Parametric interdependence, learning-by-doing, and industrial structure

William Martin Tracy · M.V. Shyam Kumar ·
William Paczkowski

Published online: 24 November 2012

© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract We explore the proposition that parametric interdependence makes learning-by-doing a nondeterministic, path-dependent process. The implications of our model challenge two conventional beliefs about the relationships between industrial structure, spillovers, and learning-by-doing. First, we challenge the belief that the monopolistic industrial structure always maximizes learning-by-doing gains when there are no spillovers. Second, we challenge the belief that increasing spillovers unambiguously increases welfare when learning-by-doing drives innovation.

Keywords Learning-by-doing · Industrial structure · Spillovers · NK landscape · Parametric interdependencies

1 Introduction

Parametric interdependency exists when the optimal setting of one parameter is impacted by the current setting of other parameters. For example, what is the optimal arrangement of magnets in an electric motor? The answer depends in part on whether the motor will run on AC or DC current. A magnet setting that is optimal on an AC motor might not work at all on a DC motor. Hence, there is an interdependency between the magnet placement and motor's current; the optimal magnet placement is dependent on the current.

W.M. Tracy (✉) · M.V.S. Kumar · W. Paczkowski
Lally School of Management and Technology, Rensselaer Polytechnic Institute, 110 8th Street, Troy,
NY 12180, USA
e-mail: tracyw@rpi.edu

M.V.S. Kumar
e-mail: kumarm2@rpi.edu

W. Paczkowski
e-mail: paczkw@rpi.edu

There is evidence that such interdependencies between technology parameters are widespread. Fleming and Sorenson (2001) proposed a technique for estimating the level of interdependence in a technology covered by a patent. In their analysis, they found evidence of parametric interdependence in all of the 17,264 patents they analyzed. Given the prevalence of parametric interdependence, it is important to consider the impact of parametric interdependencies on models of innovation and learning.

The extent of interdependence among components or parameters of a technology is understood to impact that technology's development (Kauffman et al. 2000). There has been prior work considering the impact of parametric interdependencies on learning-by-doing (cf. Auerswald 2010; Auerswald et al. 2000) and incremental innovation (Zhang and Gao 2010). However, prior work in this area has not addressed the socially optimal industrial structure, nor has it addressed the optimal level of spillovers. We argue that parametric interdependencies can moderate the received wisdom in both of these areas. Our arguments are predicated on the assertion that parametric interdependencies make technological innovation a nondeterministic, path-dependent process.

To illustrate the intuition behind this assertion, reconsider the electric motor discussed above. Assume an engineer tasked with maximizing a motor's initial torque begins with the erroneous belief that DC motors produce the highest initial torque. All of her design decisions would maximize the initial torque of a DC motor. In reality, the highest initial torque motors use AC current. However, if after optimizing the motor using DC current the engineer were to simply switch to AC current, the initial torque in her DC-optimized motor would likely decrease; indeed the motor might not work at all. She would need to simultaneously change most of the motor's design parameters to see a benefit from switching to AC power. Therefore, it is unlikely that trial-and-error would lead the engineer to discover the optimality of AC power once she starts making incremental improvements along the DC power path.

The trial and error development of technologies with parametric interdependencies are thus a path-dependent process, because decisions made in early trials can limit possible technology outcomes. This path dependency also implies that the process is nondeterministic. The particular path followed can be as important as the number of trials executed. For example, if the above engineer began down the path of incrementally improving a DC power motor, then 10,000 trials might produce a less efficient motor than would 1,000 trials using the AC power motor.

Parametric interdependencies are not the only phenomenon that can render technology development a path dependent process. For example, when decision makers become too deeply focused on particular mechanisms or type of technology, they may lack the mental categories needed to incorporate alternatives. This process is illustrated in Fioretti (2006), which uses a Kohonen neural network to demonstrate how rigid mental categories can prevent firms from recognizing investments in superior technologies. In such cases, the path dependency is not driven by the nature of the technology, but rather by the rigidity of the decision makers' thinking. Insofar as the impacts of these processes are complementary, we will focus our discussion on the role of technological parametric interdependencies.

In this paper we present a model that illustrates two implications parametric interdependencies. First, we demonstrate that learning-by-doing in the monopolistic industrial structure explores fewer technology development paths than it does an

oligopolistic industrial structure. Therefore, the expected efficacy of the best specification of a technology uncovered by an oligopolistic structure is greater than that of a monopolistic structure. Second, increasing the rate and breadth of spillovers does not necessarily improve an industries development of a technology.

2 Theory

2.1 Industrial structure

The impact of industrial structure on innovation has been a topic of central interest since Schumpeter (1942) and Arrow (1959). Dasgupta and Stiglitz (1988) present a model that suggests a monopolistic industrial structure maximizes the gains from learning-by-doing when spillovers are limited. The intuition behind this view is appealing: if knowledge derived through learning-by-doing is non-decreasing in experience, then the knowledge generated by a firm with a monopolistic market share is at least as great as the knowledge that would be generated by any firm if the industry were more competitive. Notably, Dasgupta and Stiglitz's (1988) model does not allow for technological parametric interdependencies.

We develop the proposition that parametric interdependency moderates the relationship between industrial structure and learning-by-doing. Contrary to Dasgupta and Stiglitz (1988), when the underlying technology exhibits parametric interdependencies, we posit that an increase in the number of firms in the industry increases the expected efficacy of the technology developed in the long run. This is because an increase in the number of firms in an industry decreases the probability that all firms get locked into the same low-outcome technology development path.

This notion of getting locked into a low-outcome technology development path is illustrated in an example provided by Henderson and Clark (1990). Henderson and Clark examined the development of photolithographic aligners for the production of solid-state semiconductor devices. A technology start-up named Kesper Instruments was an early leader in that market. Kesper's high performing technology featured "contact" aligners as opposed to "proximity" aligners. Kesper was aware of the "proximity" approach, and in 1973 Kesper offered a "proximity" setting on their "contact" aligners. However, their aligners worked best in the "contact" settings. In the late 1970s, Cannon introduced aligners that were designed to work as "proximity" aligners. Although all the component technology in Cannon's "proximity" aligners had been available to Kesper for almost a decade, the design decisions embodied in Cannon's approach were different from that in Kesper's. Cannon's products proved to be far superior, and Kesper exited the industry in 1981.

Parametric interdependencies inhibited Kesper's ability to remain at the forefront of this technology. Had Kesper re-optimized every design decision around the premises that "proximity" aligners were superior to "contact" aligners, they might not have been so vulnerable to Cannon's technology threat. However, when their initial experiments with "proximity" aligners failed to increase the efficacy of their technology, Kesper moved on in search of other incremental innovations. If Cannon had not entered this industry, and Kesper became a monopolist, we might still be stuck with suboptimal "contact" aligner technology.

The Kesper example demonstrates that R&D driven technology development can get locked into a low-outcome technology path, but this phenomenon is even more likely with learning-by-doing driven development. Learning-by-doing typically changes fewer parameters simultaneously, increasing the odds that firms will get stuck with a globally suboptimal technology when parametric interdependencies are present. Innovations derived from learning-by-doing are also less likely to be directed; consider innovations derived from accidental process deviations on the shop floor. As such, the learning-by-doing process is more likely to lead two firms down different technology paths when parametric interdependencies are present.

The notion that different firms developing the same technology could get stuck with different globally suboptimal instantiations of the technology implies that the expected efficacy of the best technology developed by the industry will decrease as industrial concentration increases. (The expected value of one die role is lower than the expected value of the better of two die roles.) This insight can be formalized as follows:

Proposition 1 *In the long run, the expected efficacy of the technology uncovered by a monopolistic structure is less than the expected value of the technology of a baseline oligopolistic structure if parametric interdependencies are present.*

2.2 Spillovers

There is strong empirical evidence for the existence of spillovers (Bernstein and Nadiri 1989; Irwin and Klenow 1994; Jaffe 1986; Lieberman and Montgomery 1988, p. 43; Mansfield 1985). Classic microeconomic work emphasizes the spillovers' potential to reduce the incentives for R&D (e.g. Spence 1984). More recent work deemphasizes spillovers as a potential deterrent to R&D. For example, Cohen and Levinthal (1989) argue that R&D builds absorptive capacity, which enables a firm to exploit spillovers. Rivkin (2000) provides a computational model that demonstrates why complex strategies or technologies can be difficult to unilaterally imitate. Pyka et al. (2009) use a computational model to demonstrate the stability of a system in which innovating firms voluntarily choose to enable spillovers within innovation networks.

A similar progression is seen in the work on learning-by-doing. Early works (e.g. Spence 1981) examined the possibility that the learning gains from spillovers might cause firms to reduce production. More recent work focuses on non-strategic learning-by-doing, in which spillovers do not discourage production or capital investment. Jin et al. (2004) consider the impact of spillovers on non-strategic learning-by-doing. Their model suggests that increasing spillovers is an unambiguously desirable policy goal, because they both reduce shakeouts and limit the welfare implications of those shakeouts that occur. The model presented in Randon and Naimzada (2006) is congruous to this view.

Our model suggests that enhancing actionable spillovers beyond a critical point can decrease the gains from non-strategic learning-by-doing when parametric interdependencies are present. When spillovers are actionable, complete, and instantaneous, all firms will end up on the same development path. However, if spillovers are

Without Interdependencies	With Interdependencies

Absent parametric interdependencies, the technology landscape is smooth; all development paths eventually lead to the same global optimum.	Parametric interdependencies make the topology rugged; different development paths can lead to different local optima.

Fig. 1 Parametric interdependencies and path dependencies

limited then it is possible that different firms will travel down different technology paths. Following the logic of the section above, when parametric interdependencies are present, the expected efficacy of the best technology eventually discovered increases with the number of technology development paths explored. This notion is formalized as follows:

Proposition 2 *When parametric interdependencies are present in a technology that oligopolists are developing through learning-by-doing, there exists a critical level of actionable spillovers. Beyond the critical point, increases in spillovers decrease the expected efficacy of the best technology discovered.*

3 Model

3.1 Parametric interdependencies and NK landscapes

Performance landscapes are a useful tool in modeling the impact of parametric interdependencies. Kauffman (1993) introduced the concept of a “technological fitness landscape”, or technology performance landscape, as a repurposing of Sewell Wright’s (1932) concept of a fitness landscape. Fitness landscapes are $N + 1$ dimensional topologies, in which each of the first N dimensions represents a design parameter of the technology under development. The final $N + 1$ th “fitness” dimension is a measure of efficiency, or technological efficacy, associated with each possible instantiation of the technology.

For performance landscapes in which all of the non-fitness dimensions are discrete (i.e., the first N dimensions), we can conceptualize a peak as a technology specification at which changing the setting of any C (or fewer) parameters lowers the technology’s efficacy. C is a constant, whose value is typically one. When a landscape exhibits multiple peaks with various heights, the landscape can be said to be rugged or complex. Conversely, if there is only one peak on a landscape (i.e., the

global optimum), then that landscape is said to be smooth. If a landscape is rugged, it necessarily implies that there are parametric interdependencies.

It is the ruggedness of the landscape that engenders the path dependency discussed in the introduction and theory sections. This notion is illustrated in Fig. 1.

It is intractable to specify a performance landscape for every real world technology. Therefore, researchers frequently identify classes of relevant landscapes and use stochastic simulation techniques to generate thousands of landscapes within each class (see Auerswald 2010; Auerswald et al. 2000; Frenken 2006; Kauffman et al. 2000; Lobo and Macready 1999; Zhang and Gao 2010).

NK landscapes (Kauffman 1993) constitute a class of landscapes commonly used to examine the impact of parametric interdependencies on technology development. NK landscapes are a family of well-defined landscapes in which the level of interdependence among parameters can be easily adjusted. Although NK landscapes were designed to model the fitness of protean chains, Kauffman (1988) proposes using generalized NK landscapes to represent the technology embedded in products. Kauffman et al. (2000) argue that NK landscapes can be used to represent technology landscapes.¹ The proposition that NK landscapes can represent a technology landscape, a strategy landscape, or the contours of the “firm’s problem” is widely applied in the literature (cf. Almirall and Casadesus-Masanell 2010; Auerswald et al. 2000; Fang et al. 2010; Frenken 2006; Frenken and Nuvolari 2004; Gavetti et al. 2005; Ganco and Hoetker 2009; Kauffman et al. 2000; Kollman et al. 2000; Lenox et al. 2007; Levinthal 1997; Lobo and Macready 1999; Lobo et al. 2004; Marengo et al. 2000; Rivkin 2000; Rivkin and Siggelkow 2002, 2005; Zhang and Gao 2010). NK landscapes are also frequently used as a model of team interaction (cf. Carroll and Burton 2000; Lacks 2004; Solow et al. 2005).

An NK landscape maps a vector of N binary elements $(x_1, x_2, x_3, \dots, x_N)$ to one continuous fitness variable $F(\mathbf{x})$. Each element, x_i , takes a value of either one or zero, and represents one parameter in the technology’s specification. The fitness of a position on the landscape is the summation of N functions, g_i . The function g_i is dependent on the value of x_i , and the value of K other parameters, $\mathbf{x}_{-i}$. The total fitness associated with a particular location is the summation of each element’s contribution:

$$F(\mathbf{x}) = g_1(x_1, \mathbf{x}_{-1}) + g_2(x_2, \mathbf{x}_{-2}) + \dots + g_N(x_N, \mathbf{x}_{-N}) \quad (1)$$

To generate an NK landscape, each element, x_i , is randomly assigned K other elements, $\mathbf{x}_{-i}$. There are 2^{K+1} possible values of $(x_i, \mathbf{x}_{-i})$, each of which is assigned a value drawn randomly from a uniform $[0,1]$ distribution. Let $\text{RD}(x_i^*, \mathbf{x}_{-i}^*)$ equal the value of the random draw associated with a particular value of $(x_i, \mathbf{x}_{-i})$, $(x_i^*, \mathbf{x}_{-i}^*)$. Then:

$$g_i(x_i^*, \mathbf{x}_{-i}^*) = \text{RD}(x_i^*, \mathbf{x}_{-i}^*)/N \quad (2)$$

When K equals zero, g_i is only a function of x_i . Hence, there is no parametric interdependence, and the maximum fitness difference between two adjacent points is

¹ Kauffman et al. (2000) actually proposed a new landscape, called the “Ne technology landscape”. However, their Ne technology landscape is mathematically equivalent to the NK landscapes used in this paper.

$1/N$.² When K equals zero, the fitness function, $F(\mathbf{x})$, is perfectly linear, and there is exactly one local optimum.

K is literally a quantitative measure of the number of parametric interdependencies in the landscape. As the value of K increases, the greater the degree of parametric interdependency in the technology, and the greater the expected fitness difference between two adjacent points on the landscape. Also, as the value of K increases, the expected number of peaks (i.e., local optimum) increases. In the context of an NK landscape, a peak, or local optimum, refers to a location from which no single change in one parameter can increase the fitness.

It is important to note that the absolute value of the fitness associated with each point on the landscape is not interpretable in isolation. The average fitness associated with a point on the landscape is a function of K and N . As a result, it is most meaningful to view the value of each point as a percentage of the value of the highest peak on the landscape. Hence, the highest peak is given a value of 1, and a point whose value is half of the highest peak is given a value of 0.5.

3.2 A search-based model of learning-by-doing

The concept of learning-by-doing has been illustrated in many ways. First, the classic learning curve model implies that productivity is an exponential function of experience, and has been applied to the increases in productivity of those individuals who perform multiple repetitive tasks (Adler and Clark 1991; Chambers and Kouvelis 2003). Second, experiential learning-by-doing, identified as first-order learning by Adler and Clark (1991), includes incremental development of expertise resulting from the production process. Third, Arrow's (1962) paper on learning-by-doing denotes that learning arising from the act of production provides the ability to identify problems from which solutions can be generated over time.

We promote the view put forth by Young (1993), which argues that technical change is the serendipitous result of production experience. Incremental improvements lead to increased productivity due to learning-by-doing on the "shop floor" by production workers (Adler and Clark 1991; Auerswald et al. 2000; Dasgupta and Stiglitz 1988; Zollo and Winter 2002). Through continuous repetition of operating routines, employees are able to identify methods to improve their portion of production on the shop floor. The accumulation of this tacit knowledge generated by repetitive tasks leads to incremental improvement (Zollo and Winter 2002).

Learning-by-doing is frequently represented with computational models (cf. Ahrweiler et al. 2011; Péli and Nooteboom 1997). Our approach is in line with that of Auerswald (2010) and Auerswald et al. (2000). We model learning-by-doing as a local, myopic search process. From this perspective, we are unconcerned with whether the discoveries are serendipitous (cf. Young 1993) or the result of routines developed to identify improvements (cf. Zollo and Winter 2002). In either case, slight alterations embodying incremental improvements may be adapted. The initial state of a firm's

² Adjacent points on an NK landscape are points whose value is identical in all but one of the N dimensions. For example, the point (1,1,1) is adjacent to the points (0, 1, 1), (1, 0, 1), and (0, 0, 1) on a landscape where N equals three.

technology is represented as a randomly chosen point on the NK performance landscape described in section three-one. Experience, or learning-by-doing, is assumed to provide insight into the impact of a low dimensional alteration to the technology's specification. Hence, one of the first N dimensions will be selected at random and flipped.^{3,4} If the one-dimensional flip fails to increase the efficacy of the firm's technology, the alteration is rejected and the firm's position in the NK landscape remains unchanged. However, if the alteration does increase the efficacy of the firm's technology, the change is adopted. When a change is adopted, the firm effectively moves to a new point on the NK landscape. The next iteration of learning-by-doing begins from this new point, not the original point.

We believe the simple search process above captures the salient features of the learning-by-doing process described at the start of this section. First, experience provides insight into a change that might increase efficiency. An experiment is used to reject or confirm the insight. If the experiment does confirm the insight, the change is adopted and incorporated into the new starting point for future insights and experiments.

To operationalize the difference between learning-by-doing in a monopolistic structure and an oligopolistic structure, we create two search algorithms employing the above learning-by-doing search process. In the oligopolistic search algorithm, there are two firms. Each firm executes the above search process once per time period. In our initial experiments, the firms operate independently and do not share information. In the monopolistic search algorithm there is only one firm, which executes the above search process twice per time period. This difference captures the notion that the monopolist produces more and hence has more opportunities to learn by doing.

Our baseline model includes a number of assumptions. Specifically:

- Each firm starts with a different initial instantiation, chosen at random from all possible locations on the landscape.
- Search is strictly local. When firms execute the learning-by-doing search process described above, they only consider changing one parameter at a time.
- There are no spillovers. Each firm develops the technology in isolation.
- Total industry output is identical under the monopolistic and oligopolistic structures.
- Each oligopolist produces exactly half of total output.
- Each firm develops only one instantiation of the technology. Hence, the monopolist does not operate multiple factories with each factory developing a unique instantiation of the technology.

After establishing a baseline, we relax each of the first three assumptions. Relaxing the first two assumptions does not qualitatively change our findings. The implication of the third assumption is a topic of extensive analysis in the following sections.

³The proofs in the [Appendix](#) generalize the first proposition to allow search in which the maximum number of simultaneous parameter changes is less than the total number of parameters.

⁴If the value of the selected dimension was originally one, it becomes zero. Conversely, if the value was originally zero, it becomes one.

Assumption four is a point of departure from many models examining the impact of industrial structure. However, relaxing this assumption would only strengthen our findings. Assumption five is made for convenience. As demonstrated in the proofs in the [Appendix](#), assumption five does not impact our findings. The implications of assumption six are discussed in the concluding section.

4 Results and analysis

4.1 Industrial structure

The results are scaled as a percentage of the highest point on the landscape. Hence, the global peak has a value of 1. A point corresponding to an instantiation of the technology that was only 80 % as effective as the best instantiation would have a value of 0.8.

Figure 2 displays the average results of 1000 runs of the baseline model with K equal to zero, and hence no parametric interdependencies.⁵ The x -axis displays the number of time steps, and the y -axis displays the efficacy of the best technology uncovered by each structure. In this case, both the monopolistic and oligopolistic mechanisms uncover the global peak. However, the monopolistic structure reaches the peak first since it produces twice the output of each oligopolist. This finding is parsimonious with the received wisdom regarding industrial structure and learning-by-doing; when there are no parametric interdependencies and no spillovers, the monopolistic structure is superior.

Figure 3 compares the performance of the two industrial structures when K equals 4; hence, parametric interdependencies are present in the technology being developed. In this case, the monopolistic structure initially outperforms the oligopolistic structure, once again because it learns at twice the rate. However, this phenomenon is short lived; the oligopolistic structure uncovers a better technology specification in the long run.

The results displayed in Fig. 3, reflect the fact that increasing the number of firms in the industry decreases the likelihood that a stochastic learning-by-doing process will lead all firms to the same technology specification. Put another way, increasing the number of firms in the industry increases the likelihood that at least one of them will develop a better technology specification than a monopolist, such as Cannon's development of the proximity-optimized aligner technology.

Table 1 displays the frequency with which the long run performance of each industrial structure outperformed the other industrial structure as a percentage of the 1000 landscapes-trials generated for each value of K . Table 1 also shows the percentage of trials for which both industrial structures tied. Finally, Table 1 displays the

⁵We use the Mersenne Twister pseudo random number generator in these experiments. Despite their low diffusion, MTs are preferred for this type of simulation analysis because of their exceptionally long cycle length. The MT implemented in this simulation was coded by Anger Fog and distributes as a C++ library class available at www.anger.org. The instantiation of the class used was last modified on 8/3/10. With the exception of the pseudo random number generator, all other code was either part of the C++ standard ("std") library or coded by the authors.

Fig. 2 No parametric interdependencies ($N = 8, K = 0$)

Fig. 3 With parametric interdependencies ($N = 8, K = 4$)

significance with which paired t-tests reject the following: (i) the hypothesis that the mean efficacy of the technology uncovered by the monopolistic mechanism exceeds that of the oligopolistic mechanism, (ii) the hypothesis that the mean efficacy of the technology uncovered by the oligopolistic mechanism exceeds that of the monopolistic mechanism, and (iii) the hypothesis that the mean efficacy of the technology uncovered by both structures is equal.

The results in Table 1 support Proposition 1. Specifically, the significance of the paired t-tests in the “Oligopolist” column and the “Ties” column suggests that when K is greater than zero, we can reject both the hypothesis that the mean efficacy of the technology uncovered by the monopolistic mechanism exceeds that of the oligopolis-

Table 1 Percentage of times each structure uncovers a superior technology

<i>K</i>	Oligopoly	Monopoly	Ties
0	0.0 % ^a	0.0 % ^a	100.0 % ^a
2	48.4 %***	17.0 %	34.6 %***
4	57.0 %***	27.9 %	15.1 %***
6	62.2 %***	31.0 %	6.8 %***

Notes: (1) Percentages in each cell are based on 1000 trials each with 1000 time steps. (2) Asterisks in “Oligopoly” Column indicate the significance with which a paired t-test rejects the hypothesis that the technologies uncovered by the monopolistic structure have a higher expected performance. (3) Asterisks in “Monopoly” Column indicate the significance with which a paired t-test rejects the hypothesis that the technologies uncovered by the oligopolistic structure have a higher expected performance. (4) Asterisks in “Ties” Column indicate the significance with which a paired t-test rejects the hypothesis that the technologies uncovered by both structures have the same expected performance. (5) ^a Indicates that the paired t-test cannot be calculated because all runs ended in ties. (6) *** $p < 0.001$

Table 2 Identical starting technologies & the percentage of times each structure uncovers a superior technology

<i>K</i>	Oligopoly	Monopoly	Ties
0	0.0 % ^a	0.0 % ^a	100.0 % ^a
2	29.2 %***	12.0 %	58.8 %***
4	40.6 %***	17.9 %	41.5 %***
6	43.1 %***	19.3 %	37.6 %***

Notes: (1) Percentages in each cell are based on 1000 trials. (2) Asterisks in “Oligopoly” Column indicate the significance with which a paired t-test rejects the hypothesis that the technologies uncovered by the monopolistic structure have a higher expected performance. (3) Asterisks in “Monopoly” Column indicate the significance with which a paired t-test rejects the hypothesis that the technologies uncovered by the oligopolistic structure have a higher expected performance. (4) Asterisks in “Ties” Column indicate the significance with which a paired t-test rejects the hypothesis that the technologies uncovered by both structures have the same expected performance. (5) ^a Indicates that the paired t-test cannot be calculated because all runs ended in ties. (6) *** $p < 0.001$

tic mechanism, and the hypothesis that the mean efficacy of the technology uncovered by both structures is equal.

Table 2 provides full data in the model when all firms started with the same initial technology. Hence, on each landscape, the monopolist and both of the oligopolists began at the same initial location. As the results of Table 2 show, giving all firms an identical initial technology does not qualitatively change the above results. These results further confirm Proposition 1.

Following the results displayed in Tables 1 and 2, the Appendix offers a formal proof of Proposition 1. Lemmas in the Appendix prove more generalized versions of Proposition 1. One generalization increases the maximum number of parameters simultaneously changed to any number below N . Another generalization shows that the expected efficacy of the best technology developed by any firm in the industry decreases as industrial concentration increases.

Table 3 Impact of spillovers when firms start with different technologies

Nature of spillovers between oligopolist firms		
Scope of spillovers (# of parameters)	Diffusion rate (# of time periods)	Oligopolist develops superior technology (%) (Ties in parentheses)
0 of 8	NA	58.6 %*** (15.2 %)
2 of 8	0	56.5 %*** (14.5 %)
	10	61.3 %*** (15.4 %)
	35	63.7 %*** (15.6 %)
4 of 8	0	52.3 %*** (14.3 %)
	10	60.3 %*** (15.6 %)
	35	63.9 %*** (16.3 %)
8 of 8	0	44.2 % (12.5 %)
	10	52.0 %*** (14.1 %)
	35	58.2 %*** (15.2 %)

Notes: (1) “Oligopolist develops superior technology (%)” measures percent of trials in which the oligopolistic structure uncovers a technology that is superior to that of the monopolistic structure. Percent of trials that end with both structures uncovering equally effective technologies is reported in parentheses. (2) Percentages reported in cells are based on 10,000 trials, with 2000 time steps each, on landscapes where $N = 8$, $K = 4$. (3) Significance measures in this column report where a paired t-test rejects the hypothesis that the monopolistic structure locates superior technologies. (4) *** $p < 0.001$

4.2 Spillovers

There are a number of different ways to add spillovers to our model. All involve permitting an oligopolist the ability to see if its competitor’s technology is more efficient than its own. If the competitor’s technology is more efficient, the firm can copy some percentage of the competitor’s technology specification. Let X equal the number of dimensions of a competitor’s technology that are subject to spillovers. If X equals N , then spillovers are complete. If X is less than N , then spillovers are partial. Partial spillovers capture the idea that while some aspects of a technology are easy to copy, others are either protected by intellectual property law, corporate secrets, or are otherwise opaque. They also capture that the idea that some firms may lock-in to some aspects of their technology. Some moves across the technology landscape may be too far. Without loss of generality, we will assume that it is the first X dimensions that are copied when spillovers are partial. In addition to whether spillovers are partial or complete, we must also decide how quickly and frequently spillovers occur.

While many models assume spillovers are instantaneous, we also examine the effect of slower diffusion rates in addition to instantaneous spillovers. To operationalize a slow diffusion rate, we let D equal the number of consecutive periods in which an oligopolist must employ an inferior technology before copying from its rival. If D equals zero, the spillovers are instantaneous. The higher the value of D , the lower the spillover diffusion rate.

Table 3 compares the impact of various spillover models and the original monopolistic model when K equals 4 (i.e., some parametric interdependencies are present). We use comparisons to the monopolistic model as baseline in order to facilitate com-

Table 4 Impact of spillovers when firms start with identical technologies

Nature of spillovers between oligopolist firms		
Scope of spillovers (# of parameters)	Diffusion rate (# of time periods)	Oligopolist develops superior technology (%) (Ties in parentheses)
0 of 8	NA	41.2 %*** (42.5 %)
2 of 8	0	38.4 %*** (41.6 %)
	10	42.9 %*** (41.2 %)
	35	44.5 %*** (41.3 %)
4 of 8	0	34.3 %*** (41.7 %)
	10	41.8 %*** (41.7 %)
	35	45.2 %*** (40.9 %)
8 of 8	0	30.5 % (40.5 %)
	10	36.2 %*** (41.9 %)
	35	40.9 %*** (42.5 %)

Notes: (1) “Oligopolist develops superior technology (%)” measures percent of trials in which the oligopolistic structure uncovers a technology that is superior to that of the monopolistic structure. Percent of trials that end with both structures uncovering equally effective technologies is reported in parentheses. (2) Percentages reported in cells are based on 10,000 trials, with 2000 time steps each, on landscapes where $N = 8$, $K = 4$. (3) Significance measures in this column report where a paired t-test rejects the hypothesis that the monopolistic structure locates superior technologies. (4) *** $p < 0.001$

parisons across different levels of spillovers for the oligopolistic model. The data in Table 3 suggest that the oligopolistic structure’s advantage decreases as the diffusion rate increases and the scope of spillovers increase. However, the oligopolistic model with low scope of spillovers and slow diffusion can outperform the oligopolistic model with no spillovers. This is because firms can benefit from other companies’ experiments without being locked into the same development path. Table 4 replicates the results of Table 3 under the assumption that all firms start with the same technology. Again, this assumption does not substantively alter the results.

As shown in Tables 3 and 4, the addition of instantaneous spillovers generally decreases the benefit of the oligopolistic structure. This is consistent with the insights of the baseline model; the value of the oligopolistic structure is derived from the likelihood that it will uncover two different peaks. If rapid and rich spillovers increase the likelihood that the two oligopolists will converge on the same peak, then spillovers will decrease the advantages of the oligopolistic structure.

As we examine variants of the model in which spillovers have decreasing scope and slower diffusion rates, Tables 3 and 4 suggest that slow and partial spillovers are superior to no spillovers. When the interval between oligopolists incorporating spillover knowledge is long (i.e., low diffusion rate), it is possible for each oligopolist to near, or even reach, a peak before incorporating the information from spillovers. In this case, the oligopolist that located the lower peak changes its technological specification while the oligopolist that located the higher peak does not move. If the spillover scope is also low, the oligopolist that located the lower peak does not necessarily move to the peak identified by the more successful oligopolist. Hence, with very low spillover diffusion and scope the oligopolist structure has even greater

search range than it does in the baseline “no-spillover” instantiation. These results support Proposition 2.

An example best illustrates why cases of low spillover scope with a low diffusion rate can outperform cases of no spillovers. Consider a hypothetical case in which N equals 8, K equals 2, the scope of spillovers is two parameters, and information from spillovers are only assimilated if one firm’s technology has been inferior to the other firm’s technology for 35 consecutive time steps. Assume that Oligopolist 2’s technology has been inferior to that of Oligopolist 1’s for the first 35 time steps. Further assume that both oligopolists have located peaks. Oligopolist 2 would then replace the values of the first two dimensions of its technology with the values of the first two dimensions of Oligopolist 1’s technology. After this change, the learning-by-doing engendered local search *might* cause Oligopolist 2 to locate to Oligopolist 1’s peak. However, it is also possible that the local search will lead Oligopolist 2 to a new peak. If Oligopolist 2’s new peak is higher than Oligopolist 1’s peak, then the oligopolistic structure’s performance on that landscape is higher than it would have been in the absence of spillovers. If Oligopolist 2’s new peak is higher than Oligopolist 1’s peak, *and* the new peak values in the first two dimensions of the new technology differ from that of Oligopolist 1, the cycle begins again.

5 Conclusions

This paper makes two primary contributions. First, we demonstrate that when parametric interdependencies are present and spillovers are minimal, the monopolistic industrial structure does not maximize technological gains from learning-by-doing. Second, we demonstrate that when parametric interdependencies are present, there is a level of actionable spillovers beyond which the technology gains from non-strategic learning-by-doing decrease. These assertions are driven by the realization that parametric interdependencies in the underlying technology make learning-by-doing a nondeterministic, path-dependent process. As a result, the expected efficacy of the best technology uncovered by an industry increases with the number of technology development paths followed by firms in that industry. Our assertions, both of which refine the insights of previous works, are summarized in Table 5 below.

Because use most technologies exhibit some parametric interdependencies (Fleming and Sorenson 2001), our findings have policy implications. First, our model helps refine the policy implications of Jin et al. (2004). The model in Jin et al. examined the impact of spillovers on ‘shakeouts,’ or events in which a significant number of firms exit an industry. Jin et al. provides evidence that increasing spillovers reduces the number of shakeouts, possibly because spillovers allow less efficient firms to catch up with more efficient competitors. This insight motivates Jin et al.’s policy conclusion; increasing spillovers is an unambiguous policy goal. Our model’s results are congruent with this view when parametric interdependencies are minimal, or when lock-in prevents firms from fully acting on spillovers. However, the picture is less clear when the technology under development is subject to parametric interdependencies and firms are largely able to act on the information in spillovers. The ability of spillovers to delay and soften the impact of shakeouts could be counteracted by the suboptimal technologies to which they may lead.

Table 5 Summary of key findings

	Increased industrial concentration (Absent Spillovers)	Increased spillovers
Without parametric interdependencies	Enhances rate of technology development (congruent with Dasgupta and Stiglitz 1988)	Enhances rate of technology development (policy implications congruent with implications of Jin et al. 2004)
With parametric interdependencies	Lowers expected efficacy of the best technology uncovered	Broad and instantaneous spillovers lower the expected efficacy of the best technology uncovered. However the expected efficacy of the best technology uncovered with slow and partial spillovers is higher than that without spillovers

While our results suggest that increasing spillovers does not always improve outcomes, it would be imprudent to apply these insights to intellectual property policy. As noted in the bottom right quadrant of Table 5, some spillovers are better than no spillovers. Even absent intellectual property policies, spillovers are neither instantaneous nor complete. For example, spillovers can be limited by geographic constraints (Audretsch and Feldman 1996) and firms' absorptive capacity (Cohen and Levinthal 1989). Whether the natural rate of spillovers is above or below the optimal level is most likely a technology and industry specific question, which is best addressed by empirical research.

Our model has clearer policy implications for the “National Champions” model of economic development. Under the National Champions model, governments often support industrial consolidation in an attempt to grow a large domestic firm capable of competing internationally. Our model suggests that such policies might limit the gains derived from learning-by-doing; higher industrial concentration limits the number of technology paths the industry explores. The impact of policies that promote consolidation on innovation is a timely topic. China's current National Champions policies promote industrial consolidation in key industries (Guest and Sutherland 2010). Unlike previous instances of state sponsored industrial consolidation, the Chinese government also promotes target levels of R&D spending on par with that of the United States (Cao et al. 2006). Hence, the notion that industrial structure impacts the level of innovation is not as relevant in the case of Chinese state owned enterprises. The behavior of the model in this paper modestly suggests that path dependence might limit the technological development achieved by this structure. One possible solution is to limit the information flows between various work units within these National Champions. Hence, from an information flow perspective, these monopolists would function more like oligopolists. This is an area for future research.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

Appendix: Proof of Proposition 1

The simulation results in section four provide important insight into the model's behavior. However, it is also illustrative to employ formal analytical analysis and flesh out how the model's mechanics corroborate our first proposition.

Proposition 1 *In the long run, the expected value of the technology uncovered by the baseline monopolistic structure is less than the expected value of the technology of a baseline oligopolistic structure if parametric interdependencies are present.*

Proof (in six steps)

- (i) *In the long run, firms will get stuck on a peak:* Over a long enough timeline, a myopic, local, “hill-climbing” algorithm, such as our learning-by-doing algorithm, will eventually get stuck on a peak when searching any NK landscape. The peak upon which the hill-climber gets stuck may or may not be the highest peak. Once the hill-climber is stuck, additional time to search will not improve the outcome (see Kauffman 1993, pp. 33–67). Note that we simply define “long run” as a time period long enough to ensure that all of our learning-by-doing search algorithms have reached a sticking point. Assuming the slowest algorithm executes a search once per time step, the number of time steps needed to fulfill this “long run criteria” can be bounded as less than the total area of the landscape (i.e., 2^N time steps).
- (ii) *In the long run, the expected behavior of the monopolist is identical to each of the oligopolists:* As noted in step (i), all our learning-by-doing algorithms will eventually come to rest on a peak. In the baseline model, the behavior of all three firms is fully independent. The only difference between the monopolistic search algorithm and the search algorithm for *each* firm in the oligopolistic model is the number of times the algorithm executes per time step: twice for the monopolist, once for each oligopolist. If we were to redefine the units of our timeline from “time steps” into “2time-steps,” such that each 2time-step equaled one time step, the expected behavior of each oligopolist in this 2time-steps timeframe would be identical to the behavior of the monopolist in normal timeframe. Again, as noted in step (i), all our learning-by-doing algorithms will eventually come to rest on a peak, after which the passing of additional time steps does not alter the search behavior. Hence, if the timeline is sufficiently long (see note in step (i)), the long run behavior of the monopolist and the oligopolist are identical.
- (iii) *A probability distribution can be defined over all peaks:* For any generated NK landscape, we can identify all the peaks upon which our learning-by-doing algorithm could get stuck.⁶ Given that uniform probability distributions determine both the firm's initial location on the landscape and the algorithm's search behavior, we can assign each peak a probability with which the algorithm will get

⁶Following Rivkin (2000), this can be accomplished by starting the algorithm on every point on the landscape, and recording the points from which the algorithm could not move.

stuck on that particular point.⁷ Hence, we identify a probability distribution over all the possible sticking points.

- (iv) *The value of the efficacy of the monopolist's long run technology specification and that of each of the oligopolists can be treated as an independent and identically distributed (i.i.d.) random variable drawn from the distribution defined in step (iii):* This follows directly from the conclusions in steps (ii) and (iii).
- (v) *The expected performance of the oligopolist is at least as good as that of the monopolist:* Let D denote a random variable drawn from the distribution identified in step (iii). The expected performance of the monopolist is $E[D_m]$. The expected performance of the oligopolist is $E[\max D_{O1}, D_{O2}]$. D_m , D_{O1} , and D_{O2} are i.i.d. By basic probability theory:

$$E[\max D_{O1}, D_{O2}] \geq E[D_m] \quad (3)$$

Note: The above equality is strict whenever the range of the D contains more than one element with a non-zero probability. Put another way, as long as there is more than one sticking point on the landscape, the above equality is strict.

- (vi) *If there are parametric interdependencies (i.e., $K > 0$), the expected performance of the oligopolist is better than that of the monopolist:* When K is greater than zero, it has been demonstrated that it is possible to generate a landscape such that there exists at least one non-optimum sticking point for a local, myopic search algorithm (see Rivkin 2000).⁸ If we consider the expected values of the technology located by each structure *before* the landscape is generated, we have to admit the non-zero probability that the soon-to-be generated landscape will have one or more non-optimal local sticking points. Hence, the inequality in equation three becomes strict, leaving us with:

$$E[\max D_{O1}, D_{O2}] > E[D_m] \quad (4)$$

□

Lemma 1 *When there are no parametric interdependencies (i.e., $K = 0$), long run values of technologies identified by both monopolistic and oligopolistic structures are identical.*

Proof When $K = 0$, there is only one sticking point in an NK landscape, its global peak (Kauffman 1993). Hence, the long run probability that our learning gets stuck on that one peak is exactly one, and the inequality in equation three becomes an equality. □

Lemma 2 *Proposition 1 holds even if all firms start with the same initial technology.*

⁷The identification of such a distribution would be an arduous process, and the result would be specific to a particular instantiation of the landscape. However, all we need for this proof is the existence of the distribution.

⁸Note that the existence of sticking points that are not global optimums is not just possible, but likely. As K increases, the difficulty in crafting a landscape that does not include multiple local peaks increases. In all but the rarest cases, we would expect to find sticking points for a myopic, local, “hill-climber” algorithm whenever $K > 0$.

Proof Although the range will be more limited, the probability distribution described in step (iii) for the proof of Proposition 1 can still be defined. The likelihood of there being only one sticking point is greater when conditioned on the algorithm's starting location: some initial locations may lock a firm into a particular technology development path. However, when K is greater than zero, there exist landscapes with starting points that would not lock the firm into a particular development path.⁹ Hence, the argument in step (vi) of the proof of Proposition 1 is still valid. The proof of Proposition 1 can otherwise be replicated to prove Lemma 2. $\square$

Lemma 3 *As long as the scope of search is less than the dimensionality of the landscape, Proposition 1 will hold.*

Proof Consider an NK landscape on which the optimal value for all N parameters equals one. Assume that the second highest value on this landscape corresponds to the position at which all N elements equal zero. This second assumption necessarily implies that all positions that involve a mix of ones and zeros correspond to a lower value than the top ranked “all ones” position and the second ranked “all zeros” position. If a local, myopic, “hill-climber” algorithm with the ability to simultaneously move in $N - 1$ dimensions were to begin on the “all zeros” position, or wander to the “all zeros” position, the algorithm would become stuck and be unable to ever reach the “all ones” position. Although unlikely to be generated at random, the probability of such a landscape's generation is greater than zero. Hence, if we take our expectations of each structure's performance before the landscape is generated, the inequality described in equation four is still strict, and the reasoning in step (vi) of the proof of Proposition 1 still holds. The proof of Proposition 1 can otherwise be replicated to prove Lemma 3. $\square$

Lemma 4 *If parametric interdependencies are present in the underlying technology, the expected long run value of the technology uncovered by an industry increases as the number of firms in the industry increases.*

Proof Without loss of generality, we will compare an industrial structure with A firms to one with B firms, such that $A < B$. Assume that each of the A firms search faster than each of the B firms and on average, each of the A firms have a larger market share than each of the B firms. However, in light of our definition of “long run” as being a time period long enough for the slowest algorithm to have reached a sticking point, the different search speeds are immaterial. Using the variables defined in the original proof, the expected value of the technology located by the industry with A firms is then $E[\max D_1, D_2, \dots, D_A]$. The expected value of the technology located by the industry with B firms is $E[\max D_1, D_2, \dots, D_B]$. By simple probability theory, as long as $A < B$,

$$E[\max D_1, D_2, \dots, D_B] \geq E[\max D_1, D_2, \dots, D_A] \quad (5)$$

⁹This assumes that the firms are using a one-dimensional, myopic, hill-climbing search algorithm, such as those used to model learning-by-doing in this paper. There exist other search algorithms for which this would not be true. For example, if firms used a “never move, no matter what” search algorithm, then every possible starting point on every landscape would lock the firm into a particular “development” path.

Substituting Eq. (5) for Eq. (3) in the proof of Proposition 1, the proof of Lemma 4 follows the proof of Proposition 1. $\square$

Lemma 4 significantly increases the applicability of our findings. While our arguments were built comparing a monopolistic structure to a duopolistic structure, Lemma 4 asserts that our conclusions apply more broadly to the impact of industrial concentration on learning-by-doing. As with step (ii) in the original proof, the proof of Lemma 4 also highlights the role of time in this analysis. Lemma 4 does not require the market share to be distributed equally among the firms, but the more unequal the market share, the longer the timeline needed to obtain our results.

Lemma 5 *The expected time needed for a monopolist to locate a peak is less than that of a single oligopolist firm.*

Proof One of the model's principle assumptions is that the number of times the learning-by-doing search algorithm executes a search per time step is an increasing function of market share. The baseline monopolist (with 100 % market share) executed two searches per time step, while each baseline oligopolist (each with 50 % market share) executed one search per time step. As the search algorithms are otherwise identical, the expected number of search executions needed to reach a peak is identical between the monopolist and each oligopolist. Hence, the expected number of time steps needed for the monopolist to reach a peak is less than that of each oligopolist. $\square$

References

- Adler PS, Clark K (1991) Behind the learning-curve—a sketch of the learning-process. *Manag Sci* 37(3):267–281
- Ahrweiler P, Gilbert N, Pyka A (2011) Agency and structure: a social simulation of knowledge-intensive industries. *Comput Math Organ Theory* 17:59–76
- Almirall E, Casadesus-Masanell R (2010) Open versus closed innovation: a model of discovery and divergence. *Acad Manag Rev* 35(1):27–47
- Arrow KJ (1959) Rational choice functions and orderings. *Economica* 26:121–127
- Arrow KJ (1962) The economic implications of learning by doing. *Rev Econ Stud* 29(3):155–173
- Audretsch DB, Feldman MP (1996) R&D spillovers and the geography of innovation and production. *Am Econ Rev* 86(3):630–640
- Auerswald PE (2010) Entry and Schumpeterian profits. *J Evol Econ* 20(4):553–582
- Auerswald P, Kauffman S, Lobo J, Shell K (2000) The production recipes approach to modeling technological innovation: an application to learning by doing. *J Econ Dyn Control* 24(3):389–450
- Bernstein J, Nadiri MI (1989) Research and development and intra-industry spillovers: an empirical application of duality theory. *Rev Econ Stud* 56:249–269
- Cao C, Suttmeier RP, Simon DF (2006) “Knowledge invasion” and the Chinese Academy of Sciences. *Science* 312(5770):58–59
- Carroll T, Burton R (2000) Organizations and complexity: searching for the edge of chaos. *Comput Math Organ Theory* 6:319–337
- Chambers C, Kouvelis P (2003) Competition learning and investment in new technology. *IEE Trans* 35(9):863–878
- Cohen W, Levinthal D (1989) Innovation and learning: the two faces of R&D. *Econ J* 99:569–596
- Dasgupta P, Stiglitz J (1988) Learning-by-doing, market structure and industrial and trade policies. *Oxf Econ Pap* 40(2):246–268

- Fang C, Lee J, Schilling MA (2010) Balancing exploration and exploitation through structural design: the isolation of subgroups and organizational learning. *Organ Sci* 21(3):625–642
- Fioretti G (2006) Recognizing investments at the onset of recoveries. *Res Econ* 60:69–84
- Fleming L, Sorenson O (2001) Technology as a complex adaptive system: evidence from patent data. *Res Policy* 30(7):1019–1039
- Frenken K (2006) A fitness landscape approach to technological complexity, modularity and vertical disintegration. *Struct Chang Econ Dyn* 17:288–305
- Frenken K, Nuvolari A (2004) The early development of the steam engine: an evolutionary interpretation using complexity theory. *Ind Corp Change* 13(2):419–450. doi:[10.1093/icc/dth017](https://doi.org/10.1093/icc/dth017)
- Ganco M, Hoetker G (2009) NK modeling methodology in the strategy literature: bounded search on a rugged landscape. *Res Methodol Strategy Manag* 5:237–268
- Gavetti G, Levinthal D, Rivkin J (2005) Strategy making in novel and complex worlds: the power of analogy. *Strateg Manag J* 26:691–712
- Guest P, Sutherland D (2010) The impact of business group affiliation on performance: evidence from China's 'national champions. *Camb J Econ* 34:617–631
- Henderson R, Clark K (1990) Architectural innovation—the reconfiguration of existing product technologies and the failure of established firms. *Adm Sci Q* 35(1):9–30
- Irwin DA, Klenow PJ (1994) Learning-by-doing spillovers in the semiconductor industry. *J Polit Econ* 102(6):1200–1227
- Jaffe AB (1986) Technological opportunity and spillovers of R&D: evidence from firms' patents, profits, and market value. *Am Econ Rev* 76:984–1001
- Jin J, Perote-Pena J, Troege M (2004) Learning by doing, spillovers and shakeouts. *J Evol Econ* 14(1):85–98
- Kauffman S (1988) The evolution of economic webs. In: Anderson PW, Arrow KJ, Pines D (eds) *The economy as an evolving complex system*. Addison-Wesley, Redwood City, pp 125–146
- Kauffman S (1993) *Origins of order: self-organization selection in evolution*. Oxford University Press, New York
- Kauffman S, Lobo J, Macready W (2000) Optimal search on a technology landscape. *J Econ Behav Organ* 43(2):141–166
- Kollman K, Miller J, Page S (2000) Decentralization and the search for policy solutions. *J Law Econ Organ* 16:102–128
- Lacks DJ (2004) Dynamics of team member replacements from complex systems theory. *Comput Math Organ Theory* 10:335–347
- Lenox MJ, Rockart SF, Lewin AY (2007) Interdependency, competition, and industry dynamics. *Manag Sci* 53(4):599–615
- Levinthal D (1997) Adaptation on rugged landscapes. *Manag Sci* 43(7):934–950
- Lieberman MB, Montgomery DB (1988) First-mover advantages. *Strateg Manag J* 9:41–58
- Lobo J, Macready W (1999) Working Paper 99-05-037, Santa Fe Inst. <http://www.santafe.edu/media/workingpapers/99-05-037.pdf>. Accessed 15 Sept 2011
- Lobo J, Miller J, Fontana W (2004) Neutrality in technological landscapes. Working Paper, Santa Fe Inst. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.61.9296&rep=rep1&type=pdf>. Accessed 15 Sept 2011
- Mansfield E (1985) How rapidly does industrial technology leak out? *J Ind Econ* 34:217–224
- Marengo L, Dosi G, Legrenzi P, Pasquali C (2000) The structure of problem-solving knowledge and the structure of organizations. *Ind Corp Change* 9:757–788
- Péli G, Nootboom B (1997) Simulation of learning in supply partnerships. *Comput Math Organ Theory* 3:43–66
- Pyka A, Gilbert N, Ahrweiler P (2009) Agent-based modeling of innovation networks—the fairytale of spillovers. In: Pyka A, Scharnhorst A (eds) *Innovation networks*. Springer complexity series. Springer, New York, pp 101–126
- Randon E, Naimzada A (2006) Dynamics of the non linear learning curve with spillovers in a differentiated oligopoly: effects on industry structure. *J Evol Econ* 17(1):95–106
- Rivkin JW (2000) Imitation of complex strategies. *Manag Sci* 46:824–844
- Rivkin JW, Siggelkow N (2002) Organizational sticking points on NK landscapes. *Complexity* 7(5):31–43
- Rivkin J, Siggelkow N (2005) Speed and search: designing organizations for turbulence and complexity. *Organ Sci* 16:101–122
- Schumpeter JA (1942) *Capitalism, socialism, and democracy*. Harper & Row, New York
- Solow D, Piderit S, Burnetas A, Wong LC (2005) Mathematical models for studying the value of motivational leadership in teams. *Comput Math Organ Theory* 11:5–36

- Spence AM (1981) The learning curve and competition. *Bell J Econ* 12(1):49–70
- Spence M (1984) Cost reduction, competition, and industry performance. *Econometrica* 52(1):101–122
- Wright S (1932) The roles of mutation, inbreeding, crossbreeding and selection in evolution. In: *Proc Sixth Int Congr Genet*, pp 356–366
- Young A (1993) Invention and bounded learning by doing. *J Polit Econ* 101(3):443–472
- Zhang G, Gao R (2010) Modularity and incremental innovation: the roles of design rules and organizational communication. *Comput Math Organ Theory* 16:171–200
- Zollo M, Winter SG (2002) Deliberate learning and the evolution of dynamic capabilities. *Organ Sci* 13(3):339–351

William Martin Tracy is an Assistant Professor of Strategic Management at the Lally School of Management and Technology, Rensselaer Polytechnic Institute. His research focuses on firm learning and search. He holds a PhD in Management from UCLA and a BA from Swarthmore College.

M.V. Shyam Kumar is an Assistant Professor of Corporate Strategy at the Lally School of Management and Technology, Rensselaer Polytechnic Institute. His research is in the areas of firm boundaries, and mergers, acquisitions and joint ventures. Professor Kumar's research has appeared in *Strategic Management Journal*, *Organization Science*, *Managerial Decision Economics*, *The Journal of Business Research*, and *Academy of Management Best Paper Proceedings*. He holds a PhD in Strategy from University of Illinois Urbana-Champaign, an MBA in Finance and Marketing from the Indian Institute of Management, Bangalore and a BTech in Mechanical Engineering from the Indian Institute of Technology, Delhi.

William Paczkowski earned his Juris Doctorate (JD) from Albany Law School and his MBA from Union Graduate College. He is currently a doctoral student in management at the Lally School of Management and Technology, Rensselaer Polytechnic Institute. His research interests are focused on the business and managerial applications of strategy and leadership in corporations and knowledge intensive service industries.