How do firms manage knowledge sharing while avoiding knowledge expropriation in interfirm collaborations?

Bruce A. Heiman
San Francisco State University
College of Business
1600 Holloway Avenue
San Francisco, CA 94132
bheiman59@yahoo.com

Jackson Nickerson
Washington University in St. Louis
Olin School of Business
Campus Box 1133
1 Brookings Road
St. Louis, MO 63130
nickerson@olin.wustl.edu

July 1, 2002

[*]This paper is based in large part on Heiman's dissertation, Chapters 4-11,

How do firms manage knowledge sharing while avoiding knowledge expropriation in interfirm collaborations?

Abstract

Interfirm collaborations can raise a fundamental dilemma. To create value, collaborators may have to adopt a variety of practices to facilitate the transfer of knowledge, but in doing so may increase the likelihood that knowledge which is (1) beyond the scope of the collaboration, and (2) difficult to legally protect, is expropriated. How can firms manage this dilemma? The purpose of this paper is to empirically examine the veracity of a chain of propositions addressing this dilemma based on a joint knowledge-based view/transaction cost economics framework. A plausible chain of relationships is proposed and explored empirically using two unrelated datasets. The chain links two knowledge-based attributes of a collaboration—knowledge tacitness and problem solving complexity—to the use of knowledge management practices—high-bandwidth communication channels and co-specialized communication codes. These practices are economizing responses to the knowledge sharing difficulties rated by tacitness and complexity. However, increasing knowledge transparency gives rise to opportunism hazards, which must be safeguarded against via economizing governance choice. The empirical effort examines the effects that two knowledge attributes of collaborations have on governance choice, first directly and then indirectly through the intervening links. Empirical results from both datasets indicate substantial support for the proposed chain of propositions.

1. Introduction

Much technological collaboration between firms involves knowledge sharing. For example, Shuen (1993) describes a non-equity-based collaboration in the semiconductor manufacturing industry. To facilitate knowledge transfer, the firms co-located teams in a facility that allowed them to work together closely on a regular basis. Shuen observed that one collaborator (a U.S.-based firm) concentrated on obtaining knowledge that focused on deliverables directly associated with the collaboration's stated goals while the other collaborator (a Japan-based firm) focused on acquiring process-related knowledge of the U.S. firm’s new production technology. While acquisition of deliverables-related knowledge by the American firm appeared to be within the agreed-upon scope of knowledge sharing, the transfer of deep knowledge regarding the new production technology was not directly relevant to the stated goals of the collaboration. Moreover, Shuen’s description of the collaboration provides no evidence that either firms’ employees were explicitly aware of the other firm’s knowledge acquisition intentions or of the extent of knowledge ultimately transferred.

This illustration highlights a fundamental dilemma in the management of interfirm collaborations. Technological collaboration often involves the transfer of tacit knowledge to solve complex problems, which requires the adoption of practices, like co-location, that promote the transfer of knowledge. Yet, such practices and the socialization processes they invoke make regulation of the depth and scope of knowledge transferred difficult. Although some knowledge may be legally protected, for example, by patents, much knowledge remains unprotectable and thus is subject to potential expropriation by collaborators. A real possibility is that the value created by the collaboration from transferring knowledge may be eclipsed by the value of the knowledge expropriated—a loss in competitive advantage may result, if the dilemma is not well managed.

How can firms manage the transfer of agreed-upon knowledge in their collaborations while at the same time avoid the expropriation of other, not-agreed-upon, yet economically valuable knowledge? This dilemma is fundamental to collaborations involving knowledge sharing and is increasing in frequency of occurrence as well as importance as the popularity of collaborations, particularly in industries like biotechnology, new materials, and information sciences where knowledge transfer and creation are the central activities, has risen dramatically in recent years (e.g., Hagedoorn 1993). What advice can theory offer to those involved in managing this dilemma?

Two frameworks frequently called upon to inform the management of interfirm collaborations are the knowledge-based view (KBV) and transaction cost economics (TCE). These frameworks, however, typically focus on one part of the dilemma or the other but not both simultaneously. For instance, KBV scholars typically focus on the beneficial effects of hierarchy for reducing knowledge transfer problems while TCE scholars focus on the benefits of hierarchy for mitigating contracting problems.[1] Without a more unified treatment of the dilemma, managers may find it difficult to develop balanced prescriptions from these two separate theories. However, Heiman and Nickerson (2002) (hereafter referred to as HN) recently

---

[1] Knowledge-based theorists have highlighted the benefits of hierarchies for facilitating knowledge transfer (e.g., Arrow 1974; Kogut and Zander 1992;1996; Nahapiet and Ghoshal 1998). Hierarchy economizes on the costs of knowledge production by forming common language, identity, and affiliative bonds that promote knowledge transfer (Nickerson and Zenger 2002). In contrast, transaction cost economists focus on exchange hazards and argue that hierarchy provides safeguards against the opportunistic expropriation of transaction specific investments (Williamson 1985, 1991, 1996; Oxley 1997); knowledge attributes of the exchange that lead to contracting hazards, however, have not been considered. Thus, as Conner contends (1991; 139), KBV views integration of an activity within the boundaries of the firm as a "creator of a positive" whereas TCE views such integration as an "avoider of a negative."

introduced an extension to TCE analysis that jointly considers the problems of knowledge transfer and knowledge expropriation in interfirm collaborations. They argue that collaborations involving knowledge sharing can be characterized by at least two knowledge transfer attributes—the degree of knowledge tacitness and the level of problem-solving complexity. These attributes engender knowledge transfer problems due to cognitive limitations of man (Simon 1945).

In response to these problems, HN predict that managers adopt knowledge management practices (KMPs) such as (1) communication channels that are increasingly high in bandwidth and (2) communication codes that are increasingly co-specialized, as the degree of tacit knowledge and problem solving complexity increases. They argue that these KMPs provide communication channels and codes that economize on man's cognitive limitations by lowering the cost of transferring tacit knowledge and lowering the cost of communication associated with problem solving—a choice that they equate to choosing the most cost-efficient production technology with respect to the collaboration's knowledge attributes. However, the KMPs adopted on cognitive and "production" efficiency grounds give rise to contracting hazards once the behavioral assumption of opportunism is considered. They predict that the governance choice for the inter-firm collaboration is made directly with respect to these contracting hazards, an orthodox TCE assertion. Thus, HN's theory proposes that the knowledge-sharing/knowledge-expropriation dilemma is best managed by adopting KMPs that cost-effectively transfer knowledge but also by adopting governance structures that safeguard against the expropriation of knowledge that KMPs make possible. They predict that as the collaboration's knowledge attributes change from low to high levels of knowledge tacitness or problem solving complexity, the choice of KMPs changes from low- to high-bandwidth communication channels and from generic to co-specialized communication codes, and the governance mode changes from non-equity to equity based collaborations.

We utilize two distinct data sets to empirically examine the issues. Our first analysis employs the CATI database of interfirm collaborations (Duysters and Hagedoorn 1993). The database is sufficiently rich to allow us to examine the relationship between knowledge attributes of the collaboration and governance choice (i.e. equity v. non-equity based alliances) for the nearly 3,000 observations with relevant and complete data. Using logit analysis, we find substantial support for the propositions that link knowledge attributes to governance choice. Equity-based collaborations become increasingly likely, compared to non-equity based collaborations, as our proxy for knowledge tacitness increases. Similarly, equity-based collaborations become increasingly likely as our proxy for problem solving complexity increases. However, we temper the conclusion that can be drawn from these findings by the qualification that we find a significant and negative interaction effect when both high tacitness and high problem solving complexity are present. Besides this qualification, an unfortunate drawback of the CATI database is that it contains no information on the adoption of KMPs.

To ameliorate this empirical gap, we surveyed 36 current and former collaboration participants and/or managers responsible for overseeing or participating cumulatively in over 100 collaborations. While 36 filled out a questionnaire about managing interfirm collaborations, 18 of the managers consented to interviews. The in-depth interviews provide qualitative support that not only highlights the importance of the knowledge-sharing/knowledge-expropriation dilemma for managing interfirm collaborations, and suggest the importance of HN's theory for managing this dilemma. Moreover, OLS and logit analysis of survey responses provide support for all of the relationships predicted by HN.

In sum, the empirical efforts reported herein strongly suggest that the HN logic has currency. The theory indicates that collaboration managers need to consider problems stemming from both cognitive limitations *and* opportunism when designing the structure and management of collaborations. Thus, collaboration managers should consider adopting KMPs when they need to overcome cognitive limitations stemming from tacit knowledge and problem solving complexity. However, managers should enact these practices only when protective governance structures can be put in place or else managers will suffer with increased expectation of expropriation. When tacit knowledge and problem solving complexity are not present in the collaboration, costly KMPs and governance can and should be avoided.

The remainder of this work proceeds as follows: Section 2 presents a brief overview of the theory, developed more completely by Heiman and Nickerson (2002). Section 3 summarizes the empirical findings of our first study, which employs the CATI database. Section 4 discusses the in-depth interviews and analyzes the survey data from the 36 participants. Section 5 discusses implications of the findings, remarks on issues surrounding the analysis and data, and suggests future research.

2. Theory in brief

The topic of knowledge and its management is the focal point of a growing debate in the literature on the theory of the firm and on the role that knowledge plays in governance choice. Theorists espousing a knowledge-based view (KBV) of the firm (e.g., Conner 1991; Kogut and Zander 1992; Conner and Prahalad 1996; Kogut and Zander 1996) maintain that bounded rationality acts *sans* regard for opportunism as it influences managerial choice of governance for a particular transaction. In dispute, theorists relying on transaction cost economics (TCE) reasoning (e.g., Williamson 1985, 64-7; Foss 1996a, 1996b) maintain that an organization's governance mode decisions cannot be explained without relying on the additional behavioral assumption of opportunism. HN's theory (2002) grew out of this debate. Their theory is not formal, but the informal arguments offer an approach to the reconciliation of two theories previously interpreted as exclusively in tension (for an exception see Nickerson and Zenger 2002). Below we describe how HN's theory links together KBV and TCE arguments, provides definitions of key terms, and we summarize the theory's key predictions.

HN's theory proposes a four–link chain of relationships, which are diagrammed in Figure 1. As indicated in the figure, linkages A-B and B-C operate in the realm of KBV, which assumes people are well intentioned (no opportunism) but limited in their cognitive abilities to process information. Linkages C-D and D-E are based on TCE, which not only assumes limited cognitive abilities, but also allows for the possibility of opportunistic behavior. For simplicity in exposition, we describe this chain sequentially although the theory assumes that all decisions in the chain are, in essence, simultaneous.

Box A is the starting point for the chain of relationships. Knowledge attributes in Box A refer to the degree of knowledge tacitness and problem solving complexity found in a collaboration; these are treated as exogenous.[2]

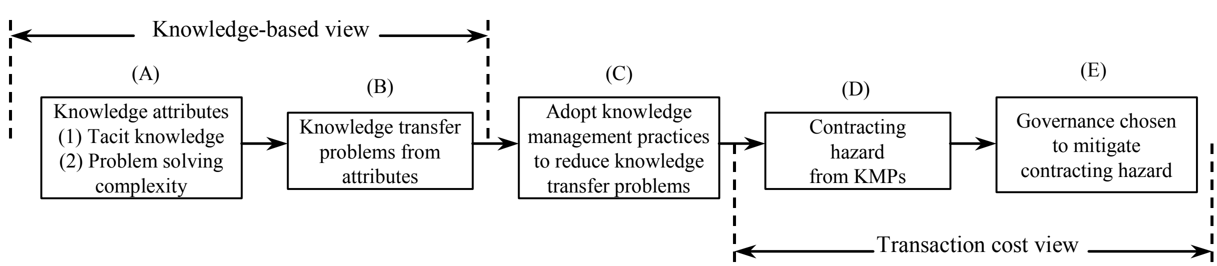

Figure 1: Proposed chain of relationships

[2] In other words, the theory is developed conditioned on the choice of entering into a collaboration and the choice of collaborator. Presumably, the theory can be used to inform both these choices as well as those under

Knowledge tacitness refers to how particular knowledge is stored with respect to three dimensions: human/physical, informal/formal, and linguistic/formulaic.[3] The cost to transfer knowledge involves, among other things, costs from verifying that the recipient accurately understands the knowledge received.[4] Such costs are a reflection of the limited rate at which individuals can assimilate knowledge. With knowledge that is represented via physical, formal, or formulaic means, verification costs are low as the recipient can verify the knowledge by inspecting its physical representation, referring to its formal representation, or logically interpreting its formulaic representation. In contrast, when knowledge is stored in humans' heads, informally, or linguistically, the recipient can only verify knowledge by transferring it back to the sender who must then provide the verification. The entire transmission sequence must be repeated if the recipient's initial understanding of the knowledge is inconsistent with knowledge received by the original sender, which incurs additional costs. Thus, verification costs are likely to increase with the tacitness of the knowledge transmitted, which increases knowledge transfer costs.

Problem-solving complexity concerns the degree to which problem solving in an exchange is complex.[5] Assuming that relevant knowledge for solving a problem is dispersed

---

primary consideration, but this is beyond of the scope of the theory and the empirical analysis reported herein.

[3] Heiman (2002) argues that three "poles" of these dimensions (human, informal and linguistic) are positively correlated with the cost of knowledge transfer because they indicate the degree of tacit knowledge, which has implications for the cost-effective method of transfer. We note that though these indicators may converge somewhat at the tacit end of the spectrum, perhaps owing the nature of tacitness itself, they are clearly distinct at the non-tacit end of the spectrum, which renders them suitable for use as measures. We can operationally specify knowledge as embodying a high degree of tacitness by looking at the extent to which it is not stored via physical, formal and/or formulaic means. Asking individuals, who possess or are aware of certain knowledge and how that knowledge is stored does not require full elucidation of the knowledge, resulting in empirically tractable dimensions. The indicators we describe provide a tractable but second-best metric for indicating the relative cost to transfer tacit knowledge.

[4] Costs of knowledge transfer that are not considered herein include the costs of indexing, maintaining, and upgrading (including supplanting) knowledge.

[5] We define complexity defined based on an adaptation of Simon's (1962) seminal contribution, which

among actors, such complexity is subject to cognitive limits of the accumulation and application of knowledge, which increase knowledge transfer costs as complexity increases. Box B indicates that knowledge transfer problems increase as the level of tacit knowledge or problem solving complexity rise.

Box C of the proposed chain of relationships indicates that KMPs can be used to economize on cognitive limitations as knowledge transfer problems increase. The KMPs represented by Box C are administrative apparatuses that serve to increase knowledge transparency between collaborators in response to the knowledge transfer problems in Box B. HN focus on two KMPs: communication channel bandwidth and communication codes.

Channel Bandwidth: The physical aspect of information channels refers to the bandwidth of communication that can be delivered by alternative means. The bandwidth of a transmission—hereafter referred to as an interaction—refers to the degree of intensity of communication among individuals. Although the choice of bandwidth falls along a continuum, for simplicity of exposition only high- and low-bandwidth channels are considered here. High-bandwidth interactions involve tight, rich, immediate interfaces between partners, and are generally relatively costly in terms of time and effort. A high-bandwidth interaction is intended to facilitate a high degree of rich context, high-affect, and high-transparency in communications between partners. High-bandwidth interactions provide easy access to knowledge by allowing, for example, physical demonstrations, immediate redundancy and restatement (rephrasing), high clarity, rich contextual cues, high interactivity, and clear emphases. As an example, von Hippel's

---

explains complexity in natural hierarchic systems and how this may be applied to organizations. Complexity exists in a collaboration if, in order for successful execution of a collaboration's mission to be achieved, distinct knowledge sets must be combined, which results in new knowledge, and the new knowledge has long-term performance implications for the collaboration. Complexity increases with the number of distinct knowledge sets relevant to solve a problem and produce new knowledge—for a more extensive elaboration see Heiman (2002).

(1994) illustration of co-location in response to problems with "sticky" information represents a polar form of high bandwidth interaction.

Low-bandwidth interactions, in contrast, are characterized by low-context communications measures: for example, email, faxes, letters, and phone calls. Low-bandwidth interactions offer little clarity of emotional affect, low or no redundancy, time-lagged (sometimes severely) queries and responses, low interactivity, very low or no contextual cues, and comparatively less clear emphases. Generally, low-bandwidth interactions provide low degrees of knowledge transparency at comparatively low cost, versus high-bandwidth interactions, which provide high transparency albeit at a high-cost.

<u>Communication codes</u>: Arrow (1974: 39) claims that in addition to the choice of communication channel, an important aspect of information costs are the costs of rendering human minds suitable to receive coded signals that comprise knowledge to be transferred. Individuals have to make an adequate investment of time and effort to be able to distinguish one signal from another and these are in part idiosyncratic and irreversible capital costs. Use of codes allows the sender to transform his knowledge into a signal and allows the recipient of a signal to understand it. The mutual adoption of codes is a form of idiosyncratic co-specialization that lowers knowledge transfer costs by increasing mutual understanding, which increases the extent of knowledge that can be shared between partner-firms. In collaborations, co-specialization represents a specific, irreversible investment by a partner-firm in its relation with another partner-firm. Another example of co-specialization is found in an international collaboration, if members of one or several partner-firms undergo cultural and language training to enhance their ability to work with and understand another partner's ways of thinking and working. A trained partner's team-members will have invested in codes to positively affect the

efficacy of knowledge transfer in the collaborative activity. HN argue that the choices of bandwidth and codes are KMPs that can be adopted to attenuate knowledge transfer problems associated with tacit knowledge or problem solving complexity.

The adoption of KMPs, HN argue, gives rise to opportunism hazards via increased knowledge transparency in collaborations. In the next part of the proposed chain (Linkages C-D and D-E), an economical governance mode is chosen (E) in response to the threat of opportunism hazards (D) resulting from (managerial) deployment of KMPs, a "transaction cost" perspective (see Williamson 1975; 1985; 1991). That is, KMPs generally increase knowledge transparency; yet increased knowledge transparency can give rise to a hazard of opportunism between collaborating firms (e.g., unverifiable, unintended over-transfer of knowledge). These hazards must be factored into the choice of governance. Since much of the research on governance of collaborations focuses on the key differences between equity and non-equity based arrangements, HN limit their discussion for the purpose of exposition to the effects that KMPs have on the choice between these two governance modes.

Equity-based governance, wherein ownership of rights to profits is shared via common stock ownership, provides incentive-aligning properties, but high-powered incentives of market-like agreements are sacrificed and bureaucratic costs incurred. Non-equity based collaborations do not involve shared ownership, but a wide variety of long- and short-term contract-types are feasible. Equity-based governance (e.g., an equity joint venture) provides comparatively stronger safeguards than non-equity, and incurs comparatively lower costs than non-equity governance when knowledge attribute intensity is deemed sufficiently high to create an opportunism hazard. Oxley (1997) does not directly address the knowledge attributes of transactions, but her work on collaborations identifies opportunism hazards in collaborations and shows that governance is responsive to these hazards using a transaction cost economics logic.

With this background we now summarize HN's proposed chain of relationships. The first two propositions reflect the over-arching relationship between knowledge attributes and governance choice:

Proposition 1: As knowledge tacitness increases in a transaction, an economizing governance mode will be chosen so as to safeguard against the hazard of opportunism.

Proposition 2: As complexity increases in a transaction, an economizing governance mode will be chosen so as to safeguard against the hazard of opportunism.

The remainder of the propositions unpacks the interior of the proposed chain of relationships:

Proposition 3a: As the degree of knowledge tacitness associated with collaborative activities increases, the costs associated with knowledge transfer increase.

Proposition 3b: As the degree of complexity associated with collaborative activities increases, the costs associated with knowledge transfer increase.

Proposition 4a: As the costs associated with knowledge tacitness in collaborative activities increase, organizations increasingly utilize the knowledge management practice of high-bandwidth communication channels to mitigate these costs.

Proposition 4b: As the costs associated with joint problem-solving in collaborative activities increase, organizations increasingly adopt the knowledge management practice of investing in idiosyncratic communication codes to mitigate these costs.

.

Proposition 5: As participants in collaborations increasingly utilize knowledge management practices to alleviate knowledge transfer costs, opportunism hazards arise.

Proposition 6: As the hazard of opportunism increases in an inter-firm collaboration, the likelihood of choosing equity-based governance to safeguard against opportunism increases.

3. Method and Results from CATI Database

Our first empirical methodology examines Propositions 1 and 2—the overarching relationship described by the proposed causal chain in Figure 1. Thus, we examine the relationship between knowledge attributes and governance choice through an analysis of an international database of inter-firm relationships, the CATI (Cooperative Agreements and Technology Indicators) database, which was developed by Duysters and Hagedoorn (1993). The CATI database contains over 9000 observations of publicly announced inter-firm relationships, with over 100 alpha-numerically coded "raw" data fields and two descriptive text fields for each relationship.[6] Publications tracked in CATI, which are the source of reported observations, consist mainly of popular (business) press and industrial trade magazines, for example, The Wall Street Journal (popular press) and Communications Weekly (industrial trade publication). Importantly, the database includes information on the governance mode adopted and the type of activity undertaken for each collaboration. While the database has a substantial number of advantages, it nevertheless has at least two important limitations. Regrettably, the CATI database provides no record of KMPs employed in inter-firm collaborations, which precludes an investigation of the interior portion of the proposed causal chain with this database. The second disadvantage is that the database manifests a variety of over-sampling issues, which shape our empirical methodology.

Sample Frame. The sample frame originates from the universe of all publicly announced alliances between 1977 and 1989, which excludes by its nature, unannounced alliances, but

[6] The database was constructed at Universiteit Maastricht from public announcements of collaborations. Universiteit Maastricht graduate students working primarily over a period of two years coded the database. The MERIT-CATI database was created during 1992 and 1993 and looks retrospectively at announcements of inter-firm collaborations for prior years as allowed by available publications for a given year. The data runs through 1989.

nonetheless represents the best, least biased, large source of data available. Data was collected solely from English language sources, suggesting that U.S., Japanese, and European partner-firms in collaborations may be over-represented, collaborations involving eastern Europe, Asia, South Asia and Latin America firms may be under-represented.

Inspection of the data shows a substantial reduction in the number of reported observations of inter-firm relationships prior to 1977, which may suggest under-reporting of collaborations prior to 1977.[7] Despite these drawbacks, the authors of CATI assert that it "produce[s] a clear picture of the joint efforts of many companies" (Duysters and Hagedoorn 1993, 1). Geographic, industry, firm size, and temporal over-sampling issues exist in the CATI database. We attempt to compensate for these potential biases with our methodology and inclusion of a large number of controls for over-sampling. The analysis also assumes that CATI data reflects firms' mostly correct governance choices. The analytical methods used are in no way particularly sensitive to the type of biases described, since our data contains adequate variation. Additionally, over-sampling affects only the constant term in logit models, which justifies the choice of logit modeling over other discrete choice methods (e.g., Probit).[8]

Coding errors. Another limitation of the CATI database is potential coding errors. Coding took place over several years with a variety of coders, which may lead not only to random errors but also to systematic errors.[9] Once the elimination of missing and questionable

---

[7] A newer version of the data that runs through 1994 is available, but it lacks important modes of non-equity governance; for example, licensing and cross-licensing are absent from all records subsequent to 1989. Data from 1990 to 1994 is deemed unsuitable for use.

[8] Thanks to Trond Petersen, U.C. Berkeley, Haas School of Business, for this insight.

[9] Several relational database-programming techniques combined with content analysis using keyword string-searches, were used to search for coding anomalies, which were corrected when found. In the course of cleaning the data, several thousand records were examined manually and the values of measures adjusted for accuracy, when appropriate.

values from all pertinent measures is completed, 2976 of 9143 observations remain useable in the sample. The following section briefly discusses the construction of variables.

Dependent variable: *Equity*. Equity, our dependent variable, is constructed as a dummy variable using values from a CATI raw-data field that lists the form of cooperation for each collaboration. This binary governance variable is constructed using six reasonably clear indicators of governance. For Equity=1, joint-equity ownership via joint venture or research corporation is required. For Equity=0, unilateral non-equity governance is the category for which the data are most unambiguously accurate. Unilateral non-equity governance consists of customer-supplier partnerships, licensing, co-makership contracts and research and development contracts.[10]

Independent variables: *Tacit*. Tacit knowledge is an inherently difficult attribute to measure. Moreover, the originators of the CATI database provide no obvious variables to indicate the extent to which knowledge to be transferred during exchange is tacit. To circumvent this measurement problem, a proxy for tacit knowledge is developed, based on the nature of the primary activity of the transaction. CATI categorizes the nature of activities based on an ostensibly ordinal "Distance to Market" measure. The four relevant categories in this measure are: (1) Basic Research, (2) Development (Applied Research), (3) Product Design, and (4) Production or Marketing. (Two additional categories, Unknown and Both Upstream and Downstream Activities, are omitted from the analysis owing to their lack of precision). We

[10] Other governance modes are not used owing to the lack of sufficient information or reliability regarding the nature of governance; they are treated as missing data. In the context of a logit model, say a Multinomial one with multiple categories of the dependent variable, restricting the analysis to cases having a subset of all the dependent variables' values gives correct estimates of the logit coefficients for the categories of the dependent variable used. The estimated logit coefficients should be very similar to those we might have obtained had more values for the dependent variable been included (Maddala, 1985). Sample selection bias is not anticipated to be a problem.

maintain that these activities differ fundamentally in the extent to which tacit knowledge is transferred. Since little confidence in Distance to Market's given rank ordering is warranted, we construct a variable that categorizes activities into those for which we expect high levels of tacit knowledge versus those we expect to display low levels of tacit knowledge.[11] Of the four activity-types captured by CATI, Development (Applied Research) involves more tacit knowledge transfer than any of the other activities discussed. Applied research "…entails alterations and enhancements to existing firm assets, production processes, and products," which contain a substantial degree of tacit knowledge (Helfat 1994, 174). Basic research, by contrast, involves the intentional distribution of knowledge in codified (physical) formats (Nelson 1959), suggesting low tacitness. Contrasting development and product design, the focus is on implementation and codification. Although actors transfer knowledge to design and integrate components, much of this knowledge is codified (e.g., technical documentation including engineering or production drawings, and specifications). Hence, knowledge transferred in the design activity is likely to be less tacit than that transferred for development.

Similarly, the exchange of knowledge in production or marketing is less likely than development to involve the transfer of tacit knowledge. Marketing knowledge is comparatively non-technical and easier to codify (store physically) than knowledge involved in development for

[11] Others have used CATI's Distance to Market measure. For example, Gulati (1995) used Distance to Market as a control for asset specificity. However, Distance to Market identifies the activity, not asset specificity. Indeed, CATI provides insufficient information to make any claim about the extent to which firms invest in co-specialized assets; such uses of Distance to Market are suspect. Oxley (1997, 75) claims that transaction involving (1) design and (2) combined upstream and downstream activities are more likely than production to involve the creation or modification of technology, which complicates the specification of contractual terms. Her analysis, however, does not consider basic research and development, while this analysis excludes her category "both upstream and downstream activities." This is done in order to avoid possible confusion over (1) the predominant activity in a transaction and (2) single cases within CATI which describe two discrete activities and/or two governance outcomes (e.g., JV, licensing). The latter data condition lacks, in particular, one-to-one mappings of activity to governance mode. This work's use of Distance to Market neither conflicts with nor contradicts Oxley's use of the data.

application. Also, routine production knowledge is typically codified unless it is in a development phase, in which case it would then be classified as such. Product design and marketing/production thus involve the use of more codified materials than applied development; that is, applied development involves more knowledge that is embedded in human's brains than product design, marketing, or production.[12] Hence, Tacit=1 for all Development (Applied Research) transactions, otherwise Tacit=0.

Complexity. Like tacit knowledge, problem-solving complexity is an inherently difficult attribute to measure. Fortunately, the creators of CATI coded a binary variable (STRAT) to indicate when new, valuable, and strategic knowledge (defined as expected to have a long-term impact on performance) is expected to result from combining distinct collaborator-contributed knowledge.[13] Manual random checking of CATI's two descriptive text fields suggests that STRAT identifies when firms announce the expectation to create new knowledge.[14] STRAT is only weakly correlated with Tacit (Pearson's $r=0.18$, $p\leq 0.05$).

---

[12] The construction of this measure is explicated in much greater detail in Heiman (2002, Ch. 5). As mentioned above, development, on average, involves comparatively more tacit knowledge transfer than any of the other three activities discussed. Less certainty exists about rank ordering among basic research, design, and marketing/production. To be conservative, inclusion of all three activities in the construction of the base case (no or low knowledge tacitness—Tacit = 0) is warranted.

[13]The CATI variable STRAT is not part of the standard CATI data typically licensed from MERIT at Universiteit Maastricht. The author expresses gratitude to Geert Duysters and John Hagedoorn, MERIT, Universiteit Maastricht for making this data available. Duysters(1999), in a personal communication described his use of STRAT: to divide the CATI dataset into distinct subsets for separate analysis; STRAT is used as a basis to divide collaborations into two sorts: those that produce a distinct new body of knowledge with long-term (strategic) value, and those that do not.

[14]The coding of STRAT was verified by randomly spot-checking CATI's two free text fields (AIM and INFO) for confirmatory evidence contained in these descriptive remarks pertaining to collaborations. While cleaning the data, the author randomly verified and corrected cases for STRAT=1 and STRAT=0. For a given record (case), if STRAT=1 the content of the text fields or the name of the collaboration must mention or imply long-term strategic value via creation of a distinct body of knowledge based on partner-knowledge contributions, often in response to a complex technical problem. Corrections to STRAT were made based on this rule and its converse for STRAT=0. Although occasional errors in coding were detected and corrected, the error rate was low (< 20%) and no systematic errors were discerned, suggesting that errors in coding of STRAT are random in character; the data is suitable for use.

Creating new, strategic (valuable in the long term) knowledge involves inherently more problem-solving complexity than other transactions. If such knowledge creation activities were not complex they could be easily and quickly replicated and thus not produce value or be of strategic interest. Complexity = 1 when STRAT = 1, otherwise Complexity=0.

Control Variables. We include two other knowledge-related factors to further isolate the effects of tacitness and problems solving complexity. *Dispersion* is N - 2 where N is the number of firms in the collaboration. Dispersion is a proxy for how "spread-out" knowledge is among different people in a collaboration. The more disperse knowledge is the more difficult knowledge sharing will may become as the cost for knowledge sharing increases. Another factors, *Ambiguity* = 0 if T = 1 and = 1 if T > 1, where T is the number of technologies mentioned in the CATI database for each collaboration. Though only 9 of the 2976 observations involve more than two technologies, many more involve precisely two technologies. Ambiguity is a proxy for the extent to which the scope of knowledge to be transferred is difficult or impossible to delimit. As the number of technologies involved in a transaction increases, firms may experience increased difficulty identifying and delimiting relevant bodies of knowledge to transfer, which raises the cost of knowledge sharing. Since additional knowledge sharing costs stemming from knowledge dispersion and ambiguity also may manifest interactive effects with our knowledge attributes, we additionally interact with these two factors with our two knowledge attributes for a total of four interaction terms.

Other control variables. *Government* is a dummy variable set to 1if a government entity is involved in the collaboration otherwise it is set to 0. *SizeLarge* (50,000 employees or more), *SizeMedium* (5,000 to 50,000 employees) and *SizeSmall* (500 to 5,000 employees) are three dummy variables indicating collaborator size based on the number of firm employees. *Infotech,* *Biotech*, and *NewMaterial* are industry specific dummy variables that are set to one if the collaboration is in the information technology, biotechnology/pharmaceutical, or new materials industries, respectively; otherwise, each dummy is set to 0. These variables control for any in-sample temporal bias due to "trendy" topics that receive frequent media mention in a given year. *Japan*, *NorthAmerica*, and *WestEurope* are dummy variables indicating the geographic home of collaboration participants. We include 12 dummies for the collaboration founding years of 1978 through 1989 to account for temporal bias (1977 data is the base year indicated by all year dummies =0). Geographic, industry, firm size, and year dummy variables allow us to control for over-sampling issues in the CATI database. Additionally, we provide one continuous measure of political hazard by country of origin of partner-firm by calculating the unweighted arithmetic mean of each partner's home-country political risk index as determined by Henisz (1998). We include this variable to control for the possibility that credibility of national institutions systematically affects the collaboration's choice of governance. Table 1 summarizes the construction of all variables discussed above. Table 2 presents summary statistics for all measures. No problematic correlation coefficients between factors manifest (using Pearson's *r*).

Methods. A cross-tabular analysis (Table 3) does not suggest any definitive conclusions about the effects of the independent variables on the outcome variable, but it shows that sufficient entries are present in each cell for a multivariate analysis to proceed using discrete choice methods. We employ a logit model since the governance outcomes are categorical in nature; logit it is an appropriate method for data with over-sampling, and coefficients are easily interpreted. The model is estimated using the STATA binary logit routine (STATA, 1997). For the logit model the probabilities are as follows:

$$\Pr(\text{Equity}_i = 1 \mid X_i) = \exp(\beta' X_i)/[1+\exp(\beta' X_i)], \tag{1}$$

where Equity_i is governance choice for case *i*, X_i is a vector of factors composed of the control and independent variables described in Table 1, and β is the vector of estimated coefficients. Eq. (1) specifies that the probability that an equity mode of governance is chosen, given the observed data values, is calculated by applying the logit function shown to the product of the transposed vector of the estimated coefficients and the vector containing the observed data.

Results. Table 4 presents maximum-likelihood estimates of parameters for the analysis of governance mode choice using each independent variable. Model (1) shows the effects of Tacit and a constant term with no control variables on $\Pr(\text{Equity}_i)=1|\text{Tacit}_i)$. Model (2) contains only the variable Complexity and a constant term. Model (3) contains both Tacit and Complexity and a constant term. The coefficient estimates for both Tacit and Complexity are highly significant ($p < 0.01$) and signed as predicted in all three models, in support of Propositions 1 and 2. Additionally, all three models have significantly more predictive power ($p < 0.01$) than a base model consisting merely of a constant. Since Models 1 and 2 are each nested within Model 3, it is possible to calculate a likelihood ratio using the following formula: $L.R. = 2(|L_0| - |L_1|)$, where L_0 is the log-likelihood of the base model (1 or 2), and L_1 is the log-likelihood of the expanded Model (3). The resulting likelihood ratio is distributed as a χ^2 statistic with $p_1 - p_0$ degrees of freedom, where p_0 is the number of estimated parameters in the base model, and p_1 is the number of estimated parameters in the expanded model. In the instance of the base Model 1 versus Model 3, the Chi-Squared statistic is a highly significant 437.26 ($p < 0.01$, d.f.=1), suggesting that Complexity adds substantial predictive power to a model consisting of Tacit alone. Similarly, for Model 2 versus 3, Chi-Squared is 20.45 ($p < 0.01$, d.f.=1) suggesting Tacit adds significant explanatory value to a model comprised of Complexity alone.

Model (4) adds Ambiguity, Dispersion, and their four interaction terms to our baseline model. Model (5) is our fully specified model including all control variables. The coefficient estimates for both Tacit and Complexity remain highly significant ($p < 0.01$) and essentially unchanged in sign and magnitude, which indicates that our control variables do not affect our parameter estimates for Tacit and Complexity.

Model (6) adds an interaction term Tacit*Complexity. We add this term to our model to explore to what extent the governance of collaborations is affected when both tacit knowledge and problem solving complexity are present. Coefficient estimates for Tacit and Complexity are identical in sign and statistical significance but are somewhat larger than in the prior models. The interaction term, however, is significant and negative. The negative sign implies that the effect of Tacit is opposite that asserted in Proposition 1 when Complexity = 1 and vice-versa: the effect of Complexity is opposite that asserted in Proposition 2 when Tacit = 1. Thus, we empirically find that a collaboration involving both tacit knowledge and problem solving complexity experience an interaction effect not predicted by the theory. Nonetheless, the empirical findings do provide substantial support the theoretical predictions.

4. Method and Results from Interviews and Survey

Although our first analysis provides support for the overarching relationship between knowledge attributes and governance choice, the analysis provides no direct measurement of KMPs, which are proposed by HN and are under consideration in this study. Given this limitation, we decided to interview and survey collaboration managers and participants from a variety of firms. We personally administered a questionnaire, which provided rich information about personal experiences regarding knowledge in collaborations. A total of 230 people were contacted for participation in the study, 36 of whom filled in the questionnaire (15.7% response rate), which consisted of a set of closed questions and a few open-ended questions. Of these, 18 people additionally agreed to semi-structured interviews—11 in-depth interviews and seven more cursory interviews.

The sample consists primarily of (1) members from a group known as the Association of Strategic Alliance Professionals (ASAP), consisting primarily of managers and participants in collaborations, and (2) participants in a European Executive M.B.A. program who are involved in one or more inter-firm collaborations. Data was gathered from October through December 1999. Potential sample bias includes biases toward respondents who work for large firms and who work for firms engaging in multiple collaborations. Additionally, the data gathered contains an Anglo-European bias; for example, no representatives of Asian partner-firms were interviewed (though two were surveyed). The results of the interview-study, therefore, may offer limited generality.[15]

Drawbacks notwithstanding, we think the sample provides high accuracy in responses, as well as good breadth, and depth of data. For instance, clarification of any question posed by any respondent was feasible because of in-person questionnaire administration. Also, owing to the personal nature of the data-gathering, as well as commitments for strict confidentiality, some of the data obtained reflects unannounced, "secret" collaborations as well as those announced to the public, a feature that most other datasets lack, and which may increase the generality of findings

---

[15] Not surprisingly, these biases are commonly reflected in extant studies involving firms engaging in collaborations (see for example Kogut 1989; Gulati 1995; Hagedoorn and Duysters 1996; Hagedoorn and Narula 1996; Oxley 1997; Duysters and Hagedoorn 1998). Aside from the relatively common biases discussed, no additional sources of bias are expected in the data.

relative to other studies. Nonetheless, the interview data is not comprised of a scientific sample, which may limit the generality of findings.

Interviews: Below we provide some qualitative evidence from our interviews. We believe this evidence suggests not only that the knowledge-sharing/knowledge-expropriation dilemma is an important concern in collaborations, but also that HN's chain of propositions is a useful tool for understanding and managing this dilemma.

Knowledge tacitness. Evidence from the interviews suggests that the presence and degree of knowledge attributes engender knowledge transfer problems. For instance, one collaboration participant mentioned serious issues arising from an "unclear understanding of key uniqueness of our partner [sic]," suggesting that some information persistently remains hidden or inaccessible. A senior pharmaceutical manager in a collaboration involving the construction of a manufacturing plant mentioned a major knowledge transfer problem owing to tacitness: "unclear definition of content and form of documentation [as well as] lack of agreed [upon] procedures." This comment suggests high degrees of informal- and human-based knowledge storage (on the formal/informal and physical/human continua, respectively), and hence, tacitness.

An American manager of a collaboration between two small design firms (the other firm was Asian) commented on the limited utility of fully articulated knowledge, in this case, drawings and other information contained in, for example, patents:

> Regarding the adequacy of knowledge available in written form, a lot had to be developed outside of what was written. For example, patents in written form were not enough for our partner to make a product. We were telling them how to build and produce things, but the drawings only told them how the final product should look, not how to make it with consistent, high quality… Our best approach was to personally show them how something is done. [sic]

Though the knowledge in this situation reflects a substantial physical component, knowledge of how to correctly produce the product remained in people's heads; tacitness is implicated as a major factor in causing problems with knowledge transfer.

Complexity. The factor of complexity also was manifest in respondents' remarks. As discussed above, complexity exists and matters when (1) new knowledge is created through interaction of distinct existing knowledge or (2) there exist long term performance implications of the interaction. A top manager in a chemical production collaboration mentioned that their collaborations have the long-term mission of reducing production costs by jointly using partner-specific technology: "both sides have to offer options to solve problems," where "options" refers to firm-specific technical/engineering knowledge necessary to jointly address cost-problems.

Collaboration participants frequently mentioned issues arising from "lack of communications," owing to cultural, and language barriers. One manager discussed "huge problems [with respect to] managing a multi-cultural, cross-border collaboration team." The same manager added that "language barriers lead to a lack of understanding of assignments and poor communications." These statements are consistent with the idea that the exchanges are complex and also illuminate a lack of co-specialized communication codes, which, if present, could ease the communication burden.

KMPs. Interview responses suggest that not only are knowledge attribute-related problems substantial, but also that KMPs are the remedial measures deployed by collaboration managers. One collaboration participant who worked for a consulting firm involved in many collaborations remarked, "In our most strategic partnerships, there is a lot of one-to-one-ness between executives and managers." The senior pharmaceutical manager from a prior example used "frequent meetings at the manager level [*sic*]," which represents high-bandwidth communication, as a simple knowledge management practice to handle his collaboration's content definition and documentation problems.

A top collaboration manager at a large multinational chemical firm noted "Personal relationships are necessary at all levels in our collaborations to keep information flowing…Compatible cultures are also essential…Joint ventures require the same intensity or closeness [with which] an SBU [strategic business unit] is internally managed." Another top manager stated, "Language barriers lead to a lack of understanding of assignments and poor communications…There is a need for face-to-face meetings and to incur the cost and time involved. Upfront meetings, translation and education help."

The most frequently mentioned knowledge management practices include exchange of personnel, co-location, frequent in-person meetings, language or vocabulary briefings, personnel matching, and training programs. These findings support the contention that KMPs exist and solve problems with knowledge attributes.

Opportunism hazards from KMPs. Collaboration participants and managers frequently recognized the existence of opportunism hazards in environments intended (via KMPs) to foster knowledge sharing between firms. For instance, a pharmaceuticals manager mentioned "problems with partners' alignment," and that partners' desire to "capture alliance knowledge" from his firm was often in tension with his firm's need to obtain knowledge from partners. One manager in specialty chemicals stated simply,

> When sharing knowledge, leaks to partners have consequences…You take in a world-class technology to a partner. The venture becomes a conduit to the partner. The partner now has access [to the technology] and can use it freely.

Safeguarding against opportunism via governance choice. Regarding the need for safeguards, one executive from a large pharmaceutical firm observed that "Anything could be important. We place a lot of attention on patents, but know-how is more difficult to protect." A manufacturing manager of a consumer products firm commented on customer-supplier relations and knowledge: "Knowledge leakage is a big issue up front; [that is,] in contracts and using exclusivity agreements, we still have to be very, very careful." Another collaboration participant remarked "Without equity we would have been much less forthcoming with our information." These statements suggest that governance choice with respect to safeguarding against knowledge leakage indeed is very important.

While not conclusive, the evidence from interviews is supportive of the proposed chain of relationships. Collaboration managers and participants acknowledge the importance of the knowledge-sharing/knowledge-expropriation dilemma and describe various approaches to managing the dilemma. These approaches appear to be consistent with the view that KMPs are adopted to relieve knowledge sharing problems stemming from knowledge tacitness and problem solving complexity. Also, governance choices appear to be made with reference to knowledge expropriation concerns, which increase with the adoption of KMPs. Nonetheless, interviews are only suggestive so we now turn to statistical analysis of our survey data.

<u>Data and Methods</u>. The data are comprised of responses to 36 surveys, which are a superset of our 18 interviews. Respondents were asked to identify one collaboration and to respond to the questions with respect to only that collaboration. The survey consists of 20 background and 39 closed questions, which required yes/no responses or one choice of a 5-point Likert scale[16], that potentially capture up to 82 distinct "raw" data fields (some questions may evoke multiple responses).

<u>Construction of variables</u>. In general, scales were constructed by additively combining scores from multiple survey items.[17] Thus, these scales, except where noted, are ordinal

---

[16] No relative intra-scale distances or unique origin information are implied by the coding of choices (Cooper and Emory 1995). The scale contains the following choices: Strongly Disagree, Disagree, Neutral, Agree, and Strongly Agree.

[17] Inter-item reliability checking (e.g., Cronbach's Alpha) is not appropriate here because we identify the latent variable by identifying different dimensions of the latent variable rather than identifying only one dimension

measures that are assumed to carry equal weight in determining the total (additive) scale, which reflects underlying, but unmeasured, continuous effects. Table 5 presents the questions that were added to develop each scale. Rather than describe each question in detail below, we ask that the reader refer to this table for the specific items comprising each variable.

Tacit is constructed by summing dichotomously coded responses from six questions (questions 1-6 in Table 5) capturing tacitness into one scale. Tacit is designed to measure the extent to which respondents consider tacitness responsible for problems with knowledge transfer. This composite, ordinal variable has a potential range of [0, 1, 2,...,6], a mean of 2.611, a standard deviation of 1.479, and is assumed to reflect an unmeasured underlying continuous factor, degree of knowledge tacitness associated with a particular collaboration. Higher levels of the scale correspond to higher levels of knowledge tacitness.

Complexity is a composite measure similarly constructed, and is comprised of three questions described in Table 5, questions 7-9. These questions identify the extent to which collaborations create valuable and new knowledge through joint problem solving. Complexity has a potential value range of [0, 1, 2, 3], a mean of 2.389, a standard deviation of 0.645. Higher values of the scale indicate greater degrees of problem solving complexity.

Hibandwidth is a composite variable constructed to detect the degree to which high-bandwidth interactions are used in the collaboration. Hibandwidth is constructed by summing the four survey questions indicated in Table 5, questions 10-13. Additive combination of the four questions leads to potential values of [0, 1, 2, 3, 4], a mean of 2.750, and a standard deviation of 1.079, with higher values corresponding to the use of higher-bandwidth communication channels.

---

several different ways.

Co-specialize is constructed from five survey questions (14-19) and assesses if the following KMPs were expected to be employed in the transfer process: frequent dialogue between partners, letting individuals work closely together to promote easy interaction, developing social relations, site visits, and common work location. Adding the recoded scores for the five questions exploring co-specialization yields a potential range of [0, 1, 2,...,5], a mean of 3.611, and a standard deviation of 1.248, with higher values corresponding to the use of more co-specialized communication codes.

<u>*Hazard*</u> captures the respondent's perceived threat of other collaborators' opportunism by summing responses from seven survey items, questions 20-26 in Table 5. The composite variable has a possible range of [0,1,2,...,7], a mean of 3.972, and a standard deviation of 1.647, with increasing values corresponding to a greater hazard. As a related control, we construct the scale *OwnHazard* from four survey items, questions 27-20 in Table 5, which measures the extent to which the respondent's own firm may act opportunistically. This scale has a potential range of [0,1,…,4], a mean of 2.861, and a standard deviation of 0.990, with higher values indicating greater levels of perceived own-firm opportunism.

<u>*Equity*</u> represents the governance mode of an alliance and is coded as Equity=1 if respondents indicated that the type of collaboration is an equity-based joint-venture otherwise it is coded Equity=0. It has a mean of 0.167 and a standard deviation of 0.990.

Summary descriptive statistics for all the composite variables are given in Table 6. No problematic, high correlations between components of constructed variables exist.

<u>Methods</u>. We examine Propositions 4a and 4b (Link B-C), which predicts that KMPs originate from high degrees of tacit knowledge and problem solving complexity, respectively. We also look at Proposition 5 (Link C-D), which predicts that opportunism concerns arise in

response to KMPs. These assertions are explored using OLS analysis owing to the measured dependent variables' reflection of an underlying (although not measured) continuous scale. We examine Proposition 6, the relationship between hazards of opportunism and governance choice (Linkage D-E), with a binary logit model since our dependent variable is binary. The covariates used in each analysis are clearly specified in the corresponding tables of results and thus are not specified here.

<u>Results: Propositions 4a and 4b</u>. Models 1-6 (Table 7, Panel A) examine Propositions 4a and 4b. Models 1–3 examine the effects of knowledge attribute intensity–Tacit (per Proposition 4a) and Complexity–on the use of the KMP measured by Hibandwidth, while Models 4 - 6 examine the effects of Tacit and Complexity (per Proposition 4b) on the KMP Co-specialize.

Model 1 shows that the coefficient for Tacit is a significant and positive, which supports of Proposition 4a. The estimated coefficient of Tacit is of moderate size (0.255) given its possible range of 0-6. Model 2, which explores if Complexity predicts Hibandwidth, indicates no statistical relationship between the two measures, which is expected; thus, Tacit predicts Hibandwidth; Complexity does not. The findings of Models 1 and 2 are further supported in Model 3, which is our fully specified model: Tacit is a significant predictor of Hibandwidth, and Complexity is not.

Models 4-6 in Table 7 examine the effects of Tacit and Complexity on Co-specialize—Proposition 4b. Model 4 reports that Tacit is not a significant predictor of Co-specialize while Model 5 reports that Complexity is a highly significant ($p<0.01$) and positive predictor of Co-specialize (signed as predicted). The estimated coefficient for Complexity in Model 5 is 0.855, comparatively large, given its range of 0-3, suggesting a strong effect. Model 6 is our full specification and yields results consistent with Models 4 and 5: the coefficient for Tacit has no effect on Co-specialize, and the coefficient for Complexity has a significant effect. Overall, Models 1 through 6 provide support for Propositions 4a and 4b. Collaborations appear to use KMPs in the face of problematic knowledge attributes that make knowledge difficult and costly to transfer. Specifically, knowledge tacitness engenders a response that includes using high-bandwidth interactions, and complexity gives rise to co-specialization. Although the construction of Tacit and Complexity in this analysis is substantially different than for the earlier analysis, we nonetheless examined our data for an interaction effect and found no such effect.

Proposition 5. Models 7 through 9 (Table 7, Panel B) examine the relationship between KMPs (Hibandwidth, Co-specialize) and their affect on the degree of perceived opportunism hazard (Hazard). In model 7, the coefficient for Co-specialize is positive and highly significant ($p<0.01$). The magnitude of the estimated coefficient of Co-specialize is 0.598, a large magnitude given its range of 0-4, suggesting a moderate effect on Hazard. In Model 8 we regress Hibandwidth in place of Co-specialize and find a coefficient that also is positive and highly significant ($p<0.01$). The coefficient estimate is large (0.853), given the potential range of Hibandwidth (0-3), suggesting that the KMP of Hibandwidth has a strong effect on the degree of opportunism hazard.

Model 9 incorporates both KMP variables into the analysis. The coefficients remain positive but only the coefficient for Hibandwidth is significant ($p<0.01$). Although the coefficient for Co-specialize is not significant in this model, we note that Goldberger (1968) argues that, for small datasets like ours, it is appropriate to look at coefficient signs and magnitudes rather than rely on statistical significance for interpretations. Thus, we conclude that Co-specialize may indeed be related to opportunism hazards, although perhaps not as strongly correlated as Hibandwidth. Furthermore, the Co-specialize's coefficient is small (0.240) relative to that of Hibandwidth (0.685), which suggests that that Hibandwidth's effect on the opportunism hazard is greater than that of Co-specialize.

Proposition 6. The results of the binary logit analysis are presented in Table 7, Panel C. Model 10, our baseline case, includes OwnHazard and a constant. Model 11 incorporates the variable Hazard, the respondent's perceived threat of other collaborators' opportunism. We focus our discussion on model 11. Although pseudo R^2 is low, adding Hazard to our base model substantially increases pseudo R^2 to 0.039 from 0.009. The coefficient estimate for OwnHazard is negative, indicting that an Equity form of governance is less likely the more the respondent's own firm is expected to act opportunistically. The coefficient estimate for Hazard is positive, indicting that an Equity form of governance is more likely the more the respondent expects collaborators to act opportunistically. These coefficients are consistent with a the view that firms do not want to safeguard against opportunism when they may act opportunistically but do want to safeguard against such opportunism when others may act opportunistically. Neither coefficient is statistically significant. Nonetheless, as described above, the signs of the coefficient estimates are informative for such a small sample.

In sum, our empirical examination of survey responses is supportive of the proposed chain of relationships. The use KMPs increases with increasing knowledge tacitness and problem solving complexity. Contracting hazards appear to increase with the use of KMPs, particularly with the use high bandwidth communication channels. Finally, use of equity-based governance appears to be related to a opportunism hazards; however, the veracity of this finding is not great.

5. Discussion and Conclusion

The goal of this paper was to provide preliminary evidence for addressing a widespread dilemma in the management of inter-firm collaborations: the tension between the need to share knowledge and the need to safeguard against uncontracted-for expropriation of knowledge. The debate is reflected in the literature by the tension between the transaction cost economics perspective (Williamson, 1975, 1985, 1991) and the knowledge-based view of the firm (Kogut and Zander 1995; Conner and Prahalad 1996; Kogut and Zander 1996; Madhok 1996). One vehicle for reconciling these two views is the chain of relationships (Figure 1) proposed by Heiman and Nickerson (2002), which considers both knowledge-based and transaction-cost factors.

The knowledge-based portion of the proposed chain asserts that intense degrees of knowledge attributes, for example, knowledge tacitness and complexity, create costly problems with knowledge transfer. These problems are subsequently addressed by deploying knowledge management practices (KMPs), for example, high-bandwidth interactions or co-specialization of communication codes. The transaction cost part of the chain asserts that KMPs, while efficacious for promoting knowledge transfer, potentially increase costs by increasing the ability of collaborators to expropriate knowledge. As a consequence of the existence of opportunism threats, collaborations seek to safeguard against opportunism by choosing more protective forms of governance (i.e., equity based in place of non-equity based) for the collaboration.

An initial empirical test of two over-arching propositions (1 and 2) was essayed, using the CATI database, a large but not very detailed dataset, with generally supportive results (with the qualification regarding a negative interaction term between Tacit and Complexity). In order to overcome the drawbacks of the first analysis, and expand the inquiry into the interior portion of the proposed chain of relationships a small, but rich and detailed dataset was constructed by personally administering questionnaires to 36 collaboration managers and personnel. Interviews of 18 respondents provided qualitative support for the proposition that the focal dilemma represents an important managerial challenge. Moreover, open-ended discussion by the interviews resonated with HN's chain of propositions. Empirical analysis of the 36 survey demonstrated support for the interior part of the proposed chain of relationships.

These findings, however, must be tempered because of several empirical and theoretical issues. First, neither dataset represents an ideal test bed for the propositions. The CATI database lacks the richness needed to examine KMPs. Our second data set is not ideal for the exact opposite reasons: although it offers a richness for assessing the propositions, the number of observations is small. Though the data in the survey sample is diverse, it remains far from a large and statistically random sample. Having many more observations without sacrificing detail-level of the survey and interviews is highly desirable. In particular, it would be desirable to have a more representative set of alliances. More detailed cases along the lines of the survey data also would allow for more stringent interpretations using statistical significance and the use of additional control variables in analyses.

Second, the measurement of knowledge tacitness and problem solving complexity are arguably subject to criticism. While we believe our survey data offers a reasonable attempt to measure these attributes, we relied on rough proxies for knowledge tacitness and problem solving complexity in the CATI database. The construction of the CATI data measures represent our best efforts to "work with what we have," and the exclusion of over 6000 of approximately 9000 total records as unusable, reflects these efforts. Results of extensive random spot-checking (and occasional corrections) of the final version of the cleaned data file, however, give us confidence that our coding scheme is reasonable given the constraints of the database.

Third, the theory we empirically examine remains pre-formal and in need of further development. While the empirical results provide substantial support for the proposed chain of relationships, the theory can be further developed in a variety of ways. For instance, it is not clear that all relevant knowledge attributes are incorporated into the theory. Our controls for knowledge ambiguity and dispersion, which are not highlighted in this paper but for which we found statistical significance, are suggestive of other knowledge attributes that could affect both KMPs and governance. Another area of concern is whether or not the two KMPs of high bandwidth communication and co-specialized communication codes are all encompassing. Could other practices also be important and can they be folded into the analysis? A final area that could benefit from further research is to unpack the dichotomy of equity vs. non-equity governance to study the wider array of governance arrangements employed to govern collaborations. For instance, CATI tracks over 20 governance modes. Although the equity dimension is important it may not be the only important dimension that can be incorporated into the theory.

These limitations notwithstanding, the theory and empirical analysis offers several insights for managers. Without guidance on how to integrate differing theoretical perspectives, managers may place too much emphasis on one perspective and not enough on another when managing inter-firm collaborations. Indeed, we believe that several stories from our interviews support this notion. Over-focus on one issue at the expense of another important issue may result in sub-optimal management of the knowledge-sharing/knowledge-expropriation dilemma and, ultimately, lead to failed collaborations. The reconciliatory approach of this work, backed by the evidence presented herein, provides a synthesis that may help managers achieve an appropriate balance between cooperation and competition.

For managers engaged in directing a firm's approach to collaboration, an awareness of the existence of knowledge attributes of transactions is the starting point for managing collaborations strategically. Managerial awareness of the (potentially impairing, costly) effects of knowledge tacitness and problem solving complexity on collaboration costs, knowledge sharing feasibility, and value creation is critical for structuring the collaboration to achieve its value creation goals. Yet, this paper highlights the idea that managers also must be aware of the darker strategic implications of the structures they choose in which expropriation concerns come to the fore.

This paper has sought to provide some preliminary direction for answering the question "How do firms manage knowledge sharing while avoiding knowledge expropriation in interfirm collaborations?" The fundamental dilemma of managing for sharing knowledge while also protecting knowledge in collaborations was explored using a sequential approach, which places knowledge-based factors at the start of a proposed chain of relationships and places transaction cost economics at its end. Empirical evidence generally supports the contentions outlined above, sometimes strongly, with none of our findings directly refuting any of our assertions. Results, while clearly encouraging, equally clearly lack a definitive character. Future (empirical and theoretical) research in the direction outlined herein seems appropriate.

Bibliography

Arrow, K. J. (1974). The limits of organization. New York, Norton & Co.

Conner, K. R. (1991). "A Historical Comparison of Resource-Based Theory and Five Schools of Thought Within Industrial Organization Economics: Do We Have a New Theory of the Firm?" Journal of Management **17**(1): 121-154.

Conner K. R. and C. K. Prahalad (1996). "A resource-based theory of the firm: Knowledge versus opportunism." Organization Science **7**(5): 477-501.

Cooper, D. R. and W. Emory (1995). Business Research Methods, New York, Irwin.

Duysters, G. (1999). Personal communication.

Duysters, G. and J. Hagedoorn (1993). "The Cooperative Agreements and Technology Indicators (CATI) System." Working Paper (MERIT, Universiteit Maastricht, Maastricht, NL).

Foss, N. J. (1996a). "Knowledge-based approaches to the theory of the firm: Some critical comments." Organization Science **7**(5): 470-476.

________. (1996b). "More critical comments on knowledge-based theories of the firm." Organization Science **7**(5): 519-523.

Goldberger, A. S. (1968). Topics in regression analysis. New York, Macmillan, Macmillan series in economics.

Gulati, R. (1995). "Does Familiarity Breed Trust? The Implications of Repeated Ties for Contractual Choice in Alliances." Academy of Management Journal **38**(1): 85-112.

Hagedoorn and Narula 1996

Heiman, B. A. (2002). Knowledge and governance in Collaborations. Dissertation, U.C. Berkeley, Walter A. Haas School of Business.

Heiman, B. A., and J. A. Nickerson (2002) "Towards Reconciling Transaction Cost Economics and The Knowledge-based View of the Firm: The Context of Interfirm Collaborations," International Journal of the Economics of Business, **9**(1): 97-116.

Helfat, C. (1994). "Firm-specificity in corporate applied R&D." Organization Science 5(2): 173-184.

Henisz, W. (1998). "The Institutional Environment for International Investment," Ph.D. Dissertation, U.C. Berkeley.

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

__________ . (1996). "What firms do? Coordination, identity, and learning." Organization Science **7**(5): 502-518

Maddala, G. S. (1985). Limited Dependent Variables in Econometrics. New York: Cambridge University Press.

Nahapiet, J., & Ghoshal, S. (1998). Social capital, intellectual capital, and the organizational advantage. Academy of Management Review, 23: 242-266.

Nelson, R. (1959). "The Simple Economics of Basic Scientific Research," The Journal of Political Economy, Vol. 67, No. 3. (Jun., 1959), pp. 297-306.

Nickerson, J.A. and T. R. Zenger (2002). "A Knowledge-based Theory of Governance Choice," manuscript, Olin School of Business.

Simon, H. A. (1945). Administrative Behavior. New York, The Free Press.

__________. (1962). "The Architecture of Complexity." Proceedings of the American Philosophical Society 106(6): 467-82.

von Hippel, E. (1994). “'Sticky information' and the locus of problem solving: Implications for Innovation.” Management Science 40(4): 429-438.

Williamson, O. E. (1975). Markets and Hierarchies: Analysis and Antitrust Implications. New York, The Free Press.

__________. (1985). The Economic Institutions of Capitalism.

__________. (1991). “Comparative economic organization: The analysis of discrete structural alternatives.” Administrative Science Quarterly **36**(June): 269-296.

__________. (1996). The Mechanisms of Governance. New York, Oxford University Press.

Table 1: Construction of variables

Dependent Variable	
Equity	=1 if equity-based, 0 if not equity-based
Independent Variables	
Tacit	Knowledge tacitness: Tacit=1 if activity is applied research & development, otherwise Tacit= 0.
Complexity	Problem-solving complexity: Complexity=1 if new, valuable, strategic knowledge is expected to result from the collaboration, else Complexity= 0.
Control Variables	
Dispersion	Knowledge dispersion: Dispersion = number of firms in the collaboration.
Tacit*dispersion	Interaction between knowledge dispersion and knowledge tacitness.
Complexity*dispersion	Interaction between knowledge dispersion and problem-solving complexity.
Ambiguity	Knowledge ambiguity: Ambiguity=1 if collaboration includes more than one technology, otherwise Ambiguity=0.
Tacit*ambiguity	Interaction between knowledge ambiguity and knowledge tacitness.
Complexity*ambiguity	Interaction between knowledge ambiguity and problem-solving complexity.
Government	Government involvement: Government=1 if a government or government agency is involved in the collaboration, otherwise Government=0.
Sizelarge	Size greater than 50,000: Sizelarge = 1 if least one collaboration partner-firm has more than 50,000 employees, otherwise Sizelarge = 0.
Sizemedium	Size between 5,000 and 50,000: Sizemedium = 1 if least one collaboration partner-firm has more than 5,000 but less than 50,000 employees, otherwise Sizemedium = 0.
Sizesmall	Size between 500 and 5,000: Sizesmall = 1 if least one collaboration partner-firm has more than 500 but less than 5,000 employees, otherwise Sizesmall = 0.
Year1978-Year1989	12 Year dummy variables for 1978-1989 (twelve variables). Base condition (all dummies=0): year is 1977.
Westeurope	Dummy variable =1 if at least one partner is based in a country in western Europe.
Northamerica	Dummy variable =1 if at least one partner is based in USA or Canada.
Japan	Dummy variable =1 if at least one partner is based in Japan.
Logpoliticalhazard	Log of average of country political risk of collaboration participants

Table 2: Summary statistics-CATI dataset

Variable Name	Mean	Std. Dev.	Min	Max
Equity	0.250	0.433	0	1
Government	0.012	0.109	0	1
Sizesmall	0.231	0.421	0	1
Sizemedium	0.421	0.494	0	1
sizelarge	0.446	0.497	0	1
Year1978	0.022	0.146	0	1
Year1979	0.035	0.185	0	1
Year1980	0.047	0.212	0	1
Year1981	0.050	0.217	0	1
Year1982	0.070	0.256	0	1
Year1983	0.062	0.241	0	1
Year1984	0.088	0.284	0	1
Year1985	0.121	0.326	0	1
Year1986	0.147	0.354	0	1
Year1987	0.101	0.302	0	1
Year1988	0.115	0.319	0	1
Year1989	0.128	0.334	0	1
Logpoliticalhazard	-1.742	0.371	-2.254	-0.198
Infotech	0.307	0.461	0	1
Biotech	0.155	0.362	0	1
Newmaterial	0.113	0.316	0	1
Northamerica	0.683	0.465	0	1
Japan	0.280	0.449	0	1
Westeurope	0.533	0.499	0	1
Complexity	0.320	0.467	0	1
Tacit	0.092	0.289	0	1
Dispersion	0.271	1.025	0	23
Ambiguity	0.078	0.268	0	1
Tacit*dispersion	0.035	0.361	0	10
Complexity*dispersion	0.087	0.697	0	23
Tacit*ambiguity	0.006	0.078	0	1
Complexity*ambiguity	0.025	0.156	0	1
Tacit*complexity	0.054	0.226	0	1

Table 3: Cross Tabulation of CATI Data

		Tacit			Complexity		
		0	1	Sum	0	1	Sum
Equity	0	93.4%	6.6%	100%	79.1%	20.9%	100%
	1	82.9%	17.1%	100%	34.7%	65.3%	100%

Table 4: Logit Analysis of CATI Dataset

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Tacit	1.071 **		0.655 **	0.484 **	0.632 **	2.878 **
	(0.130)		(0.144)	(0.159)	(0.168)	(0.238)
Complexity		1.963 **	1.902 **	1.706 **	2.033 **	2.562 **
		(0.093)	(0.094)	(0.103)	(0.118)	(0.130)
Tacit*Complexity	-1.218 **	-1.923 **	-1.969 **			-3.913 **
	(0.046)	(0.067)	(0.068)			(0.319)
Ambiguity, Dispersion and interaction terms				Yes	Yes	Yes
All other control variables					Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes
Log-Liklihood	-1641	-1432	-1422	-1378	-1272	-1191
χ^2	65.21 **	482.02 **	502.46 **	591.79 **	803.00 **	96397 **
Pseudo R^2	0.020	0.144	0.150	0.177	0.240	0.288

** $p < 0.01$
* $p < 0.05$

Table 5: Survey items used in constructing covariates

	Tacit
1	Type of knowledge involved in transfer (=1 if mainly technical, 0 otherwise.)
2	Did you expect problems with transfer due to unfamiliar jargon?
3	Almost all of the knowledge to be transferred was available in written form?
4	Knowledge to be transferred could easily be described as a set of routine, repeatable activities
5	Most of the knowledge to be transferred could be described as systematic rules
6	Most of the knowledge to be transferred could easily be transferred using common, informal non-scientific terms
	Complexity
7	Did you expect the exchange to involve solving problems jointly?
8	Did the alliance involve a project that was of long term strategic value to your firm?
9	One of the alliance's results was to produce a distinct new body of knowledge
	Bandwidth
10	To assist in knowledge transfer, we planned for frequent in-person meetings between members of each partner's team
11	To assist in knowledge transfer, we planned for frequent in-person meetings between members of each partner's team
12	We considered that we might need to act to overcome problems transferring knowledge due to cultural differences
13	I expected that technical knowledge would be transmitted mostly by personal instruction methods
14	During alliance formation, I expected that knowledge transfer would be helped if the partners would develop a common, technical language
	Co-Specialize
15	Did you plan to use frequent in-person meetings, site visits, etc. to aid in solving problems together?
16	When technical aspects of combining partner's knowledge was expected to be difficult, did you plan to provide a common work location?
17	During alliance formation I believed that successfully solving problems together would require frequent, open dialogues between partners?
18	When solving problems together, I expected that it would be necessary to work closely together in an atmosphere that promotes easy interaction
19	When solving problems together, I expected that it was necessary to develop social relations that would provide a foundation of mutual trust
	Hazard
20	During formation did you worry about the partner acquiring too much knowledge?
21	We worried about the possibility of partner firms using long distance relationship as an excuse to incompletely share necessary knowledge with us
22	The alliance posed a threat of unintentional over-transfer of knowledge to partners
23	During alliance formation, at in person meetings, our partner aggressively protected some of their knowledge from our team members
24	During alliance formation, at in person encounters, our partners never probed us about valuable knowledge lying outside the scope of our agreement
25	The problem of unintended over-transfer of our knowledge to our partner's people through close personal contact is worthy of managerial attention.
26	The problem of receiving only inadequate knowledge from an uncommunicative, uncooperative, overprotective partner is worthy of managerial attention
	OwnHazard
27	Our firm should aggressively manage issues related to the threat over unintentional over-transfer of knowledge
28	Our firm should seek to obtain as much information as possible from alliance partners
29	In this alliance, we tried to obtain as much information as possible from alliance partners
30	During alliance formation was it clear exactly which partner-knowledge would be looked for?

Table 6: Summary statistics for Survey data (n=36)

Composite Variable	Median	Mean	Standard deviation	Possible minimum	Possible maximum	Actual minimum	Actual maximum
Tacit	3	2.611	1.479	0	6	0	6
Complexity	2	2.389	0.645	0	3	1	3
Hibandwidth	3	2.750	1.079	0	4	0	4
Co-specialize	4	3.611	1.248	0	5	1	5
OwnHazard	3	2.861	0.990	0	4	0	4
Hazard	4	3.972	1.647	0	7	1	7
Equity	0	0.167	0.990	0	1	0	1

Table 7: Analysis of Survey Data (N=36)

Panel A–Ordinary least squares analysis of Propositions 4a and 4b, Linkage B-C

	Hibandwidth			Cospecialize		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Tacit	0.255 *		0.227 **	0.112		-0.077
	(0.117)		(0.134)	(0.144)		(0.148)
Complexity		0.378	0.137		0.855 **	0.937 **
		(0.280)	(0.307)		(0.298)	(0.340)
Constant	2.085 **	1.847 **	1.830 **	3.319 **	1.569 *	1.575 *
	(0.351)	(0.691)	(0.673)	(0.429)	(0.736)	(0.745)
F(1, 34)	4.72 *	1.830	2.400	0.610	8.240 **	4.170 *
R^2	0.122	0.051	0.127	0.018	0.195	0.202
Adj R^2	0.096	0.023	0.074	-0.011	0.171	0.153

Panel B–Ordinary least squares analysis of Proposition 5, Linkage C-D

	Hazard		
	Model 7	Model 8	Model 9
Hibandwidth		0.853 **	0.685 **
		(0.217)	(0.272)
Co-specialize	0.598 **		0.240
	(0.202)		(0.235)
Constant	1.814 *	1.627 **	1.222
	(0.770)	(0.640)	(0.753)
F(1, 34)	8.780 **	15.420 **	8.240 **
R^2	0.205	0.312	0.333
Adj R^2	0.182	0.292	0.293

Panel C–Binary logit analysis-Linkage D-E

	Equity	
	Model 10	Model 11
Own Hazard	-0.233	-0.492
	(0.438)	(0.511)
Hazard		0.326
		(0.338)
Constant	-0.961	-1.586
	(1.271)	(1.408)
Log Likelihood	-16.080	-15.590
χ^2	0.28 (df=1)	1.26 (df=2)
Pseudo R^2	0.009	0.039

*p<0.05, **p<0.01(1-tailed), standard errors in parentheses.