March and Olsen, 1975). A problem emerges when performance falls below expectations. Performance may be defined along various dimensions such as financial indicators (e.g., sales, return on assets), design/manufacturing metrics (e.g., defect rates, productivity), or consumer satisfaction. Expectations are set relative either to an organization's own past performance or to the performance of a suitable reference group such as other firms in the industry (Cyert and March, 1963; Greve, 1998). For example, in a field study, von Hippel and Tyre (1995) documented one firm's experiential learning in circuit board manufacturing. A problem was identified when a production step required more time and effort than expected based on the firm's prior experience with similar tasks. The solution involved a minor software change that streamlined the production step.

In practice, experiential learning is not as simple as these descriptions suggest and is indeed subject to uncertainty, ambiguity, bias, and error (Levinthal and March, 1993; Levitt and March, 1988). However, there is significant evidence of its general efficacy. The most prominent evidence is in the large body of work on organizational learning curves. These studies relate the unit cost of production to cumulative output (as a measure of experience). Studies find that firms progress along learning curves in which their performance improves at a decreasing rate with output (see Argote, 1999, for a review). Learning curves have been observed across a variety of industries (e.g., aircraft, trucking, chemicals, semiconductors), dimensions of performance (e.g., costs, recalls, customer satisfaction, survival), and levels of analysis (e.g., groups, plants, firms). Research in this domain tends to model learning curves with a power law formulation (Alchian, 1963; Argote and Epple, 1990; Lieberman, 1984), such that

$$C = AX^{\beta^x} \tag{2.1}$$

where C is cost, A is the cost corresponding to the first unit of experience, X is a cumulative experience measure, and β^x is the rate of experiential learning; $\beta^x < 0$ implies cost reduction and, hence, learning.

2.2.2 Vicarious Learning from External Knowledge

Organizations learn not only from their own experience but also from the knowledge of others (Helfat and Lieberman, 2002; Levitt and March, 1988; Miner and Haunschild, 1995). In the innovation literature, vicarious learning occurs through spillovers (i.e., unintended leakage) of technical knowledge (Griliches, 1998). In early research, Griliches (1957) studied the diffusion of hybrid corn. Subsequently, a large body of research examined research and development (R&D) spillovers across a variety of industries, including semiconductors (Appleyard, 1996) and pharmaceuticals (Furman et al., 2006).

Studies of vicarious learning, many outside the realm of R&D, posit two types of knowledge conduits between firms. First, knowledge may be acquired by observing artifacts of rivals' internal activities. Baum and Ingram (1998, p. 1000) found that Manhattan hotels learn by observing rivals' practices through visits that "start at the roof and go to the basement..." Similarly, Haunschild and Miner (1997) demonstrated that firms learn from observing rivals' choices of investment bankers, Simon and Lieberman (2010) demonstrated that magazine publishers learn from observing rivals' websites, and Gittelman and Kogut (2003) showed that firms learn by observing rivals' patents and publications.

Second, knowledge may be acquired directly from rivals' employees. The most direct of these mechanisms is employee mobility (Corredoira and Rosenkopf, 2010; Rosenkopf and Almeida, 2003) and board interlocks (Haunschild and Beckman, 1998).

Less direct are knowledge flows facilitated by industry social forums, including trade association meetings and industry conferences (Keeble and Wilkinson, 1999). Knowledge flows may also occur in less formal settings. Ingram and Roberts (2000) studied knowledge flows through the social relations of Sydney hotel managers, and Saxenian (1996) studied the technological community in Silicon Valley.

A salient feature of many of these studies is that (substantial) absorptive capacity is not always necessary. For example, in Baum and Ingram's (1998) Manhattan hotel study, general powers of observation were sufficient to facilitate learning through visits to rivals' establishments. Similarly, Simon and Lieberman (2010, p. 138) noted that magazine publishers' websites are "easily observable" and that "many websites are developed and operated by third parties who can transmit information from early adopters to late adopters."

Following standard convention in productivity studies (see Griliches, 1998, for several examples), we express the relationship between cost and external knowledge as

$$C = AK^{\beta^k} \tag{2.2}$$

where C is cost of production, K is a measure of the external knowledge pool, A is a constant, and β^k is interpreted as the rate of vicarious learning from external knowledge; $\beta^k < 0$ implies that unit production costs are smaller in contexts where the pool of external knowledge is larger.

The cost reductions realized at the intersection of experiential and vicarious learning are modeled by combining Equations 2.1 and 2.2, which reflects a simple Cobb-Douglas form, such that[2]

---

[2]Typically, Cobb-Douglas exponents are positive and less than one to reflect positive but diminishing returns. In this study, the exponents are negative.

$$C = AX^{\beta^x}K^{\beta^k} \tag{2.3}$$

2.2.3 Intersection of Experiential and Vicarious Learning

Progress in understanding organizational learning has been facilitated by the assumption of an artificial divide between experiential and vicarious learning. Many studies have focused on either experiential or vicarious learning alone, whereas others have examined them additively (Argote et al., 1990; Baum et al., 2000; Ingram and Simons, 2002). For example, Darr et al. (1995) studied vicarious learning in the pizza industry and modeled learning from own and others' production experience as independent, purely additive terms in their theory and empirical specification. Recent research has sought to relax this assumption, examining the conditions under which experiential and vicarious learning act as substitutes or complements (Baum and Dahlin, 2007; Schwab, 2007) and the temporal sequencing of experiential and vicarious learning (Bingham and Davis, 2012). Building on this research, we argue that experiential and vicarious learning are interrelated because the process of experiential learning facilitates vicarious learning.

Our theory rests on the observation that the extent of vicarious learning is a function not only of the size of the external knowledge pool (Levin and Reiss, 1988; Spence, 1984) and firms' ability to absorb external knowledge (Cohen and Levinthal, 1990) but also of firms' impetus to acquire that knowledge (Eeckhout and Jovanovic, 2002; Ethiraj and Zhu, 2008; Greve, 2005; Zemsky and Pacheco-de Almeida, 2012). This impetus is the driver of vicarious learning on which we focus.

Experiential learning is a process that embodies a problem-solution cycle (Arrow, 1962; March and Olsen, 1975). In both theoretical models and empirical research, this process is often conceptualized as occurring within the locus of a firm, decoupled from the external knowledge environment. In many formal models – for instance, of

learning on rugged landscapes (e.g. Levinthal, 1997) and of learning under uncertainty (Denrell and March, 2001; Posen and Levinthal, 2012) – firms accrue knowledge about the merits of alternatives based on their own prior experience and use it to solve subsequent problems. The empirical literature often reflects the same assumption. For example, Hoopes and Postrel (1999) studied experiential learning in a scientific software company and observed that "glitches" arose when the firm could not access internal knowledge to solve problems. Feldman (2000) examined experiential learning in staffing activities in a university residence hall and observed that solutions were found by relying on internal knowledge and not by learning from other universities' hiring practices.

Yet solutions to problems that arise in the process of experiential learning may also be found by expropriating knowledge from rivals. External knowledge serves two ends in facilitating problem solving: identifying the set of alternative solutions that may solve the problem and evaluating the efficacy of each alternative. In the former, when a problem is identified in the process of engaging in productive activity, boundedly rational firms may face significant difficulty in identifying the set of potential solutions. In many contexts, the set of alternatives that potentially solve a given problem is very large, and as such, many may be unknown to the firm. For such problems, external knowledge exposes a firm to a broader set of alternatives and, in doing so, enhances the efficacy of problem solving.

In the latter, once the set of potential solutions is identified, external knowledge may enhance the efficacy of evaluating, and selecting among, alternatives. In attempting to resolve a problem, boundedly rational firms may have only limited experience with a given alternative, which is problematic when the number of alternatives is large or when interdependence confounds cause-effect relationships (Levinthal and March, 1993). Moreover, knowledge gleaned from experiential learning tends to be biased against alternatives that are risky (March, 1996) or have performed poorly in the

past (Denrell and March, 2001). To the extent that other firms in the industry have gained knowledge about the merits of different alternatives, and where that knowledge may be accessed vicariously, the efficacy of selecting a solution from among the set of possible alternatives may be enhanced by utilizing external knowledge.

As a consequence, problems that arise in the process of experiential learning generate the impetus to seek external knowledge. In turn, the rate of learning from experience, β_i^x, is enhanced by external knowledge, k. We incorporate this in experiential learning by amending the learning model in Equation 2.3, now in log form, as follows:

$$
\begin{aligned}
c &= \alpha + \beta_i^x x + \beta^k k \,, \\
\text{where} \quad \beta_i^x &= \delta_0 + \delta_1 k + \nu_i \,.
\end{aligned}
\tag{2.4}
$$

Here, δ_0 is the overall mean learning rate, δ_1 is the contribution of external knowledge to the rate of experiential learning, i indexes firms, and ν_i represents firm learning-rate heterogeneity in the form of a random slope. In addition, the theorizing above suggests that it is the emergence of problems that trigger a search for solutions external to the firm. We incorporate problems into our learning model as follows:

$$
c = \alpha + \beta_i^x x + \beta^k k \,,
$$
$$
\text{where} \quad \beta_i^x = \delta_0 + \delta_1 k + \delta_2 p + \delta_3 pk + \nu_i \,.
\tag{2.5}
$$

where p is a problem that arises in the process of engaging in productive activity. Accordingly, we hypothesize the following.

Hypothesis 1A (H1A): Firms learn more rapidly from experience when the pool of external knowledge is larger (i.e., $\delta_1 < 0$ in Equation 2.4).

Hypothesis 1B (H1B): Problems increase the extent to which firms utilize external knowledge in the experiential learning process (i.e., $\delta_3 < 0$ in Equation 2.5).

2.2.4 Differences Between Entrants and Incumbents

The theory above posits that firms learn at the intersection of experiential and vicarious mechanisms. This learning occurs, in part, as a result of firms using external knowledge to solve problems that arise in the course of experiential learning. In this section, we argue that the impact of experiential learning as a facilitator of vicarious learning is relatively small for incumbents but is a significant driver of vicarious learning for entrants.

To clarify our arguments, consider the extent of (potential) vicarious learning, L, generated as a firm engages in an additional unit of productive experience, x. We write the marginal learning as a function of experience, $\partial L/\partial x$, as

$$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial N} \cdot \frac{\partial N}{\partial P} \cdot \frac{\partial P}{\partial x} \tag{2.6}$$

where P is the occurrence of a problem in the experiential learning process and N is the amount of external knowledge accessed in solving a problem. Holding constant the rate at which problems appear from a unit of experience, $\partial P/\partial x$, we discuss two elements of the righthand side of Equation 2.6 and highlight why we would expect each to be greater for entrants than incumbents.[3]

First, $\partial L/\partial N$ is the marginal learning from using a unit of external knowledge in the process of engaging in productive activity. The problem-solving activities underlying experiential learning tend to be repetitive in nature (Argote, 1999; Denrell et al.,

---

[3]Two additional notes are relevant. First, we examine potential learning because, given the impetus to access a unit of external knowledge, realized learning is a function of potential learning and the ability to absorb external knowledge. Second, if entrants encounter more problems in the process of engaging in productive activity, this would tend to further increase the extent to which they derive value from external knowledge and would lend additional strength to our hypotheses.

2004). As such, learning by doing that is "associated with repetition of essentially the same problem is subject to sharply diminishing returns" (Arrow, 1962, p. 155). While engaging in its early developmental activity, an entrant encounters a myriad of problems. With each attempt to solve a particular problem, the returns to engaging in effort to find new and better solutions to subsequent occurrences of the same problem decline. Thus, while solving problems encountered in experiential learning, entrants are likely to learn more from external knowledge than incumbents.

Second, $\partial N/\partial P$ is the amount of external knowledge used in solving a problem that arises in the process of engaging in productive activity. The extent to which a firm seeks outside knowledge, and thus learns vicariously, is a function of the extent to which it already harbors knowledge that might form the basis of a solution to a problem. When internal knowledge is insufficient, an organization is more likely to seek external knowledge to solve problems that emerge in the experiential learning process (Simon and Lieberman, 2010).

One would expect entrants to harbor less knowledge than incumbents. Although entrants possess pre-founding knowledge (Agarwal et al., 2004; Helfat and Lieberman, 2002; Huber, 1991), it is only a portion of the knowledge an entrant will eventually possess as it matures into incumbency. This expectation, in part, reflects the observation that the prefounding knowledge of a team of limited size is unlikely to encompass the breadth of knowledge necessary to inform all of the activities of the firm. Additionally, prefounding knowledge by definition excludes organizationally embedded knowledge stored in routines (Nelson and Winter, 1982). This knowledge is only built over time through repeated interactions across multiple members of the organization (Feldman, 2000), as well as through interactions across diverse functional units (Zahra et al., 2000). For example, Epple et al. (1996) studied learning curves across multiple production shifts in the same plant. They concluded that knowledge embedded in the organization, its routines, and its technologies exists over and above that possessed

by individuals. Therefore, entrants are more likely to draw on external knowledge when confronted with a problem that emerges in the experiential learning process because they typically lack stocks of organizationally embedded knowledge.

In sum, the discussion above suggests that the extent of learning at the intersection of experiential and vicarious processes is greater for entrants than for incumbents. As such, we hypothesize the following.

> **Hypothesis 2A (H2A)**: Entrants' experiential learning rate increases more than incumbents' when the pool of external knowledge is larger (i.e., δ_1[entrant] $< \delta_1$[incumbent] in Equation 2.4).

Moreover, because we argued that entrants access more external knowledge in solving problems that arise during experience (i.e., the $\partial N/\partial P$ term), we further hypothesize the following.

> **Hypothesis 2B (H2B)**: The extent to which problems induce the utilization of external knowledge in the experiential learning process is greater for entrants than for incumbents (i.e., δ_3[entrant] $< \delta_3$[incumbent] in Equation 2.5).

2.3 Industry

We test our hypotheses in U.S. commercial banking post deregulation. We chose this industry because it is geographically fragmented with localized competition, highly regulated, and characterized by substantial firm-level learning. Fragmentation allows us to compare discrete markets characterized by different levels of local external knowledge within the same industry. We define a competitive market as a state because the unit of observation in the Federal Deposit Insurance Corporation (FDIC) regulatory system is a state-level insurance certificate.[4] Intense regulation

---

[4]A certificate is required for each state in which a bank operates and covers all branches of a bank operating within that state.

and strict reporting requirements allow us to obtain quarterly cost data for the full census of banks, including all entrants.

Considerable learning occurs in commercial banking. However, firms do not typically compete and win on the basis of technological innovation. As a consequence, experiential learning, as opposed to mechanisms such as R&D, is an important driver of change in firm performance. To gain a deeper understanding of learning by new entrants in commercial banking, we interviewed a number of executives at start-up banks.

A central theme of founders' comments was the important role of post entry learning. Cost reduction appears to be the main goal of their learning efforts. One founder identified the need to learn how to efficiently utilize staff and streamline numerous operational processes. Another founder identified the need to learn how to maintain a sufficiently diversified loan portfolio and to reduce loan default rates. These comments reflect standard wisdom in banking texts. For example, Saunders and Cornett (2003) argued that bank performance relies heavily on improvements in areas such as operational expenses, risk management (e.g., credit, interest rate, information technology), deposit generation, and regulatory compliance.

On the surface, the need for significant learning after entry seems surprising given that new banks are typically founded by industry veterans. However, founders often noted that the knowledge of the founding team is typically limited relative to the breadth of knowledge required, and established banks possess organizationally embedded knowledge beyond the reach of individual employees. For example, one founder, formerly a corporate executive in charge of lending at a major incumbent bank, detailed his start-up's highly inefficient loan drafting process. In hindsight, he realized that the inefficiency was due to difficulties in identifying and replicating complex organizational routines for managing the myriad provisions in modern loans.

In addressing these types of problems, banking executives pointed to the lack of

internally available solutions and, as such, the need to learn from the knowledge and experience of other banks. The bankers often discussed their use of interpersonal relationships to learn from others, which is facilitated by extensive regulation in banking. For example, one executive examined public regulatory filings to identify nearby banks that excelled in residential mortgage loans, then identified those banks at which he had acquaintances that might be willing to share knowledge.

In sum, commercial banking provides uniquely rich data from which to examine the intersection of experiential and vicarious learning. In the next sections, we discuss our model, data, and results.

2.4 Empirical Model

Our empirical model consists of two stages. In the first stage, we model an industry cost frontier to collect measures of efficiency for each firm, both entrants and incumbents, in each quarter. In the second stage, we model learning by considering a firm's efficiency (from the first stage) as a function of its experience and the knowledge of rivals.

2.4.1 Stage 1: Firm Efficiency

We follow convention in the banking literature by modeling cost efficiency as a stochastic cost frontier using a translog cost function (Berger, 2003; Knott and Posen, 2005; Knott et al., 2009; Mester, 1993). The stochastic frontier model assumes that the log of firm i's cost in quarter t, c_{it}, differs from the cost frontier, c_{min}, by an amount that reflects its cost inefficiency. In particular, the translog cost function represents the total cost for firm i in quarter t as a function of a firm's chosen output levels and a given set of input prices:

$$c_{it} = \phi_0 + \sum_j \phi_{1j} y_{it}^j + \sum_h \phi_{2h} w_{it}^h + \frac{1}{2} \sum_j \sum_j \phi_{3jj} y_{it}^j y_{it}^j$$

$$+ \frac{1}{2} \sum_h \sum_h \phi_{4kk} w_{it}^h w_{it}^h + \sum_j \sum_h \phi_{5jk} y_{it}^j w_{it}^h + u_{it} + e_{it} \,, \qquad (2.7)$$

where c_{it} is the log of observed firm cost, y_{it}^j is a vector of log output levels (j indexes output elements), w_{it}^h is a vector of log input prices (h indexes input elements), u_{it} is log cost efficiency with truncated normal distribution, and e_{it} is the normally distributed error term. This translog cost function allows us to incorporate a complex array of bank inputs and outputs and to accommodate trade-offs in market strategies (product mixes and prices) and in operational strategies (input mixes and scale).

We collect estimates of the expected value of firm-quarter cost efficiency in Stage 1, $E(u_{it}|e_{it})$, which we label as $u_{i,t}$. $u_{i,t}$ reflects the log of firm i's costs relative to a firm on the cost frontier. For brevity, we refer to $u_{i,t}$ as cost, rather than cost efficiency, because cost efficiency is functionally cost after controlling for differences in scale and input/output mix. As such, lower values of $u_{i,t}$ correspond to increased efficiency – i.e., lower cost. The estimate of u_{it} serves as our Stage 2 dependent variable testing our learning rate mechanisms.

2.4.2 Stage 2: Tests of learning mechanisms

In the second stage, we model learning by examining increases in efficiency, or, equivalently, reduction in a firm's cost relative to the firm at the cost frontier. We transform the log cost variable $u_{i,t}$ with the monotonic function $g(u_{i,t}) = ln(e^{u_{i,t}} - 1)$.[5] We then model learning, as represented in Equation 2.5, while controlling for time-varying firm- and market-level attributes, such that:

---

[5]We do so because, by construction, e^u approaches 1 as experience approaches infinity. However, the functional form of the learning curve model (Equation 2.1) requires that relative costs asymptote to 0 (Argote, 1999), which is satisfied by $e^{g(u)}$.

$$g(u_{i,t}) = \beta_0 + \beta_i^x x_{i,t} + \beta^k k_{i,t} + \gamma_1 S_{s,t} + \gamma_2 F_{s,t} + \alpha_i + \epsilon_{i,t}, \tag{2.8}$$

where β_i^x is the coefficient of learning from experience, β^k is the coefficient of vicarious learning unaided by experiential learning, s is the state in which firm i operates, and $\epsilon_{i,t}$ is the error term. $S_{s,t}$ and $F_{s,t}$ represent vectors of state- and firm-level controls.

To study the impact of external knowledge on the rate a firm learns from experience, we allow the learning coefficient β_i^x from Equation 2.8 to vary across firms such that:

$$\beta_i^x = \delta_0 + \delta_1 k_{i,t} + \delta_2 p_{i,t} + \delta_3 p_{i,t} k_{i,t} + \nu_i, \tag{2.9}$$

where the coefficient on the extent to which external knowledge enhances experiential learning is given by δ_1, and the coefficient on the extent to which external knowledge is used to solve problems is given by δ_3. The inclusion of firm-specific components, intercepts α_i (in Equation 2.8) and slopes ν_i (in Equation 2.9), modeled as random effects, account for unobserved heterogeneity in both initial costs (or, in the case of incumbents, costs at the start of the observation window) as well as rates of learning from experience. We substitute Equation 2.9 into Equation 2.8 and estimate the resulting random-coefficients model using a maximum-likelihood estimator. The random coefficients model is a panel regression technique that extends a random effects model so that not only the intercept but also the coefficients are allowed to vary per panel. We make the standard assumption that the covariance matrix of the random parameters ν and α, $\Sigma_{\nu\alpha}$, is bivariate normal, and we impose no additional constraints.

The coefficients δ_1 and δ_3 in Equation 2.9 inform our hypotheses: $\delta_1 < 0$ would support the hypothesis that firms learn more rapidly from experience when the pool of external knowledge is larger (H1A), $\delta_3 < 0$ would support the hypothesis that problems increase the extent to which firms utilize external knowledge in the experiential

learning process (H1B), $\delta_1[\text{entrant}] < \delta_1[\text{incumbent}]$ would support the hypothesis that the experiential learning rate of entrants increases more than incumbents' when the pool of external knowledge is larger (H2A), and finally, $\delta_3[\text{entrant}] < \delta_3[\text{incumbent}]$ would support the hypothesis that entrants benefit more than incumbents from using external knowledge to solve problems encountered in the experiential learning process (H2B).

2.5   Data and Measures

The data for the study came from the FDIC Research Database of quarterly financial data for all commercial banks filing the "Report of Condition and Income" (call report). The FDIC assigns a unique certificate number to each bank that enters a given state. A single bank cannot have the same certificate number for branches that operate in different states. We have therefore taken the bank (certificate number) as our basic unit of analysis. We examined all banks in each state plus the District of Columbia for the period Q1 1984 to Q1 1998. This data set contains over a half million firm-quarter observations that are used in the first-stage estimation model. In particular, we observed quarterly data for a comprehensive set of 1,594 entrants for an average of seven years each and for a comprehensive set of 12,652 incumbents that operated during this period.

As the Stage 1 model follows the design of Knott and Posen (2005) and Knott et al. (2009), we omit here a detailed discussion of the cost, input, and output variables in the translog cost function.

For the Stage 2 model, we constructed a number of explanatory variables. We constructed a measure of external knowledge, $k_{i,t}$, using the cost efficiencies calculated from the Stage 1 model. In particular, we used the average log cost of firms in state s (other than firm i) and multiplied this value by -1 to represent external

knowledge available to firm i.[6] As such, we examine the productivity implications of cost as a proxy for knowledge itself. Cost can be viewed as an outcome of employing knowledge in production, $C = f_1(K)$. Thus, knowledge is the functional inverse such that $K = f_1^1(C)$ (Griliches, 1998). This functional relationship reflects a wide body of literature. It is the principle modeling assumption in the theoretical work in Spence (1984) and in the empirical work of Levin and Reiss (1988), which employed Federal Trade Commission data to estimate models in which firms invest in knowledge expressly to reduce costs. This assumption is also central to Griliches' (1998) models of R&D production functions, and Nelson and Winter (1982) assumed that firms invest to generate knowledge in the form of superior production practices that lower unit costs.[7]

Our measure of experiential learning, $x_{i,t}$, is given by the cumulative number of quarters in which a bank has competed in the market. Accordingly, we represented a firm's experience as $x_{i,t} = ln(t + 1)$ so that at time 0 a firm has no experience. This temporal operationalization of experience is widely used in the learning curve literature (Argote et al., 1990; Darr et al., 1995; Lieberman, 1984).[8] We defined a problem, p, as occurring if loan output dropped for two consecutive quarters. This measure reflects a problem that arises when performance is assessed relative to a firm's own past performance.

We defined an entrant as a new commercial banking institution. Banks that entered within the years 1984 - 1997 are considered entrants in our sample, resulting

---

[6]Multiplying by -1 preserves the more natural convention that a larger $k_{i,t}$ represents more external knowledge.

[7]Ideally, we would measure external knowledge directly. This is possible in industries where technical knowledge is codified and embodied in patents. In industries where patents are rare and nontechnical knowledge dominates (e.g., banking), directly measuring knowledge is problematic. An alternative to using an outcome of knowledge production (e.g., cost) is to employ an input, ω, to the knowledge production function, $K = f_2(\omega)$. We examine firm experience as ω in our robustness analysis.

[8]Data limitations preclude constructing a cumulative output measure of experience for the entire history of all incumbent firms because our data extend back only eight years prior to the start of the sample window.

in an entrant subsample of 1,594 banks and 42,207 firm-quarter observations. To provide a sharp delineation between entrants and incumbents, we defined incumbents as those firms founded prior to 1975 (i.e., at least 10 years prior to the earliest entrants). This results in an incumbent subsample of 10,711 banks and 490,958 firm-quarter observations.

A number of control variables are included in the model. To control for branch scale effects not already accounted for in the first-stage estimation, we included two variables: (a) *branch_count*, the number of branches operated by the bank; and (b) *branch_scale*, the average size of a branch measured in terms of total output in thousands of constant 1996 dollars. We controlled for corporate structure by including (c) a *holding_company* indicator variable, as well as a number of measures of the size of the holding company; (d) *hc_certificates*, the number of additional banks (certificates) held by the holding company; (e) *hc_branches*, the number of additional branches in the bank holding company system beyond the number of branches in the observation certificate; and (f) *hc_states*, the number of additional states in which the holding company operates banks. We controlled for economic differences across markets by including variables representing quarterly data on demand: (g) *population*, (h) housing start *permits*, and competition with (i) a Herfindahl index of industry concentration, *HHI*.

2.6   Results

2.6.1   Stage 1: Firm Efficiency

Detailed results from our Stage 1 frontier estimation are available upon request. Here, we wish to highlight important descriptive characteristics of firm costs (i.e., $e^{u_{i,t}}$). The mean cost (of all firms) over the sample period is 17.2% above that of a firm on the cost frontier. With regard to entrants, in Figure 2.1, we plot mean and

Figure 2.1: Descriptive Data on Entrant Cost by Age (Relative to Incumbents)
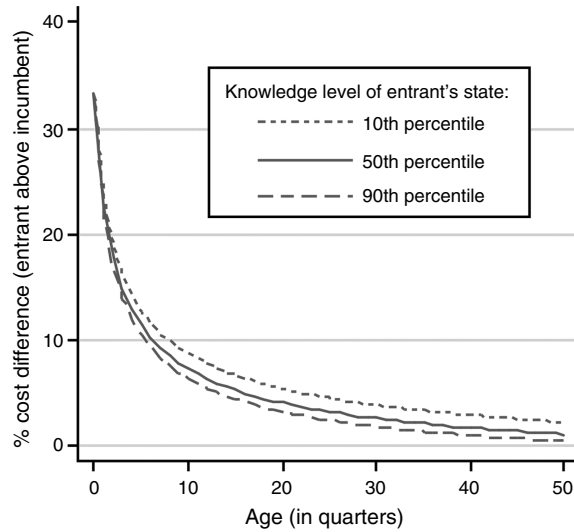
median entrant cost across entrant age (relative to incumbents). Entrants are initially much less efficient than incumbents. In the first year after entry (i.e., the average of the first four quarters in Figure 2.1), mean entrant cost is 57% higher than that of incumbents. But entrants rapidly become competitive, with cost only 4% higher than incumbents in the third year post entry.

To provide a better sense of incumbent learning versus entrant learning, as given by increased efficiency (i.e., a decrease in cost), we follow longitudinal microdata studies and compute a transition matrix of firm efficiency (see Table 2.1). The transition matrix relates a firm's efficiency in one quarter (a row in Table 2.1) to its efficiency in the next quarter (a column in Table 2.1). Table 2.1 is broken up by quintiles. As an example, the row labeled 3 in Table 2.1 indicates that 21.33% of firms in the third-most efficient quintile of all firms ascend to the second-most efficient quintile in the next quarter. Table 2.1 also shows the relative efficiency of firms at entry ("Entry" row) and at exit, either through mergers or failure ("Exit" column).

Table 2.1: Transition Tables of Incumbents and Entrants Within Three Years of Entry

			Next quarter quintile					
			Low cost			High cost		
			1	2	3	4	5	Exit
		Entry	2.33	1.89	2.46	5.62	87.70	0.00
This quarter quintile	Low cost	1	79.68	16.26	2.35	0.71	0.28	0.73
		2	17.68	58.04	19.66	3.28	0.54	0.80
	High cost	3	2.19	21.33	54.29	19.38	1.87	0.94
		4	0.55	3.16	20.83	60.70	13.54	1.22
		5	0.45	0.69	2.22	17.79	77.15	1.70

Table 2.1 highlights three notable patterns. First, relative firm efficiencies are fairly stable, as indicated in the high percentage of firms remaining in the same quintile (the boxed entries). Second, firms with lower relative efficiency (i.e., in a higher quintile) are more likely to exit, as expected. Third, firms typically have low relative efficiency at entry, with 87.7% in the lowest quintile in the first quarter post entry.

2.6.2 Stage 2: Tests of Learning Mechanisms

Using the log of firm costs, $u_{i,t}$, from our Stage 1 model, we tested the learning rate hypotheses. Summary statistics for the data employed in these tests are presented in Table 2.2.

Table 2.3 presents the results, split in two horizontal parts. Panel B represents the cost model of Equation 2.8. Panel A represents the experiential learning model of Equation 2.9. These equations are jointly estimated.

Before turning to the hypothesized effects, we will focus attention on three baseline models (Table 2.3, Models 1-3). Model 1 (panel B) isolates the control variables. Bank size effects (branch_scale and branch_count) are negative and significant, indicating that larger banks learn faster. Though high correlations preclude interpreting corporate effects individually (i.e., *holding_company*, *hc_certificates*, *hc_branches*,

hc_states), the net effect of holding-company variables is reduced cost, which is consistent with learning from corporate parents (because we control for scale). Market structure, measured by the Herfindahl index (HHI), is negative and significant, implying that higher market power (lower competition) enables higher revenue per unit cost. Market size effects (population and permits) are highly correlated. We considered the net effects of population and housing starts in Model 1 and find that costs generally decrease with demand, consistent with the explanation that firms learn from the increased output.

We examined experiential learning in Table 2.3, Model 2 (panel A) with the inclusion of cumulative experience. The negative and significant coefficient on the experience term (the constant term in panel A) supports the learning curve expectation that costs decrease with experience.

We also estimated two firm-level random effects: ν, which reflects a firm-specific rate of learning from experience, and α, which reflects a firm-specific cost intercept. For each, we estimated the standard deviation, which is significant at the $p < 0.001$ level. We also estimated $\rho_{\nu\alpha}$, the correlation between ν and α, which is positive and significant, reflecting the intuition that efficient firms proceed more rapidly along their learning curves. We examined vicarious learning (unaided by experiential learning) in Table 2.3, Model 3 (panel B), with the inclusion of external knowledge. The coefficient on $k_{i,t}$ is negative and significant, indicating that costs are lower in more knowledge-abundant contexts.[9]

---

[9]In Model 3, which includes an external knowledge term in the vicarious learning model (i.e., Table 2.3, panel B), HHI turns positive. As seen later in the models separating entrants from incumbents (Table 2.4), this change is driven by the entrant subsample and is consistent with the possibility that entrants require time to adjust their practices to respond to more- intense competition (Knott and Posen, 2005).

Table 2.2: Stage 2 Data Summary

	Panel A[a]			
	Mean	SD	Min	Max
(1) $u_{i,t}$	0.159	0.153	0.008	4.591
(2) $g(u_{i,t})$	−1.958	0.639	−4.884	4.581
(3) entrant	0.079	0.270	0.000	1.000
(4) $x_{i,t}$	5.326	0.995	0.000	6.763
(5) $k_{i,t}$	−0.162	0.056	−0.765	−0.076
(6) $p_{i,t}$	0.178	0.382	0.000	1.000
(7) HHI	0.059	0.058	0.006	0.825
(8) branch_count	5.843	27.285	1.000	2,215.000
(9) branch_scale	36.760	210.393	0.029	17,472.773
(10) holding_company	0.304	0.460	0.000	1.000
(11) hc_certificates	3.507	10.269	0.000	87.000
(12) hc_branches	33.930	140.092	0.000	3,244.000
(13) hc_states	0.333	1.342	0.000	15.000
(14) population	6,160.359	5,099.057	453.690	32,987.676
(15) permits	31,681.281	35,979.476	0.000	3.15e+05

	Panel B														
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
(1) $u_{i,t}$	1.00														
(2) $g(u_{i,t})$	0.87	1.00													
(3) entrant	0.16	0.14	1.00												
(4) $x_{i,t}$	−0.27	−0.26	−0.81	1.00											
(5) $k_{i,t}$	−0.33	−0.43	−0.20	0.28	1.00										
(6) $p_{i,t}$	0.07	0.10	−0.06	0.05	−0.02	1.00									
(7) HHI	0.12	0.14	0.19	−0.17	−0.37	−0.04	1.00								
(8) branch_count	−0.00	−0.00	−0.03	0.06	−0.09	−0.00	0.16	1.00							
(9) branch_scale	0.01	0.01	−0.01	0.02	−0.03	0.00	0.03	0.01	1.00						
(10) holding_company	0.01	0.03	−0.03	0.03	0.01	0.02	−0.02	0.14	0.05	1.00					
(11) hc_certificates	0.06	0.08	−0.04	0.03	−0.04	0.03	−0.05	0.06	0.03	0.52	1.00				
(12) hc_branches	0.04	0.06	−0.02	0.03	−0.07	0.04	0.07	0.23	0.04	0.37	0.63	1.00			
(13) hc_states	0.05	0.06	−0.01	0.04	−0.04	0.04	0.07	0.22	0.05	0.38	0.71	0.79	1.00		
(14) population	0.12	0.14	0.17	−0.16	−0.32	−0.04	0.33	0.12	0.06	0.00	−0.02	0.04	−0.01	1.00	
(15) permits	0.15	0.18	0.20	−0.26	−0.41	−0.05	0.27	0.09	0.01	0.01	0.03	0.07	−0.00	0.73	1.00

[a]$N = 533,165$.

2.6.2.1 Tests of the Intersection of Experiential and Vicarious Learning

Hypothesis 1 is tested in Table 2.3, Model 4 (panel A), with the inclusion of external knowledge $k_{i,t}$ in the experiential learning model. The coefficient estimate is −0.37. It is significant at the $p < 0.001$ level. For brevity, in the remainder of the analysis, statistical significance is at the $p < 0.001$ level unless otherwise noted. This negative coefficient estimate supports the contention of Hypothesis 1A that firms learn more rapidly from experience when the pool of external knowledge is larger.

Table 2.3: Experience and Problems as Facilitators of Vicarious Learning

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Panel A: Experiential Learning Model of Equation (2.9)						
$k_{i,t}$ (external knowledge)				-3.679e-01*** (1.814e-02)	-3.555e-01*** (1.799e-02)	-3.363e-01*** (1.799e-02)
$p_{i,t}$ (problem)					1.908e-02*** (2.050e-04)	9.549e-04 (6.551e-04)
$p_{i,t}k_{i,t}$						-1.142e-01*** (3.920e-03)
σ_v		3.338e+00*** (2.702e-02)	3.214e+00*** (2.556e-02)	3.215e+00*** (2.553e-02)	3.188e+00*** (2.524e-02)	3.186e+00*** (2.522e-02)
Constant term (i.e. estimate of δ_0)		-4.183e-01*** (3.221e-02)	-2.627e-01*** (3.103e-02)	-3.056e-01*** (3.111e-02)	-3.410e-01*** (3.085e-02)	-3.398e-01*** (3.084e-02)
Panel B: Cost Model of Equation (2.8)						
$k_{i,t}$			-4.159e+00*** (1.867e-02)	-2.244e+00*** (9.628e-02)	-2.244e+00*** (9.550e-02)	-2.233e+00*** (9.543e-02)
hhi	-4.488e-01*** (1.979e-02)	-1.518e-01*** (2.190e-02)	1.916e-01*** (2.101e-02)	1.942e-01*** (2.100e-02)	1.839e-01*** (2.084e-02)	1.805e-01*** (2.082e-02)
branch_count	-8.846e-04*** (4.242e-05)	-7.581e-04*** (4.481e-05)	-7.143e-04*** (4.292e-05)	-7.027e-04*** (4.291e-05)	-6.722e-04*** (4.258e-05)	-6.721e-04*** (4.254e-05)
branch_scale	-2.650e-04*** (6.043e-06)	-2.575e-04*** (6.142e-06)	-2.530e-04*** (5.885e-06)	-2.530e-04*** (5.883e-06)	-2.488e-04*** (5.837e-06)	-2.479e-04*** (5.833e-06)
holding_company	-6.677e-02*** (2.333e-03)	-1.178e-02*** (2.535e-03)	-2.216e-04 (2.426e-03)	-7.305e-04 (2.426e-03)	-2.699e-03 (2.406e-03)	-2.974e-03 (2.405e-03)
hc_certificates	2.254e-03*** (1.443e-04)	7.038e-04*** (1.525e-04)	1.849e-04 (1.460e-04)	2.009e-04 (1.460e-04)	1.528e-04 (1.448e-04)	1.641e-04 (1.447e-04)
hc_branches	3.016e-05** (9.899e-06)	1.157e-04*** (1.055e-05)	1.239e-04*** (1.010e-05)	1.258e-04*** (1.009e-05)	1.073e-04*** (1.002e-05)	1.061e-04*** (1.001e-05)
hc_states	-7.911e-03*** (1.168e-03)	-6.543e-03*** (1.272e-03)	1.298e-03 (1.218e-03)	1.169e-03 (1.218e-03)	1.324e-03 (1.208e-03)	1.565e-03 (1.207e-03)
population	-1.576e-05*** (8.498e-07)	7.140e-06*** (1.806e-06)	3.048e-05*** (1.752e-06)	2.999e-05*** (1.755e-06)	3.065e-05*** (1.743e-06)	3.007e-05*** (1.741e-06)
permits	4.128e-08 (3.921e-08)	-1.310e-06*** (5.125e-08)	-7.794e-07*** (4.910e-08)	-7.951e-07*** (4.909e-08)	-6.673e-07*** (4.872e-08)	-6.070e-07*** (4.873e-08)
σ_α	3.819e-01*** (3.754e-04)	1.919e+01*** (1.568e-01)	1.850e+01*** (1.484e-01)	1.850e+01*** (1.483e-01)	1.835e+01*** (1.466e-01)	1.835e+01*** (1.465e-01)
$\rho_{v\alpha}$		4.017e-02*** (4.153e-04)	4.082e-02*** (4.228e-04)	4.092e-02*** (4.239e-04)	4.104e-02*** (4.246e-04)	4.099e-02*** (4.239e-04)
Constant	-1.739e+00*** (7.618e-03)	3.271e-01+ (1.841e-01)	-1.463e+00*** (1.776e-01)	-1.248e+00*** (1.779e-01)	-1.062e+00*** (1.765e-01)	-1.049e+00*** (1.765e-01)
Log-likelihood	-271799.6	-201157.2	-177435.3	-177232.9	-172945.1	-172526
Observations	533165	533165	533165	533165	533165	533165

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

The test of Hypothesis 1B is given in Table 2.3, Models 5 and 6 (panel A), which includes problems, $p_{i,t}$, and their interaction with external knowledge, $p_{i,t}k_{i,t}$, in the experiential learning model. Here, $p_{i,t}$ is defined as two consecutive declining quarters of loan output (we test other definitions of $p_{i,t}$ in the robustness analysis). The coefficient on $p_{i,t}$ is positive and significant in Model 5. This implies, erroneously, that a problem decreases the rate of experiential learning. Of course, our theory suggests a model of experiential learning that does not account for the interaction of problems and external knowledge is mis-specified, as it seems unlikely that firms engage in experiential learning without drawing on external knowledge to solve problems.

In Model 6, the coefficient estimate on $p_{i,t}k_{i,t}$ is -0.114 and significant, while the coefficient on $p_{i,t}$ is now non-significant. At the mean level of external knowledge, the net contribution of a problem to experiential learning is significant at -0.0676. This supports Hypothesis 1B, that problems increase the extent to which firms utilize external knowledge in experiential learning.

2.6.2.2   Tests of Entrant vs. Incumbent Learning

Hypothesis 2 is tested in Table 2.4. We split the data into entrant and incumbent subsamples to examine the possibility that entrants possess a distinct advantage in employing external knowledge to solve problems that emerge in the experiential learning process.[10]

---

[10]We used a split entrant/incumbent sample to facilitate explication of the results. In additional robustness analysis, we estimate the Hypothesis 2 tests using the full sample with an entrant dummy interaction. Results are robust.

Table 2.4: Entrant and Incumbent Learning

	Entrant Models				Incumbent Models			
	(1e)	(2e)	(3e)	(4e)	(1i)	(2i)	(3i)	(4i)
Panel A: Experiential Learning Model of Equation (2.9)								
$k_{i,t}$ (external knowledge)		-4.611e-01***	-4.506e-01***	-4.004e-01***		-1.024e-01***	-9.869e-02***	-9.023e-02**
		(9.819e-02)	(9.737e-02)	(9.769e-02)		(2.963e-02)	(2.935e-02)	(2.934e-02)
$p_{i,t}$ (problem)			8.173e-02***	3.502e-02***			1.811e-02***	3.353e-03***
			(2.806e-03)	(8.218e-03)			(1.898e-04)	(6.129e-04)
$p_{i,t}k_{i,t}$				-2.208e-01***				-9.345e-02***
				(3.652e-02)				(3.690e-03)
σ_v	5.118e-01***	5.098e-01***	5.054e-01***	5.058e-01***	3.592e+00***	3.592e+00***	3.559e+00***	3.557e+00***
	(1.133e-02)	(1.128e-02)	(1.118e-02)	(1.118e-02)	(2.971e-02)	(2.971e-02)	(2.933e-02)	(2.930e-02)
Constant term	-5.394e-01***	-6.295e-01***	-6.416e-01***	-6.316e-01***	2.848e-02	1.560e-02	-2.596e-02	-2.672e-02
(i.e. estimate of δ_0)	(1.372e-02)	(2.347e-02)	(2.328e-02)	(2.333e-02)	(3.674e-02)	(3.693e-02)	(3.659e-02)	(3.657e-02)
Panel B: Cost Model of Equation (2.8)								
$k_{i,t}$	-2.979e+00***	-1.659e+00***	-1.575e+00***	-1.622e+00***	-4.261e+00***	-3.699e+00***	-3.654e+00***	-3.604e+00***
	(8.663e-02)	(2.933e-01)	(2.910e-01)	(2.910e-01)	(1.827e-02)	(1.637e-01)	(1.622e-01)	(1.621e-01)
hhi	1.624e+00***	1.661e+00***	1.645e+00***	1.640e+00***	-1.912e-01***	-1.910e-01***	-2.022e-01***	-2.050e-01***
	(7.496e-02)	(7.530e-02)	(7.459e-02)	(7.456e-02)	(2.156e-02)	(2.156e-02)	(2.136e-02)	(2.135e-02)
branch_count	4.978e-04+	5.243e-04+	5.900e-04*	5.752e-04*	-6.793e-04***	-6.770e-04***	-6.467e-04***	-6.465e-04***
	(2.894e-04)	(2.897e-04)	(2.869e-04)	(2.867e-04)	(4.084e-05)	(4.084e-05)	(4.047e-05)	(4.045e-05)
branch_scale	-2.779e-03***	-2.784e-03***	-2.678e-03***	-2.678e-03***	-2.311e-04***	-2.310e-04***	-2.271e-04***	-2.264e-04***
	(1.053e-04)	(1.054e-04)	(1.047e-04)	(1.046e-04)	(5.452e-06)	(5.452e-06)	(5.403e-06)	(5.400e-06)
holding_company	6.612e-02***	6.774e-02***	6.862e-02***	6.837e-02***	-8.189e-03***	-8.224e-03***	-1.021e-02***	-1.045e-02***
	(1.629e-02)	(1.630e-02)	(1.615e-02)	(1.614e-02)	(2.308e-03)	(2.308e-03)	(2.287e-03)	(2.285e-03)
hc_certificates	2.928e-03**	3.013e-03**	3.131e-03**	3.229e-03**	1.705e-04	1.709e-04	1.281e-04	1.337e-04
	(1.135e-03)	(1.136e-03)	(1.125e-03)	(1.125e-03)	(1.383e-04)	(1.383e-04)	(1.371e-04)	(1.370e-04)
hc_branches	1.958e-04**	2.009e-04***	1.447e-04*	1.427e-04*	1.398e-04***	1.400e-04***	1.222e-04***	1.214e-04***
	(5.979e-05)	(5.982e-05)	(5.926e-05)	(5.924e-05)	(9.698e-06)	(9.698e-06)	(9.612e-06)	(9.605e-06)
hc_states	2.885e-02***	2.820e-02+	2.907e-02***	2.863e-02***	-2.145e-03+	-2.157e-03+	-2.039e-03+	-1.816e-03
	(7.435e-03)	(7.439e-03)	(7.367e-03)	(7.364e-03)	(1.170e-03)	(1.170e-03)	(1.160e-03)	(1.159e-03)
population	5.164e-05***	5.388e-05***	5.304e-05***	5.253e-05***	-1.704e-05***	-1.723e-05***	-1.679e-05***	-1.721e-05***
	(2.695e-06)	(2.739e-06)	(2.727e-06)	(2.726e-06)	(1.771e-06)	(1.772e-06)	(1.756e-06)	(1.753e-06)
permits	-2.842e-06***	-2.716e-06***	-2.428e-06***	-2.370e-06***	-1.718e-07**	-1.703e-07**	-3.302e-08	1.836e-08
	(1.435e-07)	(1.461e-07)	(1.451e-07)	(1.453e-07)	(5.260e-08)	(5.260e-08)	(5.214e-08)	(5.215e-08)
σ_α	1.157e+00***	1.154e+00***	1.151e+00***	1.151e+00***	2.076e+01***	2.076e+01***	2.057e+01***	2.056e+01***
	(2.274e-02)	(2.266e-02)	(2.256e-02)	(2.256e-02)	(1.728e-01)	(1.728e-01)	(1.706e-01)	(1.704e-01)
$\rho_{v\alpha}$	3.533e-01***	3.604e-01***	3.608e-01***	3.603e-01***	3.012e-02***	3.013e-02***	3.013e-02***	3.009e-02***
	(1.074e-02)	(1.109e-02)	(1.109e-02)	(1.107e-02)	(3.220e-04)	(3.222e-04)	(3.214e-04)	(3.209e-04)
Constant	-1.499e+00***	-1.271e+00***	-1.254e+00***	-1.262e+00***	-2.638e+00***	-2.567e+00***	-2.344e+00***	-2.324e+00***
	(3.936e-02)	(6.327e-02)	(6.289e-02)	(6.288e-02)	(2.112e-01)	(2.121e-01)	(2.103e-01)	(2.101e-01)
Log-likelihood	-35045.18	-35035.79	-34620.92	-34605.04	-117363.6	-117360.2	-112856.8	-112540.9
Observations	42207	42207	42207	42207	490958	490958	490958	490958

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

The model numbers in Table 2.3 are appended with an "e" (entrant) and an "i" (incumbent). For ease of exposition, we drop the e and i when we refer to both

subsamples.

In Table 2.3, Model 1, we examined the effects of external knowledge on cost efficiency (in panel B), which accounts for vicarious learning unaided by experience. The coefficient on this external knowledge is 2.98 for entrants and 4.26 for incumbents. The two coefficients and their difference are significant. As such, in learning from external knowledge, incumbents learn more than entrants when unaided by experiential learning.

Hypothesis 2A is tested in Table 2.3, Model 2 (panel A), with the inclusion of external knowledge $k_{i,t}$ in the experiential learning model. This coefficient is -0.461 for entrants, which is larger (i.e., more negative) than the corresponding coefficient of -0.102 for incumbents. A Wald test indicates the difference between the two is significant, supporting Hypothesis 2A that entrants' experiential learning rate increases more than incumbents' when the pool of external knowledge is larger.

To provide a sense of the magnitude of the contribution of external knowledge to entrants versus incumbents, in Table 2.5 we report the marginal effect, $M_{i,t}$, on vicarious learning of a unit change in external knowledge $k_{i,t}$ (i.e., $\partial g(u_{i,t})/\partial k_{i,t}$) based on Model 2 in Table 2.4. Following our model and results, we decomposed this marginal effect into that unaided by experience (i.e., the $k_{i,t}$ term in Equation 2.8) and that absorbed in the process of experiential learning (i.e., the $k_{i,t}$ term in Equation 2.9). We see that the marginal effect of vicarious learning unaided by experience is 122% larger (i.e., the inverse of the 0.45 entrant-to-incumbent ratio in the table) in magnitude for incumbents than for entrants (-3.70 versus -1.66). This result conforms to the conventional wisdom that entrants are likely to benefit less from vicarious learning because, relative to incumbents, they lack prior related knowledge. Conversely, the marginal benefit of vicarious learning that arises in the process of experiential learning is 110% larger for entrants than for incumbents (-1.20 versus -0.57). In total, the contribution of external knowledge for incumbents is 49% larger than for entrants

32

(-4.27 versus -2.86). Finally, we find that for entrants, 42% of the benefit they achieved from external knowledge came as a by-product of the experiential learning process, compared to 13% of the benefit for incumbents.

Table 2.5: Marginal Effect on Vicarious Learning from a Unit Change in External Knowledge

	Marginal benefit of vicarious learning			
	Unaided by experience	Facilitated by experience	Total	Facilitated by exp. as share of total
Entrant	-1.66	-1.20	-2.86	0.42
Incumbent	-3.70	-0.57	-4.27	0.13
Entrant/ incumbent	0.45	2.10	0.67	3.14

In Table 2.4, Model 3, we examined the effects of problems on cost efficiency. The coefficient estimate on $p_{i,t}$ is significant at 0.0817 for entrants and 0.0181 for incumbents. Hypothesis 2B is tested in Table 2.4, Model 4, which includes problems interacted with external knowledge, $p_{i,t}k_{i,t}$, in the experiential learning model. The coefficient estimate of $p_{i,t}k_{i,t}$ is -0.221 for entrants and -0.093 for incumbents. Both are significant, and a Wald test indicates the difference between the two is also significant. The net contribution of a problem to experiential learning rates, at the mean level of external knowledge, is -0.098 for entrants and -0.053 for incumbents. Thus, the contribution of problems to accessing external knowledge in the experiential learning process is twice as large for entrants than for incumbents. This supports Hypothesis 2B, that entrants benefit more than incumbents from using external knowledge to solve problems encountered in the experiential learning process.

To illustrate the role of external knowledge in the experiential learning process, in Figure 2.2 we plotted entrant learning curves across knowledge contexts. We used the estimates in Table 2.4, Model 4 to generate predicted median costs for entrants relative to incumbents. We focused on the impact of being in a state in the top or bottom 10th percentile of external knowledge. The graph highlights the observation

Figure 2.2: Predicted Entrant Learning Curves in Different Knowledge Contexts
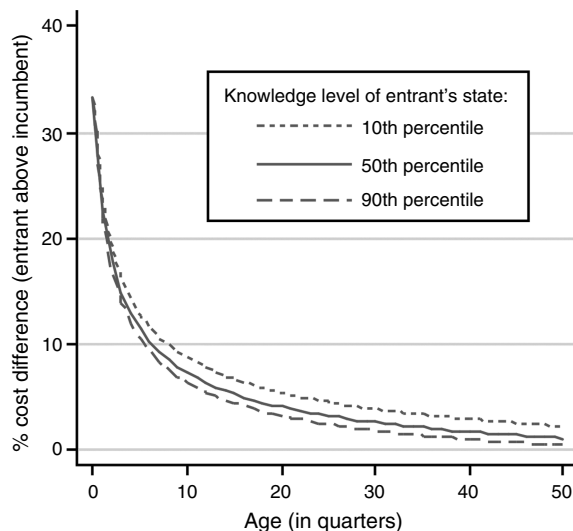
that a 10-quarter-old entrant firm in a state in the 90th percentile for knowledge is developmentally equivalent to 15-quarter-old firm in a 10th percentile state.

2.6.2.3 Robustness Analysis

We conducted extensive sensitivity analyses to rule out two broad sets of concerns, the first reflecting alternative measures and estimation techniques and the second reflecting alternative explanations. A brief overview of the sensitivity analyses follows (detailed analyses available on request). Our results are robust to each of these sensitivity tests.

In one set of analyses, we tested the sensitivity of our results to alternative measures and estimation techniques: (1) To address the possibility that the error terms in the entrant and incumbent models of our study contain common shocks that bias the estimates, we constructed a model that jointly estimates entrant and incumbent learning. (2) Firm output is often used in the organizational learning literature as a proxy for organizational knowledge (Darr et al., 1995; Ingram and Simons, 2002). Accordingly, we examined a model that uses cumulative loan output (rather than cost) to construct a measure of external knowledge and a model that uses cumulative

34

loan output (rather than time) to construct a measure of own experience. (3) Our definition of a problem reflects a firm's own past performance, but prior research has also defined problems relative to others' performance (Greve, 1998). Hence, we examined a market share-based definition of a problem, which reflects not only own firm past performance but also the past performance of other firms.

In the other set of sensitivity analyses, we addressed alternative explanations that may account for our estimated results. The alternative explanations reflect the following issues: (4) Incumbents may exhaust the external knowledge pool and, therefore, exhibit reduced returns to seeking external knowledge. We observed that incumbents actually learn more overall from external knowledge than do entrants. Moreover, firms founded earlier do not exhibit reduced vicarious learning, in part because the pool of external knowledge is growing over time. (5) Entrants in more knowledge-rich markets may be superior learners, independent of the size of the pool of external knowledge. To examine this possibility, we ran models that included controls for market knowledge at founding, as well as state and Consolidated Metropolitan Statistical Area fixed effects. (6) The time pattern of entrant learning may not reflect an intrinsic characteristic of entrants, but rather the fact that entrants are initially inefficient and small. To examine this possibility, we compared entrants' learning from experience to small, inefficient incumbents' learning from experience. (7) Inferior entrants may exit rapidly, biasing the estimation results. To examine this possibility, we employed a Heckman selection model to account for exit. (8) Firms may be heterogeneous in resources available to engage in the absorption of external knowledge. To examine this possibility, we controlled for return on assets as a proxy for resource availability. (9) The estimates may capture other mechanisms, such as mean regression, because we do not directly observe vicarious learning (e.g., a specific piece of knowledge transferred between firms). To examine this possibility, we observed that vicarious learning should be uniquely sensitive to the proximity

and similarity of those from which one learns. We developed additional measures, including the distance to a firm's nearest 30 neighbors, as well as similarity in total assets, employee count, and cost.

2.7 Discussion

We started with a puzzle. Entrants are often viewed as suffering from a liability of newness because at founding, they rarely possess the knowledge and capabilities necessary to compete and survive. They can overcome this liability by learning directly from their own experience and vicariously from the knowledge of incumbents. But how can entrants learn from external knowledge when they lack the prior related knowledge that forms the basis of absorptive capacity?

Our solution to this puzzle started with the observation that experiential learning and vicarious learning, although often treated as distinct processes in the organizational learning literature, are in fact interdependent processes. The process of experiential learning is fundamentally one of problem solving (Arrow, 1962; March and Olsen, 1975), and many solutions to problems are stimulated by learning about what other firms do. We theorized and empirically demonstrated that problems encountered in the process of experiential learning are an important trigger of the search for external knowledge. As such, vicarious learning occurs as a by-product of the ongoing process of experiential learning. Further, we argued and demonstrated empirically that this effect is particularly strong for entrants, because they lack the prior knowledge with which to solve problems encountered in the process of engaging in productive activity.

This paper contributes to our understanding of vicarious learning. Influential research on absorptive capacity (Cohen and Levinthal, 1990), and the broader research agenda on organizational capabilities, has focused significant research attention on the ability to learn vicariously. But the ability to learn vicariously and the size of

the external knowledge pool taken together are insufficient to explain the extent of vicarious learning. Our study suggests that the impetus to exploit external knowledge deserves more attention than it has so far received.

Recent research has begun to address this gap. Two ideas have surfaced in the literature. Baum and Dahlin (2007) argued that firms shift their reliance from external to internal knowledge depending on the proximity of recent performance to aspiration levels. Simon and Lieberman (2010) argued that firms are less likely to seek external knowledge when internal knowledge is available. The former reflects the process by which firms identify a problem (requiring a solution), and the latter reflects the extent to which internal knowledge is available to solve the problem. Our theory extends these two ideas by arguing that the experiential learning process generates the problems, and the lack of prior related knowledge drives firms to seek outside knowledge for solutions – with both mechanisms particularly salient for new entrants.

Our theory implies that prior related knowledge plays two interdependent roles in experiential learning. On one hand, prior knowledge engenders the ability to absorb external knowledge. On the other hand, prior knowledge reduces the need to use external knowledge to solve problems. The net effect of prior knowledge on external knowledge acquisition depends on which of these roles is more important in a given context. Absorptive capacity is central to vicarious learning in many technical and scientific domains. Yet in other domains, the issue may not be absorption, but rather that external knowledge may be too vast and diffused among firms to be relevant absent problems to which it may be applied.

Consider the following example in which the impetus to get external knowledge dominates the role of absorptive capacity. We interviewed the chief executive officer of an entrant bank we call Bank Nouveau. He recounted an early problem in the unexpectedly high number of employees needed to analyze credit. The bank initially sought solutions to this problem within the firm. However, with limited

internal knowledge, the problem soon triggered a search for external solutions. In social conversation with executives from another bank, the founder of Bank Nouveau discovered that its loan analysis reports were far more complex than necessary because they included extraneous information. By assimilating external knowledge in the experiential learning process, Bank Nouveau reconfigured its loan analysis reports to reduce cost and improve quality. A key feature of this example is that Bank Nouveau did not require significant absorptive capacity to learn vicariously (beyond that held by any competent bank manager). Instead, all that was required was a specific problem. The problem both generated the impetus to get external knowledge and directed the search for external knowledge. In such situations, a lack of prior related knowledge enhances, rather than reduces, the extent of vicarious learning.

This paper also contributes to our understanding of experiential learning. In particular, the theory and results inform our understanding of the sources of heterogeneity in experiential learning rates, which is known to occur across firms and also across new plants in the same firm (Argote and Epple, 1990). Existing explanations for heterogeneity in learning rates center on organizational factors such as R&D and capital intensity (Lieberman, 1984), firm specialization (Schilling et al., 2003), and employee task environments (Wiersma, 2007). In contrast, we focus our attention on the knowledge environment in which a firm is embedded – organizations in more knowledge-rich contexts learn more rapidly from their own experience.

Whereas the literature on experiential learning is extensive, learning by new entrants has received relatively little attention. Yet, as Ethiraj et al. (2005, p. 28) argued, "Productive knowledge ...(is)... a result of endogenous learning-by-doing processes." To our knowledge, our study is the first to construct and examine learning curves for a complete census of new entrants in an industry. Our results demonstrate just how rapidly entrants learn (in commercial banking) – efficiency improved by 30% over the first two years. Our study also points out that the process of experiential

learning is differentially impacted by the knowledge context in which it occurs – this knowledge context is much more salient for entrants than it is for incumbents.

This study is not without limitations. First, our data are uniquely comprehensive at the population level, such that we are able to conduct a detailed analysis of learning and observe high-level problems and the outcomes of solving them (i.e., efficiency gains). However, we do not have data on the actual problem-solving event. That is, we do not observe the interfirm knowledge flows that engender solutions to problems arising in the experiential learning process. We address this with broad and deep sensitivity analyses and extensive interviews with bank founders. Future research may fruitfully examine problem solving at the intersection of experience and external knowledge more directly. Second, the U.S. commercial banking context is well suited for this study because, like most industries, banking is neither completely static nor highly turbulent. As such, experiential learning, rather than R&D, is an important driver of performance improvements. However, more research is needed to understand the role of entrant learning in contexts marked by radical innovation and substantial investments in R&D.

In sum, post entry learning is an important determinant of entrant performance. The efficacy of such learning is predicated not only on internal organizational processes but also on the knowledge environment in which an organization is embedded. The issues surrounding learning by new entrants, and the relationship between experiential and vicarious learning, remain a fertile and important line of inquiry for organizational theorists and strategy scholars.

CHAPTER III

Exploiting the Exploration of Others: Learning through Collaboration in the Wireless Telecommunications Industry

3.1 Introduction

In many industries, particularly those that are technology-intensive, substantial economic activity surrounds multi-firm, innovation-oriented collaborations such as R&D consortia (Sakakibara, 1997) and standardization of emerging technologies (David and Shurmer, 1996). In these contexts, collaborating firms typically hold knowledge complementary to that of a given firm, and, as such, knowledge exchange may be central to firms realizing potential gains from collaboration (Zeng and Chen, 2003). However, prior work suggests firms have little reason to contribute knowledge to this end. Indeed, scholars have observed that inter-firm conflict and concerns over free-riding or unintended knowledge leakage dissuade firms from contributing (Kale et al., 2000; Oxley and Sampson, 2004; Greenstein, 1992; Hardin, 1982), while other researchers argue firms contribute largely to facilitate the market adoption of their own knowledge (West, 2001; Shapiro and Varian, 1999). In this study, we propose that firms contribute knowledge to achieve complementarities with the knowledge held by other participating firms. In particular, a firm's knowledge contributions facilitate

the absorption of collaborators' knowledge and in turn enhance the firm's subsequent performance. We call this a *collaboration-specific absorptive capacity* (CSAC).

A large management literature has developed around Cohen and Levinthal's (1990; 1989) observation that a firm's ability to absorb external knowledge is largely a function of its internal stocks of related knowledge. Subsequent studies have examined in finer detail the capabilities comprising absorptive capacity, such as knowledge assimilation or knowledge exploitation (e.g., Zahra and George, 2002; Todorova and Durisin, 2007; Lane and Lubatkin, 1998). Though these studies have extended Cohen and Levinthal's earlier work in several important ways, the basic concept remains the same: a firm's internal knowledge (or capabilities) facilitates external knowledge absorption through internal processes, isolated from the external knowledge that is potentially absorbed.[1]

Within a multi-firm collaboration, however, firms are not limited to using their knowledge to drive internal processes of external knowledge absorption. Instead, the knowledge a firm contributes to a collaboration may also drive absorption of external knowledge, albeit through a more instrumental and targeted process than theorized in extant conceptions of absorptive capacity. Under the process of building CSAC developed in this study, a firm's knowledge contributions provide a context with which to engage in *knowledge integration* activities fundamental to the development of valuable collective knowledge. Moreover, some of the learning from this integration is private to a firm because, as we will theorize, the knowledge contributed to a collaboration is typically interdependent with other private knowledge residing within the firm. Hence, knowledge contributions enable a firm to utilize the output of collective knowledge integration efforts, which in turn enables the firm to better utilize its interdependent private knowledge.

---

[1]Dyer and Singh (1998) theorize on partner-specific absorptive capacity and two-way learning through inter-organizational routines. In contrast, the absorption in this study is not partner-specific and is one-way in that a recipient firms' knowledge contributions are theorized to enable vicarious learning beyond that achievable by other collaborators who contribute less.

This study provides an important complementary explanation to prior work that examines more broadly the provisioning of collective knowledge (or other good). In particular, prior work tends to assume that potentially provisioned knowledge elements are independent of one other, which leads to conflict when these elements are distributed across multiple firms and there are more elements than can be adopted in a given text. Alternatively, other work focuses on a single piece of knowledge that many firms (or individuals) are equally adept at provisioning, which leads to free-riding problems. In contrast, we argue that knowledge, particularly in systemic technologies comprising multiple components (Winter, 1987; Teece, 1986), tends to have many distinct components that are interdependent and diffused across many firms. For brevity, we say that this knowledge exhibits *distributed interdependence.* The key implication of distributed interdependence is that firms contribute towards knowledge integration at the group level in order to facilitate external knowledge absorption at the individual level.

While knowledge contributions may benefit a firm by facilitating external knowledge absorption, they also have the downside of potentially revealing valuable information on a firm's strategy and intent to partnering firms. Unlike the internal knowledge, capabilities, and processes that underlie prior conceptions of absorptive capacity, knowledge contributions are freely observable to collaborators. Thus, high levels of knowledge contribution by a firm may reveal to other firms its high expectations of the value embodied in the collaborative output. In this manner, knowledge contributions may uncover elements of a firm's technology strategy – that is, information that the firm is "placing its bets" in the technological areas underlying the collaboration. Accurate signals of this form may lead other firms to re-allocate additional effort to this more fruitful direction, thereby increasing subsequent competition in these technological areas. Moreover, an inevitable side effect of knowledge contributions is that it contains information about a firm's technological approach

within the collaboration and therefore give cues on the capabilities that best enable a firm to utilize knowledge produced from the collaboration. Signals of this type are particularly beneficial when there are many firms collaborating on technologies, in which case firms may have access to a wide variety of external knowledge contributed by other collaborators but may face considerable difficulty in discerning which from among these best facilitates innovation.

As the preceding discussion suggests, knowledge contributions have both a positive and negative effect on the value of collaborative output towards a firm's subsequent innovative activity. Hence, we predict a firm's knowledge-contributing efforts increase and then decrease the performance of its *follow-on innovations* that build from the knowledge output of the collaboration. The positive effects – that is, the CSAC that enables a firm to understand the implications of this knowledge output – dominate at low levels of effort. In contrast, the negative effect – revealing a focal firm's innovation strategy – dominates at high levels of effort.[2]

The focus of this study is on multi-firm rather than dyadic collaborations for two reasons. First, mechanisms outside of those theorized in this study tend to exert a strong influence on participation, and its nature thereof, in a dyadic alliance, such as trust (Gulati, 1995; Das and Teng, 1998), or contractual mechanisms (Mowery et al., 1996). This likely reflects both (a) an ability to implement such mechanisms in a dyadic versus a multi-party collaboration – for example, mutual trust may be difficult to foster among many firms – and (b) the importance of each firm's participation in a dyadic collaboration – if one firm chooses not to contribute, then there is effectively no collaboration. Second, the proposed downside of knowledge contribution – that is, revealing a firm's intentions or strategy – is not applicable to dyadic collaborations

---

[2]The assumption embedded within this negative performance effect is, as in much of the innovation literature, that firms' innovations compete with one another. If an innovation's performance is assessed using a non-competitive measure, such as a functional attribute (e.g., gas mileage), other innovations would have no effect on performance. If however (as in this study), innovation performance is competitive (as is likely the case with patent citation counts), performance should decline with competing innovations.

since a firm's contribution and effort level are always highly visible when there is only one other partner.

This study adds to our understanding of the factors influencing the extent to which firms contribute knowledge to a multi-firm collaboration. In particular, a theory is advanced wherein firms contribute knowledge to build CSAC, which in turn serves to increase subsequent performance. The knowledge underlying the absorption process theorized here is openly observable to collaborating firms (and often to non-collaborating firms), which implies that, unlike in most prior research, there is a potential drawback to building absorptive capacity too rapidly.

This chapter is organized as follows. In the next section, we review prior work on collective goods provisioning and absorptive capacity. From this backdrop of prior work, we develop the theoretical arguments for CSAC, we elucidate the role of firm heterogeneity to provide additional evidence of this knowledge absorption mechanism, and we also examine the role of heterogeneity on knowledge absorption depth and breadth within a collaboration. Next, we describe the empirical setting for this study: the 3GPP (3rd Generation Partnership Project), for which we have detailed firm-level data on knowledge contributions to many (over 1000) collaborations from 1998 to mid-2012. We then present empirical analysis and results. Finally, we close with a discussion of contributions made by the study.

3.2 Literature Review

We begin with a review of prior research on collective (or public) goods provisioning to provide a context for theorizing on CSAC, which is the private learning benefit gained through contributions to a collective knowledge pool. The main impediments to collective goods provisioning identified in this work are free-riding and conflict. Following this is a brief overview of work on open source software, which serves to highlight the difficulties in multi-firm collaboration more generally by comparing and

contrasting mechanisms that overcome difficulties in this specific context.

3.2.1 Under-provisioning of Collective Goods Due to Free-Riding

Mancur Olson's *The Logic of Collective Action* (1965) anchored a "standard result of the theory of public goods" (Bergstrom et al., 1986, p. 25), whose basic argument is that collective (public) goods – that is, goods that are non-rivalrous and non-excludable, such as clean air or street lights – are under-provisioned through voluntary contributions. Olson demonstrates that under-provisioning occurs because (a) other agents (individuals or organizations, depending on the context) cannot be excluded from utilizing a given agent's contribution to a collective good (i.e., the non-excludability noted above), and (b) contributors receive only a fraction of the benefits of the good. Collective goods are under-provisioned even when potential contributors (e.g., individuals or organizations) have aligned interests and are increasingly under-provisioned when misaligned interests among agents creates conflict. Moreover, under-provisioning tends to be more of an issue – and indeed collective goods might not be provisioned at all – among a large group of organizations (or individuals) (Hardin, 1982; Olson, 1965).[3] In such cases, selective incentives – that is private rewards for actions by individual agents – may be necessary to induce the provisioning of a collective good.

Olson's under-provisioning logic suggests agents may have incentives to engage in *free-riding:* since agents cannot be excluded from using a collective good, they may refrain from providing it in the expectations that another agent will. Free-riding behavior has been demonstrated in numerous experimental studies (e.g., Andreoni,

---

[3]While the detrimental effects of larger size is the more normative view, there is some nuanced debate in the literature. Hardin (1982) argues it is not the number of agents that is most relevant. Rather, it is the minimal subgroup (possibly a single agent) that would benefit from provisioning the good in absence of other agents. Chamberlin's (1974) formal analysis suggests (somewhat common) conditions under which provisioning of collective goods may increase with group size. In particular, when the collective good is a normal economic good and is "inclusive" (i.e., consumption is non-rivalrous), collective goods provisioning is non-decreasing with group size. Note that in this study, a collective good is used as an input to innovations, often among rivals, and as such, is not inclusive.

1988; Isaac and Walker, 1988). Moreover, as Bliss and Nalebuff's (1984) formal model demonstrates, free-riding is especially problematic when agents have little information on the cost or benefits of collective goods to other firms, in which case they engage in a strategic waiting game to benefit from the contributions of others.

3.2.2 Under-provisioning of Collective Goods Due to Conflict

An important notion rarely discussed in prior work is that some types of collective goods, knowledge being a prime example, possess two characteristics that tend to mitigate under-provisioning through free-riding but engenders under-provisioning through conflict. First, there may be more different collective goods that solve a common underlying problem than can be supported by the environment at hand. Second, collective goods can be non-rivalrous in consumption but nevertheless rivalrous in benefits because collective goods (knowledge, for instance) are often embedded in a competitive context. Thus, when agents differ in the collective good that benefits them the most, free-riding diminishes because agents have incentives to support their preferred variant, but conflict increases because agents cannot agree on the collective good(s) to provision.

To illustrate, consider scientific knowledge as a public good. Suppose multiple university scientists each create a distinct piece of knowledge – they each discover and publish a distinct solution to the same scientific problem – but firms cannot adopt all of them (e.g., due to scale economy or compatibility considerations). Knowledge of this sort is not rivalrous in consumption since one firm's use of a given solution does not affect others firms' ability to do so, but it is rivalrous in benefits since firms better at applying a given solution to product markets will gain a competitive advantage. Conflict therefore arises when firms are heterogeneous in the knowledge they are best able to utilize. This scientific knowledge example stands in contrast to traditional conceptions of public goods as embodied by clean air, for example. Clean air is not

46

likely to suffer from conflict-driven under-provisioning because there are no alternative solutions to the underlying problem (a human's need for oxygen), and the benefits of clean air are non-competitive (i.e., people do not compete in their enjoyment of clean air).

The negative effects of conflict on collective goods have been demonstrated in the technology standardization literature (Greenstein and Stango, 2007), though concepts developed in this body of work apply to a broader set of collaborative settings (Farrell and Simcoe, 2012). Formal (Farrell and Saloner, 1988; Farrell and Simcoe, 2012; Simcoe, 2012) , descriptive (Weiss, 1993), and empirical (Simcoe, 2012) results support the basic conclusion that firms with conflicting interests cause under-provisioning (e.g., as delays) in formal standardization efforts. Scholars have also focused considerable attention on how firms overcome this conflict. One oft-cited mechanism is an organization exerting power – gained from mechanisms such as market presence, an installed base of captive consumers, or technological superiority – to force adoption of their preferred variant of a collective good. Examples include Qualcomm's attempts to steer the wireless telephone industry to its CDMA technology (West, 2001), Microsoft's often successful and frequently contentious efforts to control various elements related to its operating system environment such as its internet browser and ActiveX controls, and Sun's de facto control over Java evolution (Shapiro and Varian, 1999). Firms also overcome conflict through some form of compromise. For instance, firms may exchange side payments to align interests, or if collective goods provisioning in a given context is a sequential, ongoing process, firms may engage in logrolling by trading support for each other's interests at different points in time (Shapiro and Varian, 1999).

3.2.3 Provisioning of Open-source Software

Given this substantial theoretical and empirical support for the difficulty in provisioning collective goods, the proliferation of open source software – public domain software developed through voluntary contributions by individuals and firms – is, on the surface, somewhat puzzling. It is particularly surprising given the many open-source projects in which rival firms contribute substantially (West, 2003; Dahlander and Magnusson, 2008). As such, the open source phenomenon has attracted substantial scholarly attention (e.g., Von Hippel and Von Krogh, 2003; Weber, 2004; Lerner and Tirole, 2002, to name a few).

A partial explanation for this puzzle is that open-source is not particularly vulnerable to conflict or free-riding. Two related attributes mitigate conflict. First, open-source software (and all software) is substantively embodied in codified form (e.g., as textual lines of code), in contrast to many other technological outputs in which much of the knowledge is tacit, with perhaps only "blueprints" or a physical product serving to codify the knowledge. Hence, contributors can more readily and definitively assess each others' contributions, which lessens disagreements arising from subjective judgments. Moreover, disparate firms (or individuals) may contribute to open source projects with minimal communication (Hertel et al., 2003), since less tacit knowledge is necessary to integrate among different contributions than would be required in other contexts.[4] Second, open-source software tends to be highly modular so that each contribution is less likely to draw objections from users relying on interdependent components. At a basic level, software is amenable to modularization since it is highly codified, which facilitates partitioning software functions into modules. Moreover, software scientists have devoted considerable effort to developing modular design techniques (CITE), and open source project architects have strong

---

[4]Large open source projects may be coordinated (and governed) by a "beneficial dictator" (e.g., Linus Torvalds for Linux) and a small set of "lieutenants" underneath, which is still minimal relative to the coordination that typifies collaborations in other technological areas.

incentives to modularize the design since contributors are diffuse.

Several mechanisms serve to overcome free-riding in open-source projects. First, modularity in open-source reduces free-riding by providing a structure and incentives to divide effort (Baldwin and Clark, 2006; Lerner and Tirole, 2002). Second, open-source offers individuals incentives to contribute. Open source a particularly attractive vehicle for individuals to signal their capabilities in the job market since contributions are available for scrutiny by other firms (by definition) and are also more likely to have broad rather than specialized application (Lerner et al., 2006). Also, individuals contribute to open source for personal enjoyment because their contributions are of their choosing, in contrast to project assigned to them as employees of a firm, and are therefore more likely to match their interests and abilities (Lakhani and Wolf, 2005). Third, several firm level incentives to contribute have been identified in the open-source and related literature. For instance, firms may have an interest in improving open source projects that are complementary to its products or services (Spencer, 2003; Lerner et al., 2006). Cisco's contributions to open source internet code, such as that found in Linux, potentially enhance the value of its network equipment business. Firms may also identify good programming talent through connections to the open source community (Lerner et al., 2006). Finally, firms may contribute as a strategy maneuver to prevent other firms from appropriating the gains to the contributed knowledge (Pisano, 2006).

Another threat to open source provisioning is firms making proprietary enhancements and commercializing the upgraded software and shutting other contributors out of the continued evolution of the project. On this front, legal innovations have helped combat this form ex-post privatization and, as a result, served to sustain voluntary contributions as open source projects mature (Osterloh and Rota, 2007). In particular, "copyleft" licenses such as the ubiquitous GNU Public License (GPL), require revisions to open source codebases to remain open source.

3.3 Theory and Hypotheses

A central distinction we make from prior work on collective goods provisioning surrounds the nature of knowledge and its implications for how knowledge is provisioned in a collaborative context. This prior research tends to consider knowledge or other collective good as one or more separable components that are sequentially and independently provisioned. For example, the problems of free-riding and conflict noted above are typically assumed to arise from provisioning a single good.[5] Another example is formal models used in work on research consortia and technology standardization that depict either a one-shot (Sakakibara, 2003) or sequential (Simcoe, 2012) game. Qualitative research often focuses attention on collaborators making one key decision – for example, adopting Qualcomm versus another wireless technology standard (West, 2001), or standardizing the width of a country's railway system (Friedlander, 1995).

In contrast, we observe that, particularly in increasingly pervasive systemic technologies (Teece, 1986; Winter, 1987), knowledge held across collaborating firms may be interdependent, which for brevity we call distributed interdependence. Distributed interdependence forms the basis for our theory. In particular, this implies that gains may be realized by combining knowledge held by different collaborators. Moreover, since substantial tacit knowledge typifies innovation-oriented contexts, knowledge integration of this sort requires active collaboration among member firms. Accordingly, we theorize a collaboration-specific absorptive capacity (CSAC) in which a firm contributes knowledge to collaboration-wide knowledge integration efforts, the results of which enables the firm to realize private gains.

---

[5]A rare counterexample is (Henkel, 2004), where sequential provisioning ameliorates free-riding.

3.3.1 Prior Conceptions of Absorptive Capacity vs. Collaboration-Specific Absorptive Capacity

In theorizing CSAC, we build on existing conceptions of absorptive capacity (see Figure 3.1). Much of this research has viewed knowledge outflows (e.g., sharing or spillovers) and inflows (i.e., absorption) as arising through independent processes. In contrast, we theorize a process in which knowledge outflows in the form of voluntary knowledge contributions are the impetus to absorbing external knowledge from collaborating firms.

Cohen and Levinthal (1989; 1990) make the basic proposition that internal stocks of relevant knowledge underlie a firm's ability to absorb external knowledge. Subsequent research has made considerable progress in understanding the micro-mechanisms within the firm that underlie absorptive capacity (see Figure 3.1). For example, Zahra and George (2002) and Todorova and Durisin (2007) decompose the firm capabilities underlying absorptive capacity – identifying external knowledge, recombining new and existing knowledge, and incorporating new knowledge in firm operations – and they theorize social integration mechanisms as an important moderator of a firm's absorptive capacity. Lane and Lubatkin (1998) focus on observable indicators of a firm's internal processes, such as compensation practices and organizational structures, and they argue that similarity in these characteristics influences a recipient firm's ability to absorb knowledge from a source firm. Finally, Posen and Chen (2013) demonstrate, both theoretically and empirically, that the amount of external knowledge a firm absorbs is a function of not just the ability to comprehend and utilize outside knowledge but also the internal impetus to get it.

Figure 3.1: Prior conceptions of absorptive capacity in relation to this study
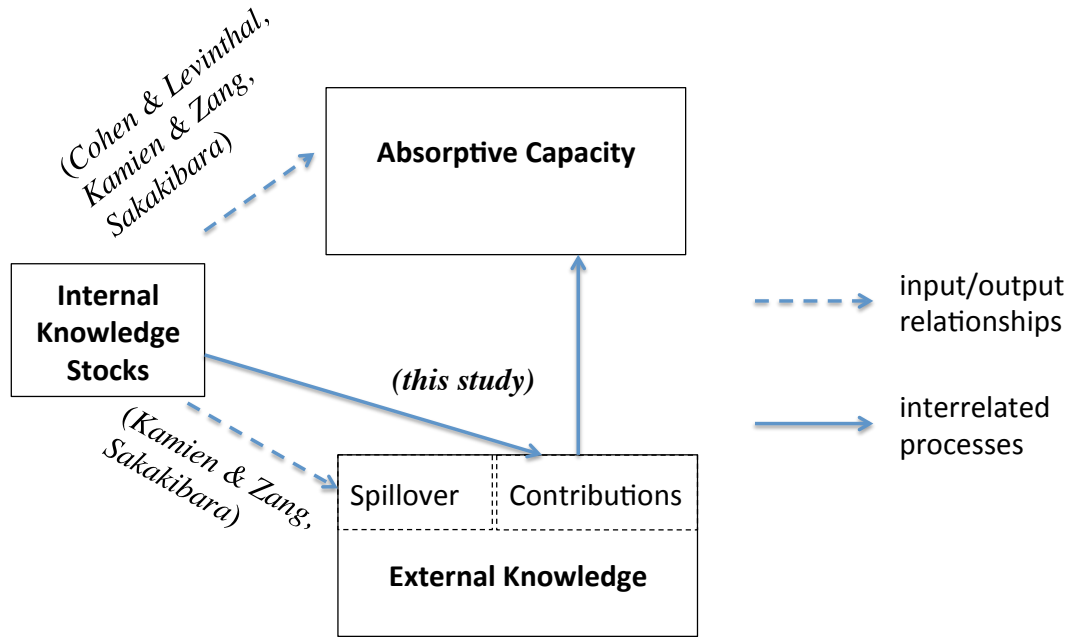
As in our study, research in industrial organization economics has argued that knowledge inflows through absorptive capacity requires knowledge outflows (Kamien and Zang, 2000; Sakakibara, 2003). A crucial distinction, however, is that knowledge outflows are not intentional contributions a firm makes to enhance external knowledge absorption, as we theorize, but unintentional spillovers that are an unavoidable and costly side effect of conducting internal, absorption-enhancing R&D.

Closer to our study is research that suggests participation in wider "community" may be necessary for a firm to absorb external knowledge, particularly in rapidly evolving, technology-intensive contexts. Powell and collaborators (1996, p. 119) observe: "Passive recipients of new knowledge are less likely to appreciate its value or to be able to respond rapidly. In industries in which know-how is critical, companies must be expert in both in-house research and cooperative research...." Similarly, Henderson and Cockburn (1998) examine firms collaborating with the academic scientific community on basic research and applying the knowledge gained to downstream innovations within the firm. They observe that firms in principle could conduct this same

research in-house as a means to better understand and utilize the publicly conducted research efforts of others, but their empirical findings are that firms collaborating more heavily in the basic science community (i.e., co-authoring with university scientists) have higher downstream innovative productivity. The implication here, as corroborated by in-house scientists they interviewed, is a firm cannot rely solely on internal research to build sufficient absorptive capacity. Rather, "a firm's researchers need to be active participants in the construction of publicly available research results, despite the issues of appropriability that such active collaboration raises" (Cockburn and Henderson, 1998, p. 163).

In our study, we build on this line of inquiry by theorizing how participation in a community of collaborating firms enhances absorptive capacity. As expressed by Dahlander and Magnusson (2008, p. 632), "our understanding about how absorptive capacity is manifested in terms of the tactics firms use to access external knowledge ... is limited," and we address this concern by observing that absorptive capacity is not necessarily limited to internal, passive processes of capturing knowledge spillovers. In particular, we raise the possibility that external knowledge may also be absorbed through ongoing interactions with other firms.

We examine multi-firm collaborations, which we define as cooperative arrangements where multiple firms pool knowledge contributions in an attempt to achieve collective gains (Zeng and Chen, 2003). As such, we restrict attention to multilateral collaboration, and not, for example, unilateral collaborations, such as licensing agreements, whose theoretical issues lie outside the scope of the study. Firms contribute by purposively sharing internal stocks of knowledge with other collaborating firms. As we theorize below, it is the process of contributing this knowledge and integrating it with other firms' contributions engenders absorptive capacity (see Figure 3.1).
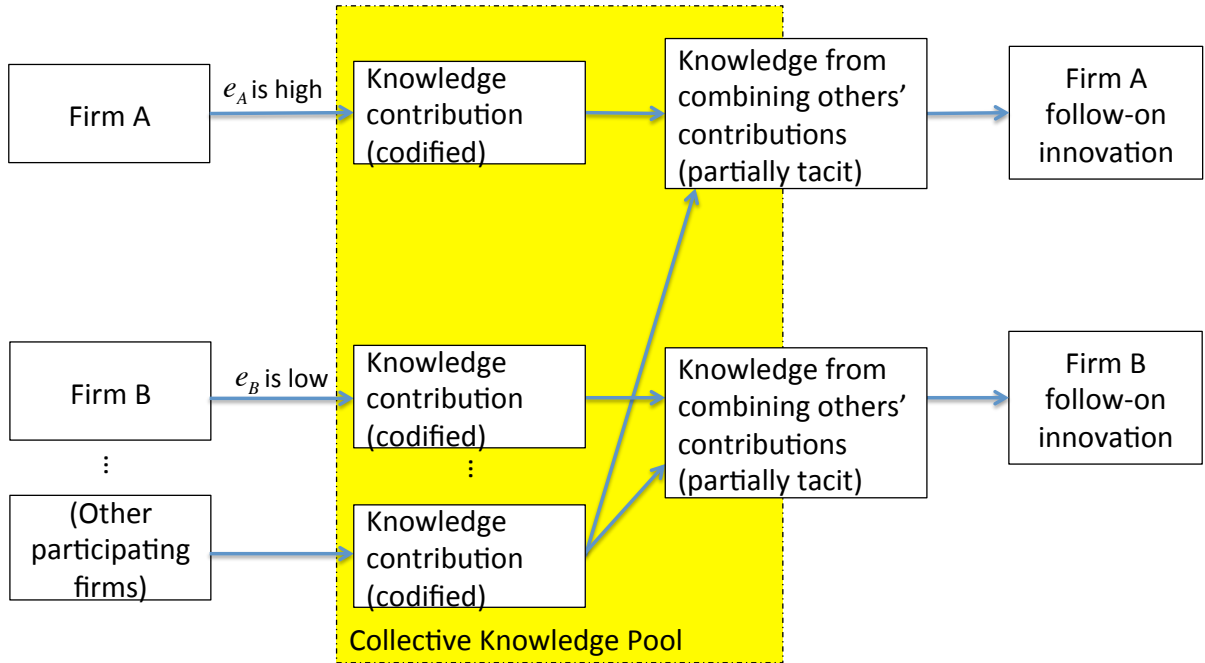
3.3.2 Collaboration-Specific Absorptive Capacity as a Process

To clarify subsequent theoretical discussion, Figure 3.2 illustrates the process underlying CSAC. The intuition behind this figure is that collaborating firms are likely to initiate any collaboration through communications interpretable by other firms – that is, codified knowledge. Examples include technical specifications, research whitepapers, or an embodiment of knowledge in physical form (e.g., a deployed semiconductor manufacturing process or clay model of a car). The levels of codified knowledge contributed are represented by e_A and e_B for representative firms A and B. Strategic considerations aside, this level is limited by the extent to which a firm holds knowledge relevant to the collaboration. For the sake of example, A contributes heavily to the collaboration (e_A is high) and B contributes relatively little (e_B is low).

Next, tacit knowledge typically surrounds these codified contributions (though in open-source software for example, such tacit knowledge may be very minimal). While we illustrate only a single sequence of codified followed by tacit knowledge, this collaborative knowledge integration process is likely to involve repeated iterations and an interplay between codified contributions and tacit knowledge co-development (Nonaka, 1994) that ultimately produce a coherent piece of collective knowledge. Finally, this collective knowledge is applied to subsequent innovations developed within individual collaborating firms.[6] Since A contributes heavily to the collaboration, its actions are more noticeable than B's to other participants, and as such, A is prone to having other firms notice and imitate its strategies with respect to the technology underlying the collaboration.

---

[6]For a more formal description of this process, see the Appendix, which presents a toy model that illustrates the mechanisms underlying the theory we develop below.

Figure 3.2: Collaboration-specific absorptive capacity

3.3.3 Mechanisms Underlying Collaboration-Specific Absorptive Capacity

Recall our claim that the basis for CSAC is distributed interdependence. Distributed interdependence, as we will argue, implies firms have incentives to contribute because (a) a piece of knowledge may only be useful when a firm contributes and integrates this knowledge with that of other firms, and (b) a contributing firm is likely to appropriate more of the value created through inter-firm knowledge integration than do non-contributing firms. We elucidate these claims below.

Why does contributing knowledge and integrating it with that contributed by other firms create value? In other words, what prevents a firm from realizing the same gains by annexing other firms' contributions wholly within the firm? The basis for our ensuing discussion is that distributed interdependence, coupled with the presence of tacit knowledge (Polanyi, 1966), implies firms may hold knowledge that they can only minimally utilize under their own effort. As we will argue, tacit knowledge,

which is typically pervasive innovation-oriented settings (Senker, 1995), makes knowledge integration – that is, coordination across knowledge boundaries – necessary to realize combining gains among diffuse, interdependent intra-organizational knowledge elements (Grant, 1996; Nonaka, 1994; Kogut and Zander, 1992). While most of the literature has focused on boundaries within a firm (e.g., between individuals), a few studies have begun stressing the importance of knowledge integration in an inter-firm collaborative context (Tiwana, 2008; Ireland et al., 2002). The need for inter-firm knowledge integration in the presence of tacit knowledge is strongly evidenced in real-world settings. Indeed, even in technology licensing where both sides of a transaction would benefit from a fully codified knowledge exchange, substantial knowledge integration is usually required to transfer tacit knowledge so that the licensee can effectively utilize a licensor's technology (Arora, 1996). Likewise, codified knowledge exchange within a multi-firm collaborative context is also likely to require the support of tacit knowledge exchange and attendant knowledge integration efforts.

Knowledge integration can in principle utilize a broad set of "arms-length" coordinating devices such as rules or routines (Grant, 1996), but such mechanisms are unlikely to be sufficient for knowledge integration within innovation-oriented collaborations marked by substantial tacit knowledge. Instead, such collaborations tend to entail substantial technological novelty and complexity, high unpredictability, and inter-organizational learning that push the level of coordination required beyond that addressable through arms-length mechanisms. As a result, knowledge integration in these settings must utilize more communication-intensive forms of coordination (Grant, 1996; Carlile, 2004). For instance, firms in the wireless standards context of the present study collaborate closely through frequent physical meetings around the world (see Section 3.4 for details). SEMATECH took this a step further, creating a physical presence where researchers from member firms collaborated side-by-side on R&D activities.

Moreover, two other reasons suggest that close collaboration is necessary. First, close collaboration helps firms jointly discover the most valuable interdependencies. Indeed, Henderson and Clark (1990) echo this idea in observing that a fundamental challenge in new product architectures is discovering the design interdependencies that matter. When the constituent knowledge elements that comprise the architecture reside in many firms, coordination difficulties and the resulting need to jointly engage in the co-discovery process is all the more acute. Scholars have also identified a related issue within work groups and organizations, wherein finding "who knows what" – termed a transactive memory – is a challenge that has a strong bearing on firm's performance (Ren et al., 2006; Moreland, 1999). This issue is particularly salient in a multi-firm collaborative context where limited inter-firm communication (relative to intra-firm communication) constrains the level of transactive memory among a group of member firms. Moreover, since novel contexts tend to reduce transactive memory (Ren et al., 2006), close collaboration plays a more important role in uncovering valuable interdependencies within innovation-oriented collaborations.

As an illustration of this distributed search for interdependence, CDMA technology made it cost-effective to build a "soft handover" phone that could communicate with multiple towers simultaneously, but this created numerous hard-to-discover interdependencies that straddled the expertise of technology companies collaborating on CDMA systems development. For example, while the basic soft handover technology exploited how CDMA codes radio transmissions, the novel paradigm of having a single caller's phone signal arrive at multiple cell towers simultaneously also had a profound and ex-ante unknown effect on many parts of the landline network supporting the radio network. As a result, firms with highly diverse expertise engaged in collaborative knowledge integration to discover and resolve these system interdependencies.

Second, close collaboration may be necessary for the development of the common

language necessary for effective knowledge integration, particularly in the presence of tacit knowledge. In particular, novel knowledge that typifies innovation-oriented contexts often sits outside any established "common lexicon" used to exchange knowledge (Carlile, 2004; Carlile and Rebentisch, 2003). Indeed, a common lexicon enables diverse perspectives to contribute towards a collective effort (Page, 2008). Carlile (2004) illustrated a common lexicon through a clay model of a car that allowed, for instance, engine designers to exchange knowledge with body designers. Knowledge integration enables collaborating firms to amend an existing lexicon so that is sufficiently rich to allow for the exchange of newly created knowledge. For example, when CDMA (code-division multiple access) cellular phone technology was new, "few engineers outside Qualcomm understood how the technology worked" (West, 2001), and part of the difficulty was in characterizing CDMA with the existing lexicon (and associated concepts) of the cellular phone industry. Particularly problematic were the "codes" in CDMA technology that in some ways displaced "frequency" (radio frequency band), a concept that was well-understood and used heavily by the vast majority of engineers in the cellular phone industry. Much of the lexicon (and associate concepts) surrounding frequencies was difficult to apply to codes, so Qualcomm had to work closely with other firms to revise the existing common lexicon (and associate concepts) for CDMA systems development.

Recall that our discussion thus far builds on the notion that firms may be unable to utilize some of their knowledge without complementary knowledge from other firms. As such, a firm's contribution and subsequent close collaboration with others increases the value of the collective knowledge pool, both for itself and other firms. If enlarging the value "pie" in this fashion gave rise to symmetric benefits, it is not always clear whether firms have incentives to contribute.[7] We argue, however, that the benefit derived from a firm's contributed knowledge is likely to be *asymmetric* in

---

[7]For instance, if a firm's contributions resulted in a product option with both higher quality and cost, it might not increase its own (or any firm's) profits.

favor of the contributing firm. The basic idea for this claim is that a firm is likely to have private knowledge whose value depends on the success of integrating contributed knowledge with that of other firms. We turn now to clarifying this claim.

Figure 3.3: Dependency between contributed and private knowledge integration
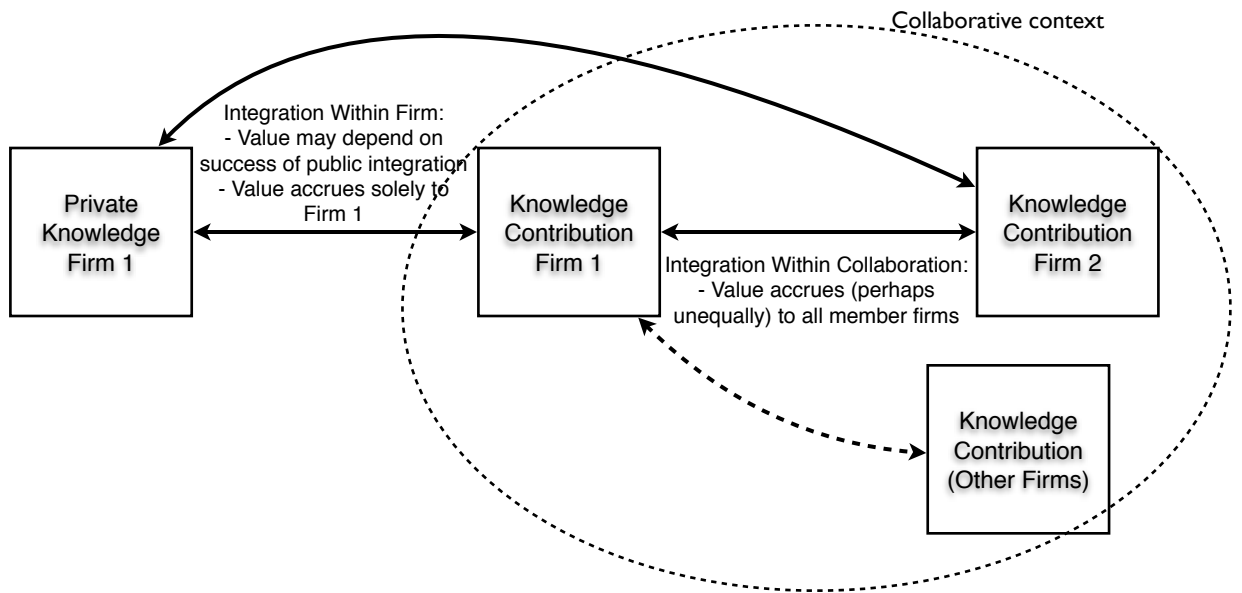
Figure 3.3 depicts the knowledge elements relevant to this argument, namely Firm 1 and 2's knowledge contributions to the collaboration and Firm 1's private but related stock of knowledge.[8] All firms potentially benefit from integration of Firm 1 and 2's contributed knowledge. As such, if Firm 1 realized no benefit beyond that available to other firms, it might have relatively small incentives contribute or, going a step back, develop its knowledge. Yet we contend this is rather unlikely because knowledge is developed in the context of a firm's existing knowledge stocks. Indeed, a firm's existing knowledge stocks are the underpinnnings of a firm's "capacity for action" (Iansiti and Clark, 1994, p. 560) that enables problem-solving and, as a consequence, generates new knowledge (Iansiti and Clark, 1994). As such, a firm's contributed knowledge is likely to be interdependent with a broader set of knowledge contained

---

[8]Following the discussion above, Firms 1 and 2 contribute because their respective knowledge contributed would have little value absent knowledge integration within the collaboration.

59

within the firm, which we represent as Firm 1's private knowledge in Figure 3.3. In this fashion, the value of the private knowledge depicted in Figure 3.3 likely depends on the successful integration of Firm 1 and 2's contributed knowledge.

We illustrate this by way of a stylized example. Suppose Gamma Motors develops advanced brake technology that enhances vehicle safety, but it requires integrating knowledge of competitors who are more capable in other automotive systems (e.g., electrical. engine) on which the brake innovation depends. Absent this coordinated knowledge integration, the brake technology has no value - it adds cost to the car but provides no benefit. With this integration, the collective knowledge arising from the combined expertise of several firms would create value for all firms since consumers are willing to pay a premium for the brake safety innovation. At the same time, Gamma Motors has a selective incentive to contribute its brake technology because it was developed from Gamma Motors' core R&D knowledge and capabilities, which not only produced the "blueprints" for the brake technology itself but also, quite naturally, tailored the brake technology to Gamma Motors' leading edge (and private) manufacturing capabilities. As a result, Gamma Motors is able to manufacture the brake technology more cheaply than its competitors.

Thus, a firm have incentives to contribute because knowledge contributions enable a firm to learn from knowledge integration efforts and, in turn, enhance a firm's follow-on innovations. Here, we make an additional observation that knowledge contributions serve not only as an external conduit to integrate knowledge and learn from other member firms, but also as a means to promote broader internal organizational attention (Ocasio, 1997) and buy-in to the firm's efforts within the collaboration and subsequent efforts to make use of the knowledge output of the collaboration. In particular, knowledge contributed to a collaboration is subject to scrutiny by other firms. Therefore, such knowledge is likely to draw widespread attention across units in the firm to ensure, for example, that the firm maintains its reputation (Fombrun and

Shanley, 1990) and does not involuntarily leak its core knowledge assets (Kale et al., 2000). Moreover, passing the scrutiny of other firms enhances the internal legitimacy of a firm's knowledge contributing efforts. Visibility and legitimacy issues may be particularly salient in larger, bureaucratic organizations where activities within an organizational unit are not as visible and may compete for legitimacy among many organizational activities (Kostova and Zaheer, 1999). The increased internal attention serves to further enhance the follow-on innovation benefits from contributed knowledge.

3.3.4 Limits to Developing Collaboration-Specific Absorptive Capacity

The above discussion points to the benefits a firm sees from contributing knowledge to a collaboration. Yet at higher levels, an increase in knowledge contributions is potentially detrimental to a firm because it may reveal valuable information about the collaboration to other firms. This information is particularly salient when collaborative efforts are exploratory (e.g., research-oriented) (Lavie and Rosenkopf, 2006) and hence marked by considerable uncertainty. Indeed, firms tend to observe and learn from others when faced with uncertainty (Festinger, 1954; Mansfield, 1961). While prior research in inter-firm collaborations typically focus on leakage of proprietary knowledge (Hamel, 1991; Kale et al., 2000; Oxley and Sampson, 2004), the information revealed by effort is readily observable by members of the collaboration. The notion that easily observed knowledge may trigger vicarious learning is reflected in numerous studies within the management literature where firms' imitative efforts are driven not by the private know-how of other firms, but rather by the visible artifacts of that know-how (Baum and Ingram, 1998; Simon and Lieberman, 2010; Haunschild and Miner, 1997).

A firm's knowledge contribution potentially reveals two valuable aspects of a collaboration. First, high levels of contribution reveal a firm's high expectations on the

value of collaborative knowledge (i.e., it will enable valuable follow-on innovations). Of course, high expectations are not informative if they have the same likelihood of being correct as low expectations. However, high effort in contributing knowledge is likely to reflect a more accurate assessment of collaborative knowledge quality because it likely reflects prior internal effort at understanding the technological issues that underlie the collaboration. Consistent with this notion, other studies (e.g., Türüt and Ofek, 2012) have also assumed that signals indicating a high value for productive inputs are more accurate than are signals indicating a low value. Hence, high effort tends to be an accurate – and therefore valuable – signal to firms with erroneously low expectations on the quality of collaborative knowledge.

Second, a high level of knowledge contributions from a firm is likely to give cues on the capabilities germane to utilizing the collaborative knowledge for subsequent innovation. For instance, suppose a firm makes many contributions that specify small components for the system design. Other firms might infer that miniaturization is the key capability necessary for utilizing collaborative knowledge in follow-on innovations. The notion here is conceptually that of "reverse engineering," in which freely observable artifacts – which in this study is knowledge from a collaboration rather than observable product features – enable a firm to "decode the substantive technology" (Kogut and Zander, 1992, p. 393) that underlies it. Assuming time-compression diseconomies are not prohibitive (Dierickx and Cool, 1989), other firms may then reverse engineer the relevant capabilities of the focal high-effort firm.

In sum, a firm's knowledge contributions enhance not only common but private learning benefits from collaboration, which increases the performance of follow-on innovation. Beyond a certain level, however, knowledge contributions lead to an information revealing effect that diminishes follow-on innovation performance. The learning benefits of knowledge contributions tend to exhibit diminishing returns to effort because a firm can learn only as much as other firms know. On the other

hand, the information revealing effect of knowledge contributions are only salient, for example, beyond the mean contribution level of all firms, and in this capacity knowledge contributions may in fact exhibit increasing effects of effort (intuitively, large effort is highly noticeable while small effort is not). This leads to the following prediction:

> **Hypothesis 1**: The performance of a firm's follow-on innovations increases then decreases with the amount of knowledge a firm contributes to a collaboration.

3.3.5   Inter-firm Heterogeneity and Knowledge Absorption

Firm heterogeneity is essential to the theory of knowledge absorption offered in this study since there is conceptually no knowledge to be absorbed in interactions between identical firms. As such, we examine whether firm heterogeneity increases the effect of knowledge contributions on follow-on innovations. As discussed below, this prediction, if supported, provides further evidence that knowledge contributions lead to learning (knowledge absorption) and are therefore not solely a reflection of firms advancing their vested interests.

Firm heterogeneity enables a richer set of knowledge combinations comprised of the knowledge generated by the collaboration. Extending the single-firm notion of novel resource combinations as a source of superior performance (Penrose, 1959; Galunic and Rodan, 1998; Kogut and Zander, 1992), a focal firm may combine its own knowledge contributions with the disparately held and complementary knowledge resources of partnering firms in novel ways that no single firm can achieve on its own (Das and Teng, 2000), thereby enhancing the potential for the collaborative knowledge to lead to valuable firm-level innovations.

Yet the existence of a richer set of potential knowledge combinations is insufficient for firm learning to occur: as discussed in the previous section, firms must interact

with other firms to access their knowledge. To this end, prior research has associated differences in resources (and knowledge) to a "cognitive distance" between two firms (Nooteboom et al., 2007). Moreover, this prior work argues that inter-firm cognitive distance increases the efficacy of learning by interaction in that the individual people who carry out the interaction are stimulated to learn through the more diverse knowledge and perspectives of their counterparts.[9]

In sum, the preceding arguments suggest firm heterogeneity enhances a firm's knowledge absorption through its knowledge contributions, which in turn increases the performance of its follow-on innovations. We therefore hypothesize the following:

> **Hypothesis 2**: As firm heterogeneity increases, knowledge contributed to a collaboration is more beneficial to the performance of a firm's follow-on innovations.

3.3.6 Scope of Knowledge Absorption

Under CSAC as theorized above, knowledge absorption has two dimensions. On one hand, firms may *broaden* their existing knowledge base, building on the knowledge held by other firms. As with traditional conceptions of absorptive capacity, knowledge broadening in a collaborative context occurs as internal knowledge facilitates acquisition of external knowledge, albeit through collaboration-level processes rather than processes internal to a firm. On the other hand, firms may also *deepen* their existing knowledge base (e.g., Katila and Ahuja, 2002) through knowledge contributions that trigger interactions with other firms. For example, knowledge integration efforts in the collaboration may uncover potential modifications or extensions of a

---

[9]Beyond a certain point, cognitive distance may reduce the learning benefits of interaction because firms may be too different to achieve the mutual understanding necessary to engage in effective knowledge combining. However, this distance is perhaps rarely observed in multi-firm technical collaborations because firms often collaborate only when they have some degree of technology overlap to begin with. For example, the majority of firms in the present context are in the wireless industry. Moreover, the actual unit of collaboration within the present context is further subdivided into narrower technological subfields – otherwise, collaboration would be unmanageable given the diversity of technologies embodied in a wireless architecture.

firm's contributed knowledge that enhance its value to the firm.

Depth and breadth of search have natural interpretations within the collaborative context examined in this study: a firm deepens its knowledge base when its follow-on innovations build on its own contributions to the collaboration, and a firm broadens its knowledge base when it builds on the contributions of other firms. As such, a multi-firm collaborative context allows us to examine the efficacy of knowledge deepening and broadening search, and their determinants thereof.

We focus on the impact of firm heterogeneity on the performance of knowledge deepening and broadening follow-on innovations within the learning effect. The arguments in the preceding section suggest a greater diversity of knowledge is likely to be garnered from a firm's knowledge contributions when firms are more heterogeneous. Therefore, we predict that increased firm heterogeneity is likely to benefit follow-on innovations that broaden rather than deepen a given firm's knowledge base:

> **Hypothesis 3**: Firm heterogeneity is more beneficial to knowledge broadening follow-on innovations than to knowledge deepening follow-on innovations.

3.4 Context

The empirical context of the study is 3GPP, an umbrella organization comprising several regional standards development organizations (referred to in 3GPP as "organizational partners"). The activities within 3GPP are conducted by approximately 400 (at present) member firms (referred to in 3GPP as "Individual Members"), which consists of the major manufacturers and service operators as well as a number of smaller firms developing 3GPP-based wireless cellular technologies. Member firms collaborate in producing a standard that seeks to establish technical interoperability for 3GPP-based wireless data and voice services.

The focus of this study is inter-firm collaboration and the ex-post innovations that build from this collaboration. The 3GPP context is well-suited to examining this

concept for three reasons. First, 3GPP can be considered a "laboratory" of frequent inter-firm collaborations that exhibit substantial heterogeneity in group composition across collaborations and individual firms' effort levels within a collaboration. Second, fine-grain data on firm level effort within inter-firm collaborations are rarely available. In the 3GPP context (as described in more detail below), we observe micro level data on firm effort levels within collaboration, e.g., the documents (and email discussions) each firm contributes.[10] Third, "risk set" of potential collaborators – and therefore potential follow-on innovators – is known (namely, the set of member firms), which allows us to account for selection bias.[11]

---

[10]Emails are not currently used in the empirical analysis. Beyond being difficult to code, there is also the issue of how to combine effort represented as document contributions with effort represented as email contributions.

[11]The knowledge generated within our empirical context could in principle lead to innovations by non-members, but the incidence is quite low in the sample. Firms that are not members of 3GPP may of course cite collaborative knowledge within 3GPP, but 3GPP is a relatively closed ecosystem in that nearly all innovations building on its knowledge base come from its member base. Only 3.5% (114/3260) of observations in our sample are from non-members, and most of those are from a single firm. Aside from this firm, there are only 0.7% (23/3260) of observations from non-members.

Figure 3.4: Relationship between 3GPP work flows and ex-post innovations
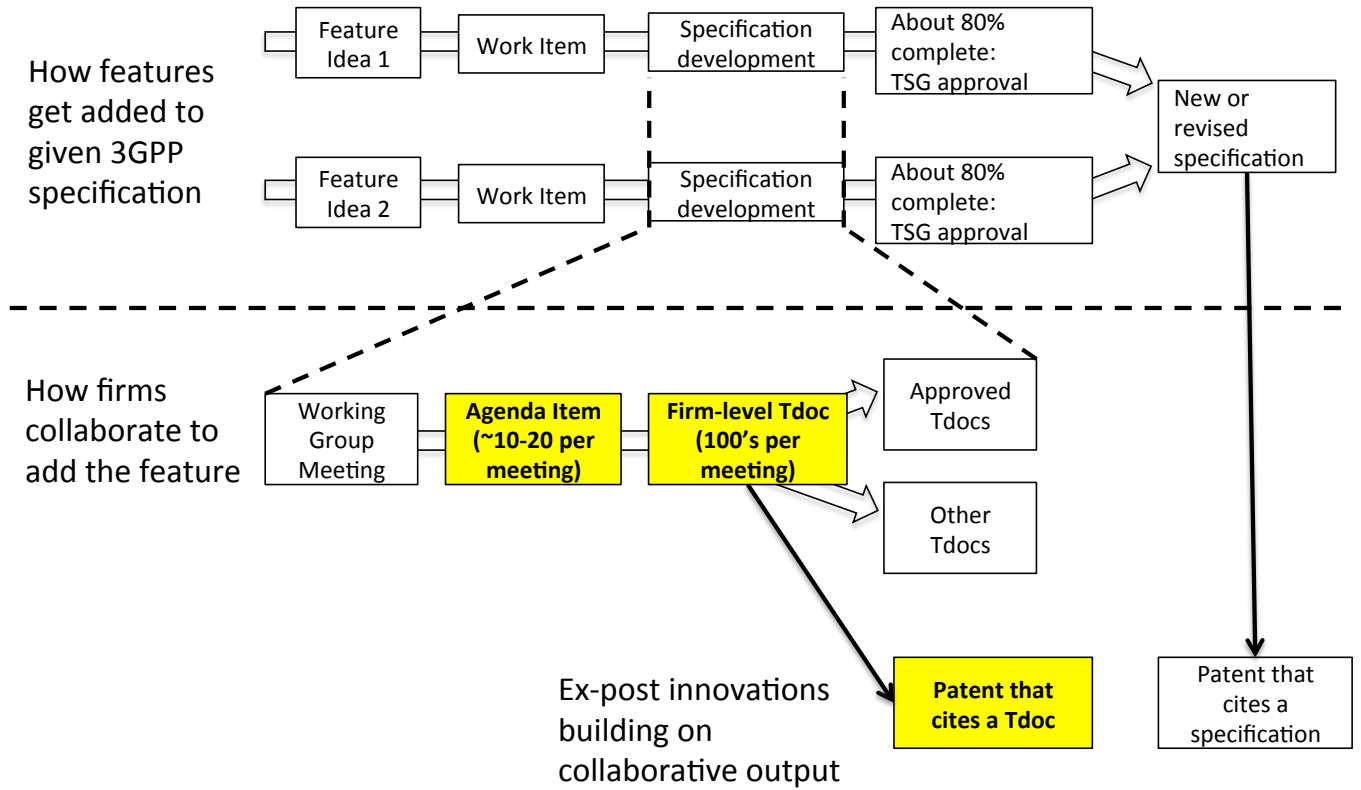
To explicate the points made above, we turn to a more detailed discussion of the work flow within 3GPP, as depicted in Figure 3.4 (portions adapted from Leiponen, 2008). The top of the figure provides the broader setting in which firm level technical contributions are developed by showing the flow from technical concept – that is, ideas for product features – to standardization. Feature ideas may come from, for instance, opportunities perceived by member firms or demand from the market. Ideas approved within 3GPP are then formalized as work items. Work item completion then continues within the specification development phase, which contains the actual technical work within 3GPP, organized into Technical Specification Groups (TSG) corresponding to different technical subfields of wireless system design. Once specification development on a given work item nears completion, it is submitted to TSG for approval and subsequent inclusion into a new or revised standards specifi-

cation. These specifications (at the top right of the figure) define what is "in the standard," though the knowledge underlying the specification is codified in firm level documentation, as we describe below.

The bottom of Figure 3.4 illustrates how firms collaborate in specification development. Work within each TSG is organized into working groups (WG) that meet several times per year. Meetings are organized into agenda items to solve specific technical issues. Firms submit technical documents, or Tdocs, in advance of each meeting for discussion, debate, and possible approval. A firm may contribute Tdocs to some, none, or all agenda items, and all firm representatives that are present at the meeting may participate in discussions for any agenda item. Moreover, further discussion may take place within an email exploder organized by 3GPP. Note there is not a stark bright line between approval or non-approval and inclusion in a specification: while knowledge within formally approved Tdocs tends to flow more readily into a revised (or new) specification, other Tdocs may also contain knowledge that ultimately shape specifications.

The highlighted boxes in the bottom of Figure 3.4 represent the focal items of the empirical portion of this study. We take an agenda item from a single meeting to be a discrete inter-firm collaboration. A single Tdoc submitted by a single firm represents a unit of effort in participating within a collaboration.[12] Finally, ex-post innovations building on this collaborative knowledge are represented by patents that cite a Tdoc.[13]

The data for this study, which corresponds to the elements described in Figure 3.4 , have been collected from custom-developed scripts that "scrape" the 3GPP website (www.3gpp.org). Data on patents have also been obtained through the

---

[12]Two issues are worth noting: (a) alternatively, we may normalize effort by the content (e.g., length) of a Tdoc, and (b) while most Tdocs are authored by a single firm, some are authored by multiple firms.

[13]It is worth noting that patents also cite the end specification, as shown in Figure 3.4, but it is often difficult to trace firm-level ownership of text contained within specifications, as suggested by the preceding discussion in this section.

Harvard Patent Network Dataverse database (Lai et al., 2011) and extended to mid-2012 using Google's bulk patent download site (http://www.google.com/google-books/uspto.html). Finally, financial data for firms in the sample were obtained through COMPUSTAT.

3.5 Empirical Model

Innovation performance is modeled by the following equation:

$$p_{ij} = \beta_1 + \beta_2 e_{ij} + \beta_3 e_{ij}^2 + \beta_4 f_{ij} + \beta_5 b_{ij} + \beta_6 e_{ij} b_{ij} + \beta_7 e_{ij} f_{ij} + \beta_8 b_{ij} f_{ij} + controls + \epsilon_{ij}$$

where:

- p_{ij} is the innovation performance of firm $j's$ follow-on innovation that builds on knowledge from collaboration i.

- e_{ij} is firm $j's$ effort contributing to a collaboration i

- f_{ij} is firm distinctness (the average difference between firm j and other firms in collaboration i)

- b_{ij} is a dummy set to 1 if the follow-on innovation is knowledge-broadening (i.e., a firm cites a knowledge contribution from a different firm) or 0 if it is knowledge deepening (i.e., a firm cites knowledge that it contributes).

Support for the Hypothesis 1 would be indicated by β_2 positive and significant, and β_3 negative and significant. Support for Hypothesis 2 would be indicated by β_7 positive and significant. Support for Hypothesis 3 would be indicated by β_8 positive and significant.

3.6 Data

3.6.1 Dependent Variable

We measure p_{ij} as the citation count of a firm j's patent that cites a Tdoc in collaboration i. Multiple patents by firm j that cite Tdocs in collaboration i are counted as separate observations.[14] Patent citations are subject to exposure effects, in which an older patent will have more citations simply because it has had more time to accumulate them. To account for this, we employ a "fixed effects" approach (Hall et al., 2001) in which the raw patent citation counts are adjusted by the average number of citations received by all patents within the same patent category (defined by Hall and collaborators) and application (or grant) year. The fixed effects adjustment therefore indicates impact relative to other patents in the same cohort. Another feature of the fixed effects adjustment is that it accounts for changes in citation behavior (within a patent category) across time. Table 3.1 shows the average citations received by patents in the "Computers and Communications" category, to which nearly all (over 99%) of the patents in our sample belong. The manner in which we include fixed effects adjustment in our empirical model is described in more detail below.

---

[14]This is necessitated by control variables that are patent-specific, such as patent lag.

Table 3.1: Average citations received by patents in the "Computers & Communications" category, as defined in (Hall et al., 2001) (as of July 2012)

App. Year	Average Citations
1998	21.496
1999	17.351
2000	12.606
2001	8.230
2002	6.199
2003	4.193
2004	2.880
2005	1.905
2006	1.166
2007	0.743
2008	0.388
2009	0.239
2010	0.186
2011	0.109

3.6.2 Explanatory Variables

Effort e_{ij} is measured as the number of Tdocs contributed by firm j to collaboration i.

The heterogeneity measure is based on the technological positions of firms, as defined by a vector v representing the number of patents issued to a firm in each patent class (Jaffe, 1986). The technological distance between two firms A and B is given by:

$$T_{AB} = 1 - \frac{v_A \cdot v_B}{\|v_A\| \, \|v_B\|}$$

Thus, T_{AB} is zero for identical firm positions and 1 for maximally distinct firms.[15]

We proxy firm heterogeneity with *firm distinctness*, f_{ij}, which is given by the average technological distance between firm j and other firms in collaboration i.

---

[15]In a multi-industry sample, using raw patent classes might be problematic since there are too many (and hence distance is skewed towards 1). However, in the 3GPP wireless context used in this study, all patents in the sample are from 14 patent classes.

Compared to a collaboration-level measure such as average pairwise distance, firm distinctness has the advantage of accounting for differences in technological positions within a collaboration.

3.6.3 Controls

One empirical concern is that there may be endogeneity in effort. A potential cause of this is firms contribute more to collaborations when they are more knowledgeable in the technological fields underlying the collaboration. We control for this by computing the technological distance (*firm pat distance*) between the firm and the technological subject matter treated in the collaboration. The former is simply the firm's technological position as discussed above, and the latter is imputed from the patents citing Tdocs in the collaboration.

Since a firm's level of knowledge contributions may reflect a larger collaborative task (i.e., larger aggregate effort), we control for the total effort in a collaboration using the total number of Tdocs (*total doc count*) submitted to a collaboration.

We also control for the average time lag between the knowledge contribution and an innovation (*patent lag*), the former given as the date of collaboration (i.e., working group meeting) and the latter given as the date the innovation (patent) was filed. This accounts for the possibility that, for example, a longer lag reflects increased investment in the innovation.

We control for the patent citation truncation discussed above by including the average citations in a patent's cohort (*patent adjust*), as given in Table 3.1. While it would be preferable to scale the dependent variable by this factor, this would preclude using a count model for the empirical analysis. In additional robustness analysis, we scale the dependent variable, as discussed in the robustness section below.

Finally, we control for financial performance indicators that may affect a firm's propensity to patent and/or the quality of the patents: profitability (ROA), liquidity

(*current ratio*), and (log of) a firm's total assets (*size*) to account for the possibility that large firms are more or less proficient at patenting than small firms (all obtained from COMPUSTAT).

3.7   Method and Results

We employ a negative binomial regression model in the empirical analysis. Two reasons suggest that this model is appropriate. First, the dependent variable (citation counts) is overdispersed – the variance higher than the mean. The alternative Poisson regression model assumes equal mean and variance while the negative binomial regression models allows for variance higher than the mean. Second, the large number of patents in the data sample that have not received any citations (i.e., a zero dependent variable) suggests the need for a zero-inflated negative binomial model. However, the fixed effects adjustment for heterogeneity in exposure to patent citations also adjusts, as a by-product, for large zero counts. A Vuong (1989) test statistic for non-nested models supports this conclusion ($z = 0.00$ and $pr > z = 0.5$ in the full model).

Recall that we employ a fixed effects adjustment to account for differences in patent citation counts that arise from differences in the period over which a patent accumulates data as well as cohort and time period effects. We incorporate the fixed effects adjustment in the negative binomial model through an exposure variable (also called an offset variable). The exposure variable is a covariate with coefficient constrained to one, which therefore serves to transform the dependent variable from a count to a rate.[16] While the exposure variable is perhaps most naturally associated with a time duration (e.g, hours in the sun as an exposure variable for skin cancer counts in a population), a time duration construct may not be the best normalization

---

[16]This can be seen by observing that an additive exposure variable in the log-link function of a count model is multiplicative in the exponentiated log-link function that reflects counts. Dividing both sides of this exponentiated function by this variable transforms the equation from a count to a rate estimation.

for what the count "should be" in a given context. In the present context, the adjustment count given in Table 3.1 is a construct that reflects a suitable baseline for actual counts, and as such, we use this adjustment count for the exposure variable.

Table 3.2 provides the descriptive statistics for the empirical model. Results from the empirical model are given in Table 3.3.

<div align="center">Table 3.2: Summary statistics</div>

	Mean	S.D.	Min	Max	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
p_{ij}	1.749	5.508	0.000	87.000	1.00									
e_{ij}	2.339	3.345	0.000	24.000	-0.03	1.00								
f_{ij}	0.701	0.098	0.070	0.957	-0.01	0.19	1.00							
b_{ij}	0.774	0.418	0.000	1.000	0.00	-0.31	-0.09	1.00						
total doc count	36.732	30.527	2.000	246.000	-0.05	0.53	-0.08	-0.08	1.00					
firm pat dist	0.359	0.213	0.000	0.904	0.00	0.13	0.22	-0.02	0.10	1.00				
firm adjust	1.427	1.988	0.186	17.351	0.48	-0.05	0.01	0.04	-0.11	0.01	1.00			
ROA	0.075	0.163	-1.479	0.496	-0.14	0.01	-0.02	-0.01	-0.05	0.04	-0.22	1.00		
current ratio	2.241	1.377	0.000	9.408	-0.04	-0.07	-0.22	0.05	-0.05	0.03	-0.10	0.32	1.00	
size	9.852	1.655	3.300	11.515	0.02	0.18	0.24	-0.08	0.07	-0.04	0.01	-0.18	-0.56	1.00

Table 3.3 presents preliminary results on the tests for the hypotheses. Model 1 includes just the control variables. Notably, *patent lag* is positive and significant, consistent with the idea that firms are utilizing the patent lag time window to better understand how to incorporate collaborative knowledge in follow-on patents. Firm size is also associated with greater citation counts, which may reflect the greater market influence of larger firms. Finally, profitability decreases patent citation counts, which may indicate firms are engaging in strategic patenting behavior. That is, firms with more slack resources are better equipped for "patent portfolio races" (Hall and Ziedonis, 2001), and are therefore strategically trading quality for quantity.

Models 2 through 7 progressively introduce the explanatory variables into the specification. Model 2 indicates knowledge contributing level does not have an effect on follow-on innovations. But, as Model 3 and H1 suggest, this is because knowledge contribution is misspecified in Model 2. Indeed, support for H1 would be indicated by knowledge contributing level having an inverse-U effect on follow-on innovations. This is the case in Model 3: the coefficient on e_{ij} positive and significant, and the coefficient on e_{ij}^2 is negative and significant.

The introduction of the interaction term between knowledge contributions and the knowledge broadening dummy, in Model 6-8, lends additional support for our arguments for the benefits of contributing knowledge (H1). In particular, with the inclusion of this term, the coefficient on knowledge contributions reflect firms that build on their *own* contributions, which is not significant. Interestingly, however, the net coefficient for firms that build on *others'* contributions (the sum of the knowledge contribution term and the interaction term) is significant (at the 5% level) in Model 8, our full model. This is consistent with firms building CSAC – that is, they learn from others in the process of integrating their contributions to the collaboration – but is not consistent with, for example, firms furthering their own interests in the collaboration. However, our results do not imply that the latter phenomenon does

not occur. Indeed, our results do not pick up a firm's private knowledge that both embodies a firm's vested interests and precedes the firm's knowledge contributions, since we only examine the ex-post effects of these contributions.

We now turn our attention to Hypotheses 2 and 3. Model 7 does not support H2 since the coefficient on $e_{ij}f_{ij}$ is positive, but not significant. Moreover, Model 8 does not support H3 since the coefficient on $b_{ij}f_{ij}$ is not significant.

Table 3.3: Inter-firm heterogeneity and firm's knowledge contributions as facilitators of follow-on innovations

Dependent variable: Firm j's aggregate citation count of patents that cite the output of collaboration i.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
e_{ij} (knowledge contributed)		2.956e-02	1.442e-01***	1.489e-01***	1.246e-01**	7.197e-02	1.012e-01	1.018e-01
		(2.150e-02)	(3.744e-02)	(3.835e-02)	(4.042e-02)	(6.448e-02)	(8.766e-02)	(9.100e-02)
e_{ij}^2			-9.791e-03***	-9.917e-03***	-9.015e-03***	-7.275e-03*	-7.086e-03*	-7.084e-03*
			(2.639e-03)	(2.649e-03)	(2.676e-03)	(3.154e-03)	(3.166e-03)	(3.168e-03)
f_{ij} (firm distinctness)				3.284e-01	2.992e-01	2.878e-01	1.284e-01	1.032e-01
				(5.865e-01)	(5.867e-01)	(5.863e-01)	(6.680e-01)	(1.225e+00)
b_{ij} (broadening innovation)					-2.391e-01+	-3.888e-01*	-3.743e-01+	-3.686e-01
					(1.339e-01)	(1.973e-01)	(1.992e-01)	(3.076e-01)
$e_{ij}b_{ij}$						4.682e-02	4.408e-02	4.416e-02
						(4.482e-02)	(4.503e-02)	(4.514e-02)
$e_{ij}f_{ij}$							1.270e-01	1.299e-01
							(2.580e-01)	(2.839e-01)
$b_{ij}f_{ij}$								2.967e-02
								(1.209e+00)
patent lag	8.821e-04***	8.867e-04***	8.773e-04***	8.857e-04***	8.932e-04***	8.906e-04***	8.873e-04***	8.871e-04***
	(9.204e-05)	(9.190e-05)	(9.063e-05)	(9.186e-05)	(9.199e-05)	(9.179e-05)	(9.199e-05)	(9.212e-05)
total doc count	-3.468e-03+	-5.099e-03*	-5.004e-03*	-5.334e-03*	-4.506e-03+	-4.396e-03+	-4.588e-03+	-4.591e-03+
	(1.953e-03)	(2.273e-03)	(2.266e-03)	(2.343e-03)	(2.390e-03)	(2.387e-03)	(2.415e-03)	(2.419e-03)
firm pat distance	3.474e-01	3.516e-01	3.083e-01	2.655e-01	2.596e-01	2.661e-01	2.597e-01	2.595e-01
	(2.450e-01)	(2.448e-01)	(2.434e-01)	(2.547e-01)	(2.540e-01)	(2.539e-01)	(2.541e-01)	(2.542e-01)
ROA	-2.688e+00***	-2.696e+00***	-2.600e+00***	-2.578e+00***	-2.542e+00***	-2.516e+00***	-2.521e+00***	-2.523e+00***
	(5.211e-01)	(5.229e-01)	(5.211e-01)	(5.226e-01)	(5.215e-01)	(5.217e-01)	(5.217e-01)	(5.246e-01)
current ratio	1.663e-02	1.569e-02	8.420e-03	-1.066e-03	9.934e-04	1.126e-03	-1.408e-03	-1.368e-03
	(5.287e-02)	(5.275e-02)	(5.193e-02)	(5.427e-02)	(5.413e-02)	(5.401e-02)	(5.413e-02)	(5.416e-02)
size	1.447e-01**	1.350e-01**	1.204e-01*	1.209e-01*	1.242e-01*	1.248e-01*	1.219e-01*	1.219e-01*
	(4.959e-02)	(4.975e-02)	(4.903e-02)	(4.892e-02)	(4.893e-02)	(4.885e-02)	(4.911e-02)	(4.912e-02)
constant term	-1.353e+00*	-1.258e+00*	-1.212e+00*	-1.143e+00+	-1.006e+00+	-8.748e-01	-8.759e-01	-8.806e-01
	(6.055e-01)	(6.054e-01)	(5.956e-01)	(6.064e-01)	(6.106e-01)	(6.225e-01)	(6.214e-01)	(6.499e-01)
Log-likelihood	-2144.471	-2143.514	-2136.54	-2136.382	-2134.77	-2134.224	-2134.104	-2134.103
Observations	1497	1497	1497	1497	1497	1497	1497	1497

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

To probe H2 and H3 further, we examine the possibility that the effects of firm distinctness are dichotomous. That is, when other firms are sufficiently dissimilar, learning occurs, but at some point of increasing similarity learning quickly dissipates. Accordingly, we substitute the firm distinctness measure f_{ij} with a dummy variable d_{ij} that is set to one for firm distinctness levels above the median. Table 3.4 presents the results of the same tests as Table 3.3, except that this discrete measure of firm heterogeneity is substituted for the continuous measure.

The full model (Model 8) lends support for H2 since the coefficient on $e_{ij}d_{ij}$ is positive and significant (at the 10% level). Nevertheless, caution must exercised in interpreting the coefficients in count (or, in general non-linear) estimation models (Hoetker, 2007). Accordingly, we computed marginal effects at the means of the covariates, and the coefficient on $e_{ij}d_{ij}$ remains positive and significant at the 10% level. The magnitude of the marginal effect is 0.823 and the standard error is 0.423.

H3 is also supported in the alternative specification of Table 3.4. The coefficient on $b_{ij}d_{ij}$ is positive and significant, and an examination of marginal effects analogous to that performed in H2 also show significance (at the 5% level).

Table 3.4: Inter-firm heterogeneity (as a dummy variable) and firm's knowledge contributions as facilitators of follow-on innovations

		Dependent variable: Firm j's aggregate citation count of patents that cite the output of collaboration i.							
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
e_{ij} (knowledge contributed)			2.956e-02	1.442e-01***	1.137e-01**	8.818e-02*	2.856e-02	1.288e-01	1.325e-01
			(2.150e-02)	(3.744e-02)	(3.890e-02)	(4.100e-02)	(6.510e-02)	(8.475e-02)	(8.457e-02)
e_{ij}^2					-9.791e-03***	-8.645e-03**	-7.687e-03**	-5.712e-03+	-4.944e-03	-4.900e-03
					(2.639e-03)	(2.668e-03)	(2.705e-03)	(3.179e-03)	(3.156e-03)	(3.160e-03)
d_{ij} (firm distinctness dummy)					-3.524e-01**	-3.563e-01**	-3.619e-01**	-4.485e-01***	-6.655e-01***
					(1.247e-01)	(1.245e-01)	(1.244e-01)	(1.329e-01)	(1.591e-01)
b_{ij} (broadening innovation)						-2.482e-01+	-4.170e-01*	-3.622e-01+	-4.859e-03
						(1.337e-01)	(1.976e-01)	(1.984e-01)	(2.476e-01)
$e_{ij}b_{ij}$							5.271e-02	4.325e-02	5.437e-02
							(4.485e-02)	(4.452e-02)	(4.470e-02)
$e_{ij}d_{ij}$								4.362e-01+	4.643e-01+
								(2.373e-01)	(2.375e-01)
$b_{ij}d_{ij}$									1.979e+00*
									(8.184e-01)
patent lag		8.821e-04***	8.867e-04***	8.773e-04***	8.381e-04***	8.459e-04***	8.428e-04***	8.359e-04***	8.448e-04***
		(9.204e-05)	(9.190e-05)	(9.063e-05)	(9.195e-05)	(9.208e-05)	(9.185e-05)	(9.189e-05)	(9.176e-05)
total doc count		-3.468e-03+	-5.099e-03*	-5.004e-03*	-3.309e-03	-2.460e-03	-2.320e-03	-3.136e-03	-3.492e-03
		(1.953e-03)	(2.273e-03)	(2.266e-03)	(2.337e-03)	(2.382e-03)	(2.380e-03)	(2.406e-03)	(2.408e-03)
firm pat distance		3.474e-01	3.516e-01	3.083e-01	5.321e-01*	5.271e-01*	5.369e-01*	4.963e-01+	4.399e-01+
		(2.450e-01)	(2.448e-01)	(2.434e-01)	(2.568e-01)	(2.560e-01)	(2.558e-01)	(2.559e-01)	(2.563e-01)
ROA		-2.688e+00***	-2.696e+00***	-2.600e+00***	-2.775e+00***	-2.731e+00***	-2.704e+00***	-2.725e+00***	-2.815e+00***
		(5.211e-01)	(5.229e-01)	(5.211e-01)	(5.293e-01)	(5.283e-01)	(5.287e-01)	(5.295e-01)	(5.331e-01)
current ratio		1.663e-02	1.569e-02	8.420e-03	3.056e-02	3.139e-02	3.144e-02	1.137e-02	-1.038e-02
		(5.287e-02)	(5.275e-02)	(5.193e-02)	(5.342e-02)	(5.324e-02)	(5.314e-02)	(5.352e-02)	(5.342e-02)
size		1.447e-01**	1.350e-01**	1.204e-01*	9.840e-02*	1.014e-01*	1.018e-01*	8.691e-02+	7.756e-02
		(4.959e-02)	(4.975e-02)	(4.903e-02)	(4.972e-02)	(4.967e-02)	(4.958e-02)	(4.976e-02)	(4.967e-02)
constant term		-1.353e+00*	-1.258e+00*	-1.212e+00*	-8.254e-01	-6.676e-01	-5.110e-01	-3.026e-01	-6.374e-02
		(6.055e-01)	(6.054e-01)	(5.956e-01)	(6.135e-01)	(6.188e-01)	(6.323e-01)	(6.363e-01)	(6.406e-01)
Log-likelihood		-2144.471	-2143.514	-2136.54	-2132.559	-2130.818	-2130.128	-2128.452	-2125.42
Observations		1497	1497	1497	1497	1497	1497	1497	1497

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

3.8 Robustness Analyses

In this section, we discuss potential issues with the main empirical analysis and robustness checks to address these. First, citation counts are found to be representative of a patent's "impact" (e.g., market value) (Hall et al., 2005), but there is variation in how it is incorporated as a performance measure. As in this study, some studies simply use the raw citation count (e.g., Gittelman and Kogut, 2003), which discounts the value of the patent itself (i.e., absent forward citations) to zero. Others use some form of citation-weighting to give a positive value to the patent (e.g., Trajtenberg, 1990; Sampson, 2007). In earlier model specifications, we conducted alternate analyses using a citation-weighted count as the dependent variable (we weight the patent and citations equally), and results were robust to the alternative specification.

Second, firms self-select into the group that patents within each collaboration. To address this in our model specifications, we employed a Heckman selection model (using OLS in the second stage) in which we include the pool of potential patenters (i.e., any firm that has a follow-on innovation) as first stage observations for every collaboration. Since this alternative specification is no longer a count model (so that the exposure method above is not applicable), we applied a "fixed-effect" scaling to the dependent variable to adjust for citation truncation issues. Results were robust to this alternative specification.

3.9 Discussion

In many technology intensive industries, multi-firm collaborations are an important means by which firms conduct their productive activities. At the core of this study is the notion that a firm's knowledge contributions facilitate absorption of knowledge from collaborating firms, which serves to increase the firm's follow-on innovation performance. A key distinction from prior conceptions of absorptive capacity

is that absorption is facilitated by observable knowledge contributed to a collaboration and not by internal knowledge-building efforts. In particular, knowledge contributions may reveal to partnering firms the quality of the knowledge produced by the collaboration and the capabilities necessary to exploit this knowledge, which decreases the follow-on innovation performance of a focal firm as other firms become better equipped to develop competing innovations. We find support for these predictions using extensive data on collaborative development efforts within the 3GPP wireless standards organization.

This study contributes to refining knowledge-based explanations of inter-firm collaboration. Indeed, prior work has observed that "a key feature of knowledge-based explanations of alliances is imprecision around the concepts of organizational learning, knowledge sharing, and knowledge transfer" (Grant and Baden-Fuller, 2004). We make some headway towards addressing this issue by elaborating the role of knowledge integration in a multi-firm collaborative context. In particular, we conceptualize knowledge sharing within a collaboration as a codified contribution followed by subsequent integration efforts that typically involve tacit knowledge. Under this conceptualization, outbound contributed knowledge acts as the impetus for inbound knowledge transfers from the collaboration and subsequent learning within the firm.

The theory developed in this study has implications for the emergence of performance differentials within multi-firm collaborations. In particular, heterogeneity in relevant internal knowledge stocks leads to heterogeneity in knowledge contributed to the collaboration, which in turn gives rise to differences in the amount of knowledge absorbed from partnering firms and ultimately results in inter-firm performance differentials. In contrast to most prior work on multi-firm (and dyadic) collaborations, where performance advantages arise from access to partners' core, protected resources and capabilities (Lavie, 2006; Kale et al., 2000; Sakakibara, 2002), the theory in this study suggests inter-firm performance differentials may arise from freely observable

knowledge contributions. Observable knowledge contributions may be the stronger source of performance differentials in collaborations among many firms, in which case firms' proprietary knowledge may be more highly guarded due to the difficulties in establishing trust- or familiarity-based relationships within a large set of firms.

This study contributes to reconciling cooperative and competitive aspects of multi-firm collaborations. As Jorde and Teece (1990, p. 76) note, "very little literature addresses how cooperation among competitors can promote competition." This study raises one possibility: an underlying competition to reap ex-post benefits may emerge in cooperative contexts, and the driver of these benefits is knowledge contributions to facilitate knowledge absorption within a collaboration. The notion that competition may emerge within cooperation adds a pro-competitive argument to broad policy concerns over potential anti-competitive (e.g., collusive) behavior within multi-firm collaborations (Anton and Yao, 1995; Goeree and Helland, 2012).

In this study, we theorize that knowledge contributions underlie an important but under-examined process of knowledge absorption within multi-firm collaborations. Since prior work on multi-firm collaborations typically focuses on ex-ante characteristics of the collaboration such as knowledge diversity or experience among member firms, this study represents an early effort at examining activity within a collaboration. As such, there is substantial opportunity for future work to shed light on the underlying processes by which firms collaborate and their implications for subsequent firm performance.

CHAPTER IV

Gambling on the Past: Risk Taking as a Learned Response to Problem-Solving

4.1 Introduction

A wide variety of "problems," such as technology discovered to be dangerous (asbestos), increased competition, or industrywide shocks (a fuel shortage for SUV makers), may force a firm to replace existing practices with ones more suited to its newfound circumstances. While management scholars have stressed the importance of tackling problems to a firm's long-run organizational learning and performance (e.g., Nickerson and Zenger, 2004; Nickerson et al., 2012; Cyert and March, 1963; Arrow, 1962), problem-solving also comes with substantial risk because firms lack the knowledge to respond effectively, leading them to search across, and perhaps temporarily adopt, alternatives that can vary widely in quality (March and Simon, 1958; Amburgey et al., 1993; Nickerson and Zenger, 2004). A seemingly obvious means to reduce this risk is through anticipatory exploration – that is, exploration prior to a problem – to accumulate a "knowledge inventory" (Levinthal and March, 1993) that guides post-problem solution-finding and subsequent adaptation. We demonstrate that there are conditions under which this line of reasoning is not fully correct because anticipatory exploration engenders endogenous risk seeking in a firm's subsequent

problem-solving efforts.

One prevalent theme in extant work is that firms' lack of knowledge makes adaptation in the face of such problems risky. For example, research within the behavioral theory of the firm suggests that insufficient knowledge on available alternatives may result in performance below an organizational aspirational level, which leads firms to engage in an uncertain process of experimentation and search to raise performance above aspirations (Miller and Leiblein, 1996; March and Shapira, 1987). Research on firm entry provides another example, wherein a lack of pre-entry knowledge increases the binary risk of organizational failure (Klepper and Simons, 2000; Helfat and Lieberman, 2002; Dencker et al., 2009). Here, a lack of knowledge not only clouds the decision to enter but also stifles its ability to learn and adapt conditional on entry (Dencker et al., 2009).

We offer the alternative possibility that adaptation is risky because firms inherently gravitate to riskier alternatives when a problem forces change, even when firms are risk-neutral. The basis for this claim is the endogenous emergence of risk preferences in risk-neutral search. While prior work finds that, absent problems, agents (e.g., firms) exhibit apparent risk aversion under risk-neutral search (March, 1996; Denrell and March, 2001), we find agents exhibit apparent risk-seeking in reaction to a problem. In particular, we focus not on policy choices adopted in a static environment, but rather on the once-discarded choices to which a firm turns after a problem renders an existing policy decision unviable. It is in this latter regime where an increase in anticipatory exploration may result in increased risk-seeking within a firm's problem-solving efforts.

In this study, we seek to demonstrate that these two sources of risk in adapting to problems – insufficient knowledge and a search bias toward risky alternatives – gives rise to a tradeoff in anticipatory exploration. On one hand, anticipatory exploration reduces adaptation risk by augmenting a firm's knowledge inventory (Levinthal and

85

March, 1993) that can be used to solve a problem. On the other hand, anticipatory exploration also increases adaptation risk by biasing risk-neutral agents towards riskier solutions to a problem. In other words, anticipatory exploration reduces performance variability *across* different alternatives a firm selects in its attempts to solve the problem but increases performance variability attributable to variation *within* each alternative selected. Therefore, we find that exploration may either increase or decrease riskiness in responding to problems, depending on a firm's strategies and the environment in which it operates.

We elucidate our theory using a multiarmed bandit model (Gittins and Jones, 1979). This model has been used to examine the tradeoff between exploration–exploitation in a wide variety of fields, including economics and statistics (Berry and Fristedt, 1985; Gittins, 1979; Robbins, 1952), computer science (Holland, 1975; Sutton and Barto, 1998), and internet engineering (White, 2012). In the management literature, March and his collaborators have used the bandit model extensively to examine organizational learning (e.g., March, 1991, 1996, 2003, 2010; Denrell and March, 2001).

We apply the bandit model to analyzing a simple organization where managers are faced with the task of choosing one of many alternatives under uncertainty. In each period, the organization may choose to continue with the previously selected alternative or select a different one. The payoff for choosing a given alternative reflects a stochastic draw with alternative-specific mean and variance. Within this process of sequential decision-making, the organization seeks to maximize cumulative returns over time. Due to uncertainty, the organization makes its choice in each period based on its imperfect beliefs about the expected returns to each of the alternatives. In the absence of perfect knowledge, the organization employs a strategy that allocates effort between (a) exploiting the alternative currently believed to be superior and (b) exploring alternatives that currently seem less promising in the hopes of identifying

an alternative superior to the currently preferred alternative.

We assume a simple model in which a problem emerges after a period of stability. That is, a firm's adopted alternative (i.e., the alternative it most frequently selects) becomes unviable due to a change in the circumstances under which the firm operates (e.g., videocassette technology is obsolete after the emergence of DVD). Moreover, we add the stringent assumption that knowledge accumulated on other alternatives prior to the problem does not degrade afterwards. Even with this assumption, we find that, under suitable conditions, the increase in adaptation risk from endogenous risk seeking outstrips the decrease in adaptation risk from greater knowledge on alternative policy choices.

To illustrate our theory and model of risky adaptation in response to a problem, consider a pharmaceutical firm searching for its next major drug to develop. The firm is concerned that a narrow focus on exploiting one of the first promising prospects it finds might lead to difficulties in adapting to environmental change. In particular, if the firm encounters a problem with the drug and is forced to quickly adopt a replacement, its best bet may be to revisit the alternatives originally discarded in favor of this drug, but this would entail risky experimentation because of the firm's exploitive search strategy. As such, the firm decides to explore fairly broadly and eventually finds a drug to take to clinical trials – call this drug A. Suppose further that a problem subsequently renders drug A unviable (e.g., a side effect emerges during late stage trials). Armed with the knowledge from prior exploration, the firm is able to determine, with low-risk post-problem experimentation, what it thinks to be an effective alternative suited to the new environmental regime. Despite this, the firm finds, to its surprise, that adaptation is risky not because of volatile experimentation across different alternatives, but because of high risk (i.e., variable performance) in the development path within the new drug it has adopted.

Thus, the pharmaceutical firm finds anticipatory exploration to be a double-edged

sword for adapting to a problem, both enhancing knowledge that guides adaptive efforts and biasing the unit towards choosing riskier alternatives post-problem. How might this bias towards risky alternatives arise? Consider that the firm receives noisy feedback from anticipatory exploration across different alternatives. The unit is likely to see unusually good feedback from at least a few of the risky (i.e., high variance) alternatives, while the unit will see neither unusually good nor unusually bad feedback on low risk (low variance) alternatives. Hence, the R&D unit is likely to believe that one of the high variance alternatives is the best contingency to a future problem.

The remainder of the chapter is organized as follows. In the next section, I describe the canonical multi-armed bandit model and how we adapt it to analyze riskiness in problem-solving. I then use the model to analyze the main claim of the paper – that increased exploration prior to the emergence of a problem may in fact increase risk in finding a solution to the problem. Following this, I examine the sensitivity of the model to assumptions on how firms form beliefs on the alternatives and the strategies they employ for exploring different alternatives. Finally, I conclude with a discussion of the contributions made by this study.

4.2   Model

The term "multi-armed bandit" refers to the analogy with a bank of slot machines each with an arm that gives a noisy reward when pulled. In the conventional bandit model, an agent performs a trial by selecting one of these arms in a given time period. Arms are heterogeneous in payoffs (i.e., average rewards). Through repeated trials on this bank of slot machines over time, a player's objective is to "maximize ... (cumulative) winnings by concentrating ... plays on the best levers" (Sutton and Barto, 1998). The salient point of the bandit model is that information on the merit of an arm is only obtained by selecting an arm. The bandit model has been applied widely to practical applications such as project budget allocation in a large

organization and routing algorithms in communication networks. The bandit model has also been used as a device to examine various issues in economics (Bergemann and Valimaki, 1996) and strategy (Posen and Levinthal, 2012).

The multiarmed bandit can be mapped to an organizational context by viewing firms as adopting exactly one from among a set of possible policy alternatives at any given time period. The reward (performance) a firm receives choosing an arm i is given by a mean payoff for the alternative plus an alternative specific noise term, i.e.

$$r_i = p_i + \eta_i, \tag{4.1}$$

where p_i is a payoff drawn from a fixed underlying distribution and η_i is a Gaussian noise term. The payoffs, p_i, are independently drawn for each arm of each firm prior to the first search iteration and remain stable throughout all time periods (i.e., p_i is a "true" payoff for alternative i that is drawn from a Bayesian prior probability distribution, which is specified below). The noise term η_i is drawn independently each time a firm selects arm i to represent noisy measurements of the underlying true payoff p_i.

We assume firms are boundedly rational as they engage in sequential decision-making within the multiarmed bandit model. Two constructs are central to this decision-making process. First, firms hold beliefs – that is, assessments – on the merits of the N alternatives (or arms) from which they must choose. At each time period t, firms select one of these N arms based on their beliefs, and they update their beliefs based on the outcome of this selection. We set beliefs of arm i for each firm to be the average reward from all prior trials of that arm. Thus, the belief $q_{i,t}$ after an arm i is selected at time t is given by:

$$q_{i,t} = \frac{1}{k} \sum_{\tau=t_1}^{t} r_{i,\tau} \tag{4.2}$$

where the sum ranges over all time periods up to (and including) t in which arm i was selected, and k is the number of times i has been selected.

Second, firms are assumed to employ a search strategy. One choice would be an optimal strategy that maximizes cumulative rewards, assuming such a strategy exists. While Gittins and his collaborators (e.g., Gittins and Jones, 1979; Gittins, 1979) have proved the existence of an optimal strategy for the multiarmed bandit (under certain conditions), it is computationally intractable to implement for all but a limited set of assumptions (Gans et al., 2007). Moreover, research suggests that, in practice, firms (or individuals) often use much simpler heuristics when confronted with bandit-like decision problems (Meyer and Shi, 1995). Indeed, a heuristic commonly assumed in a diverse bandit literature is one in which firms simply choose the alternative that provides the highest expected reward in the current time period (e.g., Gans et al., 2007).

This simple heuristic is known as the greedy algorithm (Sutton and Barto, 1998). However, since the extent to which firms exploit good alternatives is central to this study, a strategy that allows for tuning the degree of exploration is needed. The canonical method for modifying the greedy algorithm to incorporate exploration is the ϵ-greedy algorithm, and I adopt this modified strategy for the analysis below. In the ϵ-greedy algorithm, agents choose a random alternative with probability ϵ and are otherwise greedy (i.e., choose the arm for which they hold the highest belief). Hence, $\epsilon = 0$ represents a fully exploitive (greedy) strategy in which a firm always selects the alternative believed to be best and $\epsilon = 1$ represents the opposite extreme whereby a firm chooses an alternative randomly. Importantly, ϵ-greedy strategies consider only the *rewards* of alternatives, not their *variances* – in this sense, it represents search by risk-neutral agents.

To simplify our analysis of the effects of problems on firm policy choices, we construct a task environment where firms must select an alternative post-problem

that is different from the one they adopted (i.e., selected most often) pre-problem. In our model, we implement this by partitioning time periods into a pre-problem interval in which the environment is fixed for $I = 50$ periods, and a post-problem interval in which firms cannot select the pre-problem alternative they adopted. We simulate $I + 2$ periods because 2 periods in the post change environment are required to observe post change performance risk.

Our main theoretical concern is whether environmental change affects performance risk (i.e., performance variability) through a lack of knowledge on good alternatives – that is a lack of experience in trialling these alternatives – or the underlying variability within each selected alternative. As such, we define measures of reliability corresponding to both possibilities. First, *inexperience risk* is defined as:

$$\rho_e(t) = \sum_{\tau=2}^{t} \left| p_{i(\tau),\tau} - p_{i(\tau-1),\tau-1} \right| \tag{4.3}$$

where p_i is as in Equation 4.1. $\rho_e(t)$ reflects the notion that when firms lack information on which alternative is best, they are likely to select different alternatives that vary widely in payoff from one period to the next. Second, we define *selection risk* as:

$$\rho_v(t) = \sum_{\tau=2}^{t} \left| \eta_{i(\tau),\tau} - \eta_{i(\tau-1),\tau-1} \right| \tag{4.4}$$

where η_i is as in Equation 4.1 . $\rho_v(t)$ reflects the notion that when firms select noisy (i.e., high variance) alternatives, rewards vary considerably from one period to the next even if they select the same alternative.

The opportunity structure of the environment is defined by initializing the payoffs, p_i, of a 10-arm bandit model. p_i are drawn from a uniform $U(-1, 1)$ distribution. Moreover, we assume firms know the distribution from which payoffs are drawn and

set their initial beliefs on all alternatives, $q_{i,0}$, to zero, the mean of this distribution.[1] Finally, we introduce heterogeneity in the risk (variance) of each alternative by setting half of each firm's alternatives to be low variance, i.e., η_i is deterministic (i.e., zero with zero variance), and setting the other half of each firm's alternatives to be high variance, i.e., η_i is drawn independently for each arm and each time period from a $\mathcal{N}(0,4)$ distribution. Thus, firms are (on average) risk neutral if the average variance of the alternatives they select is two. The variance of arms are set independently of variance so that, for example, firms would be risk neutral if they were able to identify the best alternative with certainty[2]

All simulations of the model are seeded with 25,000 firms, each facing a distinct environment initialized in this fashion.

4.3 Analysis

We employ the multiarmed bandit model described above to study the sources of performance reliability degradation under environmental change. Doing so requires setting a number of parameters. First, we define the opportunity structure of the environment by drawing the payoffs, p_i, from a uniform $U(-1,1)$ distribution. Second, we assume firms know the distribution from which payoffs are drawn and set their initial beliefs on all alternatives, $q_{i,0}$, to zero, the mean of this distribution. Third, we seed each simulation of the model with with 25,000 firms, each with 10 alternatives from which to choose. The environment is set independently for each firm – that is, the payoffs for each arm of each firm are drawn independently and identically from a uniform $U(-1,1)$ distribution.

The analysis comprises a test of our main proposition on the effects of exploration

---

[1]Another possibility commonly employed is to assume agents know the underlying distribution of payoffs and possess initial beliefs that are randomly drawn from this distribution.

[2]Doing so clarifies the mechanisms underlying our theory. Our main findings in Figure 4.1 are robust to continuous distributions on the variance alternatives (e.g., variance of alternatives drawn from $U(0,2)$) [do this]

on post-problem performance risk and tests of the sensitivity of our main results to a firm's method for forming beliefs and a firm's strategy for exploring different alternatives.

4.3.1 Main Experiment

Our main experiment examines the effects of increased exploration on the performance risk that firms face as they respond to a problem with an adopted alternative. Per our theory, we decompose risk into inexperience risk and selection risk in the analysis that follows.

In Figure 4.1, we compare performance risk immediately pre- and post-problem as a function of exploration (i.e., as a function of the ϵ-greedy parameter ϵ). We examine low and moderate levels of exploration where the tradeoff between inexperience and selection risk is most evident. Figure 4.1a shows the change in risk from the last period pre-problem (t = 50) to the first period in which post-problem risk is defined (t = 52). In Figure 4.1a, we focus on pre-problem risk. The solid line in this figure indicates performance risk increases with exploration. The increase in inexperience risk ("o" markered line) embodies a basic compromise of greater exploration: performance variation increases as firms gather more information across all alternatives through exploratory trials. Note here that we have decomposed inexperience risk into a pure ("o" markered line) and an interaction component ("x" markered line), the latter reflecting additional inexperience risk (i.e., performance variation in selected alternatives) induced by noise in each alternative. The effects of this interaction have largely disappeared after 50 periods pre-problem. Finally, the increase in selection risk (triangle markered line) reflects risk aversion in alternatives selected by the (risk-neutral) firms in this experiment. Risk aversion is most pronounced at low levels of exploration, which we discuss in greater depth in Sections 4.3.1.1 and 4.3.1.3.

We now turn attention to post-problem risk (Figure 4.1b), which is at the core of

our theorizing. All firms are greedy post-problem – that is, they select the alternative believed to be best – which eliminates the contribution of random exploration to risky adaptation and allows us to isolate our analysis on the competing effects of inexperience and selection risk. Figure 4.1b shows that the post-problem risk associated with firms' actions declines as exploration increases from $\epsilon = 0$ to $\epsilon = 0.1$. As the markered lines indicate, this is due to a decrease in inexperience and (to a smaller extent) selection risk that dominate the increased interaction between the two. The decrease in inexperience risk is simply a result of firms holding more information across all alternatives. We defer discussion of post-problem selection risk to Sections 4.3.1.1 and 4.3.1.3.

Corroborating the central claim of this study, post-problem risk increases as exploration increases from $\epsilon = 0.1$ to $\epsilon = 0.4$. Increased exploration within this range leads to a continued decline in inexperience risk, which is consistent with gaining more knowledge about all alternatives.[3] However, increased exploration also leads to increased selection risk (triangle markered line), which dominates overall performance risk at higher levels of exploration (solid line). Hence, as we theorized, pre-problem exploration increases post-problem risk seeking. We analyze this in greater detail in Section 4.3.1.3.

4.3.1.1 Selection risk as variance of selected alternatives

The preceding results show that riskiness in adapting to environmental change is driven by two competing effects. On the one hand, exploration reduces inexperience risk by enhancing a firm's knowledge across alternative choices. This result confirms a simple corollary to an issue identified in prior research: if a lack of experience (knowledge) on alternatives makes adaptation risky, then exploration should reduce

---

[3]While the interaction between inexperience and selection risk is non-trivial in this range of exploration, it is relatively flat and hence does not contribute significantly to an overall rise or decline in performance risk.

Figure 4.1: Performance risk pre and post problem

(a) pre-problem risk (t = 50) (b) Post-problem risk (t = 52)
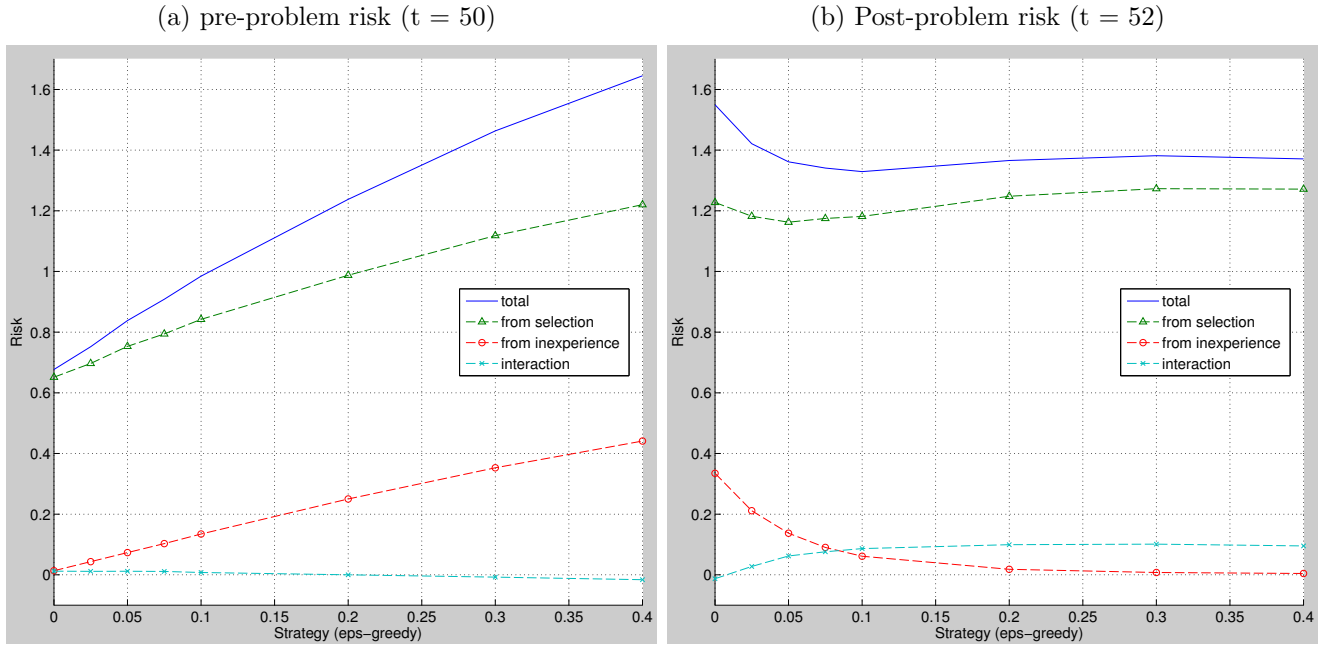
risk associated with inexperience. On the other hand, exploration increases selection risk, which dominates the knowledge-based reduction in inexperience risk gained at higher levels of exploration. In the discussion below, we examine more closely why pre-problem exploration increases post-problem (and pre-problem) selection risk.

Selection risk directly reflects the variance of the alternatives a firm selects (i.e., the variance of η_i in Equation 4.1, where i is the alternative selected by a firm in some time period). As such, we graph the variance of selected alternatives as a function of exploration, averaged across the firms in the experiment (see Figure 4.2a). As Figure 4.2a shows, prior to environmental change (solid line), firms are highly risk-averse at low ϵ but become less so as ϵ increases. Post-problem, firms become progressively more risk seeking at higher levels of exploration (from $\epsilon = 0.1$ to $\epsilon = 0.4$). This is consistent with the increase in selection risk that drives the overall increase in post-problem performance risk (i.e., as shown in Figure 4.1b).

Figure 4.2 illustrates a risk-reward tradeoff of adaptation: increased pre-problem

Figure 4.2: Pre- and post-problem variance and performance of alternatives selected by firms
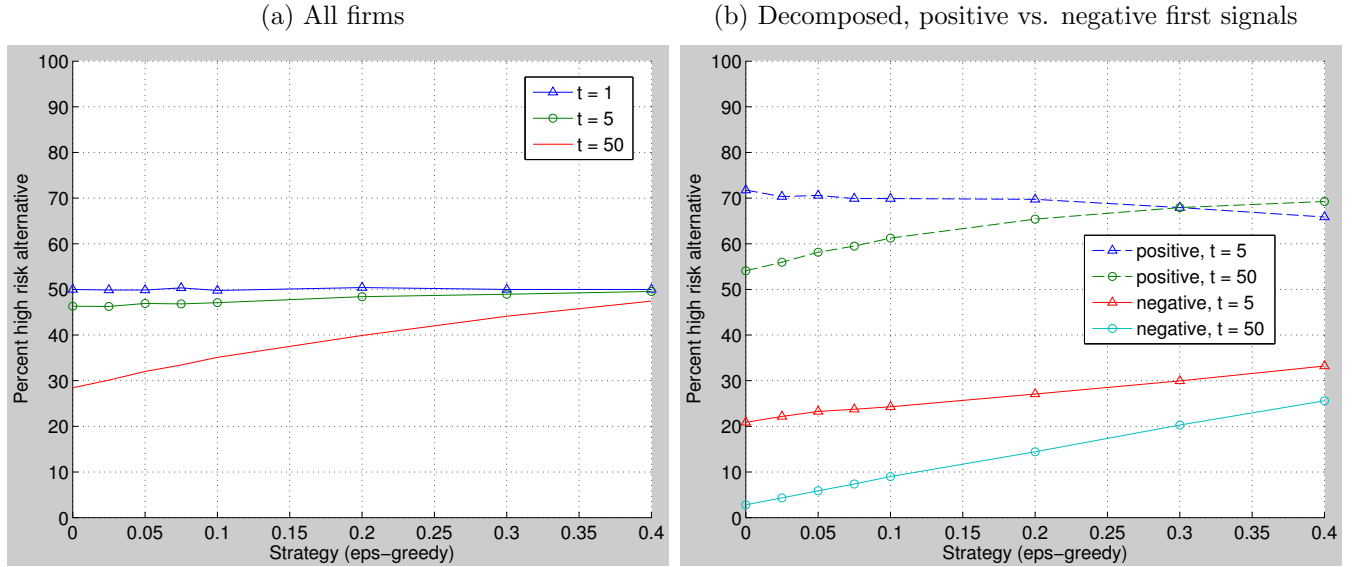

(a) Variance
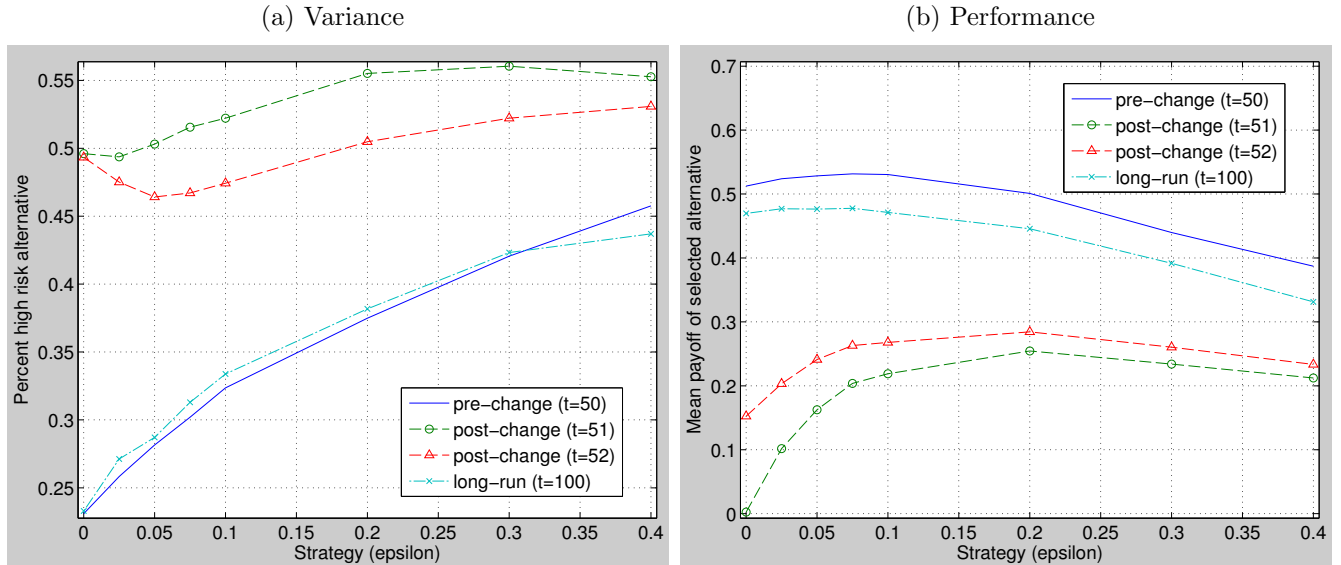
(b) Performance

exploration engenders both riskier adaptation (cf. Figures 4.1b and 4.2a) and higher performance (cf. Figure 4.2b) post-problem. This increased performance lends additional support to our assertion that increased pre-problem exploration should lead to better information on the best alternatives to select post-problem.

4.3.1.2 Variance of alternatives selected pre-problem

The prior section connected selection risk to the variance in a firm's selected alternatives. In this and the next section, we unpack how (risk-neutral) firms form risk preferences (i.e., prefer low or high variance alternatives) within a search process that seeks to maximize cumulative rewards. We begin by analyzing the pre-problem case and proceed by considering how frequently firms' high variance alternatives are selected (since this determines risk preference). Figure 4.3a shows the percentage of times high variance alternatives are tried (on average) across strategies and across three different points in time of the experiment. Firms are initially risk-neutral (i.e., they select high variance alternative approximately 50 percent of the time) across

Figure 4.3: Endogenous risk aversion pre change

(a) All firms

(b) Decomposed, positive vs. negative first signals

strategies since, with no samples from which to form informed beliefs, firms are select-
ing randomly with respect to variance. However, the figure also shows firms become
progressively more risk averse over time, and more so at lower levels of exploration.

How does this risk aversion emerge? To see how this arises, we decompose the
high variance alternatives according to whether the first signal – that is, r_i from the
first trial – is negative or positive (Figure 4.3b). Since high variance alternatives
are equally likely to have positive and negative first signals, the average of the two
lines in this Figure (for a given t) gives the total percentage of times a high variance
alternative is selected. The salient feature here is the asymmetry in the fraction of
high variance trials over time. This asymmetry appears because exploitation entails
rapid decisions on which alternatives are inferior: early negative signals lead firms
to perceive an alternative as bad, so in an effort to exploit better alternatives, the
firm decreases (perhaps prematurely) trials on this alternative. Consistent with this
notion, Figure 4.3b shows that negative first signal alternatives are rapidly ignored
by the firm, particularly at lower levels of exploration. Since negative initial signals
typically reflect negatively biased beliefs on the payoff of an arm, negative belief biases

on high variance arms tend to persist over time. In contrast, positive initial signals, which typically reflect upwardly biased beliefs on alternatives, self-correct through repeated trials. The net effect is that firms on average hold negatively biased beliefs on high variance arms (and, of course, unbiased beliefs on low variance arms), which leads to firms favoring low variance arms. In other words, firms appear risk averse prior to environmental change.

4.3.1.3 Variance of alternatives selected post-problem

We now focus attention on the variance of alternatives a firm selects immediately post-problem, which determines post-problem selection risk. This variance is a function of a firm's beliefs on the merits of different alternatives, as formed through the firm's sequential decision-making strategy prior to environmental change. Simply put, a firm will select a high (low) variance alternative immediately post-problem if it believes a high (low) variance alternative to be best. Hence, we and compare a firm's maximum beliefs within its high and low variance subsets of (non-modal) alternatives (because, by construction, the modal alternative cannot be selected post-problem).

Figure 4.4 shows the maximum beliefs firms hold on their high and low variance set of alternatives as a function of exploration. Note that for all levels of exploration ϵ (except near 0), firms believe a high variance alternative is best, or, equivalently, the maximum belief on firms' high variance sets are higher than the maximum belief on their low variance sets. Moreover, the difference in beliefs between high and low variance sets increases with exploration, which is important to explaining why post-problem selection risk increases with exploration (see Figure 4.1b).

Why do firms believe the best alternative is high variance, and increasingly so with exploration? To gain insight on this, we decompose the maximal belief on the low and high variance set of alternatives into (a) the contribution from a single alternative – specifically, the alternative with highest true payoff – and (b) the contribution from

Figure 4.4: Beliefs on low/high variance alternative subsets
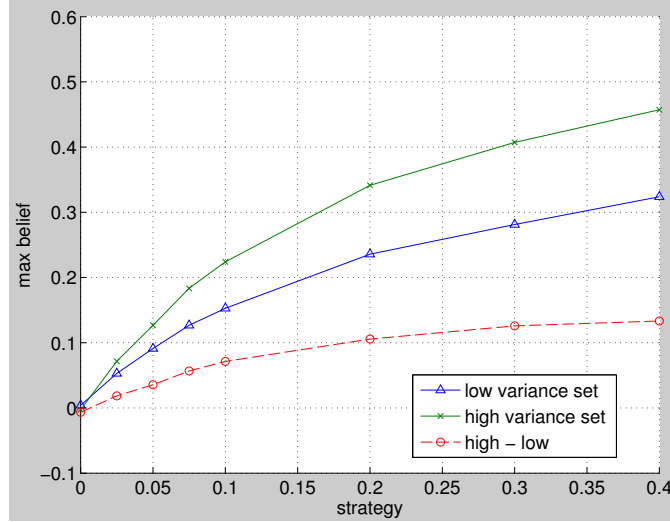
inclusion of the rest of the (non-modal) alternatives.

As the triangle-markered line in Figure 4.5b indicates, the main driver of increasingly disparate beliefs on the high versus low variance set is that, within the high variance set, there is a large gain from inclusion of multiple alternatives, and this gain grows with exploration.[4] The substantial increase is due to an *aspiration effect* of exploration, a *diversity effect* of exploration, and their interaction. The aspiration effect refers to the notion that exploration uncovers increasingly superior alternatives, which elevates the level at which beliefs on subsequently chosen alternatives cease to self-correct. The diversity effect is a high maximum belief on a set of alternatives when measurements on each individual alternative are noisy, making it highly likely that at least a few measurements are positively biased.[5] Within the high variance set, inclusion of additional alternatives has a large impact on maximum beliefs because these two effects interact: even a moderate quality (i.e., in true mean) alternative

---

[4]The single alternative contribution is stronger for the low variance set (i.e. x-markered lines in Figure 4.5), which is expected since negatively biased signals from the single high variance alternative is slow to self-correct. However, this difference in the single alternative contribution is dominated by the effect of including additional alternatives into beliefs on the high variance set.

[5]This reflects a simple statistical phenomenon. For example, the expected belief from single trials of two zero mean, unit variance alternatives is greater than zero since the probability of positive belief on at least one alternative is 3/4.

Figure 4.5: Belief decomposition

(a) Low variance set decomposition   (b) High variance set decomposition
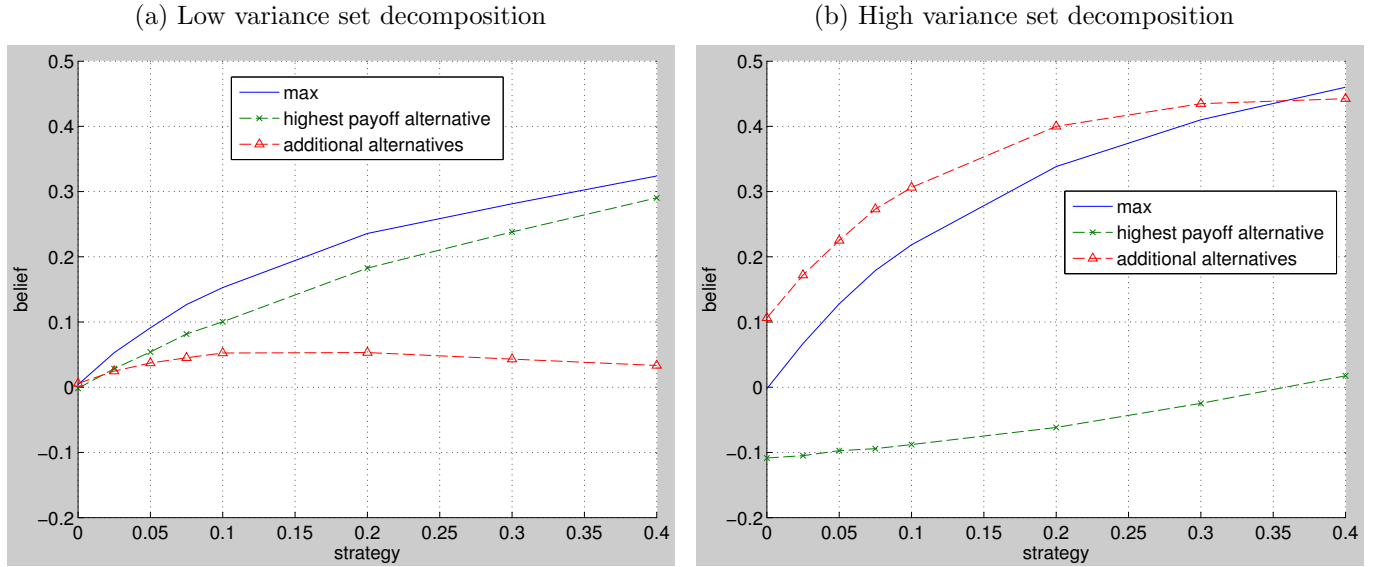
within the high variance set may contribute a large diversity effect (i.e., its inclusion raises maximal belief significantly) because positive signals self-correct to aspiration levels, not actual quality. In contrast, within the low variance set, inclusion of additional alternatives contributes little to maximal beliefs since measurements on these alternatives are (by definition) not noisy. This is confirmed by the triangle-markered line in 4.5a.

4.4 Sensitivity Analyses

In this section, we conduct sensitivity analyses for the main results (e.g., Figure 4.1b). In particular, we examine sensitivity along the two key dimensions of firms' behavior in the model: their beliefs and their firm's strategies.

4.4.1 Sensitivity to belief formation

We examine sensitivity of our main results to the means by which firms form beliefs. Recall earlier that our main model makes the assumption, frequently seen

in bandit model applications, that agents employ average updating to form beliefs (see 4.2). Here, we employ a *constant updating* heuristic (e.g., see Sutton and Barto, 1998) given by:

$$q_{i,t} = \beta r_{i,t} + (1 - \beta)q_{i,t-1} \tag{4.5}$$

While, with average updating, the impact of rewards on beliefs declines with successive trials of a given alternative, the impact remains the same with constant updating. As such, the constant updating heuristic is well-suited to nonstationary environments.[6]

Figure 4.6: Sensitivity of performance risk to belief formation heuristic

(a) Pre-problem (t=50)   (b) Post-problem (t = 52)
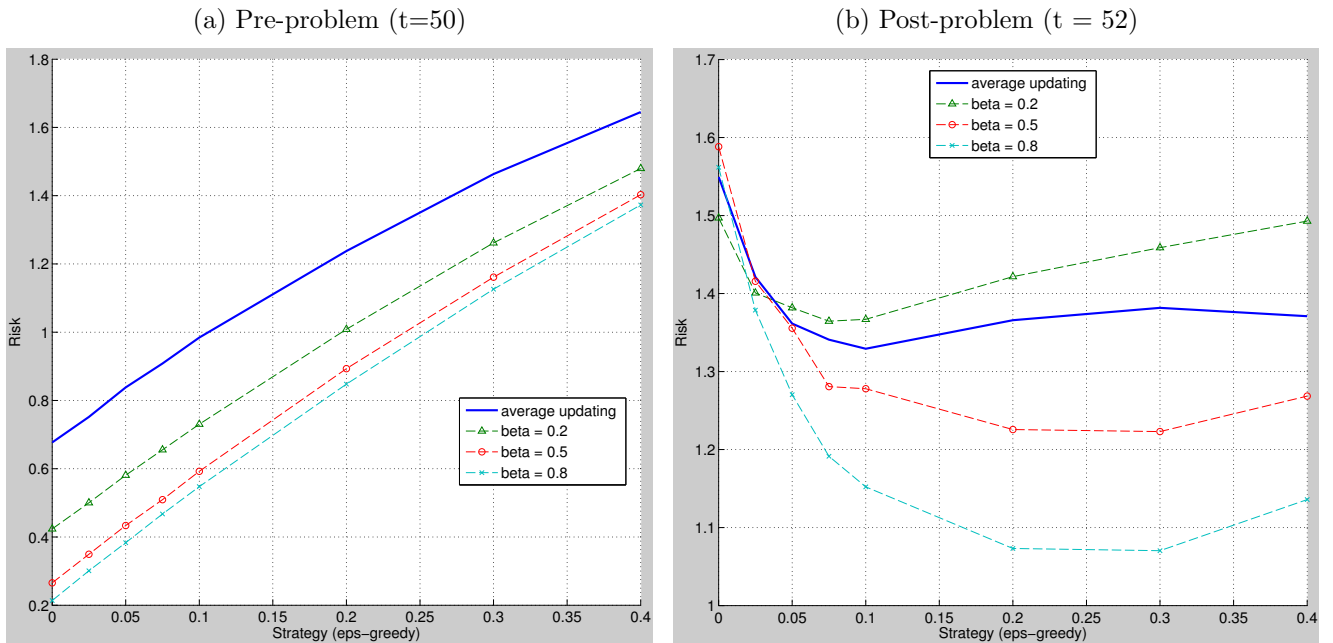
Under constant updating, firms exhibit less performance risk, both pre- and post-problem emergence, when recent rewards have a higher impact on beliefs (see Figure 4.6). The intuition underlying both cases is that reduced performance risk arises from

---

[6]Strictly speaking, the environment we model is nonstationary, but we only introduce shock at one time interval, as opposed to a turbulent environment where change is more pervasive. As such, average updating is a relatively sensible choice for our model.

negative signals on recently selected high variance alternatives, which diminishes the perceived quality high variance alternatives relative to low variance alternatives. In contrast, under average updating, a recent negative signal has a more limited effect on beliefs since its contribution is diluted by prior trials on the corresponding alternative.

We focus first on performance risk prior to the emergence of a problem. As seen in Figure 4.6a, firms have less performance risk with constant updating as compared to average updating, and performance risk decreases when recent rewards are weighted more heavily in forming beliefs (i.e., as β increases). To understand this, we focus on the set of high variance alternatives. Recall from the main results that performance risk is driven by the net effects of inexperience (from variance in mean quality across alternatives chosen) and selection (from noise variance within alternatives chosen). Pre-problem inexperience risk is driven by exploration, which is fairly constant across different belief formation mechanisms. As such, we focus on the differential effects of belief formation on selection risk.

Selection risk – that is, variance in selected alternatives – is lower for constant updating than average updating because the more that beliefs reflect recent trials, the less likely it is for high variance alternatives to be selected. To see why this is so, we examine how the frequency of high variance trials evolves over time. For convenience, we reproduce the corresponding result from the main analysis (Figure 4.3b) side-by-side with the result for $\beta = 0.2$ in Figure 4.7.

For average updating, the tendency against selecting high variance alternatives is driven mostly by early negative signals, which leads firms to believe the corresponding alternative is worse than it really is and in turn inhibits self-correction through future trials (see Figure 4.7a). Early positive signals, on the other hand, lead a firm to believe an alternative is better than it really is and therefore continue self-correction through repeated trials. Moreover, subsequent negative signals have little impact on biasing beliefs downward because later trials have a smaller effect on beliefs, as discussed

Figure 4.7: Pre-problem evolution of high variance trials

(a) Average updating

(b) Constant updating, $\beta = 0.2$
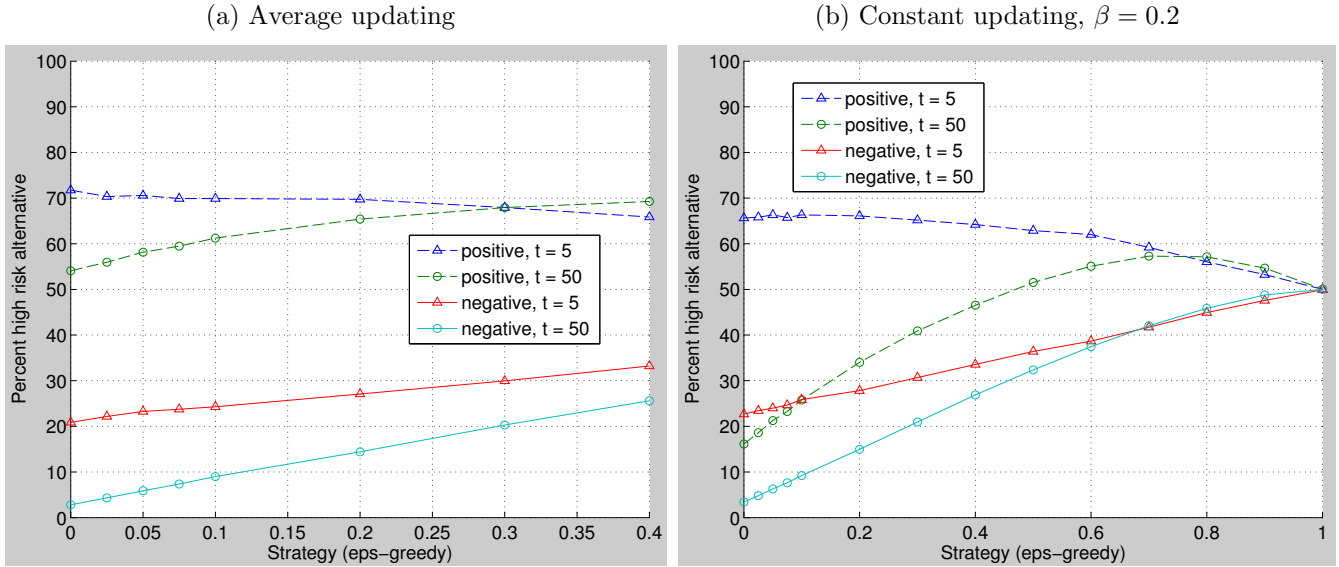
above.

In contrast, the bias against selecting high variance alternatives is driven by *both* early negative and early positive signals, making apparent risk aversion stronger. While early negative signals operate much in the same manner as the average updating case, early positive signals differ markedly, as seen by the circle-markered line in Figure 4.7b, which suggests alternatives with early positive signals also exhibit a negative belief bias over time. The reason for this is that early positive signals do not carry the "inertia" that they do in the average updating case for insulating an alternative against downward belief biases from later negative signals: indeed, beliefs on any (high variance) alternative can be strongly biased by subsequent negative signals because their effects do not diminish over time.

We now focus attention on post-problem performance risk (Figure 4.6b). At a basic level, we see that performance risk is qualitatively similar to average updating case. Two observations are worth noting, and will be discussed below: (a) the inflection point at which exploration increases risk differs (and is lowest for $\beta = 2$) and (b) performance risk decreases as the impact of recent trials on beliefs increases

(i.e., increasing β). the Since post-problem inexperience risk is relatively invariant to belief formation (at low exploration, it is largely driven by early random selection of alternatives, and at higher exploration, it is very low), we can restrict attention to the effects of belief formation heuristics on selection risk (i.e., from noise variance within alternatives chosen).

Post-problem selection risk is driven by the same underlying mechanisms as pre-problem selection risk, though its manifestation differs. In particular, an increase in the importance of recent trials to belief formation weakens the aspiration effect, where, as discussed in Section 4.3.1.3, increasingly superior alternatives uncovered from exploration sets a higher "bar" at which subsequently selected alternatives cease to self-correct. Alternatives tend to fall further below aspiration levels when recent trials have a larger impact on beliefs, which can be illustrated with a numerical example. Suppose aspiration level (i.e., highest belief outside of a focal alternative) is 2, and the belief on a focal alternative that has been selected several times is 2.1. Under average updating (or small β in constant updating), the next trial on the focal alternative has little impact on its belief, so it drops to at worst (for example) 1.9 and ceases to self-correct. As the most recent trial increases in importance, however, the next trial has a large impact on beliefs, so the drop is potentially much higher (to 1.0, for example).[7]

As such, the net effect of β on post-problem selection risk (and performance risk) is a combination of aspiration effect weakening, which tends to be stronger at low exploration levels and higher β, and the aspiration effect itself, which is stronger at high exploration levels. This explains observations (a) and (b) noted above.

---

[7]Since post-problem risk preference is driven by differentials in beliefs of high and low variance alternatives, a second possible driver of post-problem risk seeking is a decrease in beliefs on low variance alternatives. This decrease in beliefs is more acute at lowβ because convergence to true values of superior low variance alternatives is slow in the first few trials (and indeed initially slower than with average updating).

4.4.2 Sensitivity to strategy

We examine sensitivity of the results to the search strategy employed by firms. Under the ϵ-greedy strategy used in the main analysis, choices believed to be near-best and choices believed to be the worst are selected with equal probability. One obvious extension of this to weight selection probability continuously according to its perceived merit, and these types of choice mechanisms are known as softmax strategies. The most common implementation of the softmax strategy utilizes a Boltzmann distribution to generate choice probabilities as follows:

$$m_i = \frac{e^{q_i/\tau/\tau_S}}{\sum_i e^{q_i/\tau/\tau_S}}$$

where m_i is the probability of selecting alternative i, q_i is the belief on alternative i, τ is a strategy parameter reflecting extent of exploration, and τ_S scales τ for convenience.

ϵ-greedy and softmax differ qualitatively in how they explore different alternatives, so there is no exact equivalence from a given setting of the ϵ parameter to the τ parameter in the ϵ-greedy and softmax strategies. Nevertheless, for the purposes of comparison, we define the extent of exploration as the fraction of times that an alternative chosen differs from the one chosen in the prior period. Thus, any ϵ/τ parameter setting for ϵ-greedy/softmax can be mapped to a level of exploration. In this section, exploration parameters across the two strategies are mapped (or interpolated) to an exploration level according to Table 4.1.

We first examine sensitivity of performance risk to strategy prior to the emergence of a problem, as shown in Figure 4.8a. Since inexperience component of performance risk reflects exploration level across the two strategies, and is therefore nearly identical in both cases, we may focus on selection risk (i.e., noise variance within each selected alternative) as the driver of this increase.

As the figure suggests, the process leading to risk aversion under softmax is quali-

Table 4.1: Translation tables from ϵ/τ to exploration level

(a) ϵ to level of exploration

epsilon	exploration level
0	0.1016
0.025	0.137
0.05	0.1716
0.075	0.2074
0.1	0.242
0.2	0.3764
0.3	0.4977
0.4	0.6038
0.5	0.6931
0.6	0.7669
0.7	0.8229
0.8	0.8634
0.9	0.8875
1	0.8956

(b) τ to level of exploration

tau	exploration level
0	0.1018
0.4	0.1398
0.6	0.1654
0.9	0.2097
1.25	0.2715
1.75	0.3614
2.5	0.4933
3	0.569
3.25	0.6001
3.75	0.6534
4	0.6751
5	0.7422
6	0.7841
10	0.8518
15	0.8755
20	0.884
25	0.8883
30	0.8901

tatively similar that in the ϵ-greedy strategy, the latter of which is described in Section 4.3. Nonetheless, the modest increase in selection risk (or decrease in risk aversion) at intermediate exploration levels reflects the notion that, as discussed above, softmax departs from ϵ-greedy by selecting superior, but currently not believed to the best, alternatives more frequently than inferior (and not believed to be the best) alternatives. As such, we would expect the the alternative that is actually the best, 50% of which are in the low and high variance set, to be identified more reliably by softmax, which would tend to counter the risk aversion process.

We now turn attention to performance risk after the emergence of a problem, as shown in Figure 4.8b. Since post-problem inexperience risk is relatively invariant to strategy (in both cases, it is largely driven by early random selection of alternatives at low exploration levels and it is low at high exploration), we can restrict attention to the effects of strategy on selection risk (i.e., from noise variance within alternatives chosen).

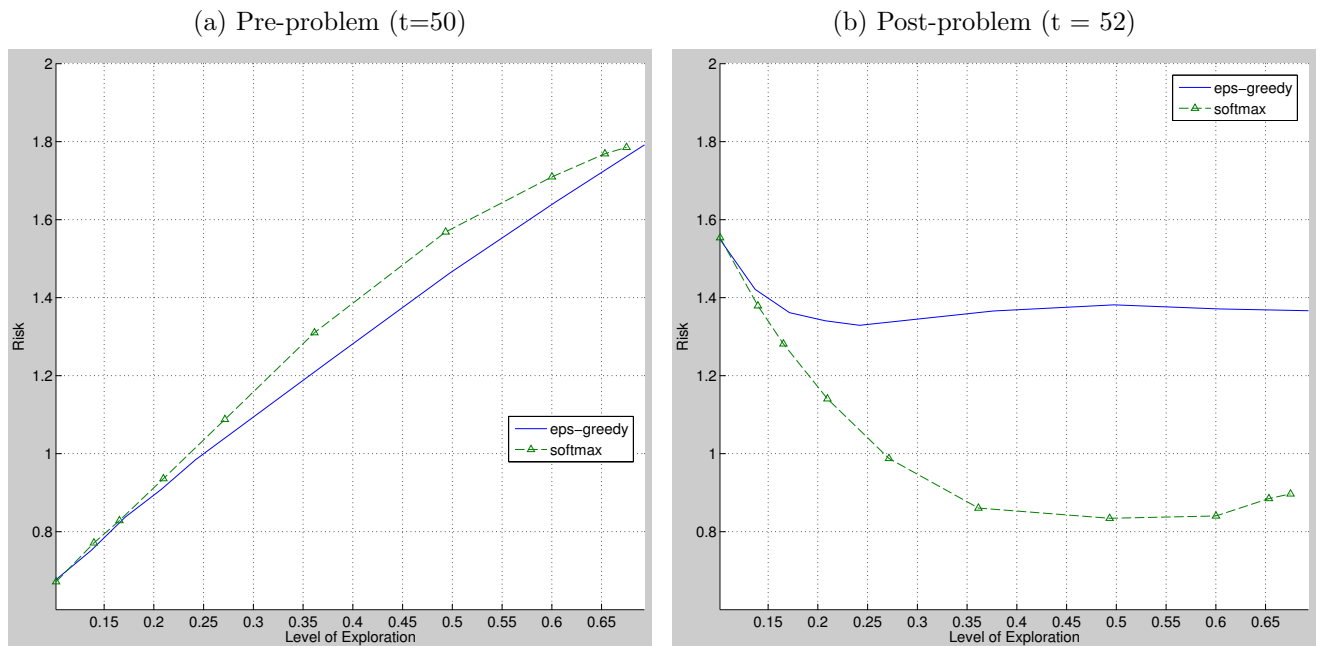Figure 4.8: Sensitivity of performance risk to strategy

(a) Pre-problem (t=50) (b) Post-problem (t = 52)

Figure 4.8b shows that selection (and performance) risk is lower for softmax, par-

107

ticularly at higher levels of exploration. However, our basic result still holds: beyond some level of exploration, performance risk increases with increased anticipatory exploration. In the discussion below, we explain these results by focusing on non-modal alternatives since the modal alternative cannot be selected after the emergence of a problem.

Post-problem selection (and hence performance) risk is the net result of two effects. First, in comparison to ϵ-greedy, softmax beliefs on the best (non-modal) low variance alternative are higher and increase more with exploration, which tends to drive down post-problem selection risk. This is a reflection of softmax favoring – and hence self-correcting – higher quality alternatives, whereas in ϵ-greedy, high quality alternatives are on equal footing with inferior alternatives once they are not perceived to be best. Second, an additional decline post-problem selection risk arise because beliefs on high variance arms are lower for the softmax strategy. This in turn is a consequence of the aspiration effect increasing over time (it must since higher quality alternatives are uncovered over time) and the softmax strategy sifting through non-modal alternatives more quickly than the ϵ-greedy strategy. As Figure 4.9 indicates, the best non-modal high variance alternative is, on average, selected considerably sooner in softmax than in ϵ-greedy, which is consistent with analyses indicating faster convergence of the softmax algorithm under suitable conditions (White, 2012).

4.5   Discussion

Why is organizational adaptation to problems risky? Common wisdom would suggest that the adoption of new alternatives with which a firm has little knowledge comes inherently with risk. Extant research has confirmed this notion through more thorough examinations from a variety theoretical lenses such as behavioral theories of the firm (March, 1991) and evolutionary economics (Tushman and Anderson, 1986). While it seems obvious that anticipatory exploration across the space of alternatives

Figure 4.9: Time periods when best (non-modal) high variance alternatives chosen

(a) First time

(b) Last time
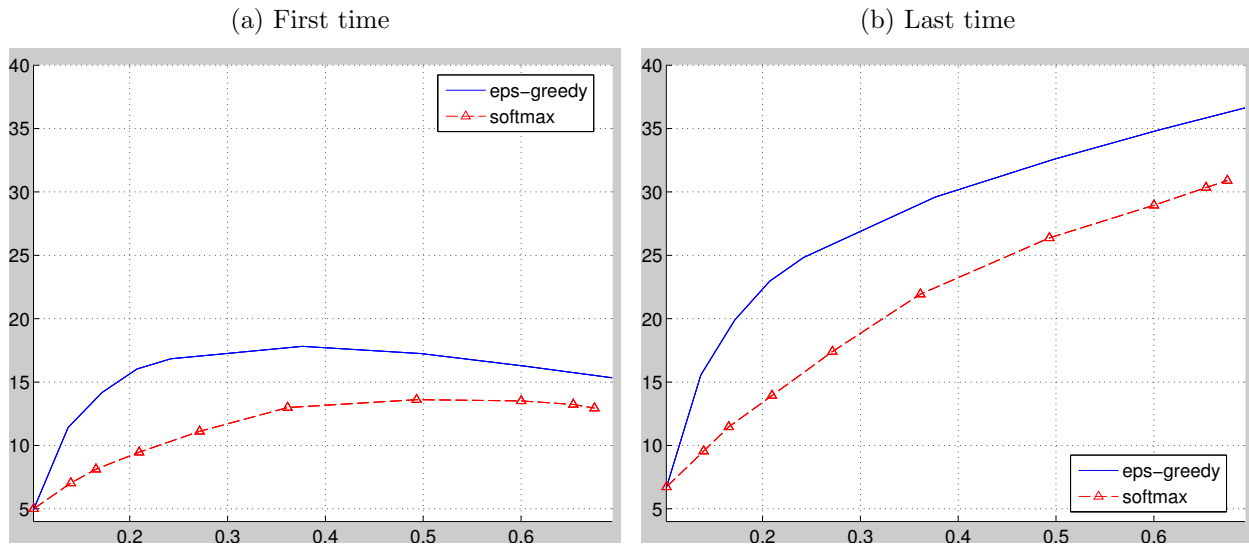
would decrease the risk associated with adaptation to future problems, we argue in this study that this intuition is not completely correct. While, consistent with the preceding intuition, this exploration augments the knowledge that guides firms towards selecting good alternatives and avoiding poor ones, we find that this exploration also biases firms towards the selection of risky solutions to problems.

In so doing, we contribute to understanding the different forms of risk associated with problem-driven organizational change. This is particularly salient when the two forms of risk examined in this study – that due to an insufficient breadth of knowledge and that due to endogenous risk seeking – have dramatically different implications for the firm. Consider our earlier example of a pharmaceutical firm whose leading drug is found to have an unacceptable side effect. In solving this problem, the firm might be willing to accept the fact that low anticipatory exploration leaves the firm vulnerable to volatile experimentation across replacement drugs that vary in performance. However, the pharmaceutical firm might be far more concerned about higher volatility *within* the replacement drug that, according to our findings, is potentially caused by higher levels of anticipatory exploration. That is, the firm might

not be willing to accept the maladaptive consequences of endogenous risk seeking, where it has a high chance of selecting a single drug whose performance, irrespective of its mean performance, has high variance across different patients (e.g., a single wrongful death from a drug trial might have drastic repercussions for the firm). A promising line of future work would be to further analyze the asymmetric implications of different types of adaptation risk.

APPENDIX

APPENDIX A

Toy Model of Collaboration

To illustrate the learning and information revealing effects of knowledge contributing efforts to collaboration, I present a toy model of collaboration in which ex-post innovations from collaborating firms engage in Cournot competition in the market for innovations. The primary intent is to show that, under this model, profit-maximizing firms with minimal limits on information are sufficient to produce the hypothesized results.

For ease of exposition, I assume two firms A and B where, without loss of generality, A exerts higher effort in contributing knowledge, as in Figure 3.2, for the calculated example (though I assume many firms for the formulation). I assume effort is a monotonic function of each firm's expected gains from the collaboration (which are exogenously given). Each firm's information limitation is that it has no (ex-ante) knowledge on other firm's expected gains nor do firms know the underlying distribution of expected gains.

Focusing on complementary gains to knowledge production (and thus ignoring additive gains), I assume collaborative knowledge K_{COL} is the sum of pairwise products of effort:

$$K_{COL} = \sum_{i \neq j} e_i e_j \tag{A.1}$$

Conceptually, this can be considered a knowledge pool that firms may draw upon given some ability to do so. For two firms this is simply given by $K_{COL} = e_A e_B$. I assume a linear inverse demand function:

$$D^{-1} = a - bQ \tag{A.2}$$

where Q is the total quantity of innovations produced by firms in the collaboration. To simplify, I assume quantities are scaled so that $b = 1$. The performance of firm i's innovations, π_i, is given by:

$$\pi_i = q_i(D^{-1} + v_i) \tag{A.3}$$

where q_i is the quantity and v_i is the (cost-adjusted) quality of firm i's innovations that build on collaborative knowledge K_{COL}. Here, we assume for simplicity that innovation is costless, but that collaborative knowledge can be exploited by a firm to achieve a "negative" cost in the form of enhanced quality.

We further decompose quality $v_i = v_i^K + v_i^O$, where v_i^K represents the quality gained through firm i's effort contributing to the collaboration and v_i^O is the quality gained through observing other firm's efforts contributing to the collaboration. v_i^K reflects both a firm's ex-post incentives to innovate based on perceived quality of the collaborative output and a firm's capability to utilize collaborative output (which should increase with effort). v_i^O can thus be considered adjustments along these two dimensions through observing other firms: observation may reveal the underlying quality of the collaborative output and the capabilities beneficial to utilize the collaborative knowledge.

In this study, I theorize effort contributing knowledge gives firms a context to interact with other firms and in effect realize potential collaborative knowledge benefits.

I model this by making collaborative gains increase (at a diminishing rate) with effort bounded above by K_{COL}, i.e., $v_i^K = e_i K_{COL}/(e_i + 1)$. For two firms, this reduces to $v_A^K = e_A^2 e_B/(e_A + 1)$ and, similarly, $v_B^K = e_B^2 e_A/(e_B + 1)$.

Moreover, I theorize effort by firms may reveal to other firms the underlying value of collaborative knowledge (or capabilities that enable a firm to benefit from this underlying value). Following prior work (Türüt and Ofek, 2012), we assume asymmetry in revealed information: high effort reveals more accurate information than low effort. This is sensible in that the process of contributing more knowledge to a contribution should give a firm a more accurate assessment of the potential embodied in the collaboration, and under the assumption knowledge flows exclusively from more knowledgeable to less knowledge firms (Jovanovic and Rob, 1989; Jovanovic and MacDonald, 1994; Knott et al., 2009), we might expect higher effort firms to provide signals to low effort firms, but not vice-versa. To capture this, I define v_i^O as follows:

$$v_i^O = \sum_{j \neq i} (e_j - e_i)^2 H(e_j - e_i),$$

where

$$H(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

and the squared term captures the sensible notion that the visibility of the signal exhibits increasing returns. For the two firm case in the example, this reduces to $v_A^O = 0$ and $v_B^O = (e_A - e_B)^2$.

Substituting Equation A.2 into Equation A.3 and solving first-order conditions, the Nash-Cournot equilibrium is given by:

114

$$q_A = \frac{a}{3} - \frac{(e_A - e_B)^2}{3} + \frac{2\,e_A{}^2\,e_B}{3\,(e_A + 1)} - \frac{e_A\,e_B{}^2}{3\,(e_B + 1)} \tag{A.4}$$

$$q_B = \frac{a}{3} + \frac{2\,(e_A - e_B)^2}{3} - \frac{e_A{}^2\,e_B}{3\,(e_A + 1)} + \frac{2\,e_A\,e_B{}^2}{3\,(e_B + 1)} \tag{A.5}$$

Equations A.4 and A.5 may be inserted into Equation A.3 to determine performance of A's innovations as a function of e_A and e_B. To illustrate, Figure A.1 graphs the performance of A as a function of e_A for $a = 2$ and $e_B = 0.5$. Note that performance increases initially as firm A learns from contributing knowledge, but at higher levels, the information revealing effect dominates, and performance declines, consistent with the theory in this study. Moreover, A's performance increases with effort to a point somewhat past B's effort level ($e_A \simeq 0.8$), which conforms to the intuition that others' effort serves as a signal only when it is "noticeably" larger than own effort.

Figure A.1: Results of toy model: Innovative performance as a function of knowledge-contributing effort

115
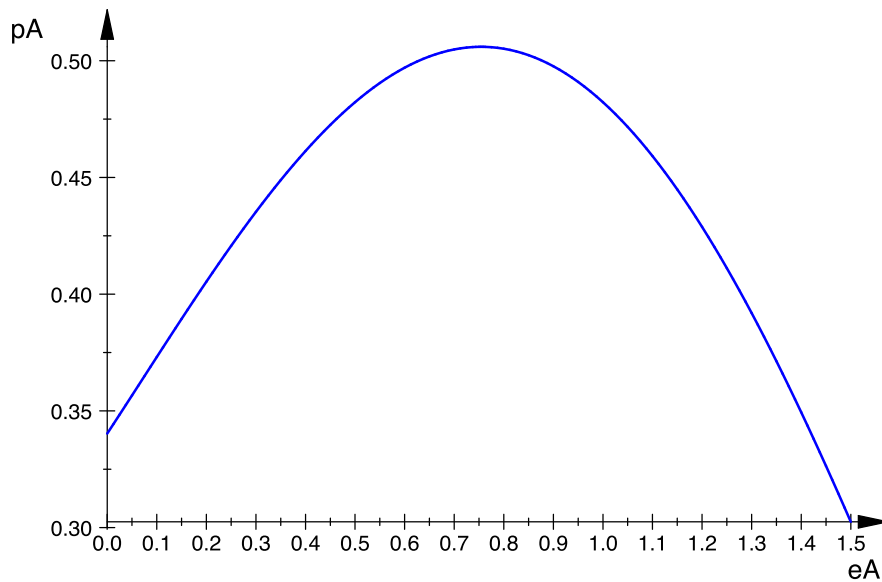
BIBLIOGRAPHY

BIBLIOGRAPHY

Agarwal, R., R. Echambadi, A. Franco. 2004. Knowledge transfer through inheritance: Spin-out generation, development, and survival. *Academy of Management Journal* **47**(4) 501–522.

Alchian, A. 1963. Reliability of progress curves in airframe production. *Econometrica* **31**(4) 679–693.

Amburgey, T., D. Kelly, W. Barnett. 1993. Resetting the clock: The dynamics of organizational change and failure. *Administrative Science Quarterly* **38**(1) 51–73.

Andreoni, J. 1988. Why free ride?: Strategies and learning in public goods experiments. *Journal of Public Economics* **37**(3) 291–304.

Anton, J., D. Yao. 1995. Standard-setting consortia, antitrust, and high-technology industries. *Antitrust Law Journal* **64**(1) 247.

Appleyard, M. M. 1996. How does knowledge flow? Interfirm patterns in the semiconductor industry. *Strategic Management Journal* **17**(Winter) 137–154.

Argote, L., D. Epple. 1990. Learning Curves in Manufacturing. *Science* **247**(4945) 920–924.

Argote, L. 1999. *Organizational Learning: Creating, Retaining, and Transferring Knowledge*. Kluwer Academic Publishers, MA.

Argote, L., S. Beckman, D. Epple. 1990. The persistence and transfer of learning in industrial settings. *Management Science* **36**(2) 140–154.

Arora, A. 1996. Contracting for tacit knowledge: The provision of technical services in technology licensing contracts. *Journal of Development Economics* **50**(2) 233–256.

Arrow, K. 1962. The economic implications of learning by doing. *The Review of Economic Studies* **29**(3) 155–173.

Baldwin, C. Y., K. B. Clark. 2006. The architecture of participation: Does code architecture mitigate free riding in the open source development model? *Management Science* **52**(7) 1116–1127.

Baum, J. A. C., K. B. Dahlin. 2007. Aspiration performance and railroads' patterns of learning from train wrecks and crashes. *Organization Science* **18**(3) 368–385.

Baum, J. A. C., P. Ingram. 1998. Survival-enhancing learning in the Manhattan hotel industry, 1898-1980. *Management Science* **44**(7) 996–1016.

Baum, J., S. Li, J. Usher. 2000. Making the next move: How experiential and vicarious learning shape the locations of chains' acquisitions. *Administrative Science Quarterly* **45**(4) 766–801.

Bergemann, D., J. Valimaki. 1996. Learning and strategic pricing. *Econometrica* **64**(5) 1125–1149.

Berger, A. 2003. The Economic Effects of Technological Progress: Evidence from the Banking Industry. *Journal of Money Credit and Banking* **35**(2) 141.

Bergstrom, T., L. Blume, H. Varian. 1986. On the private provision of public goods. *Journal of Public Economics* **29**(1) 25–49.

Berry, D., B. Fristedt. 1985. *Bandit Problems: Sequential Allocation of Experiments*. Chapman & Hall, London.

Bingham, C. B., J. P. Davis. 2012. Learning sequences: Their existence, effect, and evolution. *Academy of Management Journal* **55**(3) 611–641.

Bliss, C., B. Nalebuff. 1984. Dragon-slaying and ballroom dancing: The private supply of a public good. *Journal of Public Economics* **25**(1) 1–12.

Carlile, P. R. 2004. Transferring, translating, and transforming: An integrative framework for managing knowledge across boundaries. *Organization Science* **15**(5) 555–568.

Carlile, P. R., E. S. Rebentisch. 2003. Into the black box: The knowledge transformation cycle. *Management Science* **49**(9) 1180–1195.

Cassiman, B., R. Veugelers. 2002. R&D cooperation and spillovers: Some empirical evidence from Belgium. *American Economic Review* **12**(4) 1169–1184.

Chamberlin, J. 1974. Provision of collective goods as a function of group size. *The American Political Science Review* **68**(2) 707–716.

Cockburn, I., R. Henderson. 1998. Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. *Journal of Industrial Economics* **46** 157–182.

Cohen, W., D. Levinthal. 1989. Innovation and learning: The two faces of R&D. *Economic Journal* **99**(397) 569–596.

Cohen, W., D. Levinthal. 1990. Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly* **35**(1) 128–152.

Corredoira, R., L. Rosenkopf. 2010. Should auld acquaintance be forgot? The reverse transfer of knowledge through mobility ties. *Strategic Management Journal* **31**(2) 159–181.

Cyert, R., J. March. 1963. *A Behavioral Theory of the Firm*. Prentice-Hall, New Jersey.

Dahlander, L., M. Magnusson. 2008. How do firms make use of open source communities? *Long Range Planning* **41**(6) 629–649.

Darr, E., L. Argote, D. Epple. 1995. The acquisition, transfer, and depreciation of knowledge. *Management Science* **41**(11) 1750–1762.

Das, T., B. Teng. 2000. A resource-based theory of strategic alliances. *Journal of Management* **26**(1) 31–61.

Das, T. K., B.-S. Teng. 1998. Between trust and control: Developing confidence in partner cooperation in alliances. *Academy of Management Review* **23**(3) 491–512.

David, P. A., M. Shurmer. 1996. Formal standards-setting for global telecommunications and information services. Towards an institutional regime transformation? *Telecommunications Policy* **20**(10) 789–815.

Dencker, J., M. Gruber, S. Shah. 2009. Pre-entry knowledge, learning, and the survival of new firms. *Organization Science* **20**(3) 516–537.

Denrell, J., J. March. 2001. Adaptation as information restriction: The hot stove effect. *Organization Science* **12**(5) 523–538.

Denrell, J., C. Fang, D. Levinthal. 2004. From t-mazes to labyrinths: Learning from model-based feedback. *Management Science* **50**(10) 1366–1378.

Dierickx, I., K. Cool. 1989. Asset stock accumulation and sustainability of competitive advantage. *Management Science* **35**(2) 1504–1511.

Dyer, J. H., H. Singh. 1998. The relational view: Cooperative strategy and sources of interorganizational competitive advantage. *Academy of Management Review* **23**(4) 660–679.

Eeckhout, J., B. Jovanovic. 2002. Knowledge spillovers and inequality. *American Economic Review* **92**(5) 1290–1307.

Epple, D., L. Argote, K. Murphy. 1996. An empirical investigation of the microstructure of knowledge acquisition and transfer through learning by doing. *Operations Research* **44**(1) 77–86.

Ethiraj, S., D. Zhu. 2008. Performance effects of imitative entry. *Strategic Management Journal* **29** 797–817.

Ethiraj, S., P. Kale, M. Krishnan, J. Singh. 2005. Where do capabilities come from and how do they matter? A study in the software services industry. *Strategic Management Journal* **26**(1) 25–45.

Farrell, J., G. Saloner. 1988. Coordination through committees and markets. *RAND Journal of Economics* **19**(2) 235–252.

Farrell, J., T. Simcoe. 2012. Choosing the rules for consensus standardization. *The RAND Journal of Economics* **43**(2) 235–252.

Feldman, M. 2000. Organizational routines as a source of continuous change. *Organization Science* **11**(6) 611–629.

Festinger, L. 1954. A theory of social comparison processes. *Human Relations* **7**(2) 117–140.

Fombrun, C., M. Shanley. 1990. What's in a name? Reputation building and corporate strategy. *Academy Of Management Journal* **33**(2) 233–258.

Friedlander, A. 1995. *Emerging Infrastructure: The Growth of Railroads*. CNRI.

Furman, J. L., M. K. Kyle, I. M. Cockburn, R. Henderson. 2006. Public & private spillovers, location and the productivity of pharmaceutical research. *National Bureau of Economic Research* .

Galunic, D., S. Rodan. 1998. Resource recombinations in the firm: Knowledge structures and the potential for Schumpeterian innovation. *Strategic Management Journal* **19** 1193–1201.

Gans, N., G. Knox, R. Croson. 2007. Simple models of discrete choice and their performance in bandit experiments. *Manufacturing & Service Operations Management* **9**(4) 383–408.

Geroski, P. 1995. What do we know about entry? *International Journal of Industrial Organization* **13**(4) 421.

Gittelman, M., B. Kogut. 2003. Does good science lead to valuable knowledge? biotechnology firms and the evolutionary logic of citation patterns. *Management Science* **49**(4) 366–382.

Gittins, J. 1979. Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society. Series B (Methodological)* **41**(2) 148–177.

Gittins, J. C., D. M. Jones. 1979. A dynamic allocation index for the discounted multiarmed bandit problem. *Biometrika* **66**(3) 561–565.

Goeree, M. S., E. Helland. 2012. Do research joint ventures serve a collusive function? IEW - Working Papers 448, Institute for Empirical Research in Economics - University of Zurich.

Grant, R. 1996. Toward a knowledge based theory of the firm. *Strategic Management Journal* **17**(Winter Special Issue) 109–122.

Grant, R., C. Baden-Fuller. 2004. A knowledge accessing theory of strategic alliances. *Journal of Management Studies* **41**(1) 61–84.

Greenstein, S., V. Stango. 2007. *Standards and Public Policy*. Cambridge University Press, Cambridge.

Greenstein, S. M. 1992. Invisible hands and visible advisors: An economic interpretation of standardization. *Journal of the American Society for Information Science* **43**(8) 538–549.

Greve, H. 1998. Performance, aspirations, and risky organizational change. *Administrative Science Quarterly* **43**(1) 58–86.

Greve, H. 2005. Interorganizational learning and heterogeneous social structure. *Organization Studies* **26**(7) 1025–1047.

Griliches, Z. 1957. Hybrid corn – an exploration in the economics of technological-change. *Econometrica* **25**(4) 501–522.

Griliches, Z. 1998. *R&D and Productivity : The Econometric Evidence*. University of Chicago Press, Chicago.

Gulati, R. 1995. Does familiarity breed trust? The implications of repeated ties for contractual choices in alliances. *Academy Of Management Journal* **38**(1) 85–112.

Hall, B. H., R. H. Ziedonis. 2001. The patent paradox revisited: An empirical study of patenting in the US semiconductor industry, 1979-1995n empirical study of patenting in the US semiconductor industry, 1979-1995. *RAND Journal of Economics* **32**(1) 101–128.

Hall, B., A. Jaffe, M. Trajtenberg. 2001. The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools. *NBER Working Paper Series* (8498).

Hall, B. H., A. Jaffe, M. Trajtenberg. 2005. Market value and patent citations. *RAND Journal of Economics* **36**(1) 16–38.

Hamel, G. 1991. Competition for competence and interpartner learning within international strategic alliances. *Strategic Management Journal* **12** 83–103.

Hardin, R. 1982. *Collective Action*. The Johns Hopkins University Press, Baltimore, MD.

Haunschild, P., C. Beckman. 1998. When do interlocks matter? Alternate sources of information and interlock influence. *Administrative Science Quarterly* **43**(4) 815–844.

Haunschild, P., A. Miner. 1997. Modes of interorganizational imitation: The effects of outcome salience and uncertainty. *Administrative Science Quarterly* **42**(3) 472–500.

Helfat, C., M. Lieberman. 2002. The birth of capabilities: Market entry and the importance of pre-history. *Industrial and Corporate Change* **11**(4) 725–760.

Henderson, R. 1993. Underinvestment and incompetence as responses to radical innovation: Evidence from the photolithographic alignment equipment industry. *RAND Journal of Economics* **24**(2) 248–270.

Henderson, R., K. Clark. 1990. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly* **35** 9–30.

Henkel, J. 2004. The jukebox mode of innovation-a model of commercial open source development. Tech. rep., CEPR Discussion Papers.

Herriott, S., D. Levinthal, J. March. 1985. Learning from experience in organizations. *American Economic Review* **75**(2) 298–302.

Hertel, G., S. Niedner, S. Herrmann. 2003. Motivation of software developers in open source projects: An internet-based survey of contributors to the linux kernel. *Research Policy* **32**(7) 1159–1177.

Hoetker, G. 2007. The use of logit and probit models in strategic management research: Critical issues. *Strategic Management Journal* **28**(4) 331–343.

Holland, J. 1975. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control & Artificial Intelligence*. U Michigan Press, Ann Arbor.

Hoopes, D., S. Postrel. 1999. Shared knowledge, "glitches," and product development performance. *Strategic Management Journal* **20**(9) 837–865.

Huber, G. P. 1991. Organizational learning: The contributing processes and the literatures. *Organization Science* **2**(1) 88–115.

Iansiti, M., K. Clark. 1994. Integration and dynamic capability: Evidence from product development in automobiles and mainframe computers. *Industrial and Corporate Change* **3**(3) 557–605.

Ingram, P., P. Roberts. 2000. Friendships among competitors in the Sydney hotel industry. *American Journal of Sociology* **106**(2) 387–423.

Ingram, P., T. Simons. 2002. The transfer of experience in groups of organizations: Implications for performance and competition. *Management Science* **48**(12) 1517–1533.

Ireland, R. D., M. A. Hitt, D. Vaidyanath. 2002. Alliance management as a source of competitive advantage. *Journal of Management* **28**(3) 413–446.

Isaac, R. M., J. M. Walker. 1988. Group size effects in public goods provision: The voluntary contributions mechanism. *The Quarterly Journal of Economics* **103**(1) 179–199.

Jaffe, A. 1986. Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value. *American Economic Review* **76**(5) 984–1001.

Jorde, T. M., D. J. Teece. 1990. Innovation and cooperation: Implications for competition and antitrust. *The Journal of Economic Perspectives* **4**(3) 75–96.

Jovanovic, B., R. Rob. 1989. The growth and diffusion of knowledge. *Review of Economic Studies* **56**(4) 569–582.

Jovanovic, B., G. MacDonald. 1994. Competitive diffusion. *Journal of Political Economy* **102**(1) 24–52.

Kale, P., H. Singh, H. Perlmutter. 2000. Learning and protection of proprietary assets in strategic alliances: Building relational capital. *Strategic Management Journal* **21**(3) 217–237.

Kamien, M., I. Zang. 2000. Meet me halfway: Research joint ventures and absorptive capacity. *International Journal of Industrial Organization* **18**(7) 995–1012.

Katila, R., G. Ahuja. 2002. Something old, something new: A longitudinal study of search behavior and new product introduction. *Academy Of Management Journal* **45**(6) 1183–1194.

Katz, M. 1986. An analysis of cooperative research and development. *RAND Journal of Economics* **17**(4) 527–543.

Keeble, D., F. Wilkinson. 1999. Collective learning and knowledge development in the evolution of regional clusters of high technology SMEs in Europe. *Regional Studies* **33**(4) 295–303.

Klepper, S., K. Simons. 2000. Dominance by birthright: Entry of prior radio producers and competitive ramifications in the US television receiver industry. *Strategic Management Journal* **21**(10-11) 997–1016.

Knott, A. M., H. E. Posen, B. Wu. 2009. Spillover asymmetry and why it matters. *Management Science* **55**(3) 373–388.

Knott, A., H. Posen. 2005. Is failure good? *Strategic Management Journal* **26**(7) 617–641.

Kogut, B., U. Zander. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science* **3**(3) 383–397.

Kostova, T., S. Zaheer. 1999. Organizational legitimacy under conditions of complexity: The case of the multinational enterprise. *Academy of Management Review* **24**(1) 64–81.

Lai, R., A. D'Amour, A. Yu, Y. Sun, V. Torvik, L. Fleming. 2011. Disambiguation and co-authorship networks of the us patent inventor database. *Harvard Institute for Quantitative Social Science, Cambridge, MA* **2138**.

Lakhani, K. R., R. G. Wolf. 2005. Why hackers do what they do: Understanding motivation and effort in free/open source software projects. Hissam S Lakhani K Feller J, Fitzgerald B, ed., *Perspectives On Free And Open Source Software*. MIT press Cambridge, MA, 3–22.

Lane, P. J., M. Lubatkin. 1998. Relative absorptive capacity and interorganizational learning. *Strategic Management Journal* **19**(5) 461–477.

Lavie, D. 2006. The competitive advantage of interconnected firms: An extension of the resource-based view. *Academy of Management Review* **31**(3) 638–658.

Lavie, D., L. Rosenkopf. 2006. Balancing exploration and exploitation in alliance formation. *Academy Of Management Journal* **49**(4) 797–818.

Leiponen, A. E. 2008. Competing through cooperation: The organization of standard setting in wireless telecommunications. *Management Science* **54**(11) 1904–1919.

Lerner, J., P. A. Pathak, J. Tirole. 2006. The dynamics of open-source contributors. *American Economic Review* **96**(2) 114–118.

Lerner, J., J. Tirole. 2002. Some simple economics of open source. *The Journal of Industrial Economics* **50**(2) 197–234.

Levin, R., P. Reiss. 1988. Cost-reducing and demand-creating r&d with spillovers. *The RAND Journal of Economics* **19**(4) 538–556.

Levinthal, D. 1997. Adaptation on rugged landscapes. *Management Science* **43**(7) 934–950.

Levinthal, D., J. March. 1993. The myopia of learning. *Strategic Management Journal* **14** 95–112.

Levitt, B., J. G. March. 1988. Organizational learning. *Annual Review of Sociology* **14** 319–340.

Lieberman, M. 1984. The learning curve and pricing in the chemical processing industries. *RAND Journal of Economics* **15**(2) 213–228.

Mansfield, E. 1961. Technical change and the rate of imitation. *Econometrica* **29**(4) 741–766.

March, J., J. Olsen. 1975. The uncertainty of the past: organizational learning under ambiguity. *European Journal of Political Research* **3** 147–171.

March, J. 1991. Exploration and exploitation in organizational learning. *Organization Science* **2**(1) 71–87.

March, J., Z. Shapira. 1987. Managerial perspectives on risk and risk taking. *Management Science* **33**(11) 1404–1418.

March, J., H. Simon. 1958. *Organizations*. Wiley, New York.

March, J. G. 2003. Understanding organisational adaptation. *Society and Economy* **25**(1) 1–10.

March, J. 1996. Learning to be risk averse. *Psychological Review* **103**(2) 309–319.

March, J. 2010. *The Ambiguities of Experience*. Cornell University Press, Ithaca, NY.

Mester, L. 1993. Efficiency in the savings and loan industry. *Journal of Banking & Finance* **17**(2-3) 267–286.

Meyer, R. J., Y. Shi. 1995. Sequential choice under ambiguity: Intuitive solutions to the armed-bandit problem. *Management Science* **41**(5) 817–834.

Miller, K., M. Leiblein. 1996. Corporate risk-return relations: Returns variability versus downside risk. *Academy Of Management Journal* **39**(1) 91–122.

Miner, A., P. Haunschild. 1995. Population-Level Learning. L L Cummings, B M Staw, eds., *Research in Organizational Behavior*. JAI Press, Greenwich, CT, 115–166.

Moreland, R. L. 1999. Transactive memory: Learning who knows what in work groups and organizations. L.L. Thompson, J. M. Levine, D.M. Messick, eds., *Shared Cognition in Organizations*. Erlbaum, 3–31.

Mowery, D., J. Oxley, B. Silverman. 1996. Strategic alliances and interfirm knowledge transfer. *Strategic Management Journal* **17** 77–91.

Nelson, R., S. Winter. 1982. *An Evolutionary Theory Of Economic Change*. Belknap Press of Harvard University Press, Cambridge, MA.

Nickerson, J., T. Zenger. 2004. A knowledge-based theory of the firm–the problem-solving perspective. *Organization Science* **15**(6) 617–632.

Nickerson, J., C. J. Yen, J. T. Mahoney. 2012. Exploring the problem-finding and problem-solving approach for designing organizations. *Academy of Management Perspectives* **26**(1) 52–72.

Nonaka, I. 1994. A dynamic theory of organizational knowledge creation. *Organization Science* **5**(1) 14–37.

Nooteboom, B., W. Van Haverbeke, G. Duysters, V. Gilsing, A. van den Oord. 2007. Optimal cognitive distance and absorptive capacity. *Research Policy* **36**(7) 1016–1034.

Ocasio, W. 1997. Towards an attention-based view of the firm. *Strategic Management Journal* **18** 187–206.

Olson, M. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard University Press, Cambridge, MA.

Osterloh, M., S. Rota. 2007. Open source software development—Just another case of collective invention? *Research Policy* **36**(2) 157–171.

Oxley, J., R. Sampson. 2004. The scope and governance of international R&D alliances. *Strategic Management Journal* **25** 723–749.

Page, S. E. 2008. *The Difference*. How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies, Princeton University Press, Princeton.

Penrose, E. 1959. *The Theory Of The Growth Of The Firm*. Wiley, New York.

Pisano, G. 2006. Profiting from innovation and the intellectual property revolution. *Research Policy* **35**(8) 1122–1130.

Polanyi, M. 1966. *The Tacit Dimension*. Anchor Day Books, New York.

Posen, H. E., D. A. Levinthal. 2012. Chasing a moving target: Exploitation and exploration in dynamic environments. *Management Science* **58**(3) 587–601.

Posen, H. E., J. S. Chen. 2013. An advantage of newness: Vicarious learning despite limited absorptive capacity. *Organization Science* **24**(6) 1701–1716.

Posen, H. E., J. Lee, S. Yi. 2013. The power of imperfect imitation. *Strategic Management Journal* **34**(2) 149–164.

Powell, W., K. Koput, L. Smith-Doerr. 1996. Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly* **41**(1) 116–145.

Reinganum, J. 1983. Uncertain innovation and the persistence of monopoly. *American Economic Review* **73**(4) 741–748.

Ren, Y., K. M. Carley, L. Argote. 2006. The contingent effects of transactive memory: When is it more beneficial to know what others know? *Management Science* **52**(5) 671–682.

Robbins, H. 1952. Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society* **58**(5) 527–535.

Rosenkopf, L., P. Almeida. 2003. Overcoming Local Search Through Alliances and Mobility. *Management Science* **49**(6) 751–766.

Sakakibara, M. 1997. Evaluating government-sponsored R&D consortia in Japan: Who benefits and how? *Research Policy* **26** 447–473.

Sakakibara, M. 2002. Formation of R&D consortia: Industry and company effects. *Strategic Management Journal* **23**(11) 1033.

Sakakibara, M. 2003. Knowledge sharing in cooperative research and development. *Managerial and Decision Economics* **24** 117–132.

Sampson, R. 2007. R&D alliances and firm performance: The impact of technological diversity and alliance organization on innovation. *Academy Of Management Journal* **50**(2) 364–386.

Saunders, A., M. Millon Cornett. 2003. *Financial Institutions Management: A Risk Management Approach*. Irwin/McGraw-Hill, New York.

Saxenian, A. 1996. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Harvard University Press, Cambridge, MA.

Schilling, M. A., P. Vidal, R. E. Ployhart, A. Marangoni. 2003. Learning by doing something else: Variation, relatedness, and the learning curve. *Management Science* **49**(1) 39–56.

Schumpeter, J. 1934. *The Theory of Economic Development*. Harvard University Press, Cambridge, MA.

Schwab, A. 2007. Incremental organizational learning from multilevel information sources: Evidence for cross-level interactions. *Organization Science* **18**(2) 233–251.

Senker, J. 1995. Tacit knowledge and models of innovation. *Industrial and Corporate Change* **4**(2) 425–447.

Shapiro, C., H. Varian. 1999. *Information Rules: A Strategic Guide to the Network Economy*. Harvard Business Press, Cambridge, MA.

Simcoe, T. 2012. Standard setting committees: Consensus governance for shared technology platforms. *American Economic Review* **102**(1) 305–336.

Simon, D. H., M. B. Lieberman. 2010. Internal and external influences on adoption decisions in multi-unit firms: The moderating effect of experience. *Strategic Organization* **8**(2) 132–154.

Spence, M. 1984. Cost reduction, competition, and industry performance. *Econometrica: Journal of the Econometric Society* **52**(1) 101–122.

Spencer, J. W. 2003. Firms' knowledge-sharing strategies in the global innovation system: Empirical evidence from the flat panel display industry. *Strategic Management Journal* **24**(3) 217–233.

Stinchcombe, A. 1965. Organizations and social structure. JG March, ed., *Handbook of Organizations*. Rand-McNally, Chicago, 153–193.

Sutton, R. S., A. G. Barto. 1998. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA.

Teece, D. 1986. Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy* **15**(6) 285–305.

Tiwana, A. 2008. Do bridging ties complement strong ties? An empirical examination of alliance ambidexterity. *Strategic Management Journal* **29**(3) 251–272.

Todorova, G., B. Durisin. 2007. Absorptive capacity: Valuing a reconceptualization. *Academy of Management Review* **32**(3) 774–786.

Trajtenberg, M. 1990. A penny for your quotes: Patent citations and the value of innovations. *RAND Journal of Economics* **21**(1) 172–187.

Türüt, Ö., E. Ofek. 2012. Innovation strategy and entry deterrence. *Journal of Economics and Management Strategy* **21**(3) 563–631.

Tushman, M., P. Anderson. 1986. Technological discontinuities and organizational environments. *Administrative Science Quarterly* **31**(3) 439–465.

Von Hippel, E., G. Von Krogh. 2003. The private-collective innovation model in open source software development: Issues for organization science. *Organization Science* **14**(2) 209–223.

Von Hippel, E., M. J. Tyre. 1995. How learning by doing is done: Problem identification in novel process equipment. *Research Policy* **24**(1) 1–12.

Vuong, Q. 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* **57**(2) 307–333.

Weber, S. 2004. *The Success of Open Source*. Harvard University Press, Cambridge, MA.

Weiss, M. B. H. 1993. The standards development process. *StandardView* **1**(2) 35–41.

West, J. 2001. Qualcomm's standards strategy. *Proceedings of the 2nd IEEE Conference on Standardization and Innovation in Information Technology*. Boulder, CO, 62–76.

West, J. 2003. How open is open enough? *Research Policy* **32**(7) 1259–1285.

White, J. 2012. *Bandit Algorithms for Website Optimization*. O'Reilly, Sebastopol, CA.

Wiersma, E. 2007. Conditions that shape the learning curve: Factors that increase the ability and opportunity to learn. *Management Science* **53**(12) 1903–1915.

Winter, S. 1987. Knowledge and competence as strategic assets. David Teece, ed., *The Competitive Challenge –Strategies For Industrial Innovation And Renewal*. Ballinger, Cambridge, MA, 159–184.

Zahra, S., G. George. 2002. Absorptive capacity: A review, reconceptualization, and extension. *Academy of Management Review* **27**(2) 185–203.

Zahra, S., R. Ireland, M. Hitt. 2000. International expansion by new venture firms: International diversity, mode of market entry, technological learning, and performance. *Academy Of Management Journal* **43**(5) 925–950.

Zemsky, P., G. Pacheco-de Almeida. 2012. Some like it free: Innovators' strategic use of disclosure to slow down competition. *Strategic Management Journal* **33** 773–793.

Zeng, M., X. P. Chen. 2003. Achieving cooperation in multiparty alliances: A social dilemma approach to partnership management. *Academy of Management Review* **28**(4) 587–605.