Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Financial Economics 80 (2006) 5–33

JOURNAL OF Financial ECONOMICS

www.elsevier.com/locate/jfec

Do alliances promote knowledge flows?☆

Benjamin Gomes-Casseres[a,*], John Hagedoorn[b], Adam B. Jaffe[a]

[a]Brandeis University, Waltham, MA 02454, USA
[b]University of Maastricht, The Netherlands

Received 2 October 2003; received in revised form 19 July 2004; accepted 30 August 2004
Available online 28 September 2005

Abstract

We explore the role of interfirm alliances as a mechanism for sharing technological knowledge. We argue that knowledge flows between alliance partners will be greater than flows between pairs of nonallied firms, and less than flows between units within single firms. Using patent citations as a proxy for knowledge flows, we find results that are consistent with these expectations. We then explore how firm characteristics affect knowledge flows within alliances and find positive effects due to technological, geographic, and business similarities between partners. We use alliance data from MERIT, patent data from the USPTO, and firm data from Compustat.

JEL Classification: Firm organization and market structure: Markets vs. hierarchies; Vertical integration (L22)

Keywords: Alliances; Technology transfer; Patents; Knowledge; Multinational companies

1. Introduction

We explore the role of interfirm alliances as a mechanism for sharing technological knowledge. Because firms often form alliances to promote technology sharing, we compare the extent and nature of knowledge flows between alliance partners to analogous flows among pairs of firms that are not allied. To provide greater context for this comparison,

☆We thank Megan MacGarvie and Mehmet Beceren for research assistance, and attendees at the NBER 2002 Alliance Conference, the Academy of International Business 2002 and 2003 Annual Meetings, and two anonymous referees for their suggestions.
*Corresponding author. Tel.: +1 781 736 2264; fax: +1 781 736 2269.
E-mail address: bgc@brandeis.edu (B. Gomes-Casseres).

0304-405X/$ - see front matter
doi:10.1016/j.jfineco.2004.08.011

and because it is interesting in its own right, we also compare these interfirm knowledge flows with flows within multinational companies (MNCs).

While some knowledge flow across organizations (or units of a single organization) is no doubt accidental or involuntary, intentional flows will be greater when they are in the interest of both parties. Knowledge sharing should thus be facilitated when different units have similar interests, and/or when contractual mechanisms reduce the scope for opportunism in the exchange. In general, we expect that knowledge flows will be the smallest between firms that have only arm's length relationships with each other, because much of the flows will be involuntary spillovers. At the other extreme, an MNC will often attempt to align the interests of its distinct units so as to maximize the incentive and ability of each to share knowledge within the firm. We view interfirm alliances as an intermediate environment, in which interests are only partially aligned and opportunistic behavior only partially regulated. Hence, the extent of knowledge flow between alliance members should be intermediate between that observed in the pure spillover environment of nonallied firms and that realized in the integrated environment of different units of a single firm.

We use the well-known Cooperative Agreements and Technology Indicators (CATI) database developed by the Maastricht Economic Research Institute in Technology (MERIT) to identify pairs of firms that share one or more alliances with each other. We develop various measures of the nature or intensity of the alliance relationship between pairs of firms, such as the number of alliances they have with each other, the ages of those alliances, and whether or not the alliance agreements explicitly incorporate joint research and development (R&D) or equity investments. In order to examine knowledge flows between nonallied firms, we construct control pairs that consist of two firms that do not have any alliances in common in the database. For these controls, we use only firms that are in the CATI database, and which thus come from the same underlying universe of organizations as the firms in the allied pairs.

We match the firms in the CATI database to the NBER Patent Citations Data File (Jaffe and Trajtenberg, 2002) and then use patent citations as a proxy for the flow of technological knowledge between firms. Jaffe and Trajtenberg (2002) describe the substantial literature that uses patent citations as a proxy for knowledge flows. As will be made more precise below, we look at the number of citations to patents of Firm Y that are contained in the patents applied for in each year by Firm X. Controlling for the total number of citations made in that year by Firm X's patents, and the total stock of potentially citable patents of Firm Y, we interpret any variable that increases this citation flow as being associated with an increased flow of technological knowledge from Firm Y (the cited firm) to Firm X (the citing firm).

The relationship of patent citations to unobserved flows of technological knowledge can be thought of by analogy to citations in academic articles, though there are important limitations to this analogy. Like bibliographic citations, patent citations are supposed to indicate previous work on which the current work builds or relies, or which embodies results that are related to those of the current work. For the purpose at hand, patent citations have the advantage that they perform a legal function related to the validity of the patent and the technology to which it applies, so that they are not contaminated by unnecessary citations to friends, colleagues, or famous people. However, they have the drawback that a significant number of citations are typically added by the patent examiner, and hence may represent previous work of which the inventor was unaware. Survey research reported in Chapter 12 of Jaffe and Trajtenberg (2002) indicates that at most half of the citations in any given patent are to previous work that the inventor either knew or used. In addition, not all inventions are patented, so there are surely flows of knowledge between firms that do not show up in a patent citation. We say that patent citations are a noisy proxy for knowledge flow. It is likely, however, that the absent or extraneous citations are largely noise. Moreover, recent research suggests that there may be some systematic patterns in extraneous citations, e.g., there appears to be more such noise in some technological fields than in others (Sampat, 2004; Alcacer and Gettleman, 2004). It is unlikely, however, that these patterns would be related to the existence or absence of the intercorporate relationships studied here, hence, aggregate flows of the citations from a relatively large number of patents can be expected to contain a meaningful signal about knowledge flow that can be used to draw inferences regarding the phenomena of interest here.

Our primary empirical results confirm the expected ordering of the magnitude of knowledge flows, being greatest within firms, less for allied firms, and much less for the control pairs of nonallied firms. The increase in citations associated with alliance formation is also correlated with the intensity of the alliance relationship, being greatest for firms with alliances of extended duration, multiple alliances, and alliances with equity or joint R&D components.

To gain insight into the circumstances under which alliances between firms seem to have greater or lesser effects, we explore a number of aspects of the technological, organizational, and geographic environments of the firm pairs. We find that the knowledge flows are greatest when the firms are close to each other along several dimensions: The alliance effect is greatest for technologically similar firms, firms in the same geographic region, and firms in the same industry. We also find that large firms (as measured by sales) appear to share knowledge within alliances more than smaller firms; even after controlling for the scale of innovative activity as measured by patent variables, the effect of alliance formation on citation frequencies increases with the size of both the citing and potentially cited firms. Finally, R&D-intensive firms seem to benefit more from alliance membership; the effect of alliance participation on citations is increasing in the R&D/sales ratio of the citing firm.

These findings have potentially important implications for our understanding of the economics of the firm. Various studies suggest that alliances can enhance the value of the firm, especially alliances that involve the transfer or pooling of technology (e.g., Chan et al., 1997). Our study provides a missing link between these studies and other research that shows that a firm's technology capability (as represented by its patents) influences its value to investors (Hirschey and Richardson, 2004). If alliances indeed promote technology flows, then the technological capability of a firm can be enhanced by that of its allies, and as a result, alliances can add value to the firm. Thus, our findings help explain both why alliances sometimes complement the internal financing of R&D (Lerner et al., 2003), and how alliances enhance the value of equity ownership ties between firms (Allen and Phillips, 2000).

The organization of the paper is as follows. Section 2 describes our conceptual approach in more detail, and relates it to key elements of the existing literature. Section 3 describes the data sources, the compilation of the samples used in the analyses, and the construction of key variables. Section 4 provides results using varying techniques and specifications. Section 5 offers concluding observations and suggestions for future research.

2. Framework: Firms, alliances, and knowledge flows

We define an alliance as any organizational structure used to govern an incomplete contract between separate firms, such that each partner has limited control (Gomes-Casseres, 1996). Alliances take different forms, from equity joint ventures to nonequity contractual arrangements. Regardless of form, these arrangements are in some way open-ended and contain gaps typical of incomplete contracts. As a result, to deal with unforeseen contingencies, the partners need to make decisions jointly, although there is no automatic convergence in their interests. The alliance structure thus enables the parties to coordinate joint work and align interests better than arm's length contracts. In other words, the alliance structure can be seen as an alternative to vertical integration for dealing with problems created by the incompleteness of contracts (Hart and Moore, 1990). This approach is consistent with recent papers on alliances as relational contracts (Garvey, 1995; Baker et al., 2002).

2.1. Firm boundaries and alliances

Traditionally, the boundary of the firm has been defined by ownership of assets with a strict distinction being made between transactions inside and outside the firm (Coase, 1937; Williamson, 1975). Even so, early theorists knew that this distinction was artificial; in a footnote, Coase notes in 1937: "[I]t is impossible to draw a hard and fast line which determines whether there is a firm or not. There may be more or less direction" (pp. 386–405).

With the rising popularity of alliances as an organizational strategy, theorists have begun to see alliances as a middle ground between the firm and the market, "blurring" the boundary of the firm (Macaulay, 1963; Richardson, 1972; Williamson, 1979; Garvey, 1995; Baker et al., 2002). Much of the theoretical and empirical research on alliances in the past two decades explores how alliances combine attributes of firms and markets. One of the areas in which alliances can potentially have an advantage over markets is in the pooling and transfer of technological capabilities among separate firms. This paper is an attempt to test this conjecture using statistical methods on a broad sample of firms.

2.2. Alliances and knowledge flows

This paper focuses on two activities typically thought to be particularly subject to problems associated with incomplete contracts, namely, technology transfer and cooperation in the development of new technology. Because of difficulties in monitoring inputs and outputs, in negotiating exchanges of value under conditions of uncertainty and asymmetric information, and in enforcing contracts in relation to intangible assets, these activities are typically conducted more efficiently, and at lower transaction costs within an integrated firm than between unrelated firms (Teece, 1986). Because alliances facilitate the governance of incomplete contracts, we expect that when firms use alliances to transfer technology and cooperate in technological development, these transaction costs would be lower than for unrelated firms (Kogut, 1988; Hamel, 1991), though they still may be higher than if the firms were fully integrated.

In sum, we expect that more technology transfer and more cooperative technological development would take place between allied firms than within a comparable pair of nonallied firms, but that the extent of these activities among allied firms would still be less than between different units of an integrated firm. Thus, we expect to find an increase in a firm's patent citation intensity as we move from patents of an unaffiliated firm, to patents of an allied firm, to the firm's own patents.

Not all alliances lead to knowledge flows and not all are intended to do so. There is much disagreement in the alliance literature about whether knowledge transfer in alliances is the rule or the exception. One side in this debate claims that alliances are all about learning, leading to the possibility that partners engage in learning races with each other or appropriate knowledge from each other opportunistically (Mankin and Reich, 1986; Khanna et al., 1998). The other side argues that precisely because of the competitive tensions that learning entails, firms in most alliances intend to co-specialize. That is, rather than the firms seeking to absorb the knowledge of the partner, each will focus on deepening its own knowledge in a way that complements the knowledge of its partner (Hennart, 2000).

In an ideal world, we would focus our study on alliances intended by their members to facilitate the exchange of technological knowledge, and we would ignore those in which no such attempt is made. However, the data offer only a crude indication as to the purpose of each alliance. We therefore use this information, together with empirical strategies discussed below, to try to focus our examination on those alliances that we intended to facilitate knowledge flow. Of course, to the extent that the data continue to be contaminated by alliances motivated by other purposes, our measures of alliance impact will understate the magnitude of that impact for the subset intended to facilitate knowledge flow.

2.3. *Organization of the multinational company*

While we view the integrated firm as a benchmark against which to compare knowledge flows in alliances, a wide range of characteristics of the integrated firm could affect internal knowledge flows. Because our data contain mostly large MNCs, it is useful to recognize the role of technology flows within MNCs. As is well known, an MNC must have a competitive advantage over local firms in order to overcome the liability of foreignness. This competitive advantage often consists of advanced and proprietary knowledge that is costly to transfer to the foreign location through pure market transactions. Otherwise, we would not observe an MNC, but rather international trade, cross-border licensing, and other forms of international business. Technology flows thus play a central role in the very existence of MNCs (see the review in Caves, 1996). A recent study by Singh (2003) that also uses patent citations as a proxy for technology flows finds that MNCs are not disadvantaged in accessing technology from host-country firms, and confirms that cross-border technology flows within the MNC are much stronger than flows between firms, even for firms located in the same country.

However, several managerial studies of the MNC emphasize that these global firms cannot be seen as a single entity with a single center of control—foreign subsidiaries have power that is distinct from and counterbalanced by the power of headquarters. This local power is based on the information, local connections, tacit knowledge, and other intangibles that are important to a firm. As a result, the MNC is best seen as a network of entities that are owned by a single parent, but not controlled unilaterally by the parent headquarters (see, e.g., Bartlett and Ghoshal, 1989).

This view of the MNC as a collection of distinct units with common ownership but not necessarily fully aligned incentives leads us to conceptualize the difference between an alliance and a firm as a matter of degree rather than of kind. The different units of an MNC may or may not always work to bring about efficient knowledge exchange, but it is reasonable to expect that they do a better job at this kind of coordination than organizational units linked only by an alliance agreement. On the other hand, some observers suggest cross-border alliances are ideal for tapping into multiple sources of knowledge in different countries. Branstetter (2002), for example, examines the role of R&D alliances and foreign direct investment (FDI) in facilitating knowledge flow between the U.S. and Japan. Still, there is little systematic evidence of the relative abilities of MNCs and alliances to transfer knowledge across countries.

The main implication of these ideas for our study is that our self-citation pairs represent much more than a simple identification of citing and cited firms. Rather, these pairs usually represent large organizations that span multiple countries and conduct technological development in multiple locations. Unfortunately, our patent information is not sufficiently fine-grained to disentangle relevant characteristics of the firms; indeed, our primary measure of geographic location of the entity is the home base of the firm. Simple statistics of the source of invention of the firms in our sample confirms that the vast majority of inventions are made in the home country of the firm. While many might debate what exactly is the home country of a firm such as IBM, the fact remains that even for this most global of firms, 60% of its patents originate in the United States. A related implication is that firms from different regions may behave differently, both in their self-citations and in their alliance citations.

2.4. Geographic boundaries and technology diffusion

Aside from drawing on research about alliances and MNCs, we also use ideas from the literature on technology diffusion across geographic space. A number of studies find that technology flows to nearby locations are greater and faster than flows to more distant locations (Jaffe et al., 1993; Jaffe and Trajtenberg, 1996). A few recent studies of technology diffusion also evaluate the role of organizational linkages. These studies usually argue that organizational linkages may reduce communication costs and thus promote the flow of knowledge across locations (Mowery et al., 1996; Branstetter, 2002).

From another angle, the alliance literature has long been concerned with the effect of geographic differences between partners in an alliance. Much of the work in business alliances stems from an interest in international alliances, the arena in which joint ventures and other inter-firm arrangements first gained popularity (Stopford and Wells, 1972; Contractor and Lorange, 1988). This literature sees alliances as a way to facilitate cross-border technology transfer—a claim that has long influenced host-government foreign investment and industrial policy strategies.

2.5. An integrated approach

Notwithstanding the fact that each of the streams of research described so far is well established, they are seldom combined in empirical work. To our knowledge, there are no studies that test the interaction of the organizational and geographic variables just described. Fig. 1 helps one visualize the range of interactions between these variables.

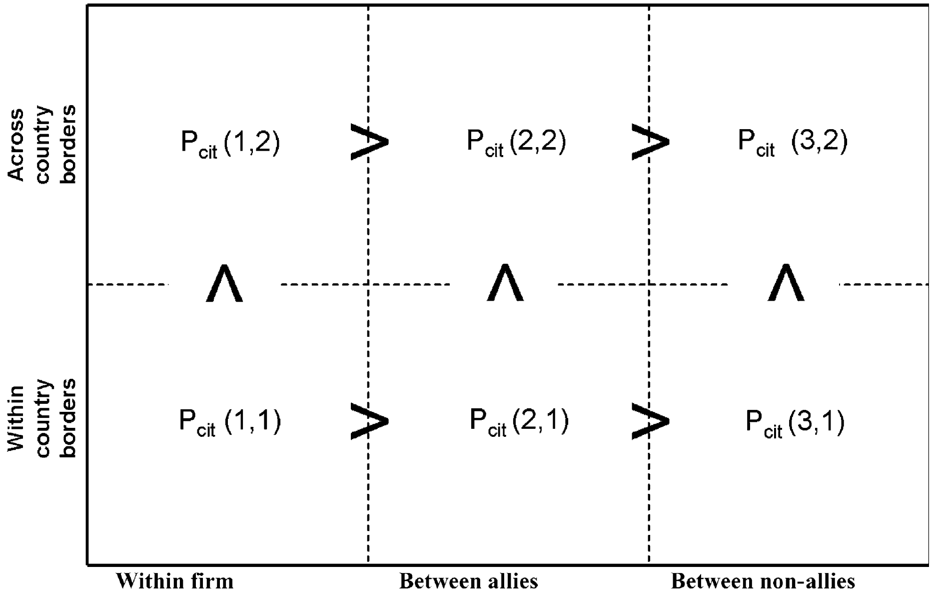

Fig. 1. Hypothesized intensity of technology flows across organizational and geographic boundaries. Inequality signs indicate expected ordinal rankings of probability of citation of pairs with different characteristics; numbers in parentheses refer to row and column of each cell. For example, in the lower-left cell, Pcit (1,1) represents the probability of citation in pairs in that cell, i.e. pairs within one firm, and within one country. The inequality sign immediately to the right indicates that this probability is expected to be greater than that for pairs that are allies, but still within one country [Pcit (2,1)].

The figure shows our two basic hypotheses. First, self-citations will be more likely than alliance citations, which in turn will be more likely than citations among unrelated pairs. Second, within-country citations will be more likely than cross-country citations, regardless of organizational form. This leads to our expectations about the ordinal ranking of citation probabilities across geographies and organizational firms as shown in Fig. 1.

What is noteworthy and potentially of importance to managers is that the ranking between diagonal pairs of cells in Fig. 1 is not determined by the ordinal rankings shown. For example, will a firm prefer within-country, cross-firm citations over cross-country, within-firm citations, or, what is the best avenue to acquire or share a given technology? As a concrete example, consider the choice that an Intel unit in the United States might have if it needs a technology that is available from a local partner as well as from its own foreign subsidiary. Will it be more costly (in organizational terms) to obtain that technology from the local partner (through an alliance) or from the Intel subsidiary (across geographic borders)? The answer is likely to be an empirical question, and one that has not been addressed so far.

A broad way to conceptualize these ideas is to hypothesize that any dimension of distance that increases communication costs will reduce or delay technology diffusion, that is, firms that are geographically, technologically, organizationally, economically, or otherwise more distant will tend to communicate less, all else equal. The formation of an alliance and the intensification of an alliance relationship between two firms can be thought of as reducing the organizational distance between the firms. Of course, this is simply a restatement of the basic hypotheses laid out above. The interesting question is whether—and in what direction—these distance effects interact with one another. One might imagine that alliance formation would be particularly useful to overcome handicaps deriving from distance along other dimensions. If the purpose of alliances is to gain efficient access to knowledge that would otherwise be hard to obtain because of geographic or technological distance, then we would expect the augmentation in communication associated with alliance formation to be greatest for firms that are distant along these dimensions. We find, however, the opposite—the benefit of alliance formation is greatest for firms that are already proximate along other dimensions.

3. Data, samples, and variables

We use data from three major sources: (1) The CATI database, which provides the members and certain characteristics of alliances; (2) the NBER patent database, which offers the patents and patent citations of firms; and, (3) the Standard and Poors Compustat database, which contains R&D, sales, and other economic information about firms. We analyze citations on an annual basis from 1975 to 1999. Our three data sources do not share a common firm index, so combining these three sources of information requires matching the firms in the different databases on the basis of the firm names as recorded in each. Our primary sample consists of all of the firms and alliance pairs that we are able to match between the CATI and patent databases, as described below.

To reduce complexity in the merging procedure as well as in the interpretation of results, we use only alliance and patent data from information technology sectors, which include computers, communications, semiconductors, and related materials—a previous industry study suggests that alliances and MNCs played important roles in the international transfer of knowledge in these fields during the period under study (Gomes-Casseres, 1993).

3.1. Linking patent and alliance data

The CATI and NBER databases are combined by matching the firm names. We begin with 2,637 firms in the alliance data set and 175,116 patenting organizations (“assignees”). The names of alliance members are matched to names of USPTO assignees using both a computer algorithm and case-by-case reviews. First, we use an algorithm to strip the text strings of punctuation and capitalization and to identify word-for-word matches between the names found in the two datasets. This algorithm succeeds in matching roughly a quarter of the names. We then modify the algorithm so that it matches the first three words of the names, then the first two, and finally the first word of the names. We then examine each of these matches manually, which results in matching about half of the names. Finally, we match the remaining names by manually searching for key words. In the end, we succeed in matching 1,785 alliance names to patent assignees (68% of allied entities). These firms represent 4,839, or 77%, of the 6,290 alliance relationships. The matched firms are involved in a disproportionate share of alliance relationships because larger firms are more likely to patent (and therefore be found in the patent database), and also tend, on average, to be involved in more alliance relationships. In some cases, there is more than one name in either the alliance data or the patent data that appears to match a single name in the other data set. In these cases, we add the patent and citation counts or alliance memberships of the multiple names into an aggregated entity. For example, several names from the alliance data set are combined to form an aggregated entity called "GE" that is then matched to patent data. This means aggregating the "GE" entity names appearing in the original alliance data; General Electric Company, GE/Calma, GE/Storno, as well as the USPTO assignees; General Electric Company, GE Chemicals, Inc., GE Company, General Electric Canada Inc., General Electric CBR SA, General Electric CGE SA, General Electric CGR SA, and Calma Company. The inclusion of the USPTO assignee "Calma Company" reflects the fact that while there is no assignee called "GE/Calma," there are USPTO assignees that match "GE" and "Calma" separately. In such cases where there is no exact match for what appears to be a combination of names matching assignees, the amalgamated name is matched to both patent assignees.

Once this procedure is complete, we drop firms that do not have at least one patent in an information technology (IT) field. Patents in these fields fall into classes corresponding to technologies used in Communications, Computer Hardware & Software, Computer Peripherals, Information Storage, and Semiconductor Devices (subcategories 21–24 and 46 in Hall et al., 2002). Once firms without patents in these classes are eliminated, we are left with 975 firms.

It is natural to wonder to what extent our sample of IT alliances and their member firms are representative of the universe of alliances. First, to the extent that the universe contains many alliances that are not motivated by technology transfer, we do not wish to represent that universe. However, it is reasonable to ask whether the IT alliance sample is representative of high-technology R&D alliances in other sectors such as pharmaceuticals, aerospace, or advanced materials. The labor-intensive nature of the matching process (combined with the finite nature of resources) precludes the construction of a comprehensive match between the firms in the CATI database and the patent database. Without such a match, we cannot really compare the firms and alliances in the IT sector to those in other sectors. Nonetheless, the IT alliances represent 40% to 50% of all alliances in the database over the period, so they are clearly an important fraction of the overall story. We can also compare IT alliances and their members to all high-tech alliances in the database. There are not large differences between the IT alliances that we use and the other high-technology alliances in the database with respect to the average number of firms per alliance (2.2 for both groups), the number of alliances in which the average firm participates (3.1 for firms in IT alliances; 2.8 for other high-tech firms), the fraction involving equity (20% for IT; 23% for other high-tech), or the fraction of alliances involving a U.S. company (74% for IT; 66% for other high-tech). While this does not confirm, of course, that the IT alliances behave in the same way as other alliances, it does suggest that they are a large, and apparently representative, component of the overall story.

3.2. *Panel data on firm pairs over time*

The 975 firms in the matched alliance/patent database are key in the creation of a database in which each observation is a unique firm pair. Because we are interested in the direction of technology flow, each alliance relationship generates two directional pairs in the database, e.g., IBM-Apple with IBM as citing firm and Apple as cited firm, and Apple-IBM, with the reverse. To examine a firm's citations to its own patents, we create a "self-pair" for each firm (e.g., IBM-IBM). Finally, we generate a random sample of nonallied pairs from the same set of firms, drawn from pairs that have no alliances in any year; these control or nonallied pairs are used to measure knowledge flows through market transactions. The number of control pairs is equal to the number of allied pairs.

Our final sample contains over 7,000 pairs of citing and cited companies that are one of three types: Nonallied pairs (about 3,000), allied pairs (about 3,000), and self-pairs (975). From these 7,000 pairs, we construct a panel data set consisting of annual observations for the period 1975 to 1999, with alliance formation, citation, and other variables defined by year. The total sample includes 160,901 pair-year observations, but this set includes a large number of pairs for which there are no citations. The panel is unbalanced, because in some years either the citing or cited firm does not have any patents, thus there could not possibly be any citations between them.

There are both conceptual and practical difficulties with analysis based on this full sample, given approximately three-quarters of the observations have zero citations. First, as discussed above, we want to limit ourselves to those alliances that intended to foster knowledge flow; a sample in which so many of the observations indicate no flow is problematic for this purpose. Second, as mentioned above, patent citations are at best a noisy indicator of knowledge flow. Extracting meaning from differences between zero and small numbers of citations is therefore difficult. (Indeed, preliminary analysis using both probit and negative binomial models to examine the zero-rich complete dataset yields unstable results.)

The rest of this paper therefore uses the main analysis sample that contains 40,898 observations for which there is at least one citation between the firms in the pair, and the results in this paper are to be interpreted as estimates of the effects of the independent variables on knowledge flows, conditional on there being a minimal level of flow. This formulation also facilitates construction of the models in the familiar log-log form, in which effects of the independent variables are multiplicative and the coefficients can be interpreted as elasticities. The means and standard deviations of the most important variables are shown in Table 1; definitions of all variables are summarized in the Appendix.

The NBER patent database contains identifying information that can be used to match a subset of the patenting firms to the Compustat data file. This match is based on the ownership of each of the patenting entities in a pair according to the Directory of Corporate Affiliation in 1989 (Jaffe and Trajtenberg, 2002). Because Compustat contains mostly U.S. firms and the CATI database contains firms from all countries (including many small firms), the merged subsample is considerably smaller than the main sample. Of the 829 firms used in the analysis sample, 128 are successfully matched to Compustat data (15%). Of the 7,583 allied and control pairs in the analysis sample, 1,160 consist of firms that are matched to Compustat. This set of successfully matched pairs produces a panel data set over time that consists of 8,917 pair-year observations (22% of the pair-year observations in the main panel set).

3.3. Variables: patent citations and related measures

We use a number of variables derived from the patent data. The main independent variable is the number of citations from citing to cited firm in a given year. An important control variable based on the patent data is the total number of citations made by the citing firm in a given year; this variable controls for the overall size of the citing activity of

Table 1
Descriptive statistics for citation pairs, by relationship between paired firms

The main analysis sample, shown in Panel A below, consists of 8,077 directional pairs of firms in which we measure the number of citations in each year from the citing firm to the cited firm. Of these pairs, 5,116 are firms that have no alliance with each other in the sample, 2,467 are firms that have at least one alliance, and 494 are pairs such that the firm cites itself. The main panel set contains a total of 40,898 pair-year observations. A second sample is constructed by merging Compustat data. This sample had 1,160 directional pairs, as shown in Panel B below. Means and standard deviations of variables are shown by subsample. The variables shown are: number of citations from citing to cited firm in year *t*; total citations of citing firm in year *t*; patent stock of cited firm in year *t*; percent of patents of cited firm that are less than three years old; percent of patents of cited firm that are between 3 and 7 years old; technological proximity of the citing and cited firms; a dummy equal to one if firms in the pair are from the same geographic region (U.S., Europe, Japan, Rest of World); a dummy equal to one if firms had an alliance in or before year *t*; a dummy equal to one if the firms had an alliance within 3 years of *t*; a dummy equal to one if the firms had an alliance 3–5 years before *t*; a dummy equal to one if the firms had an alliance more than 6 years before *t*; and, a dummy equal to one if the firms had an intensive alliance (see text, Section 4.1).

	All pairs		Nonallied Pairs		Allied pairs	
	Mean	SD	Mean	SD	Mean	SD
Panel A: Main analysis sample						
No. of citations from citing to cited firm in *t*	12.5	63.1	5.2	13.0	17.1	41.8
Total citations of citing firm in *t*	802	1,717	644	1,424	1,249	2,232
Patent stock of cited firm in *t*	1,380	1,934	1,187	1,667	1,976	2,361
% Cited stock <3 years old in *t*	34%	20%	33%	20%	34%	19%
% Cited stock 3–7 years old in *t*	36%	13%	36%	14%	36%	12%
Technological proximity	52%	28%	45%	24%	52%	23%
Firms in pair are from same region	59%	49%	50%	50%	63%	48%
1 if firms had alliance in or before *t*	15%	47%			100%	0%
1 if alliance <3 years before *t*	32%	35%			45%	50%
1 if alliance 3–5 years before *t*	5%	22%			16%	37%
1 if alliance 6+ years before *t*	16%	36%			49%	50%
Intensive alliance	7%	25%			21%	41%
Number of directional pairs	8,077		5,116		2,467	
Number of pair-year observations	40,898		27,686		13,212	
Panel B: Compustat sample						
No. of citations from citing to cited firm in *t*	16.5	44.0	8.0	19.9	26.7	60.0
Total citations of citing firm in *t*	1,066	1,985	758	1,605	1,436	2,310
Patent stock of cited firm in *t*	1,573	1,983	1,318	1,703	1,879	2,237
% Cited stock <3 years old in *t*	35%	18%	33%	18%	36%	18%
% Cited stock 3–7 years old in *t*	35%	10%	36%	11%	34%	9%
Technological proximity	46%	24%	43%	24%	49%	23%
Firms in pair are from same region	86%	35%	84%	37%	88%	33%
1 if firms had alliance in or before *t*	45%	50%			100%	0%
1 if alliance <3 years before *t*	23%	42%			50%	50%
1 if alliance 3–5 years before *t*	7%	26%			15%	36%
1 if alliance 6+ years before *t*	21%	40%			46%	50%
Intensive alliance	10%	30%			21%	41%
Number of directional pairs	1,160		610		550	
Number of pair-year observations	8,917		4,874		4,043	

the citing firm. Similarly, we control for the size of the citable stock of patents by including the size of the stock of previously granted patents of the cited firm in a given year. With these controls in place, we interpret the coefficients on the other independent variables as effects on the likelihood that a given citation made by the citing firm will be made to a given patent of the cited firm (Jaffe and Trajtenberg, 1996).

Two additional control variables based on the patent data are the age of the citable patent stock and the technological similarity of the patent portfolios of the firms in a pair. Previous studies find that firms cite patents that are three to seven years old more often than younger patents or older patents (Jaffe and Trajtenberg, 1996): As such, we include variables that measure the percentage of patents in these age categories in the citable stock of patents. Several previous studies also find that firms tend to cite other firms that are active in similar technology domains. Thus, we use a measure of technological proximity that is by now fairly common in the literature. First introduced by Jaffe (1986), this measure is defined as the angular separation of the patent class distribution vectors of the two firms, which is the scalar product of these vectors normalized by their scalar products with themselves, so that the measure takes the value of unity for any two identical vectors. The distribution vectors are defined over the approximately 500 patent classes defined by the USPTO (Jaffe and Trajtenberg, 2002).

Since patent citations are only a proxy for the knowledge flow between firms, we must consider whether artifactual sources of variation in citation rates bias the results. A complete set of year dummies is used in all regressions, in order to control for possible changes over time in the average propensity to cite. The most important remaining source of bias in the data derives from a peculiarity of the rules for citation. In order to achieve patent protection for an invention in different countries, the inventor must apply separately to different patent offices. For example, the most important and valuable inventions will typically be protected by a U.S. patent, a Japanese patent, and a patent granted by the European patent office. Our data include only U.S. patents and citations made to other U.S. patents. Because important inventions, regardless of origin, are likely to be patented in the U.S., the limitation to U.S. patents is probably not a serious one. If, however, the U.S. patent examiner determines that a citation to a previous invention is necessary, in some cases, that obligation can be satisfied by citing either the U.S. patent document covering that invention or a foreign patent document covering the same invention. Unfortunately, the citations to foreign patent documents are not in the computerized patent data. It seems likely that the origin of the invention affects the likelihood that a citation, if made, will be made to the foreign patent document rather than to the U.S. document; this introduces a bias in our data, because the data are limited to citations to U.S. documents. In particular, the likelihood that a Japanese company would cite the Japanese patent document rather than the U.S. document is probably higher than for a U.S. company. This means that comparisons of citation probabilities for companies from different countries cannot be used to reliably measure the magnitude of knowledge flows within and between different countries. (Indeed, preliminary tests comparing citation patents of U.S., Japanese, and European firms suggests that U.S. firms cite other U.S. firms more often than Japanese and European firms cite other firms in their own regions.)

In our analysis, we include dummy variables for the home country of the citing and cited firms. Because of the foreign patent citation problem, the coefficients on these variables cannot be interpreted in terms of effects on knowledge flows. What interests us most, however, is the extent to which the effects of alliances interact with these geographic variables. We cannot think of any reason why the bias associated with our inability to count citations to foreign patent documents should differ in any way as between allied and nonallied firms—one might expect that the bias is purely a result of the patent strategies and histories of the firms. Hence, our strategy is to treat the geographic variables as controls and to not interpret their first-order effects; however, we do interpret the effects of geographic variables when they appear as interactions with alliance variables.

More generally, because the behavior underlying the relationship between patent citations and knowledge flows is complex, there will always be the possibility that results based on patent citations are biased because the independent variable affects citation behavior rather than knowledge flow. In general, we believe that most variations in citation practices are not correlated with the economic variables of interest here, and hence they do not introduce bias. Even if such a bias does exist, however, it still does not undermine the results regarding the differential behavior of allied and nonallied firms, unless the correlation of economic variables and citation practices extends beyond the individual behavior of the citing firm to encompass some kind of connection to the allied firm or to the propensity to form alliances. Because we think that such connections are implausible, we are confident that the results regarding the effects of alliances on citation frequencies are unlikely to be biased by the peculiarities of the citation process itself.

3.4. Variables: measures of alliance

Our measures of alliance activity between citing and cited firms come from the CATI database developed at MERIT. This database now covers over 20,000 national and international interfirm agreements formed between 1970 and 1999 (the first phase of data collection is described in Hagedoorn, 1993; Duysters and Hagedoorn, 1993). For every alliance, we have the year of formation, but the data rarely report a year of termination, even though it is well known that most alliances end or are replaced after a few years. As a consequence, we expect that the effect of an alliance on knowledge flows between two firms will be felt in the first years after the alliance is formed, and then will erode over time. In our analysis, we attempt to measure this effect by including dummy variables that indicate when an alliance is two years old or younger, when it is three to five years old, and when it is six or more years old.

We also know from the data when two firms have more than one alliance, either in the same year or in successive years. Such repeated alliances can be interpreted in two ways, both of which represent a deeper relationship between the firms than if the pair only had one alliance. The first interpretation is that successive alliances indicate a repeated game and a live relationship contract. When alliances are repeated in this way, the firms can be expected to have a stronger and longer-term underlying relationship than if their alliances were one-off deals. The second interpretation is that having more than one alliance in successive years or in the same year indicates a multipoint relationship, such that the firms engage in cooperation on several fronts or in several businesses. This too is a proxy for a deeper relationship than otherwise (Doz and Hamel, 1998). To measure the effects of these kinds of relationships, we use a set of dummy variables to indicate when a pair has one to two alliances, three to six alliances, or seven or more alliances. In some tests, we also use a dummy variable that simply measures whether the firms have any alliance at all by the year of citation.

One additional set of dummy variables that measure the nature of the alliance is drawn directly from the CATI database. For most observations, the database indicates certain features of the alliance structure, in particular, whether the alliance structure includes joint R&D or whether the agreement involves equity investment from either firm into a joint venture or into its partner. These measures are not mutually exclusive—an alliance can have either, both, or neither characteristics. We use dummy variables to measure each of these possibilities.

A broad way to think about all these alliance variables is that they are attempts to measure the intensity of the relationship between the partners. We find, in fact, that the various indicators of intensity have effects in the same direction—the more intensive the alliance, as measured by various proxies, the higher the citation frequencies between partners. Based on these findings, we construct a composite variable to separate more-intensive from less-intensive alliances, and we use that variable in further tests.

3.5. Variables: firm characteristics

Aside from patent citation variables and the organizational relationship variables, we use variables that measure various characteristics of the firms in a pair. One of these is the national origin of a firm; our measure of nationality is simple, but it has powerful effects. For each firm, the CATI data define the firm's nationality, that is, the home country of the firm's management headquarters. As mentioned above, these home-country designations correspond closely to where the firm has the bulk of its innovative activity, as measured by USPTO patent data. So for all intents and purposes, our nationality variables can be considered to have been drawn either from the CATI or USPTO data.

In our tests, we use regional aggregations of the home-country variables, i.e. we identify separately firms in the United States, Europe, Japan, and the rest of the world (emerging markets, Australia, Canada, and such). It is worth recognizing here something that is obvious but that has implications for our analysis: The Europe variable aggregates firms from different home countries, while the U.S. and Japan variables represent firms from single countries. Based on these regional variables, we define a dummy variable that indicates whether the two firms in a pair are from the same or different regions (e.g., Europe–Japan or U.S.–Japan).

Other firm characteristics are derived from Compustat, namely, aggregate sales, R&D expenditures, and the main SIC categories of each firm. Sales is used as a measure of overall firm size, and the R&D-to-sales ratio is employed as a measure of the R&D intensity of the firm. When merging the Compustat and CATI/NBER data, we drop any firms for which Compustat does not report R&D expenditures; this is not a common issue for firms in our industries. The three-digit SIC of each firm is used to construct a dichotomous variable that indicates whether the two firms in an allied or control pair are in the same industry.

4. Models and results

We begin with tests comparing the frequency of citation in pairs of nonallied firms, in pairs with alliances of varying degrees of intensity, and within firms. We then examine how the observed alliance effect varies with attributes of the firms and of firm pairs.

4.1. Base models: effects of organizational linkages

Tests of our basic hypotheses about the effects of organizational linkages on knowledge flows are presented in Table 2. These base models examine how the log of the citation flow

Table 2
Effects of alliance relationship on citations between firms in a pair

To estimate the effects of various types of alliance relationships, we use the whole sample of 40,898 pair years and OLS regressions. The dependent variable is the log of the number of citations from citing firm to cited firm in year t. Independent variables are: total citations of citing firm in year t; patent stock of cited firm in year t; percent of patents of cited firm that are less than 3 years old; percent of patents of cited firm that are between 3 and 7 years old; technological proximity of the citing and cited firms; a dummy equal to one if firms in the pair are from the same geographic region (U.S., Europe, Japan, Rest of World); a dummy equal to one if the pair is a self pair (i.e., firm citing itself); a dummy equal to one if firms in the pair had an alliance in or before year t; a dummy equal to one if the firms had an alliance within 3 years of t; a dummy equal to one if the firms had an alliance 3–5 years before t; a dummy equal to one if the firms had an alliance more than 6 years before t; a dummy equal to one if the firms had 1–2 alliances before t; a dummy equal to one if the firms had 3–6 alliances in or before t; a dummy equal to one if the firms had 7 or more alliances in or before t; a dummy equal to one if the firms had an R&D alliance before t; a dummy equal to one if the firms had an equity alliance before t; and, a dummy equal to one if the firms had an intensive alliance (see text, Section 4.1). The models shown use different sets of dummies to describe aspects of the alliance between allied firms. Standard errors are in parentheses. Full set of year dummies are included in the regressions, but are not shown here.

	Dependent variable: log of no. of citations from citing to cited firm in t				
	Model 1	Model 2	Model 3	Model 4	Model 5
ln(Total citations of citing firm in t)	0.154	0.157	0.161	0.148	0.160
	(0.010)***	(0.010)***	(0.009)***	(0.010)***	(0.009)***
[ln(Total citations of citing firm in t)]^2	0.025	0.024	0.024	0.025	0.024
	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***
ln(Patent stock of cited firm in t)	−0.169	−0.164	−0.157	−0.166	−0.160
	(0.010)***	(0.010)***	(0.010)***	(0.010)***	(0.010)***
[ln(Patent stock of cited firm in t)]^2	0.052	0.051	0.050	0.051	0.051
	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***
% Cited stock <3 years old in t	0.051	0.019	0.029	0.021	0.030
	(0.022)**	(0.022)	(0.022)	(0.022)	(0.022)
% Cited stock 3–7 years old in t	0.399	0.390	0.388	0.387	0.389
	(0.033)***	(0.033)***	(0.033)***	(0.033)***	(0.033)***
Technological proximity	1.421	1.409	1.387	1.364	1.395
	(0.017)***	(0.017)***	(0.017)***	(0.018)***	(0.017)***
Same region	0.184	0.180	0.179	0.162	0.180
	(0.008)***	(0.008)***	(0.008)***	(0.008)***	(0.008)***
Self-citation	1.131	1.135	1.141	1.208	1.137
	(0.018)***	(0.018)***	(0.018)***	(0.019)***	(0.017)***
1 if firms had alliance in or before t	0.060			−0.049	
	(0.009)***			(0.012)***	
1 if alliance <3 years before t		0.124			
		(0.012)***			
1 if alliance 3−5 years before t		0.191			
		(0.019)***			
1 if alliance 6+ years before t		−0.035			
		(0.012)***			
1 if 1–2 alliances in or before t			0.004		
			(0.010)		
1 if 3–6 alliances in or before t			0.315		
			(0.018)***		
1 if 7+ alliances in or before t			0.629		
			(0.040)***		
1 if joint R&D in an alliance before t				0.113	
				(0.010)***	

Table 2 (*continued*)

	Dependent variable: log of no. of citations from citing to cited firm in t				
	Model 1	Model 2	Model 3	Model 4	Model 5
1 if equity in an alliance before t				0.150 (0.011)***	
Intensive alliance					0.355 (0.016)***
Constant	−2.335 (0.053)***	−2.306 (0.053)***	−2.308 (0.053)***	−2.338 (0.053)***	−2.306 (0.053)***
R-squared	0.65	0.65	0.65	0.65	0.65

**Significant at 5%; ***Significant at 1%.

varies with patent control variables, technological and geographic proximity, and measures of the organizational relationship between the firms in the pair.

These results are robust across the models and are generally consistent with our hypotheses. The coefficients on the alliance variables are on average positive and statistically significant at 95% or 99% levels and their relative sizes are for the most part consistent with our hypotheses regarding the effects of organizational form.

Because the alliance variables are all dummies, the coefficients are average percentage differences between a given group and the reference group, which in this case is the group of nonallied pairs. For example, the coefficient on the alliance dummy in Model 1 implies that firms that have at least one alliance have, on average, 6% more citations than firms with no alliances, after controlling for the other variables. The coefficient on self-citation implies that firms cite themselves approximately 131% more than they cite firms with which they have no relationship. By comparison, the alliance effect seems small; however, Models 2 and 3 show that the effect of alliance varies substantially by alliance age and intensity. Model 2 implies that firms with alliances that are three to five years old have 19% more citations than nonallied firms, and Model 3 implies that firms with seven or more alliances have 63% more citations.

Further perspective on the magnitude of these effects comes from comparing them to the effects of technological and geographic distance, measured here by our proximity variables. The coefficient on technological proximity is approximately 1.4 in all of the models. This means that a hypothetical pair of firms with identical technology profiles (proximity of unity) has 140% more citations than a pair whose technology distributions are orthogonal (proximity of zero). The mean technological proximity is about 0.5; the interquartile range is also about 0.5. Thus, the difference in citation frequency between a pair with seven or more alliances and a nonallied pair (about 63%) is similar to the difference between a nonallied pair at the 25th percentile in technological proximity and a nonallied pair at the 75th percentile in technological proximity (half of 140% or 70%).

The effect of geographic proximity is measured by the dummy that indicates whether the firms in a pair are based in the same region or not; at about 18%, this effect is considerably smaller than the effect of technological proximity. It is likely, however, that this estimate is biased downward by the fact that our measure of geographic location (home country) is a crude one.

Self-pairs have, by definition, technological proximity of unity, with the two firms being in the same region. This means that if we compare rates of cross-firm citation to rates of self-citation, including the inherent technological and geographic proximity of a firm to itself, the differences are large. Adding together the pure self-citation effect (115%), the difference attributable to technological proximity of unity rather than the mean of 0.5 (70%), and the difference attributable to being in the same region rather than different regions (18%), we find that firms are about three times as likely to cite themselves as are nonallied firms of typical technological proximity originating in different home countries.

The coefficients of control variables are also consistent with our expectations. There is a strong scale effect, which we model with linear and quadratic terms for the number of citations of the citing firm and the number of patents in the patent stock of the cited firm. Patents that are three to seven years old are cited more often than younger or older patents, and pairs of firms in the same region cite each other more often than pairs in different regions. The orders of magnitude of the patent-age effect and the regional effect are similar to that of the alliance effect.

Model 4 examines whether the type of alliance affects citation behavior. It shows that firms linked by one or more equity alliances and firms that have one or more alliances that explicitly involve joint R&D are, in each case, 10 to15% more likely to cite each other than if they were not allied at all. These effects are additive, implying that pairs with an equity alliance and joint R&D have 25% more citations than nonallied firms.

Taken together, the results of Models 1–4 indicate that increasing the length, depth, and focus of the alliance relationship all enhance knowledge flow, as measured by citations. In order to explore further how this alliance effect varies in different circumstances, it is convenient to derive a single measure of the intensity of the alliance relationship that can then be interacted with other variables in a way that is relatively straightforward to interpret. Theory does not tell us the best form for such a variable. But based on the results in Models 1–4, we construct a summary measure that is defined as the sum of the R&D and equity dummies multiplied by a dummy indicating whether the pair has three or more alliances. This variable is zero for any pair-year combination that has no previous R&D or equity alliances, or for which the total number of alliances as of that year is fewer than three; this variable is unity for firms that have at least one previous R&D or equity alliance, or which in a given year have had three or more alliances. This dichotomous variable thereby divides the data into pairs that represent either intensive alliances or all others. Model 5 in Table 2 shows that, all else equal, pairs of firms that are intensively allied have 36% more citations than the rest of the pairs in the sample. This represents a value that is clearly of economic significance, while remaining, as we would expect, well below the approximately 130% increase associated with citation within the firm.

Table 3 explores the robustness of these results across different estimation methods. The data comprise an unbalanced panel of (directionally distinguished) pairs over time. Ordinary least squares (OLS) estimation is therefore potentially inefficient, and potentially biased if there are unobserved pair characteristics that are correlated with the independent variables. To explore this possibility, Table 3 presents random effects and fixed effects estimation results for selected models. In general, the random effects results are statistically and qualitatively similar to the OLS results, though there is some tendency toward quantitatively smaller alliance effects using random effects estimation. The only exception is Model 4, in which the random effects approach eliminates the effect of joint R&D and

Table 3
Comparison of results with alternative estimation methos

Because the stochastic structure of the panel data is unknown, we explore the sensitivity of the results to alternative estimation approaches: OLS (errors assumed independent and homoskedastic across time and firm pairs), random effects (observations for a given firm-pair have common variance, but are independent of regressors), and fixed effects (an unobserved firm-pair effect may be correlated with regressors). The first three columns apply these three approaches to the model in which alliance activity is measured by the number of alliances (Model 3 in Table 2); the next three apply the three methods to the model with a summary measure of alliance intensity (Model 5 in Table 2); the last two columns apply to the model with R&D alliances and equity alliances (Model 4 in Table 2). In all cases, variables that do not vary over time for a given firm-pair are dropped in the fixed effects version; for this reason there is no Fixed-Effects version for Model 4. As in the original regressions, the sample size is 40,898 in all these models and the dependent variable is the log of the number of citations from citing firm to cited firm in year t; see Table 2 for a further description of the OLS models. Standard errors in parentheses. Full set of year dummies are included, but are not shown here.

	Dependent variable: log of no. of citations from citing to cited firm in year t							
	Alternative estimations of Model 3, Table 2			Alternative estimations of Model 5, Table 2			Alternative estimations of Model 4, Table 2	
	OLS	Random effects	Fixed effects	OLS	Random effects	Fixed effects	OLS	Random effects
ln(Total citations of citing firm in t)	0.161	0.130	0.220	0.160	0.128	0.218	0.148	0.121
	(0.009)***	(0.010)***	(0.012)***	(0.009)***	(0.010)***	(0.012)***	(0.010)***	(0.010)***
[ln(Total citations of citing firm in t)]^2	0.024	0.026	0.038	0.024	0.026	0.039	0.025	0.027
	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***
ln(Patent stock of cited firm in t)	−0.157	−0.100	−0.051	−0.160	−0.103	−0.053	−0.166	−0.113
	(0.010)***	(0.011)***	(0.020)**	(0.010)***	(0.011)***	(0.020)***	(0.010)***	(0.011)***
[ln(Patent stock of cited firm in t)]^2	0.050	0.043	0.049	0.051	0.043	0.049	0.051	0.045
	(0.001)***	(0.001)***	(0.002)***	(0.001)***	(0.001)***	(0.002)***	(0.001)***	(0.001)***
% Cited stock <3 years old in t	0.029	0.044	0.201	0.030	0.046	0.204	0.021	0.064
	(0.022)	(0.024)*	(0.037)***	(0.022)	(0.024)*	(0.037)***	(0.022)	(0.024)***
% Cited stock 3–7 years old in t	0.388	0.454	0.603	0.389	0.456	0.605	0.387	0.465
	(0.033)***	(0.031)***	(0.039)***	(0.033)***	(0.031)***	(0.039)***	(0.033)***	(0.032)***

Technological proximity	1.387	1.203		1.395	1.209		1.364	1.218
	(0.017)***	(0.024)***		(0.017)***	(0.024)***		(0.018)***	(0.025)***
Same region	0.179	0.169		0.180	0.172		0.162	0.170
	(0.008)***	(0.012)***		(0.008)***	(0.012)***		(0.008)***	(0.012)***
Self-citation	1.141	1.077		1.137	1.069		1.208	1.077
	(0.018)***	(0.028)***		(0.017)***	(0.027)***		(0.019)***	(0.028)***
1 if firms had alliance in or before t							−0.049	0.045
							(0.012)***	(0.012)***
1 if 1–2 alliances in or before t	0.004	0.020	0.024					
	(0.010)	(0.011)*	(0.013)*					
1 if 3–6 alliances in or before t	0.315	0.280	0.134					
	(0.018)***	(0.020)***	(0.022)***					
1 if 7+ alliances in or before t	0.629	0.448	0.145					
	(0.040)***	(0.041)***	(0.042)***					
1 if joint R&D in an alliance before t							0.113	−0.020
							(0.010)***	(0.014)
1 if equity in an alliance before t							0.150	0.047
							(0.011)***	(0.017)***
Intensive alliance				0.355	0.284	0.114		
				(0.016)***	(0.018)***	(0.019)***		
Constant	−2.308	−2.317	−2.563	−2.306	−2.322	−2.563	−2.338	−2.328
	(0.053)***	(0.052)***	(0.078)***	(0.053)***	(0.052)***	(0.078)***	(0.053)***	(0.052)***
R-squared	0.65		0.58	0.65		0.58	0.65	
Number of identifiers or directionally distinct pairs		8077	8077		8077	8077		8077

*Significant at 10%; ***Significant at 1%.

greatly attenuates the effect of equity alliances. For the summary variable indicating an intensive alliance, the random effects result is about 28%, just slightly smaller than the 36% for the OLS estimate.

The fixed effects estimates do suggest that there are unobserved pair effects that are positively correlated with the tendency to form alliances. Those alliance effects that can be estimated are all smaller using the fixed effects approach, with the impact of the summary measure of intensive alliance falling to about 11%. It is still, however, statistically different from zero. Since other effects such as those of self-citation and technological proximity cannot be estimated in the fixed effects model, there is no obvious benchmark against which to compare this effect to determine whether it is big or small. Since it is, however, purely a within pair effect—estimated off of the difference in yearly citation intensity when firms change their alliance status—a roughly 10% increase appears to be economically important.

4.2. Interactions of alliance effect with firm and pair characteristics

The results of the previous subsection show a clear pattern of increased knowledge flow, as measured by citations, for firms in alliance relationships as compared to nonallied firms. The next question we ask is what kinds of firms and what kinds of alliances observe the greatest effects. A straightforward approach to answering such a question is to interact the alliance variable with variables that capture important firm and pair characteristics, as we do next. Note that we do not include the self-pairs in the estimation in this subsection, because our main focus here is the variation in alliance effects.

The results of tests with interaction terms are presented in Table 4. Because a number of the effects of interest would drop out, we do not use the fixed effects approach here. Rather, we use the random effects estimator, which in principle is more efficient, and which is also more conservative than OLS in the sense that it seems to generally give smaller alliance effects. The first column (Model 5A) merely re-estimates Model 5 in Table 3 to show that excluding the self pairs and moving to the smaller Compustat sample yields the same basic story as regressions with the main analysis sample. The effect of the intensive alliance variable is estimated at 30%, compared to 28% in the larger sample. Other estimates change only slightly with increases in the effects of technological proximity (from 1.2 to 1.7) and geographic proximity (from 0.17 to 0.36), but the qualitative picture is similar.

Model 6 includes first-order (i.e., not interacted with alliance) effects of three sets of variables: (1) Firm sizes for each of the firms, measured by the log of citing- and cited-firm sales revenues; (2) R&D intensities for each of the firms, measured by the R&D-to-sales ratios of the citing and cited firms; and, (3) market proximity of the firms, measured by a dummy that is unity if Compustat reports the firms to be in the same three-digit SIC industry. We use for this purpose whatever SIC is reported by Compustat. Indeed, many firms are assigned a three-digit SIC. For firms whose operations span multiple industries, Compustat assigns three-digit SIC numbers that sometimes are in effect two-digit SICs. That is, a company with activities in multiple sectors of the chemical industry are assigned SIC 280. Thus, as for the same region dummy, this same SIC dummy is crude. Except for market proximity, these variables do not have important first-order effects. Most are statistically insignificant. The effect of the size of the citing firm is statistically significant, but the coefficient is negligibly small (it implies an elasticity of citations with respect to size

Table 4
Variation of the alliance effect with characteristics of paired firms

To estimate how the alliance effect varies with characteristics of citing and cited firms, we use the Compustat sample of 8,917 pair-years and Random Effects regressions. The dependent variable is the log of the number of citations from citing firm to cited firm in year *t*. Independent variables are: total citations of citing firm in year *t*; patent stock of cited firm in year *t*; percent of patents of cited firm that are less than 3 years old; percent of patents of cited firm that are between 3 and 7 years old; technological proximity of citing and cited firms; a dummy equal to one if firms in the pair are from the same geographic region (U.S., Europe, Japan, Rest of World); a dummy equal to one if the firms had an intensive alliance (see text, Section 4.1); a dummy equal to one if the firms are in the same 3-digit SIC industry; the log of the citing firm's sales; the log of the cited firm's sales; the percentage of R&D in sales of the citing firm; and, the percentage of R&D in sales of the cited firm. Interaction terms of intensive alliance and the firm characteristics are used in one model. To interpret the combined effects of these interactions and enable comparison across models, the effect of an intensive alliance is calculated at sample means. The different models shown are: Model 5A uses the list of variables from Model 5 from Table 2, but on the Compustat sample; Model 6 adds firm characteristics as independent variables; and, Model 7 adds the interaction terms. Standard errors are in parentheses. Full set of year dummies are included in the regressions, but are not shown here.

	Dependent variable: log of no. of citations from citing to cited firm in *t*		
	Model 5A	Model 6	Model 7
ln(Total citations of citing firm in *t*)	0.149	0.165	0.167
	(0.026)***	(0.026)***	(0.027)***
[ln(Total citations of citing firm in *t*)]^2	0.038	0.038	0.037
	(0.002)***	(0.002)***	(0.002)***
ln(Patent stock of cited firm in *t*)	−0.202	−0.212	−0.182
	(0.034)***	(0.035)***	(0.035)***
[ln(Patent stock of cited firm in *t*)]^2	0.062	0.062	0.059
	(0.003)***	(0.003)***	(0.003)***
% Cited stock <3 years old in *t*	−0.043	−0.073	−0.087
	(0.066)	(0.069)	(0.069)
% Cited stock 3–7 years old in *t*	0.631	0.623	0.600
	(0.091)***	(0.092)***	(0.092)***
Technological proximity	1.739	1.702	1.677
	(0.063)***	(0.066)***	(0.065)***
Same region	0.356	0.350	0.314
	(0.047)***	(0.048)***	(0.047)***
Intensive alliance	0.297	0.299	−2.759
	(0.032)***	(0.032)***	(0.552)***
Same 3-digit SIC		0.092	0.079
		(0.041)**	(0.041)*
Log citing firm's sales: size		−0.032	−0.032
		(0.011)***	(0.011)***
Log cited firm's sales: size		0.016	0.013
		(0.013)	(0.013)
Citing firm's R&D per sales		−0.173	−0.313
		(0.260)	(0.263)
Cited firm's R&D per sales		0.061	0.138
		(0.257)	(0.260)
Interaction of intensive alliance with:			
Technological proximity			0.425
			(0.138)***
Same region			0.336
			(0.086)***

Table 4 (*continued*)

	Dependent variable: log of no. of citations from citing to cited firm in *t*		
	Model 5A	Model 6	Model 7
Same 3-digit SIC			0.232 (0.076)***
Log citing firm's sales: size			0.096 (0.031)***
Log cited firm's sales: size			0.132 (0.031)***
Citing firm's R&D per sales			3.343 (0.865)***
Cited firm's R&D per sales			0.217 (0.867)
Calculation:			
Effect of intensive alliance at sample means	0.297 (0.032)***	0.299 (0.032)***	0.091 (0.357)
Constant	−3.017 (0.154)***	−2.890 (0.181)***	−2.862 (0.180)***
Number of identifiers or directionally distinct pairs	1160	1160	1160

*Significant at 10%; **Significant at 5%; ***Significant at 1%.

of negative 0.03). The estimated effect of market proximity is positive, though slightly smaller than those of geographic and technological proximity; firms in the same industry are about 10% more likely to cite each other, all else equal. Addition of these variables does not change the estimated magnitude of the effect of having an intensive alliance, which remains at about 30%.

Model 7 interacts the summary measure indicating an intensive alliance with all of these firm and pair characteristics. Once these interactions are added, the coefficient on the alliance-intensity variable itself loses economic interpretation, because it represents the effect of an intensive alliance when all of the interacted variables have values of zero, including the log of sales. In order to estimate the complete alliance effect in this model, one must add together the coefficient on the alliance variable, plus the coefficients on each interaction multiplied by plausible values for each of the interacted variables. The last row in the list of independent variables in Table 4 shows the result of such a calculation with all of the interacted variables set to their sample means. Of course, for the models with no interactions, the entry in this row is just the coefficient on the intensive alliance variable itself. But in Model 7, the overall calculated effect of intensive alliance drops to 10% and is no longer statistically significant. This result suggests that while, on average, the alliance effect is large and positive (we know this from Model 6), this average effect comes virtually entirely from the firms with certain attributes. This total marginal effect of an intensive alliance is calculated by setting all other independent variables to their sample means. When we use the means conditional on there being an intensive alliance instead, the marginal effect of an intensive alliance rises to 0.311, but it still has a high standard error (0.344).

All of the proximity variables have positive and significant interactions with the alliance variable. This means that proximity fosters citation, alliance fosters citation, and the combination of proximity and alliance fosters citation more than additively. Put another way, the benefit to knowledge flow from forming an alliance is greatest for pairs of firms that are also proximate, despite the fact that such proximity already fosters knowledge flows in nonallied pairs. We review the interaction coefficients first, and then offer a broad explanation.

The interaction effect between intensive alliance and technological proximity is about 43%. This can be compared to the pure alliance effect of about 33%. As before, it is useful to interpret this in terms of comparing pairs at the lower and upper ends of the interquartile range of technological proximity. Allying with a partner at the lower end increases citation frequency by about 44% (33% plus 0.25 of 43%); allying with a partner at the upper end of the interquartile range increases citation frequency by about 67% (33% plus 0.75 of 43%).

The estimated coefficients for geographic and market proximity are smaller, but because the plausible range of these variables is larger, their real impact is comparable. That is, while there are virtually no firms with technological proximity of zero or unity, the difference between "same region" and "different region" is, in fact, unity. The coefficient of 0.34 on the same-region dummy implies that a firm in an intensive alliance with a partner in its home region is twice as likely to cite its partner as it would be if it were in an equivalent alliance with a foreign partner (33% plus 34% compared to the sole effect of 33%). The effect of the interaction of same-SIC and intensive alliance is somewhat smaller at about 23%; this implies that an alliance with a firm in the same industry increases citations by a little over 50%, compared with about one-third for firms not in the same industry.

There are also interaction effects between the intensive-alliance variable and size variables, which we measure here by firm revenues (recall that sizes of patent stocks are already in the controls). The coefficients on both of these interaction terms imply elasticities of 10% to 13%, meaning that a firm in an intensive alliance will cite its partner this much more for every doubling of revenue either of themselves or their partner.

Theory or existing literature does not provide prior expectation about these interaction effects of technological proximity, geographic proximity, and firm size. We suspect that further work and modeling is needed to pin down their meaning. For now, however, we interpret these results as indicating an opening of the floodgates to knowledge flows when firms engage in intensive alliances. The managerial literature has many examples of alliances that grew beyond their initial intents because the partners encouraged the free flow of people, information, and ideas across firm boundaries (see, e.g., the case of Xerox and Fuji Xerox in Gomes-Casseres, 1996). Narrow, single-point, one-shot alliances are unlikely to lead to such broad collaboration, nor is incidental contact between unallied firms. Nevertheless, the concerted management effort involved in large, intensive alliances may well have such snowball effects, in which case firms that have a lot in common are likely to benefit more from the increased flows of knowledge than other firms; this could explain the interaction with technological and geographic proximity. Furthermore, partners that have large portfolios of businesses in which to apply each other's technology may also benefit more from broad collaboration than other firms; this could explain the interaction with citing- and cited-firm size.

The interaction between intensive alliance and R&D intensity of citing and cited firms is also interesting, and seems consistent with a major strand of the literature on organizational learning. Unlike the size effects, the R&D-intensity effects are asymmetric: R&D intensity of citing firms is strongly correlated with increased citations in intensive alliances (statistically significant coefficient of 3.3), but R&D intensity of the cited firm is not correlated at all (the much smaller coefficient is not statistically significant).

A different explanation applies to each side of this knowledge flow story. It is not too surprising that there is no interaction effect for cited-firm R&D, because the impact of this R&D is already captured in the control variables that measure the size and age of the stock of citable patents, which represent the direct output of this R&D. But the R&D of the citing firm could also have a different effect—it could help the citing firm absorb external technology, aside from its role in generating new technology within the firm (Cohen and Levinthal, 1989). As a general rule, if R&D encourages internal development more than absorption of external technology, then we would expect a negative effect on citation frequency; otherwise, we would expect the reverse. Our results suggest that at least in intensive alliances, the absorption effect of R&D dominates, while the reverse is true for nonallied pairs, weakly allied pairs, and internal to the firm. Put more strongly, participation in an intensive alliance could act to release partners from the well-known not invented here syndrome, and encourage them to devote more effort to absorbing a partner's technology.

5. Conclusion

As a whole, our analysis is consistent with expectations about how alliances affect knowledge flows; it also yields new conclusions that are potentially important to managers and raises questions that deserve further work.

Before summarizing these conclusions, we should acknowledge one limitation of our research. We find a positive correlation between the existence of an alliance and citation rates, and based on the weight and variety of the evidence, we conclude that alliances in fact promote citations. But there is a chance that these results also reflect reverse causation—a tendency for higher citations rates between two firms to lead them to form an alliance. In separate tests not shown here we explore this problem with instrumental variables models, but we find no reliable determinants of alliance formation that do not also determine citation likelihood. Even so, our results suggest that even where the effect of alliance is lower than in the tables shown here, it is generally still positive and statistically significant. In a related project, Megan MacGarvie uses a matched-pair approach to disentangle these effects. In unpublished preliminary results, she finds that, indeed, pairs that form alliances do tend to have a slightly higher-than-average citation rate than other firms, but this rate increases after alliance formation more rapidly than in matched pairs that remain unallied. In her analysis, she cannot compare these results with self-citations.

Despite this limitation, we believe that our study is consistent with expectations in the literature regarding the extent of knowledge flows in different organizational arrangements. As one moves from the loosest organizational arrangement to the tightest, the likelihood of knowledge flows increases; citation probability is higher for alliances than for nonallied firms, and is higher for partners internal to the firm than those within alliances. Our various tests with different forms of alliance relationships are also consistent with this finding: Specifically, partners that have many alliances, relatively recent alliances, and alliances involving equity and joint R&D are more likely to cite each other than those that have fewer, older, or less-involved alliances. This reinforces our conception of the difference between firms and alliance pairs as a matter of degree, not of kind, at least as regards flows of technological knowledge.

By extension, one implication of our results is that the narrow definition of the firm may not be the best unit for economic and business analyses in some situations. There is an emerging view that the legal definition of the firm is sometimes at odds with the economic definition (Rajan and Zingales, 2001). For instance, scandals over Enron's special purpose entities are a well-cited manifestation of this problem or, more positively, many large firms have come to depend on alliances for a large part of their sales or supplies. In essence, these firms create webs of relationships that are just as important to their competitive success as are some of their internal, wholly owned assets. Our conclusions suggest similar tendencies with respect to research activities. The resources that these firms tap into when developing new knowledge extend beyond their legal boundaries and beyond their home bases, even though the channels for tapping these sources may vary in effectiveness. Perhaps a firm is still a firm, but the relevant unit of inventive activity may be better defined by knowledge flow boundaries that cut across the organizational and geographic landscape.

We find interesting new results when exploring how the likelihood of knowledge flow varies with different kinds of alliance pairs. Aside from the intensity (or structure) of an alliance, knowledge flows seem to be affected by characteristics of the partners, in particular, their technological and business fit to their geographic origins and their sizes. These differential effects are potentially important to managers. We mention above the role of managerial decisions regarding alliance structure and intensity; our results show that managerial decisions regarding partner choice are equally important. We find that alliances can bridge technological and geographic differences between firms. However, we also find that similarities between partner capabilities and origins tend to promote knowledge flows within alliances; the effects of interactions between alliance and our measures of technological, geographic, and business-segment proximity are all positive. Indeed, because the effect of alliance formation on knowledge flows depends on the firms' underlying proximity and other characteristics, the results imply that mismatched firms may not succeed in increasing knowledge flows through alliance formation. These findings have direct managerial implications.

Our results leave some other important managerial questions unanswered. For a firm that seeks to enhance its technological position, is it better to select a partner similar to itself because the knowledge flows within the alliance would be enhanced? Or, is it better for the firm to select a dissimilar firm, because the potential for learning something truly new could be greater, notwithstanding the likelihood of lesser flows overall? The current study is not set up to answer this question, but we hope to address it in future work.

Finally, our study leads to the broad conclusion that alliances and knowledge flows both demand effort and investment to succeed, with a possible corollary that this success can snowball under the right conditions. We mention just above how intensive alliances yield more benefit than simpler, narrower alliances. It is well known that an intensive alliance relationship demands management time and effort (Dyer and Singh, 1998; Bamford et al., 2003). In the case of knowledge flows, this effort is sometimes counted as R&D, but it is R&D of a different kind than the traditional one as it is aimed at learning, adopting, and applying the partner's technology. Indeed, we find that R&D by citing firms (the receivers of knowledge flows) strongly encourages the flow of knowledge. Our corollary finding is that this effort can open up greater opportunities than perhaps originally foreseen; this is one interpretation of the positive interaction between firm size and alliance. Again, these results are suggestive and there is ample room for further work.

Appendix

Definitions of main variables

Variable Name	Description	Source*
No. of citations from citing to cited firm in *t*	For patents granted to citing firm in year *t*, number of citations to existing patents of cited firm	USPTO
Total citations of citing firm in *t*	Total citations in all patents granted to citing firm in year *t*	USPTO
Patent stock of cited firm in *t*	Total number of patents granted to cited firm in years 1975 to *t*	USPTO
% Cited stock < 3 years old in *t*	Percent of cited firm's patent stock that consists of patents that are less than 3 years old as of year *t*	USPTO
% Cited stock 3–7 years old in *t*	Percent of cited firm's patent stock that consists of patents between 3 and 7 years old as of year *t*	USPTO
Technological proximity	Index of technological proximity (see text)	USPTO
Same region	Dummy equal to 1 if country of citing and cited firm are in same region, where regions are defined as U.S., Europe, Japan, and Rest of World	CATI
Self-citation	Dummy equal to 1 if citing firm is the same as cited firm	
1 if firms had alliance in or before *t*	Dummy equal to 1 if pair had at least one alliance in or before year *t*	CATI
1 if alliance < 3 years before *t*	Dummy equal to 1 if pair had at least one alliance fewer than three years before year *t*	CATI
1 if alliance 3–5 years before *t*	Dummy equal to 1 if pair had at least one alliance 3 to 5 years before year *t*	CATI
1 if alliance 6+ years before *t*	Dummy equal to 1 if pair had at least one alliance 6 or more years before year *t*	CATI

1 if 1–2 alliances in or before t	Dummy equal to 1 if pair had 1 or 2 alliances in or before year t	CATI
1 if 3–6 alliances in or before t	Dummy equal to 1 if pair had 3 to 6 alliances in or before year t	CATI
1 if 7+ alliances in or before t	Dummy equal to 1 if pair had 7 or more alliances in or before year t	CATI
1 if joint R&D in an alliance before t	Dummy equal to 1 if any alliance between the firms in or before year t was structured as a joint R&D agreement	CATI
1 if equity in an alliance before t	Dummy equal to 1 if any alliance between the firms in or before year t included equity investment	CATI
Intensive alliance	Dummy equal to 1 if the firms have three alliances AND any of the alliances involved equity or joint R&D in or before t	CATI
Same 3-digit SIC	Dummy equal to 1 if the firms have same 3-digit SIC segment	Compustat
Log citing firm's sales: size	Log of citing firm revenues in year t	Compustat
Log cited firm's sales: size	Log of cited firm revenues in year t	Compustat
Citing firm's R&D per sales	R&D expenditures divided by revenues for citing firm in year t	Compustat
Cited firm's R&D per sales	R&D expenditures divided by revenues for cited firm in year t	Compustat

*Sources are described in the text: USPTO = U.S. Patent Office database of U.S. patents; CATI = Cooperative Agreements and Technology Indicators database of technology alliances developed by Maastricht Economic Research Institute in Technology; Compustat = Standard and Poors Compustat database of financial and operating data.

References

Alcacer, J., Gettelman, M., 2004. How do I know what you know? Patent examiners and the generation of patent citations. National Bureau of Economic Research Summer Institute Presentation.

Allen, J., Phillips, G., 2000. Corporate equity ownership, strategic alliances, and product market relationships. Journal of Finance 55, 2791–2815.

Baker, G., Gibbons, R., Murphy, K., 2002. Relational contracts and the theory of the firm. Quarterly Journal of Economics 117, 39–83.

Bamford, J., Gomes-Casseres, B., Robinson, M., 2003. Mastering Alliance Strategy: A Comprehensive Guide to Design, Management, and Organization. Jossey-Bass, San Francisco.

Bartlett, C., Ghoshal, S., 1989. Managing Across Borders: The Transnational Solution. Harvard Business School Press, Boston.

Branstetter, L., 2002. The roles of FDI and interfirm R&D alliances in international knowledge spillovers: evidence from Japanese multinational firms. Unpublished working paper. Columbia Business School.

Caves, R., 1996. Multinational Enterprise and Economic Analysis. Cambridge University Press, Cambridge, UK.

Chan, S., Kensinger, J., Keown, A., Martin, J., 1997. Do strategic alliances create value? Journal of Financial Economics 46, 199–221.

Coase, R., 1937. The nature of the firm. Economica 4, 386–405.

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

Contractor, F., Lorange, P., 1988. DC Heath, Lexington MA.

Doz, Y., Hamel, ., 1998. Alliance Advantage. Harvard Business School Press, Boston.

Duysters, G., Hagedoorn, J., 1993. The cooperative agreements and technology indicators (CATI) information system. Working paper, MERIT.

Dyer, J., Singh, S., 1998. The relational view, cooperative strategy and sources of interorganizational competitive advantage. Academy of Management Review 23, 660–679.

Garvey, G., 1995. Why reputation favors joint ventures over vertical and horizontal integration: A simple model. Journal of Economic Behavior and Organization 28, 387–397.

Gomes-Casseres, B., 1993. Computers: Alliances and industry evolution. In: Yoffie, D.B. (Ed.), Beyond Free Trade. Harvard Business School Press, Boston.

Gomes-Casseres, B., 1996. The Alliance Revolution: The New Shape of Business Rivalry. Harvard University Press, Boston.

Hagedoorn, J., 1993. Understanding the rationale of strategic technology partnering: Interorganizational modes of cooperation and sectoral differences. Strategic Management Journal 14, 371–385.

Hall, B., Jaffe, A., Trajtenberg, M., 2002. The NBER patent-citations data file: lessons, insights, and methodological tools. In: Jaffe and Trajtenberg, 2002 (Chapter 13).

Hamel, G., 1991. Competition for competence and inter-partner learning within international strategic alliances. Strategic Management Journal 12, 83–104.

Hart, O., Moore, J., 1990. Property rights and the nature of the firm. Journal of Political Economy 98, 1119–1158.

Hennart, J., 2000. Co-specialization and learning in alliances. Working paper presented at AIB Annual Meeting in San Juan, Puerto Rico.

Hirschey, M., Richardson, V., 2004. Are scientific indicators of patent quality useful to investors? Journal of Empirical Finance 11, 91–107.

Jaffe, A., Trajtenberg, M., 1996. Flows of knowledge from universities and federal laboratories: modeling the flow of patent citations over time and across institutional and geographic boundaries. Proceedings of the National Academy of Sciences 93, 12671–12677.

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

Jaffe, A., Trajtenberg, M., 2002. Patents, Citations, and Innovations: A Window on the Knowledge Economy. MIT Press, Cambridge, MA.

Jaffe, A., Henderson, R., Trajtenberg, M., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics 108, 577–598.

Khanna, T., Gulati, R., Nohria, N., 1998. The dynamics of learning alliances: competition, cooperation, and relative scope. Strategic Management Journal 19, 193–210.

Kogut, B., 1988. Joint ventures: theoretical and empirical perspectives. Strategic Management Journal 9, 319–332.

Lerner, J., Shane, H., Tsai, A., 2003. Do equity financing cycles matter? Evidence from biotechnology alliances. Journal of Financial Economics 67, 411–446.

Macaulay, S., 1963. Non-contractual relations in business: a preliminary study. American Sociological Review 28, 55–70.

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

Rajan, R., Zingales, L., 2001. The firm as a dedicated hierarchy: a theory of the origins and growth of firms. Quarterly Journal of Economics 116, 805–851.

Mankin, E., Reich, R., 1986. Joint ventures with Japan give away our future. Harvard Business Review, March/April, 78–86.

Richardson, G.B., 1972. The organization of industry. The Economic Journal, September, 883-896.

Sampat, B., 2004. Examining patent examination: an analysis of examiner and applicant generated prior art. National Bureau of Economic Research Summer Institute Presentation.

Singh, J., 2003. Knowledge diffusion and multinational firms: evidence using patent citation data. Unpublished working paper, Harvard University, Boston.
Stopford, J., Wells Jr., L., 1972. Managing the Multinational Enterprise. Basic Books, Part II., New York.
Teece, D., 1986. Profiting from technological innovation: implications for integration, collaboration, licensing and public policy. Research Policy 15, 285–305.
Williamson, O., 1975. Markets and Hierarchies: Analysis and Antitrust Implications. Free Press, New York.
Williamson, O., 1979. Transaction-cost economics: the governance of contractual relations. Journal of Law and Economics 22, 3–61.