Papers in Evolutionary Economic Geography

05.11

Informational complexity and the flow of knowledge across social boundaries

Olav Sorenson, Jan W. Rivkin, Lee Fleming

Utrecht University
Urban & Regional research centre Utrecht

INFORMATIONAL COMPLEXITY AND THE FLOW OF KNOWLEDGE ACROSS SOCIAL BOUNDARIES

OLAV SORENSON
London Business School
Regent's Park
London NW1 4SA
UNITED KINGDOM
osorenson@london.edu

JAN W. RIVKIN
Harvard Business School

LEE FLEMING
Harvard Business School

Scholars from a variety of backgrounds – economists, sociologists, strategists, and students of technology management – have sought a better understanding of why some knowledge disperses widely while other knowledge does not. In this quest, some researchers have focused on the characteristics of the knowledge itself (e.g., Polanyi, 1966; Reed and DeFillippi, 1990; Zander and Kogut, 1995) while others have emphasized the social networks that constrain and enable the flow of knowledge (e.g., Coleman *et al.*, 1957; Davis and Greve, 1997). This chapter examines the interplay between these two factors.

Specifically, we consider how the complexity of knowledge and the density of social relations jointly influence the movement of knowledge. Imagine a social network composed of patches of dense connections with sparse interstices between them. The dense patches might reflect firms, for instance, or geographic regions or technical communities. When does knowledge diffuse within these dense patches circumscribed by social boundaries but not beyond them? Synthesizing social network theory with a view of knowledge transfer as a search process, we argue that *knowledge inequality across social boundaries should reach its peak when the underlying knowledge is of moderate complexity*.[1] To test this hypothesis, we analyze patent data and compare citation rates across three types of social boundaries: within versus outside the firm, geographically near to versus far from the inventor, and internal versus external to the technological class. In all three cases, the disparity in knowledge diffusion across these borders is greatest for knowledge of an intermediate level of complexity.

THE TRANSFER OF COMPLEX KNOWLEDGE

Our hypotheses build on two themes in the literature. First, the transfer of knowledge from one party to another typically requires effort on the part of the recipient to fill gaps in the transmitted knowledge and to correct transmission errors. The acquisition of knowledge therefore is best seen not as the receipt of a complete, well-packaged gift, but rather as a search process. Second, social networks, and consequently the social boundaries that shape them, critically influence that search process.

[1] This version describes the intuition underlying our theoretical model and reports novel empirical results. For those interested in a more detailed description of the theoretical model, please see Rivkin (2001) and Sorenson, Rivkin and Fleming (2004).

Knowledge receipt as search

Following the lead of evolutionary economists (Nelson and Winter, 1982), we think of a unit of knowledge as analogous to a recipe. The list of 'ingredients' might include both physical components and processes. The recipe further explains how to combine these components and processes – in what order, in which proportions, under what circumstances – to achieve a desired end. Viewing knowledge as a recipe leads naturally to thinking of innovation as a search for new recipes. Following a long tradition – beginning at least as early as Schumpeter (1939) – we explicitly model innovation as a search process; inventors explore the space of possible combinations of ingredients (i.e. recipes) for new and better alternatives. In discussing this process, we adopt the idea of a fitness 'landscape' as a metaphor for the characteristics of the search space. Innovators search these landscapes for peaks and plateaus, which correspond to good recipes, useful inventions, and profitable strategies.

Once a useful innovation has been discovered, transferring its recipe, even between cooperative actors, can fail for at least two reasons: First, the recipient usually does not fully understand the original recipe, as a result of imperfections in the transfer process. She must therefore begin a search for the missing information and to correct the errors in her (imperfect) copy of the recipe. Second, the local ingredients and the experience of the recipient rarely match those of the sender; recipients may therefore need to adapt the original recipe to their own context. Knowledge recipients do not act as passive beneficiaries; they actively search, recreate, and build upon the original recipes.

In this process, the transfer of certain types of recipes is particularly difficult. For instance, knowledge characterized by causal ambiguity (Lippman and Rumelt, 1982), a high degree of tacitness (Polanyi, 1966), or difficult codification (Zander and Kogut, 1995) may resist transfer because any communication of such recipes proceeds only with many and large gaps. Our focus, however, is on the *informational complexity* of transferred knowledge. We consider a piece of knowledge complex if it comprises many elements that interact richly (Simon, 1962), and we pay special attention to the intensity of interdependence among the ingredients in a recipe.

To connect the degree of informational complexity to the characteristics of the space that inventors search, we use Kauffman's *NK* model (Kauffman, 1993; cf. Frenken and Nuvolari, 2004). N denotes the number of (binary) elements in a system while K represents the degree to which these components interact in determining the fitness of a particular configuration of components. In our context, N is the number of ingredients used in a recipe, and K is the richness of the interactions between those ingredients.

To understand better the way in which the model relates interdependence to the search process, consider two examples with $N = 3$. Fig.1 depicts a fitness landscape for a recipe with no interdependence between its components. Each vertex represents a different potential configuration; the arrows connecting them show paths toward higher fitness levels. When $K = 0$, Kauffman randomly assigns a fitness from the uniform unit distribution to each value (0 or 1) of each element. The overall fitness value for a particular configuration is the average of each element's fitness contribution. As one can see, any starting point on this landscape leads to the unique optimum (011).

Fig. 2, on the other hand, illustrates an example with $N = 2$. The value of the fitness contribution for each component then depends not just on its value but also on the values of two other components. Each component therefore can contribute any of eight ($2 \times 2 \times 2$) different fitness levels. Kauffman again randomly assigns these values from the uniform unit distribution. Even in this simple example, one can see that interdependence complicates search; depending on where one begins, an agent using a simple hill-climbing algorithm could arrive at either the global maximum (101) or a local one (000).

INSERT FIGS 1 & 2 ABOUT HERE

Complex knowledge resists transfer by making it difficult for a recipient to fill transmission gaps. On the landscapes depicted, a gap is equivalent to not knowing the correct value for the global optimum of one of the three elements. Interdependence produces two effects that undermine the recipient's attempts to regenerate the original recipe (i.e. identify the optimum). First, small errors in transmission cause large problems when ingredients cross couple in a rich manner. Second, interdependence leads to a proliferation of "local peaks." These peaks undermine improvement through incremental search because changing any single element degrades the quality of the outcome (Kauffman, 1993). As a result, searchers frequently find themselves trapped on local peaks (i.e. inferior recipes) when faced with high interdependence.

Complexity and access to a template

Success in acquiring and employing complex knowledge depends crucially on access to the original success, which serves as a *template* (Nelson and Winter, 1982: 119-120; Winter, 1995). For reasons considered below, individuals vary in their degree of access to the template. Superior access facilitates the receipt of knowledge by allowing the recipient to commence search with fewer errors and by permitting him to solicit advice from the source during the search process. Consider two actors, both attempting to assimilate a valuable piece of knowledge but who differ in their access to the template. The first has superior, though still imperfect, access to and understanding of the original, successful recipe. The second has far poorer access. How valuable is the first actor's superior access to the template during the search process? We contend that the value depends on the complexity of the knowledge being transferred.

Suppose first that the ingredients used in the recipe do not interact (i.e. $K = 0$). In this situation, the first actor's access to the template does not produce a persistent advantage. Through routine, incremental search, the second actor can reconstruct the recipe. Few local peaks threaten to trap the poorly informed recipient. As a result, both actors eventually fare equally well; search on the part of a recipient can easily substitute for high-fidelity transmission.

Next consider knowledge with an intermediate degree of interdependence. Local peaks now appear, but they remain relatively few in number. The well-informed actor begins her search near, but not precisely at, the original optimal set of ingredients. Through incremental search, she can find the proper combination of ingredients. The second actor, who begins search farther from the target and receives less guidance about the direction in which to explore, more likely becomes ensnared on some local peak, away from and inferior to the original success. Here superior template access gives the first actor an advantage the second cannot recreate through search.

Finally, imagine a piece of maximally interdependent knowledge (i.e. $K = N - 1$): ingredients depend on one another in an extremely delicate way. Local peaks now pervade the landscape and neither actor's incremental search will likely build on the original knowledge with any success. The first actor's superior access to the template thus has little value beyond the second's inferior access.

Taken together, these arguments imply that the advantage of superior but imperfect access to the template reaches its peak at moderate levels of interdependence between knowledge components (Rivkin, 2001, develops this argument further, with the aid of simulations).

Social boundaries and template access

Access to the template depends on the distribution of social relations, which provide the channels through which valuable information flows (Hägerstrand, 1953; Coleman *et al.*, 1957). These social relations do not link actors at random. Rather, sociologists have consistently noted and demonstrated that networks concentrate within the boundaries of communities and organizations. Our study tests the salience of three types of social boundaries – organizational memberships, geographic regions, and technological communities – in structuring social networks, and concomitantly influencing the flow of knowledge.

Consider organizational boundaries first. A firm attempting to replicate and build on its own prior success has better access to its knowledge than would an outside imitator, both because fellow members of an organization share codes, specialized languages, and beliefs that facilitate high-fidelity transmission (Arrow, 1974) and because strong interpersonal ties and dense social networks inside a firm provide access to the template (Granovetter, 1985). As argued above, the value of this access peaks for transmitting knowledge of intermediate complexity:

> **Hypothesis 1:** The advantage in receiving and applying knowledge that members of the same firm have over members of different firms reaches its maximum for knowledge of intermediate interdependence.

In other words, the insider's advantage over the outsider has an inverted U-shaped relation to the interdependence of the knowledge.

Actors belonging to the same geographic unit (e.g., city, country or state) as the innovator also have superior access to the template. The geographic concentration of social relations reflects a variety of factors: the greater odds that individuals in close proximity encounter one another (Festinger *et al.*, 1950), the high costs of maintaining distant ties (Zipf, 1949; Boalt and Janson, 1957), and the prevalence of local cultures (Benedict, 1934). We therefore expect that actors physically close to a source of knowledge have better access to it:

> **Hypothesis 2:** A nearby knowledge recipient's advantage in receiving and applying knowledge over a distant recipient peaks for knowledge of intermediate interdependence.

To the extent that networks localize geographically, even within a firm, organizations find it difficult to diffuse knowledge beyond its point of origin. Within a firm, then, we expect simple knowledge to spread broadly and highly complex knowledge to remain isolated within a single team or department. Knowledge of moderate complexity, however, should spread within a firm to the edges of a facility or locale, but not to geographically distant installations:

> **Hypothesis 3:** Within a firm, a nearby knowledge recipient's advantage in receiving and applying knowledge over a distant recipient reaches its maximum for knowledge of intermediate interdependence.

An analogous argument applies to technological communities (also called communities of practice, defined in terms of cognitive proximity). Actors who work in the same technological domain as an inventor have superior access to the template. Universities, trade associations, professional societies, industry consortia, and work experience foster dense social connections within such technological communities. These communities also develop common knowledge and communal languages that can facilitate knowledge transfer. Membership within a common technological community thus engenders superior access to the template, which should have its greatest impact when the target knowledge displays moderate interdependence:

> **Hypothesis 4:** The advantage in receiving and applying knowledge that a member of a technological community has over a non-member of the community reaches its maximum with knowledge of intermediate interdependence.

EMPIRICAL CORROBORATION

To test these hypotheses, we analyzed prior art citations to all U.S. utility patents granted in May and June of 1990 (n = 17,264). The data came from the Micro Patent database and NBER public access data on patents (Hall, *et al.*, 2001). As in many previous studies, we view a prior art citation as evidence of knowledge diffusion. Our statistical approach involves estimating the likelihood that a focal patent receives a citation from a future patent as a function of several factors: the interdependence of the knowledge underlying the focal patent, the status of the citing patent's inventor as an insider or outsider on some dimension with respect to the focal patent, the interaction of interdependence and insider/outsider status, and a set of control variables. The results of the estimation allow us to examine how the likelihood of insider citation compares to the likelihood of outsider citation and, crucially, whether the gap between the two probabilities peaks when the focal patent embodies moderately interdependent knowledge.

Our unit of analysis is a patent dyad, one patent issued in May or June of 1990 and one issued later that may or may not cite the first. Hence our approach conceptually follows that of other studies of the likelihood of tie formation – in this case, the likelihood that a future patent builds on the knowledge embodied in one of our focal patents. Specifically, our analysis follows Sorenson and Stuart (2001) in adopting a case-control approach to analyzing the formation of ties. We begin by including all cases of future patents, from July 1990 to June 1996, that cite any of our 17,268 focal patents: 60,999 in total. Since these citations occur, the dependent variable for these cases takes a value of "1" to denote a realized citation. In addition, we pair each of the 17,268 focal patents with four future patents that do not cite it (but that could have), with the dependent variable set to zero. From this set of 130,071, we restrict our analysis to the dyads in which both inventors reside in the United States, leaving us with a set of 72,801 dyads.

Interdependence: For each dyad, we measure the complexity of the knowledge in the focal patent, *k*, by observing the historical difficulty of recombining the elements that constitute it (Fleming and Sorenson, 2001). Though the metric involves intensive calculation, the intuition behind it is simple: a technology whose components have, in the past, been mixed and matched readily with a wide variety of other components has exhibited few sensitive interdependencies and receives a low value of *k*. The measure takes the subclasses identified in a patent as proxies for the underlying components (see Fleming and Sorenson, 2004, for survey-based validation of the measure).

We compute *k* in two stages. Equation 1 details our calculation of the ease of recombination, or inverse of interdependence, for each sub-class *i* used in patent *j*. We first identify every use of sub-class *i* on patents from 1980 to 1990. The denominator is simply the tally of the number of patents with a classification in sub-class *i*. To compute the numerator, we count the number of different sub-classes appearing with sub-class *i* on previous patents. Hence, our measure increases as a particular sub-class combines with a wider variety of technologies, controlling for the total number of applications. This term captures the ease of combining a particular technology.

$$\text{Ease of recombination of subclass } i \equiv E_i = \frac{\text{Count of sub - classes previously combined with sub - class } i}{\text{Count of previous patents in sub - class } i} \quad (1)$$

To create our measure of interdependence for an entire patent, we invert the average of the ease of recombination scores for the sub-classes to which it belongs (equation 2).

$$\text{Interdependence of patent } j \equiv K_j = \frac{\text{Count of sub - classes on patent } j}{\sum_{j \in i} E_i} \quad (2)$$

Social boundaries: Three variables capture the insider/outsider status of the potential citing inventor with respect to the holder of a focal patent. The variables reflect membership in organizational, regional, and technical communities. *Same assignee* is set to one if two patents in a dyad share a common owner and is zero otherwise. *Geographic proximity* is equal to the natural log of the distance in miles between the first inventors listed on the two patents in a dyad multiplied by negative one (so that we expect larger values to increase the likelihood of citation). *Same class* is set to one if two patents belong to the same primary technological class – a proxy for shared membership in a community – and is zero otherwise. In all three cases, we test our hypotheses by interacting *k* and its square with the proxy for the density of social networks – whether due to firm boundaries, geographic proximity, or technological similarity. The benefits of social proximity should peak for inventions of moderate complexity.

The regressions also include several control variables. *Subclass overlap* is the number of subclasses that the two patents in the dyad have in common divided by the number of subclass memberships for the (potentially) citing patent. An *activity control* estimates the typical number of citations received by a patent in the same technological areas as the focal patent (see Fleming and Sorenson, 2001). *Recent technology* is the average reference number of the patents listed as prior art, a measure of the closeness of the patent to the technological frontier. We also include counts of the number of *backward patent citations* and *backward non-patent citations*, the *number of class* memberships, and the *number of subclass* memberships for the focal patent. We report robust standard errors and correct for potential bias in logistic regression of rare events (King and Zeng, 2001).

RESULTS

The results appear in Table 1. Model 1 tests hypothesis 1 by interacting k and k^2 with *same assignee*. As expected, membership within the same firm produces the greatest diffusion advantage over outsiders for knowledge of intermediate complexity, as evidenced by the positive coefficient on k x *same assignee* and the negative coefficient on k^2 x *same assignee*. The interactions between geographic distance and interdependence in model 2 tests hypothesis 2, again showing strong support. Model 3 tests hypothesis 3 by re-estimating model 2, but only for dyads where both patents belong to the same firm. In essence it asks: Does geography still matter for knowledge diffusion within firms? In support of hypothesis 3, the results reveal that even within firm boundaries, social networks influence the flow of knowledge, with the greatest disparity between local diffusion and distant diffusion arising for knowledge of moderate interdependence. Model 4 tests the salience of technological communities. Once again, the estimates show strong support; the impact of technological community membership on citation probability peaks for intermediate *k*. Model 5 includes all three measures of social proximity simultaneously and shows that each has an independent and significant effect when estimated together, in support of hypotheses 1, 2, and 4.

As expected, the value of superior access to the template reaches a maximum for knowledge of moderate interdependence, regardless of whether the superior access comes from organizational membership, geographic proximity, or technological community membership. In addition to being significant, the effects have substantial economic import. For simple or highly complex knowledge, the insider has no greater likelihood than the outsider of attaining and building on the knowledge in a focal patent. For knowledge of moderate complexity at the gap-maximizing levels of *k*, a firm insider is 218% more likely than an outsider to transfer knowledge effectively; an inventor located in the same zip code is 160% more likely to absorb the knowledge in a region than one at the average distance (~665 miles); and a technological insider is 238% as likely as a technological outsider to build on knowledge in the class.

DISCUSSION

Our analyses considered the impact of superior access to some original knowledge on the likelihood of diffusion as a function of knowledge complexity, using three indicators of social proximity. All knowledge recipients, near and far, compete on equal footing when assimilating simple knowledge. Highly complex knowledge, on the other hand, equally resists diffusion to both classes of would-be recipients. For knowledge whose ingredients display a moderate degree of interdependence, however, superior but imperfect access to the template translates into better knowledge reproduction. Thus in our patent data, the largest gap between the ability of a close recipient to build on prior knowledge relative to the ability of a distant recipient arises when the cited patent involves moderate interdependence.

Our findings speak to the question, when does inequality of knowledge arise across social borders? One might initially suspect that highly complex knowledge, the most difficult to reproduce, would create the greatest inequality. But this intuition ignores the fact that inequality in its sharpest form requires *some* diffusion: to create the most inequity across social boundaries, knowledge must creep up to the edge of the thick patch of connections in which it originated but not beyond. This phenomenon, we have argued, most likely occurs for moderately complex knowledge. Thus, for example, one might expect industries based on moderately complex knowledge to display especially wide intra-industry dispersion in long-run financial returns.

Our argument may also shed light on a conundrum of the literature on economic geography. Explanations for agglomeration based on information spillovers assume that membership in a local community allows firms to benefit from the knowledge developed by other firms in the region, but that firms outside the region are excluded from these benefits (e.g., Marshall, 1890; Arrow, 1962). What type of knowledge would have such a characteristic? The literature to date has focused on 'tacit' knowledge, but typically uses the term simply to refer to uncodified (as opposed to uncodifiable) knowledge (an endogenous outcome of firms' decisions to invest in codification; Brökel, 2005). Our results, building on Kauffman's NK model, point to a different (presumably more exogenous) dimension: informational complexity. Industries that rely on moderately complex knowledge might be especially likely to display geographic agglomeration (for empirical corroboration, see Sorenson, 2004).

Our empirical results come from patent data alone, but the basic logic of our hypotheses applies to knowledge in general, not just the knowledge underlying inventions (cf. Wolter, 2006, for a model based on interdependence in production). Hence, future research might usefully examine these dynamics across a wide range of applications – including organizational learning, the diffusion of management practices, knowledge management, and the sustainability of knowledge-based competitive advantage.

REFERENCES

Arrow, Kenneth J. 1962. "The economic implications of learning by doing." *Review of Economic Studies*, 29: 155-173

Arrow, Kenneth J. 1974. *The Limits of Organization.* New York: Norton.

Benedict, Ruth. 1934. *Patterns of Culture.* New York: Houghton-Mifflin.

Boalt, Gunnar, and C.G. Janson. 1957. "Distance and social relations." *Acta Sociologica* 2: 73-97.

Brökel, Tom. 2005. "The spatial dimension of knowledge transfer and its economic implications." Working paper, Max Planck Institute of Economics (Jena).

Coleman, James S., Elihu Katz, and Herbert Mendel. 1957. "The diffusion of an innovation among physicians." *Sociometry* 20: 253-270.

Davis, Gerald F., and Henrich R. Greve. 1997. "Corporate elite networks and governance changes in the 1980s." *American Journal of Sociology* 103: 1-37.

Festinger, Leon, Stanley Schacter, and Kurt W. Back. 1950. *Social Pressure in Informal Groups*. New York: Harper.

Fleming, Lee, and Olav Sorenson. 2001. "Technology as a complex adaptive system: Evidence from patent data." *Research Policy*, 30: 1019-1039.

Fleming, Lee, and Olav Sorenson. 2004. "Science as a map in technological search." *Strategic Management Journal*, 25: 909-928.

Frenken, Koen, and Alessandro Nuvolari. 2004. "The early development of the steam engine: An evolutionary interpretation using complexity theory." *Industrial and Corporate Change*, 13: 419-450.

Granovetter, Mark S. 1985. "Economic action and social structure: The problem of embeddedness." *American Journal of Sociology* 91: 481-510.

Hägerstrand, Torsten. [1953] 1967. *Innovation Diffusion as a Spatial Process*. Chicago: University of Chicago.

Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg. 2001. "The NBER patent citations data file: Lessons, insights and methodological tools." National Bureau of Economic Research Working Paper No. 8498.

Kauffman, Stuart A. 1993. *The Origins of Order*. New York: Oxford University.

King, Gary, and Langche Zeng. 2001. "Logistic regression in rare events data." *Political Analysis*, 9: 137-163.

Lippman, Steve and Richard Rumelt. 1982. "Uncertain imitability: An analysis of interfirm differences in efficiency under competition." *Bell Journal of Economics* 13: 418-438.

Marshall, Alfred. 1890. *Principles of Economics*. London: MacMillan.

Nelson, Richard R., and Sidney G. Winter. 1982. *An Evolutionary Theory of Economic Change.* Cambridge, MA: Belknap.

Polanyi, Michael. 1966. *The Tacit Dimension*. New York: Anchor Day.

Reed, Richard, and Robert J. DeFillippi. 1990. "Causal ambiguity, barriers to imitation, and sustainable competitive advantage." *Academy of Management Review* 15: 88-102.

Rivkin, Jan W. 2001. "Reproducing knowledge: Replication without imitation at moderate complexity." *Organization Science* 12: 274-293.

Schumpeter, Joseph. 1939. *Business Cycles*. New York: McGraw-Hill.

Simon, Herbert A. 1962. "The architecture of complexity." *Proceedings of the American Philosophical Association* 106: 467-482.

Sorenson, Olav. 2004. "Social networks, informational complexity and industrial geography." Pp. 79-96 in D. Fornahl, C. Zellner and D. Audretsch, *The Role of Labour Mobility and Informal Networks for Knowledge Transfer*. Berlin: Springer-Verlag.

Sorenson, Olav, Jan W. Rivkin, and Lee Fleming. 2004. "Complexity, networks and knowledge flow." HBS Working Paper #04-027

Sorenson, Olav, and Toby E. Stuart. 2001. "Syndication networks and the spatial diffusion of venture capital investments." *American Journal of Sociology* 106: 1546-1588.

Winter, Sidney G. 1995. "Four Rs of profitability: Rents, resources, routines, and replication." In C. Montgomery, ed. *Resource-based and Evolutionary Theories of the Firm: Towards a Synthesis.* Boston: Kluwer.

Wolter, Kerstin. 2006. "Divide and conquer? The role of governance for the adaptability of industrial districts." *Advances in Complex Systems*, forthcoming

Zander, Udo and Bruce Kogut. 1995. "Knowledge and the speed of transfer and imitation of organizational capabilities: An empirical test." *Organization Science* 6: 76-92.

Zipf, George K. 1949. *Human Behavior and the Principle of Least Effort.* Reading, MA: Addison-Wesley.

Table 1: Rare events logit models of the likelihood of a focal patent receiving a citation from a future patent

	Model 1	Model 2	Model 3 Only same assignee	Model 4	Model 5
k	1.687••• (.302)	1.526••• (.307)	4.821••• (.359)	1.444••• (.321)	1.070•• (.362)
k^2	-.892••• (.082)	-.793••• (.074)	-4.208••• (.117)	-.704••• (.098)	-.359•• (.107)
k x same assignee	2.969••• (.515)				6.231••• (.555)
k^2 x same assignee	-3.420••• (.203)				-9.851••• (.273)
k x - ln (dist)		.835••• (.131)	6.047••• (.213)		.885••• (.139)
k^2 x - ln (dist)		-.794••• (.044)	-4.566••• (.074)		-1.025•• (.066)
k x same class				3.019•• (1.146)	5.733••• (1.032)
k^2 x same class				-1.396••• (.363)	-5.409••• (.330)
Same assignee	.343 (.280)	.389 (.276)		.432 (.281)	.172 (.278)
- Ln (dist)	.499••• (.031)	.428••• (.031)	.500••• (.066)	.499••• (.030)	.354••• (.029)
Same class	3.800••• (.306)	3.663••• (.307)	1.878••• (.394)	3.837••• (.302)	3.448••• (.299)
Subclass overlap	4.230••• (.316)	4.190••• (.317)	3.767••• (.349)	4.114••• (.316)	4.150••• (.314)
Activity control	.393 (.287)	.389 (.287)	-.746•• (.248)	.477 (.388)	.466 (.289)
Recent technology	.122 (.171)	.195 (.170)	.010 (.309)	.096 (.165)	-.024 (.151)
Backward patent citations	.002 (.013)	.013 (.013)	.025•• (.008)	-.001 (.014)	-.002 (.014)
Backward non-patent citations	.018•• (.006)	.014• (.006)	-.126••• (.037)	.019•• (.005)	.011• (.005)
Number of classes	.070 (.140)	.054 (.140)	.456 (.249)	.041 (.138)	.052 (.137)
Number of subclasses	-.016 (.045)	-.021 (.045)	.170••• (.048)	.001 (.044)	.010 (.044)
Constant	-9.224••• (.714)	-9.953••• (.703)	-7.148••• (1.208)	-9.206••• (.684)	-9.162••• (.675)
Log-likelihood	-22,262.4	-22,261.4	-2,294.1	-22,255,9	-22,251.1
N	72,801	72,801	6,497	72,801	72,801

• $p < .05$; •• $p < .01$; ••• $p < .001$. Model 3 includes only those dyads for which *same assignee* = 1.

123	w_1	w_2	w_3	$W = \frac{\sum_{i=1}^{N} w_i}{N}$
000	.8	.6	.2	.53
001	.8	.6	.9	.77
010	.8	.7	.2	.57
011	.8	.7	.9	.80
100	.4	.6	.2	.40
101	.4	.6	.9	.63
110	.4	.7	.2	.43
111	.4	.7	.9	.67

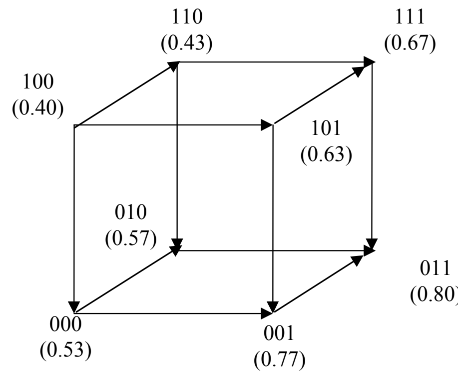

Figure 1: Landscape without interdependence (N=3, K=0). This relatively correlated landscape has only one minimum and one maximum, 100 (0.40) and 011 (0.80), respectively. The component fitness contributions come from a uniform [0,1] distribution.

123	w_1	w_2	w_3	$W = \frac{\sum_{i=1}^{N} w_i}{N}$
000	.5	.8	.6	.63
001	.6	.2	.1	.30
010	.2	.7	.3	.40
011	.8	.6	.5	.47
100	.4	.5	.2	.37
101	.9	.8	.9	.87
110	.7	.4	.1	.40
111	.9	.7	.3	.63

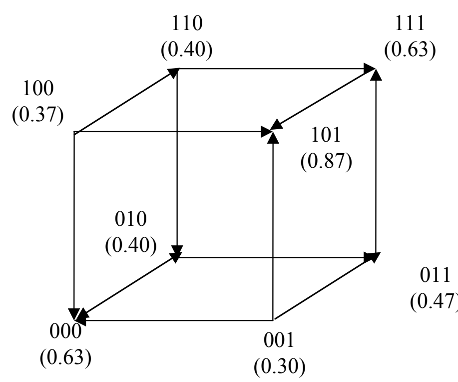

Figure 2: Landscape with maximal interdependence (N=3, K=2). This relatively uncorrelated landscape has multiple local minima, 001(0.30) and 100 (0.37), and maxima, 000 (0.63) and 101 (0.87).