Explaining observed licensing agreements: Toward a broader understanding of technology flows

By: Albert N. Link and John T. Scott

Link, A.N., & Scott, J.T. "Explaining Observed Licensing Agreements: Toward a Broader Understanding of Technology Flows", *Economics of Innovation and New Technology*, 2002, 11(3): 211-231. https://doi.org/10.1080/10438590210905

Abstract:

A lack of quantitative information on cross-firm licensing agreements constrains policy makers in their overall understanding of the innovation process and the innovative environment of firms. This paper develops a methodology for understanding the patterns of technology flows that result through licensing agreements from readily available patent data. In addition, hypotheses about firms that share technology through licensing are tested; in particular, we find that diversified firms have a higher probability of licensing their technology.

Keywords: innovation and invention | intellectual property rights

Article:

I. INTRODUCTION

Alternative sources for acquiring technical knowledge are available to the firm, and each source can augment production and related activities differently. The most obvious and most frequently studied internal source is in-house R&D. External sources of technical knowledge are more varied, and may even be more important to some firms. One external source is the federal government. Firms involved in either contracted research or Cooperative Research and Development Agreements (CRADAs) can appropriate technical information from their involvement in research. A second external source is research universities. Firms appropriate technical knowledge from universities by funding participation in university-based research, or by transferring knowledge directly or indirectly through students or faculty consultants. A third source is other firms, domestic or international. New technology is embodied in capital equipment purchased from other firms, whether they have developed the technology or simply have added value to it through technical modifications. Firms can also acquire technical knowledge directly through mergers, or indirectly through their participation in collaborative research relationships or even by observing the research outputs of other firms without any formal transaction signaling a technology transfer. Finally, firms can license technology from other firms.

The economics, management, and policy literatures are replete with studies of in-house R&D activity. Certainly, more is known about that source of technical knowledge than about any of the others mentioned above. Additionally, some studies have examined the spillover effects of R&D done by other firms on a firm's or industry's productivity. Such studies have used R&D, patents, or purchased inputs embodying the technologies developed with R&D to measure the outsiders' R&D.[1] Scholarly research is also accumulating on the use of and interaction with the external sources, particularly on firm interactions with universities and firm participation in collaborative research relationships.[2] However, there is a conspicuous lack of information about one external source in particular, namely technical knowledge obtained from other firms through licensing agreements.[3] While a paucity a publicly- or privately-available information on private-sector licensing agreements is likely the reason for such a dearth of information, scholars in general and policy makers in particular (to effectively initiate technology-based policies) need to understand patterns in licensing agreements as a part of their overall understanding of the innovation process and the innovative environment of firms.

The purpose of this paper is to set forth a general methodology for understanding the patterns of technology flows that result through licensing agreements.[4] Our methodology also provides a way to test hypotheses about the characteristics of firms that are most likely to use licensing to share technology. In Section II, we describe the methodology used to construct a database for a sample of licensing agreements among chemicals firms with U.S. patents. In Section III, we set forth a patent-based model to explain the licensing patterns observed. In Section IV, we discuss the relevance of our model for predicting technology flows, and we also provide new empirical evidence to support the hypothesis that licensing of technology and the resulting flows of technology will be more likely when firms are diversified. Finally, we conclude the paper with some summary comments in Section V.

II. CONSTRUCTING A DATABASE FOR A SAMPLE OF LICENSING AGREEMENTS

We relied on a number of electronic sources of information to identify existing licensing agreements (conducting keyword searches for "licensing" and "chemicals") to construct a database on licensing agreements. These sources included ProQuest's ABI/Inform,[5] Information Access Company's Business Index, and H.W. Wilson Company's Business Periodicals Index. We focused on the chemicals industry for the time period of 1993 to 1997. During those years,

---

[1] See Scherer (1982a, 1982b), and more recently see Siegel (1997).
[2] See Link and Rees (1991), Hall, Link, and Scott (2000), and Hagedoorn, Link, and Vonortas (2000).
[3] What limited information there is (e.g., Bozeman and Link, 1983; Fu and Perkins, 1995) suggests that licensing activity as a technology acquisition strategy is more prevalent among large firms than among small firms. See Baldwin and Scott (1987: pp. 118–120) for a review of this literature.
[4] We have shown that effective public support of innovation should be focused on sets in industries and technology areas (Link and Scott 1998b, 2001). Here we show that technologies being developed and patented are useful in sets of industries or technology areas. We conclude that public support should be aimed not piecemeal, but at the sets of areas among which technology is flowing as evidenced by licensing agreements.
[5] ProQuest's name has now changed to Bell and Howell Information and Learning.

the electronic sources had a relatively complete coverage of the industry's activities, and licensing of technology appeared to be an important activity.

The licensing agreements for chemicals technology identified for this study are listed in Table I.[6] An inspection of the table shows that there are 43 parent firms. Defining for the purpose of this study these 43 parent firms as the relevant population of firms licensing chemicals technology,[7] there are 1806 pairs of potential licensees and licensors.[8] A further inspection of Table I shows that 39 of the possible 1806 pairs are identified as having licensing agreements during the sample period, where only the presence, as opposed to the frequency, of the relationship is important for our analysis below.[9]

Table I. Licensing agreements among chemicals firms[10] (listed alphabetically by licensee)

Licensee	**Licensor**	**Licensed technology**
Allied Signal Inc.	Bayer AG	Non-ozone-depleting foam blowing agents
Allied Signal Inc.	DuPont	R-404A refrigerant
Asahi Chemical	Dow Chemical Co.	Insite technology
Boehringer Mannheim Corp.[11]	Hoffmann-La Roche & Co. (Roche Group)	Taq DNA polymerase, thermostable enzymes
Borealis	Exxon Chemical (Exxon)	Metallocene catalysts
BP Chemicals Inc. (British Petroleum)	Dow Chemical Co.	Insite singlesite catalyst technology

[6] Throughtout our sample, a ''parent'' is, as best we can achieve, a well-defined entity owning the patents on the licensed technology and making its licensing decisions largely autonomously. As a practical as well as theoretical matter, we work with parent data when dealing with companies that have subsidiaries, but not necessarily the ultimate parent for two reasons. First, in some cases, a well-defined company has U.S. patents and is in the CHI Research database—used extensively in this study as noted throughout the paper—but its ultimate parent does not, apart from the patents of its subsidiary, have patents in the CHI Research database. Second, in some other cases, the subsidiary (such as a U.S. subsidiary of a foreign firm) is a largely autonomous unit covering a highly diversified, conglomerate, ultimate parent's activity in the given technology area. We want to consider the parent company operating as an integrated and autonomous company with regard to decisions about technology development, and thus, to see the patent portfolio of a well-defined whole company in terms of the potential relationships among its patents and those of other well-defined companies operating in the specified technology areas. Formally, a license might go to or from a subsidiary, but the technology behind the license might reside elsewhere in the company. Thus, we consider parent companies that have patents in the CHI Research U.S. patent database that includes firms actively patenting in the United States. As a practical matter, the observation of a ''parent'' company is not always perfectly ''clean'', although we have endeavored to make our choices as sensible as possible. Development of a formal protocol for treating the choice of appropriate ''parent'' companies is one important issue that could usefully be addressed if policy makers implement the methodology suggested herein. Caveats aside, the information in Table I, as an empirical effort to identify licensing agreements, is unique; indeed such systematic compilations of licensing activity are rare. See Arora's (1997) important initial effort to examine licensing data.

[7] We realize that not all licensing agreements in the chemicals area have been identified but rather only those announced in the business news for which the firms involved have U.S. patent data in the CHI Research database, and additionally, probably only those in the database that involve larger firms.

[8] For 43 firms, there are 43×42 permutations (pairs of firms where order matters) of a licensee/licensor relationship.

[9] A supplemental data appendix on these firms is available from the authors.

[10] These agreements come from identified articles published from 1993 through 1997.

[11] Subsequent to the licensing agreements (after the period of our sample) shown in this table, mergers have changed the relationships among the parent firms in the table. For one example, Boehringer Mannheim merged with the Roche Group.

Licensee	**Licensor**	**Licensed technology**
Cerestar Benelux BV	Dow Chemical Company's Gas/Spec Technology Group	Shell SulFerox process
Chevron Chemical (Chevron)	BP Chemicals Inc. (British Petroleum)	Innovene technology
Chevron Chemical (Chevron)	Institut Francais du Petrole (IFP)	Eluxyl
Ciba Additives (Ciba-Geigy)	Mead Corporation	Borate photoinitiator
Daelim Industrial	Himont	Spherilene polyethylene
Dow Chemical Company (licensee and licensor)	BP Chemicals Inc. (British Pertoleum) (licensor and licensee)	BP's Innovene gas phase PE process. Dow's Insite metallocene catalyst technology.
Dow Chemical Company	Montell Polyolefins	Spheripol process technology
DSM Fine Chemicals (DSM NV)	Amoco Corp	Amoco/Chisso PP manufacturing technology
DSM Fine Chemicals (DSM NV)	BP Chemicals Inc. (British Petroleum)	Innovene technology
DSM Fine Chemicals (DSM NV)	BP Chemicals Inc. (British Petroleum)	Gas-phase polyethylene
DSM Fine Chemicals (DSM NV)	Univation Technologies (a joint venture between Exxon Chemical Company and Union Carbide Corp.)	Unipol technology
DuPont	Allied Signal Inc.	Genetron AZ-20(R-410a) refrigerant
Elf Atochem (Elf Atochem North America Inc. and also the parent—Elf Atochem—in France.)	DuPont	A refrigerant, Suva HP62 (ASHRAE designation R-404A)
Eni Chem (ENI)	BP Chemicals Inc. (British Petroleum)	Gas-phase, fluid-bed polyethylene
Exxon Chemical Company (Exxon)	Union Carbide	Cross-licensing embodied in joint venture.[12] Union Carbide's Unipol gas-phase process and Exxon's metallocene technology
Fina Inc. (Petrofina)	Phillips Petroleum Co.	Loop reactor technology
Formosa Plastics	BP Chemicals Inc. (British Petroleum)	Innovene technology
Indian Petrochemicals Corp. Ltd. (IPCL)	BP Chemicals Inc. (British Petroleum)	Fluid-bed acrylonitrile process
Indian Petrochemicals Corp. Ltd. (IPCL)	BP Chemicals Inc. (British Petroleum)	Innovene technology
Institut Francais du Petrole (IFP)	Mobil Chemical	Processes for para-xylene production
Ipiranga Quimica	Montell Polyolefins	Spherilene gas-phase technology
Lyondell Petrochemical Co.	Nissan Chemical Industries Ltd.	Slurry high-density polyethylene (HDPE) technology
Maruzen Polymer (Maruzen Co. Ltd.)	Nissan Chemical Industries Ltd.	Slurry high-density polyethylene (HDPE) technology
Mitsui Chemical	Exxon	Gas-phase metallocene technology
Montell Polyolefins	Dow Chemical Company	Rights to use Dow's Insite metallocene catalysts in Montell's Spheripol process technology

[12] Univation Technologies, a joint venture between Exxon Chemical Company and Union Carbide Corp.

Licensee	Licensor	Licensed technology
Morton International Inc.	Firmenich SA	Chemical synthesis and purification process
Phillips Petroleum Company	Chevron Chemical Company (Chevron)	Aromax catalytic reforming process
Quantum Chemicals (Millennium)	BP Chemicals Inc.	Innovene technology
Rexene (now named Huntsman Polymers Corp.)	DSM Fine Chemicals (DSM NV)	Solution phase technology
Sipsy SA	Arco Chemical Company (Arco)	Production technology for chiral glycidols and epoxides
Sumitomo Chemical	Exxon	Gas-phase metallocene technology
Tebodin BV	Hoechst AG	Catalytic scrubbing process
Texas Eastman (Eastman Chemical)	BP Chemicals Inc. (British Petroleum)	Innovene technology
Union Carbide Corporation	Exxon Chemical Company (Exxon)	Cross-licensing embodied in joint ventures (see footnote earlier). Union Carbide's Unipol gas-phase process and Exxon's metallocene technology

III. THE ANALYTICAL FRAMEWORK

A. Simple Patent-Based Probability Model of Licensing Agreements

Our patent-based model of licensing agreements specifies that, within a defined population of *n* firms, the probability of observing a licensing agreement between firm *i* and firm *j*, where firm *i* is the licensee and firm *j* is the licensor, can be predicted on the basis of the technology characteristics of firm *i* and firm *j*. Thus, our model is represented as:

$$Probability\left(license_{ij}\right) = F\left(\mathbf{X}_{ij}\right) \tag{1}$$

where $\mathbf{X}_{ij}$ is a vector of technological characteristics of the *i-j* pair of firms. Licensing agreements could be predicted if the technological characteristics—innovation strategy and funding, technological competitiveness, ability to use other firm's licensable technology—were known.[13] Such characteristics are not readily observable, although possibly discernable through extensive firm surveys and interviews (but to our knowledge no significant systematic attempt has been made). One could argue that in a competitive environment all such characteristics are determined by a fundamental set of technology characteristics that could be captured by a set of variables about the patent portfolios of the potential licensor and licensee. However, our objective in this paper is to demonstrate a methodology for explaining observed licensing agreements; prerequisite to the usefulness of the model for predicting such technology flows is that the elements of $\mathbf{X}_{ij}$ be readily observable whether or not those characteristics are primary determinants of licensing activity.

We use readily available information about each firm's patenting portfolio to create a set of instruments to represent the technological characteristics of each of the *i-j* pairs of firms. For

[13] One can view licensing activity as a form of technology adoption. See Siegel (1999) for a review of the literature on the determinants of technology adoption.

each chemicals firm, we began by constructing three patent-based variables to capture in the most parsimonious way the complementarity in the patent portfolios of a pair of firms. These variables, our simplest set of instruments of $\mathbf{X}_{ij}$, are defined in Table II and are relevant for what we call the simple patent-based probability model of licensing agreements.[14]

Table II. Explanatory variables in the simple patent-based model of licensing agreements

Variable	**Definition**
dnthrcit	Dummy variable = 1 if neither the potential licensor nor the potential licensee has citations to the other's patents, and 0 otherwise
dbothcit	Dummy variable = 1 if both the potential licensor and the potential licensee cite the other's patents, and 0 otherwise
donecit	Dummy variable = 1 when just one of the firms cites the other (thus, when either "the potential licensor cites the potential licensee while the potential licensee does not cite the potential licensor" or "the potential licensee cites the potential licensor while the potential licensor does not cite the potential licensee"), and 0 otherwise

We, of course, do not observe the underlying indicator variable (a linear combination of the instruments plus random error) that determines the probability in Eq. (1). Instead, we observe *dlic*, a dichotomous variable equaling 1 if there is an observed licensing agreement with a potential licensee as the observed licensee and a potential licensor as the observed licensor, and 0 otherwise. For the licensing agreements listed in Table I, *dlic* equals 1 in 39 of 1806 cases. Also included in Eq. (1) with the explanatory variables described in Table II are binary variables to account for the fact that within the chemicals technology area licensed technology could flow from one area segment or industry to another. Thus, a more complete specification of the simple model would control for the industries in the chemicals technology area that are occupied by a potential licensee or licensor.[15] Table III describes the industries considered.[16]

B. Statistical Results from the Simple Patent-Based Probability Models

Equation (1) was estimated as a probit model.[17] Included in the specification are the explanatory variables listed in Table II along with the industry controls described in Table III.[18]

---

[14] Values for the explanatory variables described in Table II came from patent information provided in the CHI Research database (CHI Research, 1996).

[15] We observe licensing between firms; yet our ultimate specifications will allow estimation of probabilities that licensed technology flows from one industry segment to another. Many firms are diversified, purposively combining complementary lines of business (Scott and Pascoe, 1987; Scott, 1993). Thus, when the firm as a whole is considered, as it is with licensing arrangements between whole firms with their patent portfolio data, the firm's lines of business may span much of the broad technological areas under consideration. Our ultimate model allows the identification of the probabilities of agreements across the various areas spanned by the firms.

[16] We constructed each variable by studying product category information from each firm's home page on the World Wide Web, as well as from various corporate directories available on Lexis/Nexis.

[17] Because there were only occasional instances where firm i licensed more than once from firm j, we did not consider either a count model or a multinominal probit model. We assume that the occurrence of licensing is completely determined by the systematic part of the model (the inner product of the unobserved parameters and the variables—an inner product estimated by the probit index using the maximum likelihood estimates of the unobserved parameters) and purely random error, and hence given that complete model, we have a sample of $n(n–1)$ statistically independent probabilities of licensing agreements. See Maddala (1983: p. 22).

[18] Absent from this specification are such variables as firm size and firm expenditures on R&D. We view such variables as endogenous to be determined by the underlying patent portfolios reflecting the information and technology base.

Table III. Industry controls for the firms licensing chemicals technology

Variable	Definition
dgensor and *dgensee*	Dummy variable = 1 for operations in general chemicals or intermediate or specialty chemicals n. e. c., 0 otherwise, with sor and see denoting potential licensor and potential licensee respectively
dpetsor and *dpetsee*	Dummy variable = 1 for operations in petroleum and petrochemicals, 0 otherwise, with sor and see denoting potential licensor and potential licensee respectively
dphasor and *dphasee*	Dummy variable = 1 for operations in pharmaceuticals, 0 otherwise, with sor and see denoting potential licensor and potential licensee respectively
dplasor and *dplasee*	Dummy variable = 1 for operations in plastics, 0 otherwise, with sor and see denoting potential licensor and potential licensee respectively
dfibsor and *dfibsee*	Dummy variable = 1 for operations in fibers, 0 otherwise, with sor and see denoting potential licensor and potential licensee respectively
dagrsor and *dagrsee*	Dummy variable = 1 for operations in agricultural chemicals, 0 otherwise, with sor and see denoting potential licensor and potential licensee respectively

Table IV. Probit estimates: probability of a licensing agreement between firms licensing chemicals technology

Variable	Coefficient (standard error in parentheses)
donecit	0.209 (0.212)
dbothcit	0.419** (0.172)
dpetsor	0.170 (0.149)
dpetsee	–0.017 (0.154)
dphasor	–0.306 (0.210)
dphasee	–0.173 (0.193)
dplasor	0.268*** (0.155)
dplasee	0.062 (0.157)
dfibsor	–0.269 (0.224)
dfibsee	0.028 (0.217)
dagrsor	0.548* (0.160)
dagrsee	0.108 (0.169)
constant	–2.540* (0.198)
χ^2_{12}	24.14**
Log likelihood	–176.08
Pseudo R^2	0.064
n	1806

* Significant at the 0.01-level.
** Significant at the 0.05-level.
*** Significant at the 0.10-level.
Note: *dgensor* and *dgensee* are absorbed in the intercept; thus, the constant term reflects the impact on the probit index of having both the potential licensor and the potential licensee in general chemicals. Further, *dnthrcit* is omitted and its effect is in the intercept; thus, the constant term reflects the case where both firms in the pair are in general chemicals and neither cites the other's patents.

The probit estimates are in Table IV. Controlling for the product location of the firm's operations, the probability of a licensing agreement increases with the extent of the cross-citations of the patents between the potential licensor and the potential licensee. The probability of a licensing agreement is least when neither of the firms in the pair cites the other's patents; it is greater when one but not both cites the other's patents; and it is greatest when both firms cite the patents of the other. We believe that the basic fact of complementarities in the technologies used by a pair of firms, as indicated by their cross-citations in their patent portfolios, is only a part of the explanation for why the linkages between the patent portfolios of firms are associated with the probability of licensing agreements. The fact that the potential licensor and potential licensee are actually doing R&D and establishing patent portfolios makes them better able to appropriate returns from the related technologies developed by others. As Cohen and Levinthal (1989: p. 569) have explained, "R&D not only generates new information, but also enhances the firm's ability to assimilate and exploit existing information."

C. Expanded Patent-Based Probability Model of Licensing Agreements

While the simple probability model demonstrates that the probability of a licensing agreement increases with the extent of the complementarities in the patent portfolios of the potential licensor and potential licensee, its usefulness for describing flows of technologies by means of licensing agreements is limited as indicated by the relatively low pseudo R^2 value reported in Table IV. Accordingly, an expanded patent-based probability model was estimated using the additional variables defined in Table V.

Table V. Explanatory variables in the expanded patent-based model of licensing agreements

Variable	Definition
tosorcit	''Inbound citations (hence outbound technology flow) for potential licensor'': total citations from 1990 through 1996 for the cited patents of the potential licensor that were issued from 1975 through 1996
toseecit	''Inbound citations (hence outbound technology flow) for potential licensee'': total citations from 1990 through 1996 for the cited patents of the potential licensee that were issued from 1975 through 1996
sorcitot	''Outbound citations (hence inbound technology flow) for potential licensor'': the potential licensor's total citations (1990–1996) of the patents (issued from 1975 through 1996) of the firms in our sample (including its citations of its own patents)
seecitot	''Outbound citations (hence inbound technology flow) for potential licensee'': the potential licensee's total citations (1990–1996) of the patents (issued from 1975 through 1996) of the firms in our sample (including its citations of its own patents)
secitsor	Potential licensee's citations (1990–1996) of the potential licensor's patents (issued 1975–1996)
sorcitse	Potential licensor's citations (1990–1996) of the potential licensee's patents (issued 1975–1996)
dcitsee	Dummy variable = 1 if the potential licensor cites the potential licensee's patents
dcitsor	Dummy variable = 1 if the potential licensee cites the potential licensor's patents
sornopat	Number of regular utility U.S. patents granted to the potential licensor, 1990–1996
seenopat	Number of regular utility U.S. patents granted to the potential licensee, 1990–1996
sorcil	Science linkage for potential licensor, 1990–1996: CHI Research's TECH-LINE indicator of how close a company's patents are to the scientific research base: science linkage is the average number of ''other references cited'' on the front pages of a set of U.S. patents, which are to the scientific literature, such as journal papers and scientific meetings. References to books, reports, and other non-scientific literature sources are excluded. (CHI Research, 1996)
seescil	Science linkage for potential licensee, 1990–1996

Variable	Definition
sortcit	Technology cycle time for potential licensor, 1990–1996: technology cycle time is defined as the median age in years of the earlier U.S. patents referenced on the front page of a U.S. patent. Technology cycle time is the time that has elapsed between the current patents and the previous generation of patents. (CHI Research, 1996)
seetct	Technology cycle time for potential licensee, 1990–1996
$sorpgpp_i$ and $seepgpp_i$ for i = 1 to 30	The percentage of the potential licensor's or potential licensee's patents in each of 30 SIC product groups. (CHI Research, 1996)

First, a limited expanded model, without the 30 product-group patent-percentage variables but including just the other patent variables and the industry location controls, was estimated for the entire sample of 1806 observations of potential licensors and potential licensees. Second, a similar model, adding to the first model the variable for technology cycle time, was estimated for a subset (n=1560) of the potential-licensor-licensee-pair observations for which the variable for technology cycle time was available. Finally third, the product-group patent-percentage variables are added to the first two models for full model estimations that provide whole models with excellent explanatory ability for the occurrence of a licensing agreement.

Table VI shows the estimated probit model for the full 1806 observation sample of potential licensing pairs, but without the variables to control for the potential licensor's and potential licensee's percentages of patents in each of the 30 SIC product groups into which CHI Research categorizes patents. Table VII reports the results from a reestimation of the model adding the technology cycle time for the potential licensor and the potential licensee. Although the cycle time variables are not themselves significant, the model performs better in the sense that it is even more significant statistically.[19] Our interpretative conclusions from Tables VI and VII, highlighting the relationships that have probit index coefficients greater than their standard errors in either or both specifications, are that, *ceteris paribus*, the probability of a licensing agreement:

- increases with the inbound citations (outbound technology flows) for both the potential licensor and the potential licensee; firms with patented technology that is useful to other firms are more likely to have licensing agreements.
- increases with the potential licensee's citations of the potential licensor; the citations signal the relevance of the potential licensor's technology for the technology of the potential licensee.
- includes a separable positive, constant eect given the presence of the potential licensee's citations of the potential licensor; further evidence that citations signal the relevance of the potential licensor's technology for the potential licensee.
- decreases with the licensee's outbound citations (inbound technology flows); the probability of an agreement decreases as the potential licensee's dependence on the potential licensor decreases as measured by the proportion of the potential licensee's citations that are to the patents of the potential licensor.[20]

[19] The subsample with complete technology cycle times is, by the definition of the technology cycle variable in Table V, a set of firms dealing with technologies with well-defined previous generations of patents.

[20] Estimating the model using proportions tells the same descriptive story directly. The model here (using the levels of all variables rather than their proportions) incorporates the proportions discussed and others implicitly in the *ceteris paribus* approach and does not sacrifice the observations for which the proportions would not be defined because the denominator of the proportion would be zero.

- decreases as the number of patents of the potential licensor increases; the technology of interest to the potential licensee (in a subset of the potential licensor's patents) is a less prominent part of the potential licensor's patent portfolio and less likely to receive the potential licensor's corporate resources for licensing the technology.
- decreases as the number of patents of the potential licensee increases; the potential licensee is less dependent on the technology of others in such cases.
- increases as the science linkage for the potential licensor increases; such patents are especially likely to have useful technology to transfer to other firms.
- decreases as the science linkage for the potential licensee increases; the potential licensee will be less dependent on others for technology.
- includes significant effects of several variables for industry location (chemicals firms are typically diversified into several of these industries); industry affects the likelihood of technology transferable through licensing agreements.

Tables VIII and IX show the results from an expansion of the two previous models by controlling for the distribution of the patents of the potential licensor, and for the potential licensee, across 30 SIC product-group categories.[21] The models are, as a whole, significant, and they provide excellent explanations of the probability of a licensing agreement (certainly there are too many variables in these models to estimate individual coefficients with precision). Thus, these large models estimate exceptionally well the collective effect of the variables on the probit index; hence we use them to provide probabilities of licensing as new technology indicators of the diffusion of technology through licensing based on probabilities fitted from readily available patent data.[22] From the probit model in Table VIII, we derive the probit index for each of the 1806 observations.

Table VI. Probit estimates: probability of a licensing agreement between firms licensing chemicals technology, full sample without patent percentages

Variable	**Coefficient (standard error in parentheses)**
tosorcit	0.000072 (0.000052)
toseecit	0.000083*** (0.000051)
secitsor	0.0021 (0.0015)
sorcitse	–0.00012 (0.0015)
sorcitot	0.0000076 (0.000062)
seecitot	–0.00010 (0.000084)
dcitsee	–0.145 (0.213)
dcitsor	0.370*** (0.216)
sornopat	–0.00038 (0.00047)
seenopat	–0.00056 (0.00043)
sorscil	0.248* (0.092)
seescil	–0.170 (0.135)
dpetsor	0.0083 (0.183)

[21] These categories are defined in CHI Research (1996).
[22] We use fitted probabilities from the large model to describe the historical probability of specified technology flows through licensing agreements. Our smaller models, for which the coefficients on the explanatory variables are well estimated, could be used to forecast probabilities of an agreement, or to predict such probabilities for various combinations of characteristics of hypothetical potential licensors and licensees.

Variable	Coefficient (standard error in parentheses)
dpetsee	0.057 (0.175)
dphasor	–1.349* (0.419)
dphasee	0.114 (0.231)
dplasor	0.408** (0.185)
dplasee	0.064 (0.183)
dfibsor	–0.731** (0.375)
dfibsee	0.131 (0.264)
dagrsor	0.926* (0.203)
dagrsee	0.219 (0.185)
constant	–2.869* (0.267)
χ^2_{22}	65.48*
Log likelihood	–155.41
Pseudo R^2	0.174
n	1806

* Significant at the 0.01-level.
** Significant at the 0.05-level.
*** Significant at the 0.10-level.
Note: *dgensor* and *dgensee* are absorbed in the intercept; thus, the constant term reflects the impact on the probit index of having both the potential licensor and the potential licensee in general chemicals.

Table VII. Probit estimates: probability of a licensing agreement between firms licensing chemicals technology, small sample with technology cycle time and without patent percentages

Variable	Coefficient (standard error in parentheses)
tosorcit	0.00014** (0.000061)
toseecit	0.000091*** (0.000054)
secitsor	0.0022 (0.0015)
sorcitse	0.000090 (0.0015)
sorcitot	–0.000046 (0.000070)
seecitot	–0.00013 (0.000090)
dcitsee	–0.171 (0.223)
dcitsor	0.318 (0.234)
sornopat	–0.00089*** (0.00052)
seenopat	–0.00052 (0.00046)
sorscil	0.264* (0.105)
seescil	–0.134 (0.146)
dpetsor	0.184 (0.198)
dpetsee	0.109 (0.192)
dphasor	–1.365* (0.454)
dphasee	0.077 (0.282)
dplasor	0.550** (0.227)
dplasee	0.102 (0.201)
dfibsor	–0.842** (0.418)
dfibsee	0.102 (0.275)
dagrsor	1.127* (0.228)
dagrsee	0.228 (0.212)
sortct	0.019 (0.057)
seetct	0.0047 (0.047)

Variable	Coefficient (standard error in parentheses)
constant	–3.352* (0.841)
χ^2_{24}	71.93*
Log likelihood	–135.30
Pseudo R^2	0.210
n	1560

* Significant at the 0.01-level.
** Significant at the 0.05-level.
*** Significant at the 0.10-level.
Note: *dgensor* and *dgensee* are absorbed in the intercept; thus, the constant term reflects the impact on the probit index of having both the potential licensor and the potential licensee in general chemicals.

Table VIII. Probit estimates: probability of a licensing agreement between firms licensing chemicals technology, full sample with patent percentages

Number of observations = 1806
χ^2_{80}= 122.26*
Log likelihood = –127.02
Pseudo R^2 = 0.32
The collection of patent variables, except for the technology cycle time variables, is included.
The same ten of the twelve industry variables (six for the potential licensor and six for the potential licensee) showing the production locations of the potential licensor and the potential licensee are included. As in the specifications of Tables VI and VII, *dgensor* and *dgensee* are absorbed in the intercept; thus, the constant term reflects the impact on the probit index of having both the potential licensor and the potential licensee in general chemicals.
$sorpgpp_i$ and $seepgpp_i$ for i = 1 to 30, except the $sorpgpp_{chemicals}$ and $seepgpp_{chemicals}$ were excluded and left in the intercept. Thus, there are included in this specifications 58 control variables for patent locations, 29 of the 30 SIC product group patent percentage variables for the potential licensor and the same 29 variables for the potential licensee. The percentages in the 30 categories add to 100%, and the variable dropped for the potential licensor and again for the potential licensee and left in the intercept are the variables showing the percentage of the potential licensor's or the potential licensee's patents in the CHI Research SIC product group for chemicals. Not surprisingly, that is the product group with the largest percentage of patents on average for the firms in the sample. Thus, the constant term reflects the impact on the probit index of having both firms in general chemicals and with 100% of their patents in chemicals with the effects of the included $pgpp_i$ variables to be added in to reach the effect for the exact patent distribution for the potential licensor and the potential licensee.

* Significant at the 0.01-level.

Table IX. Probit estimates: probability of a licensing agreement between firms licensing chemicals technology, small sample with technology cycle time and with patent percentages

Number of observations = 1560
χ^2_{82}= 120.15*
Log likelihood = –111.19
Pseudo R^2 = 0.35
The full collection of patent variables is included.
The same ten of the twelve industry variables (six for the potential licensor and six for the potential licensee) showing the production locations of the potential licensor and the potential licensee are included. As in the specifications of Tables VI and VII, *dgensor* and *dgensee* are absorbed in the intercept; thus, the constant term reflects the impact on the probit index of having both the potential licensor and the potential licensee in general chemicals.

sorpgpp$_i$ and *seepgpp*$_i$ for $i = 1$ to 30, except the *sorpgpp*$_{chemicals}$ and *seepgpp*$_{chemicals}$ were excluded and left in the intercept. Thus, there are included in this specification 58 control variables for patent locations, 29 of the 30 SIC product group patent percentage variables for the potential licensor and the same 29 variables for the potential licensee. The percentages in the 30 categories add to 100%, and the variable dropped for the potential licensor and again for the potential licensee and left in the intercept are the variables showing the percentage of the potential licensor's or the potential licensee's patents in the CHI Research SIC product group for chemicals. Not surprisingly, that is the product group with the largest percentage of patents on average for the firms in the sample. Thus, the constant term reflects the impact on the probit index of having both firms in general chemicals and with 100% of their patents in chemicals with the effects of the included *pgpp*$_i$ variables to be added in to reach the effect for the exact patent distribution for the potential licensor and the potential licensee.

* Significant at the 0.01-level.

From the index, we compute the cumulative normal probability, *F*. *F* is the model's estimate of the probability of a licensing agreement for the observation; the computed *F* values range from 0 to 0.5512693, with a mean value of 0.0216065. For comparative purposes, the naive probability of observing a licensing agreement is 0.0215947 (39/1806). The distribution of the values of *F* is skewed; the proportion of the 1806 observations for which the estimated probability exceeds the mean value is 0.2458472. That proportion represents 444 pairs, or 24.6 percent of the potential licensor–licensee pairs have a higher than average (mean) probability of a licensing agreement. For that 24.6 percent of the sample, the estimated probability of agreement ranges from 0.0217247 to 0.5512693, with a mean value of 0.0809248. The expanded model explains the occurrence of licensing agreements well in the sense that among the 444 observations with estimated probabilities of licensing greater than the mean probability, there are 38 of the 39 actual cases of licensing agreements.

For the 1560 observations used in the model in Table IX, there were 36 observations of actual licensing agreements. For the 1560 pairs, the estimate of the probability of an agreement, *F*, ranges from 0 to 0.6004202, with a mean value of 0.0230569, as compared to the naive probability 0.0230769 (36/1560). The proportion of the 1560 observations for which the estimated probability of a licensing agreement exceeded the mean value is 0.2358974, or 368 of the pairs. For these pairs, the estimated probability of agreement ranged from 0.0231307 to 0.6004202, with a mean value of 0.0907533. In this model, the explanation of the occurrence of a licensing agreement is "perfect" in the admittedly special sense that among the 368 pairs that have a high probability of agreement according to the model, are all 36 of the actual licensing agreements in the sample of 1560 pairs. Of course, since only a small fraction (36/1560) of the cases have agreements, the "naïve" guess that agreement would not be observed is correct 97.7 percent of the time, while our "perfect" predictor is correct only 78.7% of the time ((36 + (1560 – 368))/1560). The "naïve" guess, however, does not let one predict the cases where agreement *will* occur, and of course that is what we care about here.[23] Our model identifies every one of those cases, although it

[23] As Greene (1997: p. 893) observes: ''In general, any prediction rule of the form [we use] . . . will make two types of errors. It will incorrectly classify 0s as 1s and 1s as 0s. In practice, these errors need not be symmetric in the costs that result.'' Our rule does not incorrectly classify any of the 1s, but does identify as 1s some of the zeros. In the context of our interests, not knowing the circumstances when the 1s occur has a large cost; in our model, the errors are not symmetric in the costs that result. If one just wants to maximize the percentage of correct guesses about when one will see a licensing agreement, one could simply always guess ''no agreement here'' and be right most of the time since there are so few agreements in the population of possible agreements. That is of course true anytime the fraction of times 1s occur is very small. So, without the model, one would *never* be able to discern the cases

does as well identify the other cases where conditions imply that agreement is more likely than average yet it did not occur.[24]

IV. INTERPRETATION OF THE FINDINGS

The results presented above provide strong empirical evidence that an in-depth investigation of observable patent citation data is an effective way to explain technology flows throughout the economy through licensing agreements.[25]

To illustrate the relevance of our methodology for predicting technology flows from licensing agreements, we computed the matrices in Tables X and XI. Each cell of Table X shows the probability of a licensing agreement when the potential licensor and the potential licensee sell products in the industries indicated. For example, the third probability in the fourth row shows the average probability of a licensing agreement when the potential licensor is in plastics and the potential licensee is in pharmaceuticals to be 4.95 percent.[26] We computed the probability for each cell as follows. The probit index for the probability for a cell of the matrix is denoted *pindex**. We estimate *pindex** by using the fitted probit index values from the probit model in Table VIII. That model provides a well-fitted probit index for each pair of the full sample of 1806 pairs. To estimate the probit index, we observe that for each ith observation in the cell,

$$pindex_i = pindex^* + u_i \tag{2}$$

where u_i is random error with variance estimated by the square of $stpd_i$, where $stpd_i$ denotes the standard error of prediction for the probit index $pindex_i$.

where agreement occurs using the naive guess. The reason we want the model is to be able to find such cases. And our model is perfect on that score. Namely, Greene notes that in general there are two types of errors, and our model actually makes zero errors of the one sort. Unlike the naive guess, we can uncover the cases where licensing occurs with our model. Furthermore, given that performance, one might well think that the 0s that are predicted to be 1s in fact *are* high probability of licensing cases. A relatively high probability of agreement is, after all, a fairly low probability; therefore, we expect that there will be many cases where despite the relatively high probability no agreement occurs.

[24] Another way to observe that our model distinguishes well the cases where licensing agreement occurs is to note that the average mean probability for the 36 observations where *dlic* = 1 is 0.1613, while the average mean probability for the remaining 1524 observations where *dlic* = 0 is 0.0198.

[25] Following on the analyses of Scherer (1982a, 1982b) about R&D spillovers through purchased inputs used in production, Jaffe (1986) about spillovers from related R&D of other firms, and Link and Zmud (1987) about self-reported stocks of purchased technology, it may well be the case that external sources of technology are a more important source of productivity-enhancing innovation than in-house R&D. Scott (1993: pp. 128–131) provides an experiment demonstrating the effects of R&D spillovers across industries on total factor productivity growth and also reviews several other studies, such as the work of Geroski (1991) and Griliches and Lichtenberg (1984), that have explored such effects.

[26] The probability shown is an average for all observations where *dplasor* = 1 and *dphasee* = 1. Thus, for each observation the actual values of all variables are used. Again, that is the only sensible way to use the large model to predict probabilities, because we have high confidence in the complete model, but very low confidence in any one coefficient. Therefore we would not want to determine the predicted probabilities by creating artificial observations by setting individual variables in particular ways to generate the different predictions.

Table X. Probability of a licensing agreement for all firms (with the estimate of *pindex**, the standard error of the estimate of *pindex**, and the number of observations in parentheses)

	Licensee					
Licensor	**General chemicals**	**Petroleum**	**Pharmaceuticals**	**Plastics**	**Fibers**	**Agricultural chemicals**
General chemicals	0.0694 (–1.48, 0.0349) (*n* = 306)	0.0465 (–1.68, 0.0294) (*n* = 302)	0.0571 (–1.58, 0.0456) (*n* = 157)	0.0606 (–1.55, 0.0371) (*n* = 280)	0.0901 (–1.34, 0.0651) (*n* = 103)	0.0764 (–1.43, 0.0479) (*n* = 173)
Petroleum	0.0681 (–1.49, 0.0386) (*n* = 302)	0.0594 (–1.56, 0.0335) (*n* = 272)	0.0516 (–1.63, 0.0515) (*n* = 152)	0.0681 (–1.49, 0.0416) (*n* = 268)	0.0901 (–1.34, 0.0735) (*n* = 101)	0.0823 (–1.39, 0.0535) (*n* = 167)
Pharmaceuticals	0.0301 (–1.53, 0.110) (*n* = 157)	0.0162 (–2.14, 0.105) (*n* = 152)	0.0322 (–1.85, 0.159) (*n* = 72)	0.0212 (–2.03, 0.117) (*n* = 140)	0.0427 (–1.72, 0.195) (*n* = 51)	0.0314 (–1.86, 0.0535) (*n* = 167)
Plastics	0.0630 (–1.53, 0.0328) (*n* = 280)	0.0465 (–1.68, 0.0274) (*n* = 268)	0.0495 (–1.65, 0.0445) (*n* = 140)	0.0559 (–1.59, 0.0370) (*n* = 240)	0.0793 (–1.41, 0.0663) (*n* = 91)	0.0708 (–1.47, 0.0454) (*n* = 156)
Fibers	0.0446 (–1.70, 0.0516) (*n* = 103)	0.0233 (–1.99, 0.0273) (*n* = 101)	0.0455 (–1.69, 0.0847) (*n* = 51)	0.0329 (–1.84, 0.0533) (*n* = 91)	0.0446 (–1.70, 0.0875) (*n* = 30)	0.0465 (–1.68, 0.0717) (*n* = 57)
Agricultural chemicals	0.0901 (–1.34, 0.0501) (*n* = 173)	0.0606 (–1.55, 0.0453) (*n* = 167)	0.0778 (–1.42, 0.0644) (*n* = 86)	0.0793 (–1.41, 0.0556) (*n* = 156)	0.121 (–1.17, 0.0980) (*n* = 57)	0.102 (–1.27, 0.0727) (*n* = 90)

Table XI. Probability of a licensing agreement for undiversified firms (with the estimate of *pindex**, the standard error of the estimate of *pindex**, and the number of observations in parentheses)

	Licensee					
Licensor	**General chemicals**	**Petroleum**	**Pharmaceuticals**	**Plastics**	**Fibers**	**Agricultural chemicals**
General chemicals	0.0446 (–1.70, 0.0686) (n = 20)	0.0344 (–1.82, 0.0360) (n = 50)	0.0239 (–1.98, 0.0720) (n = 15)	0.0262 (–1.94, 0.0371) (n = 25)	— (n = 0)	0.0244 (–1.97, 0.0952) (n = 10)
Petroleum	0.0262 (–1.94, 0.0386) (n = 50)	0.0281 (–1.91, 0.0283) (n = 90)	0.0136 (–2.21, 0.0559) (n = 30)	0.0146 (–2.18, 0.0207) (n = 50)	— (n = 0)	0.0122 (–2.25, 0.0528) (n = 20)
Pharmaceuticals	0.000845 (–3.14, 0.547) (n = 15)	0.000291 (–3.44, 0.370) (n = 30)	0.0136 (–2.21, 0.981) (n = 6)	0.000208 (–3.53, 0.553) (n = 15)	— (n = 0)	0.000136 (–3.64, 0.931) (n = 6)
Plastics	0.0465 (–1.68, 0.0507) (n = 25)	0.0427 (–1.72, 0.0358) (n = 50)	0.0336 (–1.83, 0.112) (n = 15)	0.0329 (1.84, 0.0691) (n = 20)	— (n = 0)	0.0314 (–1.86, 0.0739) (n = 10)
Fibers	— (n = 0)	— (n = 0)	— (n = 0)	— (n = 0)	— (n = 0)	— (n = 0)
Agricultural chemicals	0.0526 (–1.62, 0.0368) (n = 10)	0.0559 (–1.59, 0.0398) (n = 20)	0.0367 (–1.79, 0.193) (n = 6)	0.0367 (–1.79, 0.0223) (n = 10)	(n = 0)	0.838 (–1.38, 0.00127) (n = 2)

Dividing Eq. (2) through by $stpd_i$, yields an estimable model of the average probability of each cell, with homoskedastic error. The cell probability is then estimated as the cumulative normal distribution evaluated at the coefficient from the regression, for the pairs in the cell, of *pindex/stdp* on the reciprocal of *stpd* with no constant term. The formal weighted least-squares model of the average probability for a cell is intuitive. Each pair's estimated probit index is weighted proportionately to its quality as measured by the reciprocal of its standard error of prediction. The simple, unweighted average estimated probability of a licensing agreement across the 1806 observations is 0.0216 or about 2.2 percent, while the average estimated probability across the 36 cells of Table X is 0.0591 or 5.9 percent.[27]

In contrast, each cell of Table XI (again based on the probit model in Table VIII) shows the probability of a licensing agreement when the potential licensor and the potential licensee sell products in the activities indicated *and nowhere else*—that is, when the firms are undiversified or focused. Thus, these probabilities are the evaluation of the cumulative normal probability distribution at the weighted averages of the estimated probit indices for the actual observations in the sample where the potential licensor and the potential licensee are not diversified, but completely focused in the categories indicated. Again, the formal weighted least-squares model of the cell average was used to estimate the average probit index from which the probability was then derived for each cell. The average estimated probability for the 25 cells for which there are observations of potential licensor/licensee pairs is 0.0286 or 2.9 percent.

There is a large range in the probabilities relative to one another, with rounded estimated probabilities ranging from 0.00, or 0 percent, to as much as 0.121, or 12.1 percent, across the cells in the two tables. Note that the numbers of observations in the ith column match up with those in the ith row because each pair of firms appears with one of the firms as the potential licensor and the other as the potential licensee, and then appears again with their roles reversed. Note also that all of the cell probabilities in both tables are highly significant statistically; that is, they are estimated with small standard errors for the estimated probit index for the cell.

Comparing Table X and Table XI, clearly when both firms are undiversified, the firms are less likely to be involved in a licensing agreement. The cells of Table X typically have higher estimated probabilities than the cells in Table XI because of the diversified pairs with high probability being averaged in several cells. But why do the diversified pairs typically have a higher estimated probability of reaching a licensing agreement? On one hand, the explanation is simple. Such a difference is to be expected because a pair of focused firms each have production

---

[27] The probabilities in Tables X and XI will in general be higher than the sample average naïve probability. That is because the average probability for each cell gives higher weight to probabilities with smaller standard errors of prediction for their underlying probit index and typically the high variance probit index estimates are for pairs where no licensing agreement was observed. In those high variance cases, typically there was a very negative probit index, and hence an essentially zero estimated probability. In Table X where all firms are used for the averages, note that the probabilities are in general higher than the average of the sample as a whole for an additional reason. The reason is that the observations in one cell are not mutually exclusive from those in another cell. For example, an observation in the cell where the potential licensor is in general chemicals while the potential licensee is in petroleum products can also appear in the cell where the licensor is in plastics while the licensee is in fibers. For a numerical example, suppose that in a population of 100 pairs there were just 10 with licensing agreements and the average predicted probability across the 100 observations was 0.10. If the 10 pairs with agreements (and high-predicted probabilities of agreements) appear in multiple subsets of licensor–licensee combinations, then the averages of the estimated probabilities will exceed 0.10 for those combinations.

in just a single industry in which licensed technology could be generated or used as contrasted with a pair of diversified firms where each could identify complementarities in their development and use of technology in any of several industries. On the other hand, the explanation is potentially more complex. We do not attempt in this exploratory effort to sort out all of the extant hypotheses about diversified R&D that might be applied here to explain the higher probability of licensing between diversified firms. However, we do find that this effect is associated with the diversification of the licensor, not the licensee. That is the finding, despite the logical expectation that the purchase of licensed technology may have greater value to a diversified firm for any of several reasons. The size of the expected return from using the acquired knowledge may be greater because the diversified firm recognizes potential applications in many different fields (Nelson, 1959; Link and Long, 1981); the riskiness of the application of the knowledge acquired may be less for the diversified firm (Arrow, 1962); the costs of applying the knowledge may be less for the diversified firm (Teece, 1980); and, the diversified firm may more rapidly apply the knowledge, thereby lessening the probability of preemption by other firms (Scott, 1993: pp. 112–115). The hypotheses notwithstanding, we discover that the diversification effect in our sample is associated with the licensor's diversification. That finding is consistent with diversification's effect on R&D because of recognition, risk, cost, and preemption.[28] The R&D thus affected produces licensable technology.

To describe the size and significance of the difference in the average probability of licensing for diversified pairs of firms (included with all pairs in Table X) and undiversified pairs of firms (Table XI), we use a formal model of the average probability for the different categories of diversified potential licensor=licensee pairs. We therefore consider the probit index estimated by the probit model in Table VIII, *pindex*, and two dummy variables. The first is the dummy variable *ddivsor* that takes the value 1 when the potential licensor (for the pair of firms among the 1806 observed pairs) is diversified, and 0 otherwise (i.e., when it is undiversified or focuses in one of the six chemical areas). The second dummy is *ddivsee* that takes on the value 1 when the potential licensee is diversified, and 0 otherwise. Diversification category averages for the probit index are determined by the equation equating each estimated probit index to the sum of a constant term, each dummy variable multiplied by its coefficient, and a random error term. As before, the error variance is estimated by the squared standard error of prediction for the probit index. The homoskedastic model of the category averages is then estimated, for the 1806 observations, in Eq. (3) with standard errors shown below the coefficients.

$$\begin{aligned} pindex/stdp &= \underset{(0.0599)}{-2.44}(1/stdp) + \underset{(0.0773)}{0.965}\,ddivsor/stdp + \underset{(0.0759)}{0.0657}\,ddivsee/stdp \qquad (3) \\ R^2 &= 0.0798 \\ F_{2,1803} &= 78.15 \end{aligned}$$

[28] Baldwin and Scott (1987: p. 94) review early empirical literature that on the whole does not provide empirical evidence supporting the hypothesized link from diversification to R&D activity. However, Scott and Pascoe (1987) show that when purposive diversification (nonrandom combinations of complementary lines of business) is distinguished from random diversification, R&D investments are significantly different, and overall greater, for the purposively diversified firms and total factor productivity growth increases with the R&D investments in complementary activities.

The constant term and the coefficient on the dummy variable for diversification by the potential licensor are both very highly significant statistically, with diversification by the licensor having a large positive effect on the probit index and hence on the probability of a licensing agreement. The coefficient for the dummy indicating diversification by the potential licensee is positive, but not significant.

For the undiversified pairs among the 1806 pairs of firms, the average probability of a licensing agreement is 0.007243 (corresponding to the probit index of –2.445). When the potential licensee is diversified, but the potential licensor is not, the average probability for such pairs is not significantly different and is 0.008680 (corresponding to the probit index of –2.379). In the sample of 1806 pairs, when the potential licensor is a diversified firm but the potential licensee is an undiversified firm, the average probability of a licensing agreement is significantly greater (economically and statistically) and equals 0.06944 (with –1.48 being the corresponding probit index). If both firms in the pair are diversified, then the average probability of a licensing agreement is 0.07868 (corresponding to a probit index of –1.414).

Product diversification of firms in our sample increases the probability of licensing. The pronounced effect on the probability of a licensing agreement comes when the potential licensor is diversified. For the probabilities derived from the actual averages of the fitted probit indices, if neither firm is diversified, the probability of a licensing agreement is 0.7 of 1 percent. If only the potential licensor is diversified, the probability increases by an order of magnitude to 7 percent.

V. CONCLUDING OBSERVATIONS

To date, analyses of technology flows have considered almost exclusively technology that is embodied in new capital equipment—and there is a large literature related to the adoption of computer equipment that falls under this category[29]—or embodied in new employees.[30] Previous analyses of external technology flows via licensing have been hampered by a lack of large samples of empirical data. Information about technology acquired through licensing agreements addresses some fundamental new questions. How frequently does licensing occur? Is licensing a complement or a substitute for other technology acquisition strategies? Is licensing a phenomenon specific to the technology area? Researchers have been limited in the past in their ability to consider such questions because of the lack of available data. However, as we demonstrate here, patent citation data, which are readily available, can be a valuable source of information from which to predict cross-firm licensing patterns and hence technology flows. We have derived a probability map for licensing agreements, and with the estimated probabilities of licensing patterns, we have also shown that a more diversified potential licensor has a higher probability of participating in licensing agreements.

Although our model fits the data well, showing the circumstances where licensing occurred in our sample, we cannot assume that the model would perform well predicting the probability of licensing agreements in other samples. Our method allows *historical* understanding of technology flows from licensing agreements, showing both the location of the flows and the circumstances where flows are most likely. However, further research is needed before we can

[29] See, for example, Link and Scott (1998a).
[30] See, for example, Scherer (1982a, 1982b) and Siegel (1997, 1998, 1999).

know whether the method will predict well the technology flows from licensing agreements in the future or in other industries.

References

Arora, A. (1997) “Patents, licensing, and market structure in the chemical industry”, Research Policy.

Arrow, K.J. (1962) “Economic welfare and the allocation of resources for invention”, In: The Rate and Direction of Inventive Activity (Universities — National Bureau Committee for Economic Research, Princeton University Press, Princeton).

Baldwin, W.L. and Scott, J.T. (1987) Market Structure and Technological Change (Harwood Academic Publishers, Chur, Switzerland).

Bozeman, B. and Link, A.N. (1983) Investments in Technology: Corporate Strategies and Public Policy Alternatives (Praeger Publishers, New York).

CHI Research, Inc. (1996) TECH-LINE-CD: Indicators of Technological Excellence Manual, revised.

Cohen, W.M. and Levinthal, D.A. (1989) “Innovation and learning: The two faces of R&D”, The Economic Journal.

Fu, S. and Perkins, D.S. (1995) “Technology licensors and licensees: Who they are, what resources they employ, and how they feel”, International Journal of Technology Management.

Greene, W.H. (1997) Econometric Analysis, 3rd ed. (Prentice Hall, Upper Saddle River, New Jersey).

Griliches, Z. and Lichtenberg, F. (1984) “Interindustry technology flows and productivity growth: A reexamination”, The Review of Economics and Statistics.

Geroski, P.A. (1991) “Innovation and the sectoral sources of UK productivity growth”, The Economic Journal.

Hall, B.H., Link, A.N. and Scott, J.T. (2000) “Universities as research partners”, National Bureau of Economic Research Working Paper No. 7643, April.

Hagedoorn, J., Link, A.N. and Vonortas, N. (2000) “Research partnerships”, Research Policy.

Jaffe, A.B. (1986) “Technological opportunity and spillovers of R&D: Evidence from firms’ profits and market value”, American Economic Review.

Link, A.N. and Long, J.E. (1981) “The simple economics of basic scientific research: A test of Nelson’s diversification hypothesis”, Journal of Industrial Economics.

Link, A.N. and Rees, J. (1991) “Firm size, university-based research, and the returns to R&D”, Small Business Economics.

Link, A.N. and Scott, J.T. (1998a) “Assessing the infrastructural needs of a technology-based service sector: A new approach to technology policy planning”, STI Review.

Link, A.N. and Scott, J.T. (1999) Development of an Industrial Database on Licensing Patterns, Final report submitted to the National Science Foundation, Division of Science Resources Studies, Research and Development Statistics Program, SGER Project 9615976, July.

Link, A.N. and Scott, J.T. (1998b) Public Accountability: Evaluating Technology-Based Institutions (Kluwer Academic Publishers, Norwell, MA).

Link, A.N. and Scott, J.T. (2001) "Public=private partnerships: Stimulating competition in a dynamic market", International Journal of Industrial Organization.

Link, A.N. and Zmud, R.W. (1987) "External sources of technical knowledge", Economics Letters.

Maddala, G.S. (1983) Limited-Dependent and Qualitative Variables in Econometrics, (Cambridge University Press, Cambridge, England).

Nelson, R.R. (1959) "The simple economics of basic scientific research", Journal of Political Economy.

Scherer, F.M. (1982a) "Inter-industry technology flows and productivity growth", The Review of Economics and Statistics.

Scherer, F.M. (1982b) "Inter-industry technology flows in the United States", Research Policy.

Scott, J.T. (1993) Purposive Diversification and Economic Performance, (Cambridge University Press, Cambridge, England).

Scott, J.T. and Pascoe, G. (1987) "Purposive diversification of R&D in manufacturing", Journal of Industrial Economics.

Siegel, D.S. (1997) "The impact of investments in computers on manufacturing productivity growth: A multiple-indicators, multiple-cause approach", The Review of Economics and Statistics.

Siegel, D.S. (1998) "The impact of technological change on employment: Evidence from a firm-level survey of Long Island manufacturers", Economics of Innovation and New Technology.

Siegel, D.S. (1999) Skill-Based Technological Change: Evidence from a Firm-Level Survey (W.E. Upjohn Institute, Kalamazoo, MI).

Teece, D.J. (1980) "Economies of scope and the scope of the enterprise", Journal of Economic Behavior and Organization.

U.S. Congress (1996) Office of Technology Assessment, The Effectiveness of Research and Experimentation Tax Credits (Office of Technology Assessment, Washington, DC).