Expertise, Competitive Overlap, and Partner Choice

Geoffrey Borchhardt[†] Balázs Kovács[†] Michelle Rogan[‡]

Abstract

When choosing business partners, organizations ideally prefer partners with the most relevant expertise while avoiding those who also serve their competitors. Hence, in market networks, partner choice often presents a trade-off between accessing expertise and avoiding second-order competitive overlap. We propose that as competitive overlap increases, organizations' fears of information leakage and concerns about access to resources lead them to select less expert partners. A matched sample analysis of 963,089 US patents with measures of expertise and competitive overlap constructed via a text-based, deep learning algorithm shows that the likelihood a client selects a patent law firm based on relevant expertise decreases significantly as competitive overlap with other clients of the law firm increases. However, when concerns about information leakage or access to resources are lower, in particular when the client and the law firm have a prior relationship, when the law firm is high status, and when few alternatives are available, this effect weakens. Finally, we show that when a client chooses a less expert partner, time to patent acceptance is greater and forward citations are lower, indicating that avoidance of competitive overlap may come at a significant cost.

Keywords: firm-client network, patenting, competition, expertise

[†]Yale School of Management
[‡]Imperial College London

Introduction

Market ties, relationships created by the exchange of goods, services or money, represent the predominant linkages between business firms (Baker, 1990). When choosing a business partner in market networks, firms often face a dilemma: they want to work with partners that are experts in the issue at hand but at the same time they want to avoid second order competitive-overlap - working with partners who simultaneously serve their close competitors. However, experts commonly do precisely that. Businesses that compete in the same segment or niche are likely to require very similar expertise (Baum and Singh, 1994; Podolny, Stuart, and Hannan, 1996; Dobrev, Kim, and Hannan, 2001). Clients with the same types of products rely on similar expertise of advertising firms. Similar manufacturing firms depend on strategic advice from the same consultants. Start-ups in the same industry or technology sector are funded by venture capital firms with expertise in that industry or technology. In each of these situations, the choice of partner to maximize expertise often is accompanied by competition among those also working with the same partner which increases the potential for information leakage (Asker and Ljungqvist, 2010; Aobdia, 2015) and competition for the shared partner's resources (Rogan, 2014). Across a range of settings in which expert knowledge and avoidance of competition is important, partner choice can present a fundamental trade-off between expertise and competition.

Despite the importance of this dilemma for partner choice, the trade-off between expertise and second-order competitive overlap has received limited empirical attention, and prior theory provides only limited guidance. To be sure, factors affecting partner choice in general have been the subject of considerable research to date. Scholars have argued that firms sort into partnerships based on the best fit given their needs (Mitsuhashi and Greve, 2009; Vissa, 2011). Others argued that firms seek resource rich partners (Pfeffer and Salancik, 1978; Podolny and Page, 1998; Stuart, 1998). Embeddedness researchers emphasize the role prior relationships play in partner choice (Gulati and Gargiulo, 1999; Kim, Oh, and Swaminathan, 2006; Gulati and Sytch, 2007; Uribe, Sytch, and Kim, 2020). Furthermore, a recent stream of research has begun to investigate the main effect of competitive overlap in interorganizational relationships (Katila, Rosenberger, and Eisenhardt, 2008; Rogan, 2014; Sytch and Tatarynowicz, 2014; Hernandez, Sanders, and Tuschke, 2015; Pahnke, McDonald, Wang, and Hallen, 2015; Zhelyazkov, 2018). Taken together, this research has advanced our understanding of the factors affecting partner choice, but none sufficiently considers the network dynamics surrounding partner choice, in particular how concerns for second-order competition in the network surrounding the dyad affects the importance of expertise. Investigating the expertise-competitive overlap trade-off in partner choice is important because it provides an opportunity to develop theory for the interplay between network characteristics and dyadic characteristics in shaping interorganizational networks.

We begin to address this trade-off by building and testing novel arguments for how much organizations' avoidance of second-order competitive overlap affects their prioritization of access to expertise when seeking a partner firm to handle a project. Our theory applies to market ties in settings in which organizations look for business partners to conduct a task or service for them, a common situation in industries and markets defined by expert knowledge (Baker, 1990; Sytch and Gulati, 2013; Sytch and Kim, 2020). In such markets, potential business partners offer a set of non-homogeneous skills and services administered by experienced human capital. To benefit from a partner's expertise, a firm needs to disclose private information to its partner and it often requires specific resources of its partner (e.g., expert human capital). At the same time, competitive overlap with the other firms served by the partner heightens the focal firm's concerns over access to these limited resources and potential information leakage (Asker and Ljungqvist, 2010; Rogan, 2014; Aobdia, 2015). Hence, we propose that firms' preferences for partners with the best expertise decrease with the amount of second-order competitive overlap in the prospective partner's portfolio.

To illustrate this trade-off in the empirical context of this study, the choice of patent law firms by patenting organizations, consider a firm that has just invented a new technology and is looking for a law firm to prepare and file its patent application. Ideally, the client firm would want to work with a law firm with expertise in the specific technological domain of its invention. Patent law firms employ lawyers with graduate degrees in relevant scientific and technological fields. A law firm's expertise depends on its lawyers' scientific knowledge as well as their knowledge of the patent application and examination process. But an expert in the specific field is likely to, simultaneously, be working with the client's close competitors.[1] Hence, clients face a dilemma. Choosing the most expert law firm could expose a client to knowledge leakage if that firm is also serving the client's competitors (Moeen, Somaya, and Mahoney, 2013). Yet, choosing a non-expert law firm could threaten the patent approval process, increasing the time to patent acceptance, the quality of the patent, or in some cases even leading to a rejection of the patent. This trade-off is not unique to patent law. It appears in several types of market ties, including investment bank-client relationships (Asker and Ljungqvist, 2010), advertising agency-client relationships (Baker, Faulkner, and Fisher, 1998; Rogan, 2014), commercial banking (Uzzi and Gillespie, 2002), fashion (Godart, Shipilov, and Claes, 2014), and auditor-client relationships (Levinthal and Fichman, 1988), in which the partnership involves sharing of information and resources.

The expertise-competitive overlap trade-off is likely to be more or less pressing in certain situations. For example, when few expert providers are present, the choice available to the client is constrained and so firms may be less able to avoid competitive overlap when seeking an expert partner. As one patent attorney explained, in the early 2000s there were just three or four patent attorneys with deep expertise in lithium-ion battery technologies, and the tens of firms inventing within that technological domain either had to select law firms that their competitors worked with or settle for a less-than-expert law firm (Interview notes with

[1] Legally, patent law firms are not allowed to represent directly competing innovations, and so one might assume that second-order competitive overlap is rare. However, considerable leeway exists in what constitutes a conflict of interest (Interview notes with Attorney #1, 8/14/2019). Patent prosecutors, the lawyers who prepare the patents for submission, interpret gray areas more or less favorably depending on the argument they would like to make. Additionally, conflicts of interest arise at the level of individual lawyers, which means that different lawyers at the same firm can work for competing organizations without any legal uncertainty. Further, organizational divisions to keep competing client account separate (i.e., "Chinese walls") are often not effective (Uzzi and Gillespie, 2002). Knowledge might still leak over innocuous lunch conversations, when asking a colleague for advice, or when an employee leaves the lawfirm to work with a competitor and brings existing knowledge with them.

Attorney #1, 8/14/2019)[2]. In contrast, when a prior relationship exists between firms or the partner firm is high status, the tolerance for competitive overlap is likely to be greater, and so firms are likely to continue to choose experts even when competitive overlap exists.

We investigate these second-order competitive overlap effects on firms' choices of expert partners using US patent data from the Patentsview Project by creating a matched law firm - client patent dyad for each patent, comparing the realized pair to a matched, hypothetical dyad. We use a text-based, deep learning algorithm to construct the multidimensional knowledge-space from which we construct fine-grained measures of expertise and competitive overlap. We estimate the likelihood of a law firm being selected by a client based expertise and competitive overlap while controlling for other characteristics that could affect partner choice and controlling for patent characteristics through patent fixed effects. We also examine the costs firms incur by choosing the less expert partner, by examining the effect of choosing a partner with lower expertise on the performance of the patent. Indeed, to the extent that firms do choose less expert partners to avoid competitive overlap, the expertise-competitive overlap may be potentially detrimental to the long term innovation performance of firms.

Theory

Although the trade-off between expertise and second-order competitive overlap in partner choice has received only limited theoretical or empirical attention, scholars have studied partner choice in general. In research to date, three related perspectives have been developed to explain inter-organizational partner choice.

A matching theory perspective proposes that firms sort into partnerships based on the best fit given their needs (Mitsuhashi and Greve, 2009; Vissa, 2011). Firms match based on observable characteristics and withdraw from relationships if they discover mismatches

[2] In order to gain a better understanding of the setting, we interviewed an experienced patent attorney (Attorney #1) who has more than 10 years of experience working with start-ups and established companies in the Silicon Valley. We also consulted with an ex-patent attorney who works now as an academic (Attorney #2). This project is part of a larger project in which one of the authors interviewed multiple patent examiners and inventors about the patenting process.

in unobservable criteria and better alternatives are available. Empirical research has shown that firms are more likely to choose alliance partners who have market or resource complementarities (Mitsuhashi and Greve, 2009) and more likely to withdraw when better options come available (Greve, Mitsuhashi, and Baum, 2013).

A second perspective based on resource dependence theory holds that firms form partnerships to access needed resources, and therefore they seek resource rich partners (Pfeffer and Salancik, 1978; Stuart, 1998). To manage environmental uncertainty, firms seek partners that provide predictable resources, and hence, resource complementarity is less of a concern than the quality, stability and accessibility of the resources provided. In this vein, research has found that status makes a partner more attractive because it implies that the partner has higher quality resources and has more reliable access to resources (Podolny, 1993; Stuart, Hoang, and Hybels, 1999).

A third perspective, the embeddedness perspective, emphasizes the role that prior relationships play in partner choice (Gulati and Gargiulo, 1999; Kim et al., 2006; Gulati and Sytch, 2007; Rogan and Sorenson, 2014; Uribe et al., 2020). Uncertainty about a prospective partner's quality and its willingness and ability to solve unanticipated problems in the relationship leads to a preference for ties with prior partners. Hence, firms tend to choose prior partners or partners of their current partners (Parkhe, 1993; Gulati, 1995). Indeed, firms may even sacrifice match quality to continue working with past partners (Baum, Rowley, Shipilov, and Chuang, 2005; Goerzen, 2007).

While each of these perspectives has advanced our understanding of the factors affecting partner choice, they provide limited guidance for partner choice when firms face a trade-off between expertise and competitive overlap. Matching theory suggests that firms would prioritize expertise, i.e., finding the law firm that provides the relevant knowledge needed to prosecute the firm's patent. Resource dependence suggests that the exact match on expertise would be less important than partnering with the firm that is resource-rich (i.e., large or high status firms), as long as the firm held some related expertise. The embeddedness perspective also weighs other factors more heavily than expertise, emphasizing selecting from prior, known partners rather than choosing a new partner with the best match on expertise. Furthermore, while research has considered the implications of direct competition for partner choice (Katila et al., 2008), none of the perspectives sufficiently considers the network dynamics surrounding partner choice, in particular how concerns for second-order competition in the network affects the importance of other factors in partner choice.

The theory for partner choice that we develop here builds on a recent stream of work that has begun to consider the implications of second-order competition on network dynamics. These studies go beyond the dyadic perspectives on partner choice outlined above to examine the network neighborhood surrounding the dyad. For example, Rogan (2014) observed that clients severed ties with advertising firms when competitive overlap in the advertising firms' portfolios of clients increased. The likelihood of dissolving a tie was greatest at high levels of relational embeddedness and competitive overlap, pointing to a complex interplay between a network perspective on partner choice and the previous relational embeddedness perspective. Similarly, in a study of German board interlocks by Hernandez et al. (2015), firms avoided interlock ties that would have exposed their knowledge to indirectly linked rivals, which in turn shaped the evolution of the network. Most recently, Clough and Piezunka (2020) show that firms sharing the same supplier with one of their competitors cut ties to the supplier when their competitor's performance dropped, supporting the argument that firms' decisions about partner selection extend beyond characteristics of the potential dyad to the network surrounding the dyad.

These studies have established the importance of considering the network neighborhood in theories of partner choice. However, the usefulness of this novel perspective requires consideration of how network characteristics, such as competitive overlap, interact with dyadic factors to shape partner choice. The current study begins to do so by examining a previously overlooked trade-off between maximizing relevant expertise, a dyad level predictor within matching theory, and avoiding second-order competition, a network characteristic. Thus, the perspective we develop here incorporates characteristics of the network neighborhood including the patterns of exchange among firms and the similarity of ideas in the knowledge space to develop predictions for partner choice in inter-organizational exchange.

Our aim is to advance theory regarding the effect of the expertise-competitive overlap trade-off on partner choice in interorganizational relationships. As a starting point for our arguments, we focus on market relations between firms "created by the exchange of goods, services, and money" which represent "the predominant type of linkage between business firms" (Baker, 1990, p.590).

Expertise

Though the choice of partner is influenced by several factors, in general, we expect that organizations looking to identify a business partner will weigh the relevant expertise of potential partners. As defined by Teece (2018, p.540), experts "offer unique knowledge and experience, supported by advanced methods and tools of the services provider." Expertise can be both substantive and relational (Sandefur, 2015). Whereas substantive expertise is "abstract" and "principled," relational expertise is "situated" and "contextual" (Barley, 1996, p.425, 429). Substantive expertise in legal service includes for example understanding statutes, doctrines, legal principles, relevant past cases, and legal procedures. This type of expertise often is gained through professional training, such as study for a law degree (Sandefur, 2015). Relational expertise involves understanding the social distribution of knowledge and discretion in the actual relationships through which professional work takes place (Sandefur, 2015; Sytch and Kim, 2020). A patent lawyer having prosecuted a patent related to semi-conductors will not only be more versed in the technology but will also know which prior patents to cite and what resources to access such that the patent will be accepted in a timely manner. Furthermore, they will be able to frame their applications strategically with the intent of avoiding certain patent examiners or art units, knowledge of which they will have gained from prior experience (a tactic that was confirmed to us in an interview with a patent prosecutor).

The expertise of a potential business partner therefore is a function of the technical knowledge of its employees, i.e., human capital specific to the services they provide, and employees' unique contextual knowledge gained from experience (Levitt and March, 1988; Mayer, Somaya, and Williamson, 2012). Employees with such expertise can address the specific problems and needs of clients can anticipate concerns or difficulties that clients themselves did not consider, resulting in better outcomes for the client (Mayer et al., 2012). As such, relevant expertise of potential business partners increases the quality, speed, and likelihood of success of the services which they provide. Organizations therefore generally seek business partners that possess relevant expertise (Baker and Faulkner, 1991; Mowery, Oxley, and Silverman, 1996, 1998).

Second-order Competitive Overlap

When deciding with which business partner to work, organizations do not just consider the dyadic relationship between themselves and the potential partner, but also the network neighborhood. Beyond the direct relationship, partnering with a new firm creates indirect ties that connect the focal organization to others with which the firm has worked in the past. These ties create *second order competitive overlap* when an organization is indirectly linked to one or more of its competitors via a shared partner (Rogan, 2014).

Competitive overlap affects the likelihood of tie formation between client firms and partner firms via at least two mechanisms.[3] First, competitive overlap could inhibit an organization from entering into the relationship when it worries about knowledge leaking to its competitors. Organizations, particularly in technology markets, derive some of their competitive advantage from their knowledge base (Dierickx and Cool, 1989; Peteraf, 1993). When soliciting business partners to conduct services, an organization will frequently have to share some of the knowledge from which it derives its competitive advantage. This is required so that

[3]A scope condition for these arguments is that the focal firm is aware of the ties between its competitors and the service firm. As we discuss later, this scope condition is satisfied in our empirical setting because data about the assignee-lawyer relationships are made public by the USPTO.

the business partner can deliver its service in a tailor-made fashion, applying specifically to the technology and demands of the focal organization. However, once knowledge is shared, it becomes difficult to govern the flow of information and organizations risk their knowledge becoming available to unintended third-parties (Arrow, 1974; Dushnitsky and Shaver, 2009). While there are positive outcomes associated with knowledge flow in general (Gulati and Gargiulo, 1999), organizations might be worried about their knowledge flowing to competitors (Asker and Ljungqvist, 2010; Rogan, 2014; Hernandez et al., 2015). This is further exacerbated by competitors having similar capabilities and knowledge bases and thus being more likely to be able to understand and use the focal organization's knowledge and technologies (Cohen and Levinthal, 1990; Phelps, Heidl, and Wadhwa, 2012). As such, second-order competitive overlap increases the likelihood of knowledge leaking to the organization's competitors who are then able to undermine the competitive advantage of the focal organization, affecting, for example, their innovative output (Pahnke et al., 2015).

Second, the prior allocation of limited resources to competitors might deter organizations from entering into relationships with potential business partners. Given that competitors are likely to require similar resources and services, potential business partners who are working with competitors will already have directed some of their limited resources towards them. Thus, they are less likely to devote their full attention and capabilities to a new organization (see Rogan, 2014). This could be the case for both human and physical resources. Organizations might avoid entering into relationships with potential business partners if they feel they will not be able to receive the resources, attention, and (quality of) service for which they are looking (Asker and Ljungqvist, 2010). Furthermore, a partner providing services to two competing clients may use this position to increase its power over the clients, who both need the same limited resources provided by the partner (Ryall and Sorenson, 2007; Rogan and Greve, 2015). In our setting, clients may be concerned that law firms would have greater power over clients when clients are close substitutes, i.e., when they are patenting similar technologies. Together, information leakage concerns and the potential for limited access to resources lead firms to avoid partnering with organizations with high levels of second-order competitive overlap.

To be sure, there are instances where competitive overlap could have a *positive* effect on partner choice. For example, partnering with a firm serving competitors could bring legitimacy to the focal organization, especially if the partner - and its partners - are high status (Podolny and Page, 1998). Similarly, competitive overlap could also be associated with network externalities (Katz and Shapiro, 1985) or returns to specialization (Romer, 1987), whereby working extensively with similar firms on similar projects in a narrow area of focus, provides greater benefits to partners than the risks of information leakage or limited resource access. We consider the moderating effect of these conditions on the expertise-competitive overlap in our arguments and empirical analysis.

Partner Choice Given Expertise and Competitive Overlap

While the previous sections lay out arguments for distinct effects of expertise and second order competitive overlap, these effects are interrelated. Indeed, concerns associated with high levels of competitive overlap are exacerbated when evaluating business partners with high levels of relevant expertise for two reasons. First, for a client to benefit from relevant expertise as opposed to general knowledge of the law firm, it needs to disclose more private and sensitive information to the law firm. Organizations choose business partners that are highly experienced in the exact area of the project or technology in question precisely because they want their business partner to understand the fine details that come with the task. To leverage that expertise, the client organization will have to divulge information that is more precise and likely closer to the core of its competitive advantage compared to the information that a business partner would need if they were just relying on their general (non-client or project specific) expertise. Divulging more sensitive and important information to business partners heightens clients' concerns of information leakage, since the costs are greater in that situation than when less sensitive information leaks.

Second, to benefit from relevant expertise, clients are also more dependent on specific resources, such as specific human capital, of the business partner. For clients to leverage a business partner's project-relevant expertise, they will have to get access to the specific resources of the business partner in which their relevant expertise originates. This could be a particular employee or technology. When potential business partners are working with competitors of a client, these resources might not be available because they are already being used by competitors also served by the firm, thereby decreasing the extent to which a client can leverage the relevant expertise advantage the business partner poses. As such, high levels of competitive overlap exacerbate clients' concerns of information leakage and limited access to resources when evaluating business partners with expert knowledge.

In our setting, patent prosecution, the ability of the patent law firm to apply its expertise necessitates significant information sharing by the client. Reitzig and Puranam (2009, p.767) describe the intensive information sharing process that occurs in the patent application process: "...[E]ven the most well-drafted application cannot anticipate all potential examiner reactions, and ad hoc negotiations between the applicants and the examiners inevitably occur. Supporting this negotiation from the applicant's side are R&D staff supplying or even generating additional technological information..., patent lawyers using this information, and business staff making decisions on the desirability of limiting patent scope. ...[T]he value of the efforts of each...is leveraged by the expertise of the others." In other words, for a client firm to benefit fully from the expertise of its patent law firm, it must share sensitive information during the patent application process. This places the client at greater risk of harm should information leakage occur, particularly because this information is likely to pertain to the technological core of their competitive advantage. At higher levels of competitive overlap, the potential for information leakage is greater, and so a client's concerns about information leakage are likely to be greater the more important the expertise of the patent law firm is to the patent prosecution process. Simultaneously, the client needs the law firm's attorneys with relevant expertise to facilitate the exchange of technological information. When these have already committed their time to competitors of the client, the client will be less able to leverage the relevant expertise of the law firm. Therefore, competitive overlap is likely to limit access to expert knowledge and resources, thereby decreasing the value of working with (and the likelihood of the selection of) the most expert partner.

> *Hypothesis 1: Second-order competitive overlap decreases the effect of expertise on the client's choice of law firm.*

If the expertise-competitive overlap trade-off affects partner choice as we argued, i.e., competitive overlap decreases a client's ability to leverage business partners' relevant expertise, then the effect of this trade-off on partner choice will vary with different conditions that change the relative value of expertise to the client. When expertise is less valuable than other characteristics of the partnership, it will have a weaker effect on partner choice and the expertise-competitive overlap trade-off will be less pronounced. Considering the trade-off relative to each of the three established perspectives on partner choice can elucidate the conditions under which the trade-off affects partner choice.

In addition to expertise and competition concerns, prior theory has identified embeddedness, a history of prior relationships between partners, as a key factor explaining partner choice. In embedded relationships, factors such as trust can diminish these competitive overlap concerns. Knowledge can leak in a variety of ways that are characterized by different degrees of agency. While business partners could pass along knowledge knowingly and in an opportunistic fashion, inadvertent knowledge transfers could occur by accident or when employees dealing with competing clients chit-chat during coffee-breaks (cf. Hernandez et al., 2015). A potential client's fear of any of these scenarios taking place can be assuaged when the client trusts the partner to have their best interest at heart and not to make careless mistakes. This trust can be developed through direct, prior experience with the partner (Gulati, 1995; Gulati and Sytch, 2008). Similarly, an existing prior relationship could assuage concerns of resource allocation, since the business partner has previously allocated important resources towards the client (Pahnke et al., 2015). The client thus trusts the business partner to continue to assign its resources and expert employees toward the client. We therefore expect that the role of expertise in selection of a business partner changes less with competitive overlap when choosing from potential partners with which a prior relationship exists. We therefore hypothesize:

> *Hypothesis 2a: The moderating effect of second-order competitive overlap on the effect of expertise on the client's choice of law firm is weaker for law firms with which the client has an extensive prior relationship.*

The resource dependence perspective on partner choice suggests that the relative value of specific expertise to a client could be lower when the law firm has other resources that are valued even more greatly by the client. A key resource is the status of the partner, defined as "the prestige accorded firms because of the hierarchical positions they occupy in a social structure (Gould, 2002)" (Jensen and Roy, 2008, 496). Organizations gain status by affiliation to high status organizations, and an organization's status enables better and less costly access to resources. Unlike resources that are depleted with use such as raw materials, or resources whose use is exclusive to one client at a time, such as a key expert, status is a non-rival good, whose value remains the same or even increases with use (i.e., as more organizations defer to the high status one). Even when competitive overlap is high, the focal client can still benefit from the status of the partner firm. In addition, the value of affiliation to a high status firm may be sufficiently high to outweigh the client's concerns about limited access to specific expert resources due to competitive overlap. Lastly, high status firms are generally assumed to be less likely to violate norms of behavior, such as leaking sensitive information, due to the larger penalties paid by high status actors who are discovered to have made transgressions (Graffin, Bundy, Porac, Wade, and Quinn, 2013). Hence, we expect that the expertise-competitive overlap trade-off is weaker for high status partner firms. Formally:

> *Hypothesis 2b: The moderating effect of second-order competitive overlap on the effect of expertise on the client's choice of law firm is weaker for high-status law firms.*

A matching perspective assumes that firms choose the best fit partner given the alternatives available. Although our main argument is that a client's likelihood of selecting the most expert partner is lower when competitive overlap is high, if the supply of partners is so low that the next best match on expertise is substantially worse than the expert choice, clients will exhibit greater tolerance for competitive overlap. Furthermore, if there are few potential partners from which to choose, it becomes more difficult to avoid competitive overlap. More of the available, potential partners are likely to have worked with one's competitors and therefore choosing a less expert partner is associated with negligible improvements in terms of avoiding competitive overlap. Therefore, the expertise-competitive overlap trade-off is less likely to affect partner choice when the supply of partner firms with a reasonable level of relevant expertise is low.

> *Hypothesis 2c: The moderating effect of second-order competitive overlap on the effect of expertise on the client's choice of law firm is weaker when the supply of law firms in the technology class is scarce.*

Scope conditions and generalizability

Above we outlined a general framework about how firms may trade-off expertise with competitive overlap in partner choice. We instantiated this general framework in the setting of patent applications and innovating firms choosing among law firms, but our theory is applicable more generally. Below, we discuss in detail the scope conditions and generalizability of our theory.

Our theory assumes that the ties between business partners and competitors are observable, and hence our theory will apply most to settings where competitive overlap can be observed. In our empirical setting of patent prosecution, patents are publicly available, which allows for observation of the exact history of a law firm that is under consideration as a partner - including competitors it may serve. This assumption is not unique to our empirical context but also applies in other settings where firms are aware of direct rivals (Chen and

Miller, 2012; Downing, Kang, and Markman, 2019). In settings where firms are not aware of their competitors, it is likely that competitive overlap avoidance would not be present, and we would expect that the highest expertise partner would always be chosen. The less observable competitive overlap becomes, the more likely organizations are to rely on other signals (either using these to try to obtain information on competitive overlap or dropping it from the evaluation criteria completely), and the less our theory will apply. For example, we would expect that if an industry is nascent, neither expertise nor competition would be yet clearly defined, so the effects we hypothesized would be muted.

Another scope condition of our theory is that firms must believe that selecting partners who work with their competitors is detrimental to performance. Yet, choosing partners with competitive overlap can be beneficial to firms in some cases. For example, as Rogan (2014) discusses, competitive overlap may generate benefits for firms in cases such as mutual forbearance via multimarket contact (Bernheim and Whinston, 1990; Gimeno and Woo, 1996), and in settings when positive information externalities dominate (Katz and Shapiro, 1986). We would expect that the trade-off we hypothesized in this paper would not hold, or would be weaker, in such settings.

Lastly, our theory assumes that complete avoidance of indirect ties to competitors is not a realistic option. This means that a minimal level of relevant expertise is required and costly to develop. As such, actors themselves do no not easily have the means to develop the skills for which potential business partners might be solicited. This should be the case in industries which rely on expert services in specialized markets, such as specialized legal services, auditing, and consulting. For this reason, we focused our theory and analyses on external partner choice rather than the make v. buy decision (Mayer et al., 2012; Chondrakis and Sako, 2020).[4] In other settings, where the use of in-house experts is common, we expect that the expertise-competitive overlap trade-off would be managed by using an in-house

[4]We excluded firms with in-house legal teams for patent prosecution. In our setting, only one percent of all patent applications are handled in-house (only the larger innovator companies such as IBM or Apple employ full-time in-house counsels).

expert.

Setting & Data

We test our theory on the selection of patent law firms who play an integral role in the patent application process. Focusing on the patenting process offers a few distinct advantages to study our proposed theoretical dynamics. First, it allows for the observation of interactions between organizations and third-parties that are not themselves competing in the technology space. Rather, law firms aid multiple players in the industry by providing essential services. The patenting process requires legal representation that is most frequently sought outside of one's own organization. As there is considerable variation in the areas and depth of law firm expertise, pressures of profit maximization should create a setting in which multiple organizations vie for the same third party as legal representation, creating dynamics in which one could expect some overlap in legal representation across companies. At the same time, the patenting process represents a setting in which knowledge is inherent to an organization's success and is guarded intensely.[5] Patent prosecution, the preparation, filing and processing of patents, is not a simple rubber stamping exercise and requires both legal and technical expertise. Law firms, as patent prosecutors, "need to access specialized legal resources for conducting prior art searches, [draft] patent applications, and [prosecute] these applications at the patent office" (Somaya, Williamson, and Zhang, 2007, p.923). Their attorneys are required to have at least "a bachelor's degree (or equivalent) in a technical field and [to have passed] the patent bar exam" with many also holding post-doctoral degrees (Somaya et al., 2007, p.935). As noted by Reitzig and Puranam (2009, p.767), a granted patent "represents the ability of an applicant to convince a patent examiner of sufficient novelty, inventive step

[5] As one of our interviewee, Attorney #2 wrote to us in an email on 11/12/2020: "In my view the importance of patent examination actually makes the choice of partner even more important. The client is selecting someone to work on their behalf for years in an area that requires deep engagement with the subject matter but also a great deal of independence on the attorney's part, and therefore a great deal of trust. For these reasons, the decision you are studying is one that is quite significant for clients, and more significant than many scholars might think upon first glance."

(non-obviousness), and commercial viability when judging the applicant's technical invention."

In studying patents as a stand-in for the knowledge held by any particular company, we are following a rich literature that relies on patent data to construct the knowledge space in which organizations interact and show how knowledge might transfer from one organization to the next (Mowery et al., 1996; Podolny et al., 1996; Grigoriou and Rothaermel, 2017). Because the patenting process, by definition, occurs prior to the legal protection of intellectual property, one might expect companies to be particularly aware of the sensitivity of knowledge being exposed to an actor outside of the organization. Given detailed data on the technological underpinnings of any organizations, we are able to identify competitors at the level of the patent by finding nearby patents in the technology space.[6]

The data we use stem from the publicly accessible PatentsView Initiative which provides access to information on all USPTO patents, identifying longitudinal links to patent prosecutors, assigning organizations, and inventors for the years 2003-2013. Figure A.1, in the appendix, illustrates what the front page of a patent looks like and where our data stem from. We construct the knowledge space based on a deep learning text analysis of the patent texts (a more detailed description of this will follow). To focus on firms for which one might expect consideration of competitive overlap when selecting a law firm, the sample has been restricted to patents issued by organizations with more than 50 listed patents and law firms which have worked on more than 20 patents. The resulting dataset contains 963,089 patents, assigned by 3,247 patenting organizations, and handled by 2,419 law firms.[7]

[6]One way for organizations to avoid potential threats to knowledge leakage is to handle the patenting process themselves. Organizations occasionally rely on in-house legal counsel to guide them through the patenting process. In-house law firms handle below one percent of the patents in our sample. Since our theoretical focus is on the selection of business partners external to the focal organization, we exclude these observations.

[7]In robustness checks (not shown) we further restrict our sample to organizations that are classified as US or foreign companies or corporations or exclude patents by organizations which are considered to be small entities and find that our results remain unchanged.

Empirical Strategy

The difficulty in analyzing how organizations select law firms lies in conceiving of a plausible risk set. Ideally one could identify which patent law firms were actively considered by an organization that is looking to file a patent. Such data, however, are hard to come by without being deeply embedded into the decision-making process and, even then, might be driven by subconscious influences of which the decision-maker is unaware. We follow prior literature (Rogan and Sorenson, 2014; Carnabuci, Operti, and Kovács, 2015) and use a case-control design in which lawyers observationally similar to the one used by an organization on a particular patent are matched as counterfactual patent-law firm pairs. Using a 1:1 matching process,[8] each patent i appears in two observations, one of which is a pairing with a hypothetical law firm, the other with the law firm which actually handled the patent (for an example of the patent structure, see Table A.1 in the Appendix). To be considered in the risk set, a law firm needed to have handled *any* patent over the past three years within the larger geographical division, such as Pacific or West South Central (see Census, 2010). We then use propensity score matching (Rosenbaum, 1983) to identify potential patent law firms with which the focal organization did not end up working. Patent-law firm pairs were matched based on the recent patenting activity (a proxy for size measured as the logged number of patents over the past three years), the age of the law firm (founding year decile), as well as whether the law firm has handled a patent from the same one-digit NBER category over the past three years[9] (see Appendix C). We estimated propensity scores using a logit model and kept the case with the p-scores closest to the observed pair.[10]

[8]We find equivalent results when using a 1:4 design. We chose to present the 1:1 results as at times there were not four good matches for each observed patent-law firm pair and to present more conservative specifications.

[9]By matching lawfirms on having worked in the same 1st digit NBER class, we are balancing two competing requirements. On the one hand, we do want to match on broad industry somewhat so that the matched pairs can be reasonably looked at as potential alternatives. On the other hand, we do not want to match too precisely on expertise because we need variation in expertise so that we can test our hypotheses. We believe that matching on the seven 1st digit NBER classes, such as "Chemical" or "Electrical & Electronics" achieves these dual goals.

[10]Others argue that alternative matching methods, such as coarsened exact matching (CEM), are preferable to propensity score matching since they balance directly on the underlying variables and not just the

Balance Table 1 shows summary statistics of our matching variables by group. While results from t-tests indicate a statistically significant difference for size and "NBER one digit experience", the Cohen's d normalized difference measures indicate that these differences are small. We see that the matched controls differ by less than 6 percent of a standard deviation from the observed law firms. However, to assure that any issues which could stem from the matching procedure are minimized, we control for size and age in our specifications.

—– INSERT TABLE 1 HERE —–

We model the selection of law firms using a linear probability model with patent fixed effects (see e.g., Clough and Piezunka, 2020). Much of the prior work has used conditional logit estimations to model choice (Alcácer and Chung, 2007; Sorenson and Stuart, 2008; Rogan and Sorenson, 2014; Carnabuci et al., 2015). However, at the core of our hypotheses lies an interaction term and interaction terms are problematic to interpret in logit models (e.g. Ai and Norton, 2003; Mood, 2010). Fixed effects are also critical to our estimation strategy and can lead to incidental parameter bias in logit models (cf. Clough and Piezunka, 2020). Therefore, we rely on linear probability models, whose coefficients are easier to interpret and compare across models. Robustness checks using conditional logit estimations indicate directionally equivalent results (see Appendix A.4).

Due to fixed effects at the individual patent level, all idiosyncratic characteristics of either the patent or patenting organization (i.e., the client firm) are held constant and only characteristics varying across law firms, or the patenting organization-law firm dyad, can explain the choice of law firm.[11] Our modeling strategy assumes that clients set the relationship and choose a law firm. We recognize that formation of a tie requires the law firm to agree to enter a relationship. Our interviews with patent attorneys point to client size and prior

propensity score (King, Nielsen, Coberley, Pope, and Wells, 2011). However, in our case CEM and other binning methods lead to worse balance in our matches because of skewed distributions within each bin.

[11]Price differences across law firms could also influence partner choice. While we do not observe the prices charged by a specific law firm for a specific patent, existing research (AIPLA, 2009) shows that the main determinants of patent prosecution pricing are the size of the law firm, the location of the law firm, and the type and complexity of the patent. We are controlling for these factors either via patent FEs, via the matching or the control variables.

interaction as two key factors that affect a law firm's likelihood of accepting the tie. Both are characteristics of the law firm or dyad which are either part of our matching strategy (recent patenting activity proxying for size, while fixed effects should account for the size of the patenting organization) or are explicitly modelled (such as the number of previous interactions between the patenting organization and law firm). Thus, we are confident that our approach can be used to model the selection process of patent law firms. The models report standard errors clustered at the patent level.

Variables

Dependent Variable

The dependent variable, *law firm choice*, is binary denoting whether or not the client patent-law firm pair has been observed, thereby indicating a realized choice. For each patent this variable takes on the value one when the law firm was chosen to work on the patent. It takes on the value zero for the matched law firm which had not been chosen.

Using Deep Learning to Locate Patents in the Technological Space

To measure expertise and competitive overlap, we first map out the technology space and locate the patents in the technology space. Prior research has typically used patent classification codes (Le Mens, Hannan, and Pólos, 2015) or citation overlap (Podolny et al., 1996) to map out the similarity between patents and to construct the patents' and firms' location in the technological space. We believe that while these measures are meaningful, they are less than ideal because they are not able to get at fine-grained and unbiased differences between patents: The patent classes typically contain thousands of patents and any classification-based measure, by necessity, assumes that patents listed within the same classes are perfectly identical. Patent-citations-based measures are also suboptimal because the citation process is influenced by human and institutional biases (Kuhn, Younge, and Marco, 2020; Kovács, Carnabuci, and Wezel, 2021). Rather, in line with emerging research showing that a patent's text provides a more precise and less biased measure of its technological content (Younge and Kuhn, 2016), we construct the technological space using the text of the patents. Specifically, we analyze the the patents' abstracts to locate each patent in a 38-dimensional technological space. The choice of the 38 dimensions come from the fact that there are 38 two-digit NBER technology classes (see appendix B for the list).[12] While categorization at the USPTO is binary (the patent is either assigned to a category or not), our deep learning method provides a continuous measure ranging from zero to one for each of the categories.

Deep learning refers to the process according to which the free parameters of a neural-network model are learned from data. The neural network model we use takes a textual description of a patent as input and from the text it predicts the two digit NBER classification to which the patent belongs. The model is "trained" on the data to learn the parameter values that lead to the best possible categorization performance. This occurs via an updating algorithm that proceeds by trial-and-error to maximize the out-of-sample categorization accuracy. First, we pre-process the text of the abstracts: we tokenize text and encode the words such that a numeric index is given to each unique word in our text corpus (Chollet, 2017). Then, each abstract is represented as a sequence of word indices. The neural network we use to predict NBER class probabilities from patent text consists of five layers: a word embedding layer, a biLSTM layer (which can take into account the order of words), a fully connected layer, a softmax layer, and a classification output layer. This model structure is the standard structure used in the deep learning literature for text categorization purposes (Chollet, 2017). (See Appendix C for a more detailed explanation).

We use a 10-fold training approach (Chollet, 2017). The data is split into 10 randomly selected equal-sized sample. We first train the network using eight of these subsamples as a training sample and one of them as a test sample. We train the model for three epochs, with mini batch size of 64, as we found that this specification leads to the best average

[12]We decided to rely on the NBER two-digit classification because the USPC classification and the IPC classification systems have multiple hundreds of categories, which makes the deep learning task less reliable.

prediction accuracy (61 percent), without overfitting the data. After the network is trained on these nine subsamples, we use the trained network to do an out-of-sample prediction on the 10th subsample. We repeat this approach nine more times, each time predicting a different subsample and using the other nine subsamples for training data and test data. Finally, we combine the 10 sets of predicted values and use them in the analyses.

This approach provides a more precise, continuous, and multidimensional conceptualization of the technology space than previous approaches relying on technology class alone. Patent classes may be too broad or too precise to capture meaningful technological similarities (Younge and Kuhn, 2016). Vastly removed technologies may end up in the same subcategory because, for example, they both dispense fluids (Younge and Kuhn (2016) provide the example of a catheter and a golf device). However, these are unlikely to be meaningful similarities for the purposes of predicting law firm expertise or competition between technologies. Younge and Kuhn (2016) also provide preliminary evidence suggesting that patent attorneys themselves rely on text based searches and ignore the classification ascribed by the USPTO, suggesting that text-based measures can capture the technological basis of patents more closely.[13] Another advantage of our continuous, multi-dimensional measure is that it is better at capturing the "true" position of a patent in the technology space for patents that span classes. For example, patent #12524881, a device used to make filled-food products (think of a jelly donut), is in the primary category 51, materials processing and handling. However, our algorithm provides a classification that is more evenly spread over multiple categories. Namely it assigns it a 26 percent probability of being assigned into the miscellaneous mechanical, pipes and joints, or miscellaneous electrical category, capturing the different dimensions of the product more accurately.

[13]Similarly, measures predicting technological similarity based on the patents cited might be misleading since a few patents are vastly cited in a ceremonious manner. Kuhn et al. (2020) show that the technological similarity between cited and citing patents has fallen over time and thereby systematically confounds studies.

Expertise and Competitive Overlap

As the basic building block of the expertise and the competitive overlap measures is patent-to-patent similarity, we first describe our measure for patent-to-patent similarity. First, we calculate the Euclidean distance between the two patents in the technological space, formally:

$$d(i,j) = \sqrt{\sum_{k=1..38} (position_{i,k} - position_{j,k})^2},$$

where i denotes the patent identifier and $position_{i,k}$ denotes patent i's location in the kth dimension of the technological space.

Second, because both the expertise and the competitive overlap measures rely on similarities rather than distances, we transform the patent-to-patent distance measure by applying the negative exponential transformation based on Shepard's formula (Shepard, 1987):

$$sim(i,j) = exp(-\gamma\; d(i,j)), \quad \gamma > 0$$

The transformation is similar to an exponential decay curve. Very close distances will be transformed to values close to one. Depending on the γ chosen, farther distances will approach zero values for the expertise measure more quickly. To calibrate γ, we plotted the distribution of pairwise distances of patents, and found that $\gamma = 5$ is the value that best captures our intentions with operationalizing patent-to-patent similarity. With $\gamma = 5$, a value that results in relatively high similarity values for patents that are within the same USPC subclass; medium values that are not in the same USPC subclass but in are in the same two-digit NBER category; while the similarity of the patents that are further away, i.e., not in the same two-digit NBER classes, are discounted and approach zero quickly.[14]

Expertise. With the pairwise patent similarities at hand, we operationalize the *expertise* variable by measuring how "close" the patent is to the set of patents handled by the law firm

[14] Results are robust to using different values for γ, a logistic transformation, as well as the reverse-coded Euclidean distance.

in the past three years. A law firm previously having handled a patent that is close to the focal patent indicates expertise specific to the focal technology (given that we are only able to include successful patent applications, this also indicates having succeeded with said prior patent application). Because we are interested in knowing whether the law firm has handled similar patents to the focal patent and not whether all of their previous patents are similar to the focal patent, we want to ascertain how close the closest patents are. Therefore, we calculate the distance between the focal patent and the patent that lies at the first percentile of the distribution of pairwise distances between all patents handled by the law firm in the previous three years and the focal patent. While this measure does not capture breadth or depth of expertise in the form of having worked on many similar patents, it captures the relevance of a law firm's expertise for the client's focal patent. It provides advantages over relying on the mean or median distance (since these might miss a law firm having multiple, distinct areas of expertise), or simple counts of the number of patents in the same technology category (which lacks precision). Using the patent at the first percentile (instead of the closest patent) of the distribution also accounts for extreme outliers of large firms, meaning that our measure is unlikely to be driven by a single patent of the law firm being close to the focal patent.

For example, consider patent #7615500 by NEC Electronics, titled "Method for depositing film and method for manufacturing semiconductor device". The application was filed in 2007 and was handled by the law firm McGinn IP Law Group. To assess the expertise of McGinn IP Law Group in handling patent #7615500, we compare this focal patent's similarity to all patents handled by McGinn IP Law Group in the previous three years, from 2004 to 2006. McGinn IP Law Group applied for around 1300 patents in this time period, and the similarity of these patents to the focal patent range from 0.001 to 0.88. The first percentile of the distribution corresponds to a similarity value of 0.82. As an illustration, one patent from the distribution which is at a similarity value of 0.82 to is patent #7176145. This patent also describes a method of manufacturing semiconductors, that is, a technological process closely related to the focal patent. Therefore, according to our measure, McGinn IP Law Group has a high expertise in handling this patent.

The *competitive overlap* measure is similarly created based on our construction of the technology space. Our approach follows prior research defining competition at the level of the technology, specifically the position of a patent in the technology space relative to other patents (e.g., Podolny and Stuart, 1995). Our measure of competition is based on the technology rather than market. As noted by Podolny et al. (1996, p.666) describing competition between j and i, "A like pattern of technological antecedents implies a similarity-or even redundancy-in technological competencies. The greater the overlap in technological competencies, the more that j 's pursuit of its market possibilities affects the ability of organization i to pursue its opportunities." That is, an organization's position in the knowledge space defines and constrains actions organizations can take in the market and product space (Pontikes and Barnett, 2017). As such, competition on technological dimensions is closely related interorganizational competition.

To identify a law firm's competitive overlap relative to the focal patent in the technology space, we sum the similarities between the focal patent and the patent portfolios of each client of the law firm. Note that the patent similarities here capture the role of patent-to-patent competitive pressure, and the sum of these is what Podolny et al. (1996) call "crowding."

We start by calculating pairwise similarities between the focal patent and each patent handled by any of the law firm's other clients over the three years prior to the focal year. For each client, excluding the patenting organization of the focal patent, we then take the similarity that lies at the first percentile of the distribution. We arrive at our competitive overlap measure by summing these "competitor expertise" values over all clients of the law firm:

$$competitive\ overlap = \sum_{c=1}^{n} sim(i, j_c))$$

where $sim(i, j_c)$ is the similarity between the focal patent i and a law firm's client's c first percentile closest patent j_c.[15]

Again, consider focal patent #7615500 by NEC Electronics. In the three years prior to the patent application in 2007, McGinn IP Law Group handled patents of 129 different clients. As described above, to calculate the competitive overlap, we calculate the similarity of the focal patent to the patent portfolios of each of these 129 clients, and then add up these similarity values. Consider the case of three of these 129 firms: Micron Technologies, Ricoh, and Piolax. According to our measure, Micron Technologies is a relatively close competitor to patent #7615500 because in its portfolio of patents, the one percentile closest patent to patent #7615500 is at similarity value of 0.833 and is also a semiconductor construction method related patent. The first percentile patent of Ricoh, an electronics and imaging company, is a method of manufacturing an inkjet head using silicone methods similar to those in the semiconductor industry. While this is still in the broader technological group, the patent is not as relevant to competitive overlap as the semiconductor patent by Micron. The similarity between Ricoh's patents and the focal patent is 0.41. Lastly, the law firm worked with the plastics supplier Piolax. Their first percentile closest patent is about a mechanical lid and its hinge. The similarity value of this patent to the focal patent is 0.07. If these clients were the only ones that the law firm had worked with, the values would add up to a competitive overlap value of 1.31. These examples illustrate how clients that work on very similar technologies (here: semiconductors) contribute significantly to our competitive overlap measure, but clients that work on technologies that are only broadly related (silicone for inkjets) or not at all (a lid and hinge) contribute very little or barely at all.[16]

[15] There are other ways in which our empirical strategy ensures that this measure does not simply resemble a count of the number of clients of the law firm. First, we match on law firm size, based on the number of patents handled by the law firm. Two law firms of the same size are likely to have worked with similar numbers of clients. Second, we control for the average size of those clients, based on the average number of patents handled by those clients. In combination with matching on law firm size, we thus ensure that our competitive overlap measure compares which law firm, out of two that are of similar size themselves and have handled clients of similar size, has worked with clients that are closer to the focal technology.

[16] We calculate an alternative measure for competitive overlap based on a count of patents handled by the law firm that are closer than 99 percent of other patents. We start by calculating Euclidean distances between the focal patent and all other patents of the last three years (excluding other patents of the focal organization). We then identify those that are closer than 99 percent of all patents as competitors. The competitor measure then counts the number of patents technologically competing with the focal patent that

We also note that while the expertise and the competitive overlap measures are derived from the same knowledge space, they capture distinct concepts (their pairwise correlation is 0.2). To illustrate the distinction between our expertise and competitive overlap measures, consider NEC Electronics evaluating potential law firms for patent #7615500, trying to choose between McGinn IP Law Group and, our matched control, Venable LLP. Both law firms have worked on similar numbers of patents over the previous three years (1300 vs 1100) and have similar levels of competitive overlap. For McGinn, the competitive overlap value is 26.7 based on 129 clients. Even though Venable worked with approximately twice as many clients, a larger proportion of these was unrelated to the semiconductor technology of patent #7615500 and the competitive overlap value is very comparable, 27.3. However, while McGinn has previously handled a patent that was on similar semiconductor manufacturing methods, the patent at the first percentile of the distribution for Venable is not quite as close. This patent, #7601993, is also related to semiconductors but focused on an "ohmic recessed electrode" The similarity (expertise) value for this patent is 0.64, which is lower compared to McGinn's 0.82.

Moderator Variables

Our theory proposes three moderator variables. We measure them as follows. First, to measure the level of *embeddedness* between the patenting organization and the law firm, we assess the strength of their prior relationship by counting the number of patents on which the patenting organization and the law firm have cooperated over the three years prior to the focal year (Marsden and Campbell, 1984). Because the distribution of the counts is highly skewed, we use log counts.

To measure law firm *status*, we collect data from the 2008 Am Law 100 report. This report

a law firm has worked on over the past three years. The natural log of the count variable has been included in regression models in Appendix A.5 to ensure that our results are robust. A notable difference here is that this measure defines competition at the level of the patent, and not the organization. This is important because the intensity of competitive overlap with a single organization might be more important than the number of competing clients with which a law firm has worked. The consistency of our results ensures that findings are robust to a conceptualization of competition at the level of the patent or organization.

lists the 100 most successful US law firms in terms of annual revenue, "a common measure by which law firms are evaluated and ranked within the industry" (Somaya, Williamson, and Lorinkova, 2008, p.942). Given limited access, we use the 2008 ranking as it is from the middle of our analysis' time span and as the rankings are characterized by high levels of stability over time. This status measure helps control for a potential confounding effect which would be characterized by patenting organizations being more likely to choose law firms with lower levels of expertise because these are higher in status.

We create a measure for *law firm supply* by calculating the law firm-to-patenting organization ratio based on each patent's "knowledge neighborhood". Specifically, for each patent, we calculate pairwise distances to all other patents in the three years prior to the application of the focal patent. Then, we keep those patents that are within the first percentile of distances (i.e. the 1 percent closest patents). To arrive at our measure, we divide the distinct number of law firms by the number of distinct patenting organizations within this subset. Higher values indicate that there are many different law firms available per patenting organization that have some level of relevant expertise. Lower numbers indicate that a patenting organization has to vie for scarce expert representation with many other patenting organizations.

Control Variables

To capture *specialization*, we calculate the average pairwise similarity between all patents handled by a law firm. This is a time-varying measure, calculated for each year in our data based on the three previous years' patents. Higher average similarities indicate that most patents handled by the law firm are close to one another in the technology space. We also control for the *mean client size* of a law firm which has been calculated over a three year window (using one observation per client-year). The size of the clients themselves has been measured as the logged count of their patents.

Additionally, our specifications control for measures used in our matching procedure. We control for the *size* of the law firm which is the logged count of patents handled by the law firm over the past three years and the *founding year* of the law firm.[17]

Law Firm Choice Results

Table 2 provides summary statistics of the main variables, and a correlation table is included in the appendix in Table A.2. The expertise variable ranges from zero to one with a mean of 0.39 and a standard deviation of 0.27. Thus, the regression coefficients of expertise indicate the effect size of a change from a law firm with the lowest level of expertise possible to one with the highest level possible. Our competitive overlap variable ranges from ranges from zero to 257, with a mean of 25 and a standard deviation of 30.

—– INSERT TABLE 2 HERE —–

Table 3 shows results of our main models. The first model shows a specification with main effects only and finds that both higher levels of expertise and competitive overlap increase the likelihood of a law firm being selected. The expertise coefficient of 0.207 corresponds to a six percent (0.207*0.27) increase in the likelihood of being selected for a one standard-deviation increase in expertise. A one standard-deviation increase in competitive overlap is associated with a six percent (0.002*29.9) higher likelihood of selection. In model 2, we add the interaction term and find a significant, negative interaction between expertise and competitive overlap. This indicates that expertise influences selection *less* when competitive overlap is high and provides evidence in support of hypothesis 1. Figure 1 graphs the marginal effect of expertise over values of competitive overlap reaching from zero to the 99th percentile of the distribution. We see that patenting organizations are more likely to work with law firms with higher levels of expertise when there is little competitive overlap, as evidenced by the

[17]We ran additional specifications of our models including control variables for the average time it takes a law firm to get a patent accepted, the average number of citations their patents receive, and the average number of claims that their accepted patents get granted (all calculated based on the three years prior to the focal patent). Our results remain qualitatively unchanged upon the inclusion of these additional controls.

positive marginal effect of expertise on the left side of the graph. However, this relationship turns negative for high levels of competitive overlap as indicated by the graphed line crossing zero. When competitive overlap is zero, a one unit increase in expertise is associated with a 24% decrease in the probability of selection. For a one standard deviation increase in expertise this number corresponds to about seven percent (0.244*0.27). One standard deviation above the mean of competitive overlap, the marginal effect of expertise on choice is 0.7 percent ([0.244-0.004*[24.9+29.9]]*0.27), while two standard deviations above the mean of competitive overlap it is -1.0 percent ([0.244-0.004*[24.9+29.9*2]]*0.27), indicating that at high levels of competitive overlap organizations will avoid law firms with higher levels of expertise.

—– INSERT TABLE 3 HERE —–

—– INSERT FIGURE 1 HERE —–

Table 4 shows regression results testing hypotheses 2a and 2b. We use the same set of control variables from our prior models. In model 1, we moderate the expertise-competitive overlap trade-off by introducing a triple interaction with the number of repeated interactions between the patenting organization and a law firm (to measure embeddedness). We continue to find positive main effects for both expertise and competitive overlap as well as a negative interaction between the two. The triple interaction term is positive, indicating that the negative effect of expertise at high levels of competitive overlap is ameliorated when an extensive prior relationship exists between the patenting organization and law firm, as predicted in hypothesis 2a.[18] This is shown in Figure 2. At low and medium levels of a prior relationship, we continue to see the expertise-competitive overlap trade-off, as is indicated by the negative slope of the blue, solid and the purple, dashed lines. When an extensive prior

[18]Although we did not explicitly hypothesize about the expertise-prior relationship interaction, a negative effect of expertise for relationships with an extensive history is expected. If an extensive prior relationship exists between an organization and law firm, this could create trust that makes up for a lack of expertise. Therefore, an organization is more likely to choose the law firm they know even if it is lower in expertise (thereby creating a negative coefficient for the expertise-prior relationship interaction).

relationship exists, the effect of expertise is less sensitive to changes in competitive overlap. It is barely negative at low levels of competitive overlap and turns just positive at higher levels of competitive overlap as indicated by the yellow line. Thus, we find evidence that supports hypothesis 2a.

—– INSERT TABLE 4 HERE —–

—– INSERT FIGURE 2 HERE —–

Model 2 of Table 4 shows how the expertise-competitive overlap trade-off is moderated by status as predicted in hypothesis 2b whereby the trade-off was expected to be smaller for high status law firms. We add a triple interaction with an indicator for a law firm status. We continue to find evidence for our interaction as indicated by the positive main effects of expertise and competitive overlap and expertise as well as the negative interaction term between the two. Again, we find that the triple interaction term is positive and thereby weakening the main trade-off. Figure 3 shows that the slope for high status law firms (those listed in the AmLaw 100) is close to zero. This indicates that the effect of expertise is less sensitive to changes in competitive overlap when a law firm is high status and supports hypothesis 2b.

—– INSERT FIGURE 3 HERE —–

—– INSERT TABLE 5 HERE —–

In Table 5, we show how the expertise-competitive overlap trade-off changes with law firm supply as predicted in hypothesis 2c. To illustrate this we split the sample at the median of our law firm supply measure and run regressions on the two different subsamples.[19]

[19]We chose a split-sample approach for this hypothesis since the law firm supply is calculated at the patent level and thus perfectly subsumed by our patent fixed effects. When estimating regressions that include a triple interaction with law firm supply but no main effect of law firm supply, in Appendix A.6, we find that this difference (as shown by the negative coefficient on the triple interaction term) is statistically significant. This indicates support for our hypothesis of differing slopes, i.e., that the trade-off becomes more pronounced in industries in which many alternative law firms are available.

The sample in which law firm supply is below the median is restricted to patents whose "neighborhood" in the knowledge space consists of fewer law firms per patenting organization, while the sample in which law firm supply is high is restricted to patents that are in areas of the knowledge space in which there are more law firms per patenting organizations. While both samples show the negative trade-off of our main hypothesis, interaction coefficients from Table 5 show that it is more pronounced when there are many available law firms per patenting organization in the space of technologically similar patents. In model 1, where law firm supply is below the median, the interaction coefficient is -0.002, but when law firm supply is above the median, in model 2, the interaction coefficient increases threefold in magnitude to -0.006 indicating a stronger trade-off. We illustrate this in Figure 4. When there are few law firms available per patenting organization (purple, solid line), the effect of expertise changes only gradually with competitive overlap and remains positive until after the 99th percentile of competitive overlap. However, when law firm supply is high (yellow, dashed line), the trade-off is more pronounced and the effect of expertise turns negative at levels of competitive overlap less than one standard deviation above the mean. These results add support to hypothesis 2c.

—— INSERT FIGURE 4 HERE ——

Additional Analyses: The Consequences of Choosing Partners with Lower Expertise

As we have shown, the expertise-competitive overlap trade-off affects partner choice and lowers the likelihood of firms choosing the law firm with the highest expertise. This, however, only matters if it affects firm performance. In a set of additional analyses, we test this intuition. In the setting of patent prosecution, there is abundant evidence that two key indicators, the time to acceptance of the patent and the number of forward citations a patent receives, are valid measures of patent performance and reasonable proxies of firm level innovation performance. The time to patent acceptance has important performance consequences for clients. Both faster times to acceptance as well as being the first to patent a technology are associated with higher growth and follow-on innovation (Farre-Mensa, Hegde, and Ljungqvist, 2020; Hegde, Ljungqvist, and Raj, 2020). Illustrating this point, an executive from Nokia was quoted by Reitzig and Puranam (2009, 784): "In this domain obtaining a fast grant is absolutely crucial, as you want to use the patent to add an additional protective edge to your existing product, and for that you'd rather have the patent today than tomorrow." Forward citations are also an important performance indicator because the use of the patent by other firms has both direct financial benefits such as licensing royalties and indirect benefits such as increased status or name recognition of the focal client. For our theory to have consequences for firm performance, we now proceed to show that choosing a law firm with lower expertise leads to longer patent acceptance times and fewer citations in the future.

For these analyses, because the performance measures are at the patent level, we move away from the matched dyads approach and only keep one observation per patent. The dependent variables are a patent's time to acceptance and the number of citations received. We measure these variables using the days passed between the original application and a patent's acceptance date, and the sum of all forward citations, respectively. As there is significant variation over time and between fields that could confound the relationship between expertise and our outcome measures, we add various controls and fixed effects to our models. We use year fixed effects to account for time trends and to ensure that the number of forward citations are comparable. In separate specifications, the models include patenting organization fixed effects to account for organizational variation in outcomes that is not driven by law firms and their expertise. Coefficients from these models thus can be interpreted as the effect of expertise when the same assigning organization patents a technologically similar invention, but chooses a law firm of lower(higher) expertise. Lastly, we add law firm fixed effects. Though our theory is about choosing one law firm over another, finding that higher levels of expertise are associated with better performance outcomes, even when holding the law firm constant, should further increase our confidence in the overall relationship between expertise and patent performance. We estimate ordinary least squares regressions with standard errors that are two-way clustered, at the law firm and at the patenting organization. As a robustness check, we also estimate Poisson regressions and find that our results remain qualitatively unchanged (results not shown and available upon request).

We add additional control variables to these models: First, we control for *patent breadth* as Hegde et al. (2020) suggest that patenting organizations enter into a trade-off between trying for a "quick or broad" patent. We therefore control for the number of claims made by a patent as an indicator of patent breadth. We further control for the existence of a *twin patent*, by which we mean the existence of the same patent with either the European or Japanese patent office. This is to account for confounding in terms of organizations choosing to work with law firms that have higher levels of relevant expertise when they aim to patent technologies which are of particular importance or high quality. If this is the case, they are more likely to apply for patents at other international patent offices as well.[20] Lastly we add the 38 dimensions of a patent's position in the knowledge space as control variables. In combination, these capture the patent's technology and thereby account for differences in time to acceptance or forward citations between technologies.

—– INSERT TABLE 6 HERE —–

Table 6 shows summary statistics of our control variables for the performance analysis models. We show pairwise correlations in the appendix in Table A.3. Results modeling the effect of expertise on patent-level outcomes are shown in Table 7. The coefficients of our expertise measure indicate a consistent negative, significant effect on the time to acceptance. We find that a one-standard deviation increase in expertise is associated with the process being 27 (0.27*-98) , 25, and 14 days faster for models 1, 2, and 3, respectively. Further, we

[20]The number of claims are only finalized during the patent examination process and thus both might be over-controlling or potentially opening up pathways of post-treatment bias (see e.g., Montgomery, Nyhan, and Torres, 2018). This is because they might be affected by our independent variables of interest. We ran models without controlling for the number of claims or whether a twin patent exists and find that our results remain largely unchanged.

find a positive effect of expertise on the logged number of future citations a patent receives ranging from 0.15 to 0.11. The coefficients indicate that a one-standard deviation increase in expertise corresponds to an increase in forward citations of about three to four percent.

— INSERT TABLE 7 HERE —

Discussion

Organizations often face a fundamental dilemma when choosing a business partner. On the one hand, they would prefer to work with the partner with the most relevant expertise. On the other hand, they would like to avoid partners who also work with their competitors. This poses a dilemma because an organization's competitors, by definition, are likely to require the same expertise that it does. As we argued, fear of information leakage and concerns about access to resources lead organizations to select less expert partners to avoid second-order competitive overlap, a choice that could have negative performance consequences.

In analyses of a sample of client firms selecting patent law firms across nearly one million patents, the likelihood of selecting a partner based on its expertise decreases significantly as competitive overlap increases, illustrating a trade-off between expertise and competitive overlap in partner selection. Consistent with our proposed mechanisms of information leakage and concerns about resource access, the trade-off weakens when the client firm and the law firm have a prior relationship, when the law firm is high status, and when few alternative expert law firms are available. In line with our expectation that choosing a partner with lower level of expertise will have negative downstream consequences, we showed that the time to patent acceptance is greater and forward citations are lower when a client firm chooses a patent law firm who is less expert in the technological subject matter. Hence, choosing a less expert partner may have negative performance consequences for firms.

Limitations and future research

Our study, naturally, is subject to limitations. Although patent data are advantageous for measuring expertise and competitive overlap, there are other aspects of the patent process that we are not able to observe. First, the risk set of law firms the patenting firms considered during the selection phase is not observable. Our matching approach, in which we match the law firm chosen to an equivalent law firm that is similar to the chosen one on observables but were not chosen, helps to alleviate concerns. However, it remains possible that unobserved characteristics of firms or dyads influence expertise or competitive overlap and the likelihood of selection. Settings in which the risk set during selection is known should be sought for future research. Nevertheless, to the extent that unobservable characteristics are correlated with observable ones, our matching approach provides confidence in the findings.

Second, like much prior research relying on patent data, we only analyze granted patent applications, but not rejected ones. Although data on unsuccessful patent applications are available from the USPTO's Public Pair system, these data do not include information on the original law firm that handled the applications. Observing granted patents could bias our results if success rate is related to law firm expertise. Yet, we do not think that this is a major problem because at the USPTO more than 75 percent of the submitted patent applications are eventually granted and many of the rejections occur because applicants either do not submit the revised patent applications on time or abandon them (Carley, Hedge, and Marco, 2015). The USPTO is likely to err on the side of granting patents rather than rejecting them.

Lastly, we observe only near term consequences to performance at the patent level. What we have shown in this paper is that choosing to avoid second-order competition and therefore settling for less than expert law firms has negative short term consequences. Yet, we do not claim that this practice is irrational, for example, it may be that in the long term it pays off by avoiding knowledge leakage, and it may be the case that firms that sacrifice choosing the experts in the short term may have higher survival rates in the long term. Future research on the short-term versus long-term consequences of choosing lower expertise in order to avoid second-order competition is needed to fully address these long term performance effects.

Contributions

Despite these limitations, our study offers several contributions. First, our theory extends prior dyadic perspectives on partner choice by incorporating second-order competitive overlap, a characteristic of the network surrounding the dyad. Specifically, the study highlights a fundamental trade-off between competitive overlap and relevant expertise in partner choice in market ties and argues for when this trade-off will have negative performance consequences. In so doing, it joins a stream of recent research bringing a network perspective to partner choice in interorganizational relationships (Rogan, 2014; Hernandez et al., 2015; Pahnke et al., 2015; Clough and Piezunka, 2020). More broadly, our arguments provide further insight into the forces of inertia and change in interorganizational relationships (Kim et al., 2006; Ahuja, Soda, and Zaheer, 2012). Indeed, a perspective on partner choice that considers the interplay between dyadic and network characteristics is necessary for a theory of interorganizational network evolution.

Second, our study contributes to the literature on innovation by highlighting a double-edged sword of competition for innovation (Pahnke et al., 2015). Competition can stimulate progress in technological fields (Schumpeter, 1934). But as we have shown, under certain circumstances, competition can also hinder innovation if it leads to the selection of less expert law firms, which in turn slows down the patenting process and results in lower quality patents. This pattern is likely to be even more pronounced when enforcement of intellectual property rights is low. In such cases, one might expect market relationships to be dominated by long standing ties, where partners are chosen based on trust or exclusivity, in order to avoid competitive overlap, rather than based on expertise. A cost of this security is likely to be lower innovation quality.

Third, our study also provides a methodological contribution. While our paper is not the first to apply deep learning text analyses techniques to analyze patent text (see e.g., Lee and Hsiang (2020)), we believe that we are the first to use this technique to refine measures of technological positioning, providing a more detailed approach than existing ones such as those using patent classification data. Moreover, deep learning text analysis is easy to adapt to any text data such as press releases, annual reports, or newspaper coverage, and therefore opens the potential for organization research on organizational positioning. In addition, for network research, text based analyses techniques may be particularly advantageous for research by enabling more refined measures, such as the measure of expertise in this study, compared to other measures of nodal characteristics such as industry, size or age to estimate likelihood of affiliation. This approach could be useful for disentangling partner characteristic explanations from network explanations for partner selection (Baum, Cowan, and Jonard, 2010).

In closing, our study considers constraints posed by indirect competition and how this can lead to suboptimal outcomes for firms who are not able to partner with experts. We believe that our ideas apply more widely to other settings such as organizations choosing consulting firms and auditors, entrepreneurial ventures choosing venture capital investors, or governments choosing defense contractors. In these situations, the quality of the partner chosen is both enabled and constrained by the networks of connections among firms. Connections provide the basis for developing expertise, but at the same time, they create situations of competitive friction as the network evolves (Uribe et al., 2020). Hence, our research addresses a call from Podolny and Page (1998), who argued that "researchers must counterbalance the focus on prevalence and functionality (of networks) with an equally strong focus on constraint and dysfunctionality" (1998: 73). We call for continued investigation into partner selection that considers both the choice and constraint that networks bring to organizations.

References

Ahuja, Gautam, Giuseppe Soda, and Akbar Zaheer
2012. "The genesis and dynamics of organizational networks." Organization Science 23:434–448.

Ai, Chunrong, and Edward C. Norton
2003. "Interaction terms in logit and probit models." Economics Letters 80:123–129.

AIPLA
2009. *AIPLA Report of Economic Survey, 2009.* Washington, DC: Fetzer-Kraus, Inc. American Intellectual Property Law Association.

Alcácer, Juan, and Wilbur Chung
2007. "Location Strategies and Knowledge Spillovers." Management Science 53:760–776.

Aobdia, Daniel
2015. "Proprietary information spillovers and supplier choice: Evidence from auditors." Review of Accounting Studies 20:1504–1539.

Arrow, Kenneth
1974. *The Limits of Organizations.* New York, NY: Norton Co.

Asker, John, and Alexander Ljungqvist
2010. "Competition and the Structure of Vertical Relationships in Capital Markets." Journal of Political Economy 118:599–647.

Baker, Wayne E
1990. "Market networks and corporate behavior." American Journal of Sociology 96:589–625.

Baker, Wayne E, and Robert R Faulkner
1991. "Strategies for managing suppliers of professional services." California Management Review 33:33–45.

Baker, Wayne E., Robert R. Faulkner, and Gene A. Fisher
1998. "Hazards of the Market: The Continuity and Dissolution of Interorganizational Market Relationships." American Sociological Review 63:147–177.

Barley, Stephen R
1996. "Technicians in the workplace: Ethnographic evidence for bringing work into organizational studies." Administrative Science Quarterly pp. 404–441.

Baum, Joel AC, Robin Cowan, and Nicolas Jonard
2010. "Network-independent partner selection and the evolution of innovation networks." Management science 56:2094–2110.

Baum, Joel AC, Timothy J Rowley, Andrew V Shipilov, and You-Ta Chuang
2005. "Dancing with strangers: Aspiration performance and the search for underwriting syndicate partners." Administrative Science Quarterly 50:536–575.

Baum, Joel AC, and Jitendra V Singh
1994. "Organizational niches and the dynamics of organizational mortality." American Journal of Sociology 100:346–380.

Bernheim, B Douglas, and Michael D Whinston
1990. "Multimarket contact and collusive behavior." The RAND Journal of Economics 21:1–26.

Carley, Michael, Deepak Hedge, and Alan Marco
2015. "What is the probability of receiving a US patent." Yale JL & Tech. 17:203.

Carnabuci, Gianluca, Elisa Operti, and Balázs Kovács
2015. "The Categorical Imperative and Structural Reproduction: Dynamics of Technological Entry in the Semiconductor Industry." Organization Science 26:1734–1751.

Census
2010. *Census Regions and Divisions of the United States*. U.S. Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau.

Chen, Ming-Jer, and Danny Miller
2012. "Competitive dynamics: Themes, trends, and a prospective research platform." Academy of Management Annals 6:135–210.

Chollet, Francois
2017. "Deep learning with Python." Greenwich, CT: Manning Publications CO .

Chondrakis, George, and Mari Sako
2020. "When suppliers shift my boundaries: Supplier employee mobility and its impact on buyer firms' sourcing strategy." Strategic Management Journal .

Clough, David R., and Henning Piezunka
2020. "Tie Dissolution in Market Networks: A Theory of Vicarious Performance Feedback." Administrative Science Quarterly 65:972–1017.

Cohen, Wesley M., and Daniel A. Levinthal
1990. "Absorptive Capacity: A New Perspective on Learning and Innovation." Administrative Science Quarterly 35:128–152.

Dierickx, Ingemar, and Karel Cool
1989. "Asset Stock Accumulation and Sustainability of Competitive Advantage." Management Science 35:1504–1511.

Dobrev, Stanislav D, Tai-Young Kim, and Michael T Hannan
2001. "Dynamics of niche width and resource partitioning." American Journal of Sociology 106:1299–1337.

Downing, Stephen T, Jin-Su Kang, and Gideon D Markman
2019. "What You Don't See Can Hurt You: Awareness Cues to Profile Indirect Competitors." Academy of Management Journal 62:1872–1900.

Dushnitsky, Gary, and J. Myles Shaver
2009. "Limitations to Interorganizational Knowledge Acquisition: The Paradox of Corporate Venture Capital." Strategic Management Journal 30:1045–1064.

Farre-Mensa, Joan, Deepak Hegde, and Alexander Ljungqvist
2020. "What Is a Patent Worth? Evidence from the U.S. Patent "Lottery"." The Journal of Finance 75:639–682.

Gimeno, Javier, and Carolyn Y Woo
1996. "Hypercompetition in a multimarket environment: The role of strategic similarity and multimarket contact in competitive de-escalation." Organization Science 7:322–341.

Godart, Frédéric C, Andrew V Shipilov, and Kim Claes
2014. "Making the most of the revolving door: The impact of outward personnel mobility networks on organizational creativity." Organization Science 25:377–400.

Goerzen, Anthony
2007. "Alliance networks and firm performance: The impact of repeated partnerships." Strategic Management Journal 28:487–509.

Gould, Roger V
2002. "The origins of status hierarchies: A formal theory and empirical test." American Journal of Sociology 107:1143–1178.

Graffin, Scott D, Jonathan Bundy, Joseph F Porac, James B Wade, and Dennis P Quinn
2013. "Falls from grace and the hazards of high status: The 2009 British MP expense scandal and its impact on parliamentary elites." Administrative Science Quarterly 58:313–345.

Greve, Henrich R, Hitoshi Mitsuhashi, and Joel AC Baum
2013. "Greener pastures: Outside options and strategic alliance withdrawal." Organization Science 24:79–98.

Grigoriou, Konstantinos, and Frank T. Rothaermel
2017. "Organizing for knowledge generation: internal knowledge networks and the contingent effect of external knowledge sourcing." Strategic Management Journal 38:395–414.

Gulati, Ranjay
1995. "Does Familiarity Breed Trust? The Implications of Repeated Ties for Contractual Choice in Alliances." The Academy of Management Journal 38:85–112.

Gulati, Ranjay, and Martin Gargiulo
1999. "Where Do Interorganizational Networks Come From?" American Journal of Sociology 104:1439–1493.

Gulati, Ranjay, and Maxim Sytch
2007. "Dependence asymmetry and joint dependence in interorganizational relationships:

Effects of embeddedness on a manufacturer's performance in procurement relationships." Administrative Science Quarterly 52:32–69.

Gulati, Ranjay, and Maxim Sytch
2008. "Does familiarity breed trust? Revisiting the antecedents of trust." Managerial and Decision Economics 29:165–190.

Hegde, Deepak, Alexander Ljungqvist, and Manav Raj
2020. "Quick or Broad Patents?" Working Paper, NYU .

Hernandez, Exequiel, Wm. Gerard Sanders, and Anja Tuschke
2015. "Network Defense: Pruning, Grafting, and Closing to Prevent Leakage of Strategic Knowledge to Rivals." Academy of Management Journal 58:1233–1260.

Jensen, Michael, and Aradhana Roy
2008. "Staging exchange partner choices: When do status and reputation matter?" Academy of Management Journal 51:495–516.

Katila, Riitta, Jeff D. Rosenberger, and Kathleen M. Eisenhardt
2008. "Swimming with Sharks: Technology Ventures, Defense Mechanisms and Corporate Relationships." Administrative Science Quarterly 53:295–332.

Katz, Michael L., and Carl Shapiro
1985. "Network Externalities, Competition, and Compatibility." The American Economic Review 75:424–440.

Katz, Michael L, and Carl Shapiro
1986. "Technology adoption in the presence of network externalities." Journal of Political Economy 94:822–841.

Kim, Tai-Young, Hongseok Oh, and Anand Swaminathan
2006. "Framing interorganizational network change: A network inertia perspective." Academy of Management Review 31:704–720.

King, Gary, Richard Nielsen, Carter Coberley, James E. Pope, and Aaron Wells
2011. "Comparative Effectiveness of Matching Methods for Causal Inference." Working Paper .

Kovács, Balázs, Gianluca Carnabuci, and Filippo Carlo Wezel
2021. "Categories, Attention, and the Impact of Inventions." Strategic Management Journal .

Kuhn, Jeffrey, Kenneth Younge, and Alan Marco
2020. "Patent citations reexamined." The RAND Journal of Economics 51:109–132.

Le Mens, Gaël, Michael T Hannan, and László Pólos
2015. "Age-related structural inertia: A distance-based approach." Organization Science 26:756–773.

Lee, Jieh-Sheng, and Jieh Hsiang
2020. "Patent classification by fine-tuning BERT language model." World Patent Information 61:101965.

Levinthal, Daniel A, and Mark Fichman
1988. "Dynamics of interorganizational attachments: Auditor-client relationships." Administrative Science Quarterly pp. 345–369.

Levitt, Barbara, and James G. March
1988. "Organizational Learning." Annual Review of Sociology 14:319–338.

Marsden, Peter V, and Karen E Campbell
1984. "Measuring tie strength." Social Forces 63:482–501.

Mayer, Kyle J., Deepak Somaya, and Ian O. Williamson
2012. "Firm-Specific, Industry-Specific, and Occupational Human Capital and the Sourcing of Knowledge Work." Organization Science 23:1311–1329.

Mitsuhashi, Hitoshi, and Henrich R Greve
2009. "A matching theory of alliance formation and organizational success: Complementarity and compatibility." Academy of Management Journal 52:975–995.

Moeen, Mahka, Deepak Somaya, and Joseph T Mahoney
2013. "Supply portfolio concentration in outsourced knowledge-based services." Organization Science 24:262–279.

Montgomery, Jacob M., Brendan Nyhan, and Michelle Torres
2018. "How Conditioning on Posttreatment Variables Can Ruin Your Experiment and What to Do about It." American Journal of Political Science 62:760–775.

Mood, Carina
2010. "Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It." European Sociological Review 26:67–82.

Mowery, David C., Joanne E. Oxley, and Brian S. Silverman
1996. "Strategic alliances and interfirm knowledge transfer." Strategic Management Journal 17:77–91.

Mowery, David C, Joanne E Oxley, and Brian S Silverman
1998. "Technological overlap and interfirm cooperation: implications for the resource-based view of the firm." Research policy 27:507–523.

Pahnke, Emily Cox, Rory McDonald, Dan Wang, and Benjamin Hallen
2015. "Exposed: Venture Capital, Competitor Ties, and Entrepreneurial Innovation." Academy of Management Journal 58:1334–1360.

Parkhe, Arvind
1993. ""Messy" research, methodological predispositions, and theory development in international joint ventures." Academy of Management Review 18:227–268.

Peteraf, Margaret A.
1993. “The cornerstones of competitive advantage: A resource-based view.” Strategic Management Journal 14:179–191.

Pfeffer, Jeffrey, and R Salancik, Gerald
1978. *The external control of organizations: A resource dependence perspective*. New York: Harper & Row.

Phelps, Corey, Ralph Heidl, and Anu Wadhwa
2012. “Knowledge, Networks, and Knowledge Networks:A Review and Research Agenda.” Journal of Management 38:1115–1166.

Podolny, Joel M
1993. “A status-based model of market competition.” American Journal of Sociology 98:829–872.

Podolny, Joel M, and Karen L Page
1998. “Network forms of organization.” Annual Review of Sociology 24:57–76.

Podolny, Joel M., and Toby E. Stuart
1995. “A Role-Based Ecology of Technological Change.” American Journal of Sociology 100:1224–1260.

Podolny, Joel M., Toby E. Stuart, and Michael T. Hannan
1996. “Networks, Knowledge, and Niches: Competition in the Worldwide Semiconductor Industry, 1984-1991.” American Journal of Sociology 102:659–689.

Pontikes, Elizabeth G., and William P. Barnett
2017. “The Coevolution of Organizational Knowledge and Market Technology.” Strategy Science 2:64–82.

Reitzig, Markus, and Phanish Puranam
2009. “Value appropriation as an organizational capability: the case of IP protection through patents.” Strategic Management Journal 30:765–789.

Rogan, Michelle
2014. “Too Close for Comfort? The Effect of Embeddedness and Competitive Overlap on Client Relationship Retention Following an Acquisition.” Organization Science 25:185–203.

Rogan, Michelle, and Henrich R Greve
2015. “Resource dependence dynamics: Partner reactions to mergers.” Organization Science 26:239–255.

Rogan, Michelle, and Olav Sorenson
2014. “Picking a (Poor) Partner:A Relational Perspective on Acquisitions.” Administrative Science Quarterly 59:301–329.

Romer, Paul M.
1987. “Growth Based on Increasing Returns Due to Specialization.” The American Economic Review 77:56–62.

Rosenbaum, Paul R. Rubin, Donald
1983. "The central role of the propensity score in observational studies for causal effects." Biometrika 70:41–55.

Ryall, Michael D, and Olav Sorenson
2007. "Brokers and competitive advantage." Management Science 53:566–583.

Sandefur, Rebecca L
2015. "Elements of professional expertise: Understanding relational and substantive expertise through lawyers' impact." American Sociological Review 80:909–933.

Schumpeter, J. A.
1934. *The Theory of Economic Development*. Cambridge Ma.: Harvard Economic Studiese.

Shepard, Roger N
1987. "Toward a universal law of generalization for psychological science." Science 237:1317–1323.

Somaya, Deepak, Ian O. Williamson, and Natalia Lorinkova
2008. "Gone but Not Lost: The Different Performance Impacts of Employee Mobility between Cooperators versus Competitors." The Academy of Management Journal 51:936–953.

Somaya, Deepak, Ian O. Williamson, and Xiaomeng Zhang
2007. "Combining Patent Law Expertise with R&D for Patenting Performance." Organization Science 18:922–937.

Sorenson, Olav, and Toby E. Stuart
2008. "Bringing the Context Back In: Settings and the Search for Syndicate Partners in Venture Capital Investment Networks." Administrative Science Quarterly 53:266–294.

Stuart, Toby E
1998. "Network positions and propensities to collaborate: An investigation of strategic alliance formation in a high-technology industry." Administrative Science Quarterly 43:668–698.

Stuart, Toby E., Ha Hoang, and Ralph C. Hybels
1999. "Interorganizational Endorsements and the Performance of Entrepreneurial Ventures." Administrative Science Quarterly 44:315–349.

Sytch, Maxim, and Ranjay Gulati
2013. "Markets as Networks." In The Palgrave Encyclopedia of Strategic Management, edited by Teece D.J. Augier M. London: Palgrave Macmillan.

Sytch, Maxim, and Yong H Kim
2020. "Quo Vadis? From the Schoolyard to the Courtroom." Administrative Science Quarterly 66:177–219.

Sytch, Maxim, and Adam Tatarynowicz
2014. “Friends and Foes: The Dynamics of Dual Social Structures.” Academy of Management Journal 57:585–613.

Teece, David J
2018. “Expert Services.” In The Palgrave Encyclopedia of Strategic Managementpp. 540–543. Palgrave Macmillan, London.

Uribe, Jose, Maxim Sytch, and Yong H Kim
2020. “When friends become foes: Collaboration as a catalyst for conflict.” Administrative Science Quarterly 65:751–794.

Uzzi, Brian, and James J Gillespie
2002. “Knowledge spillover in corporate financing networks: Embeddedness and the firm’s debt performance.” Strategic Management Journal 23:595–618.

Vissa, Balagopal
2011. “A matching theory of entrepreneurs’ tie formation intentions and initiation of economic exchange.” Academy of Management Journal 54:137–158.

Younge, Kenneth A., and Jeffrey M. Kuhn
2016. “Patent-to-Patent Similarity: A Vector Space Model.” SSRN Working Paper .

Zhelyazkov, Pavel I
2018. “Interactions and interests: Collaboration outcomes, competitive concerns, and the limits to triadic closure.” Administrative Science Quarterly 63:210–247.

Tables and Figures

Table 1: Balance table comparing matched control cases to chosen law firms

Variable	(1) Matched controls Mean/SE	(2) Observed lawyers Mean/SE	T-test Difference (1)-(2)	Normalized difference (1)-(2)
Law firm size	6.65	6.71	-0.06***	-0.04
	(0.00)	(0.00)		
Law firm founding year	1986.67	1986.67	-0.01	-0.00
	(0.02)	(0.02)		
Law firm NBER experience (binary)	0.86	0.88	-0.02***	-0.06
	(0.00)	(0.00)		
N	963089	963089		

Notes: ***p<.001; **p<.01; *p<.05

Table 2: Summary statistics on key variables

	(1)			
	Mean	Std Dev	Min	Max
Law firm expertise	0.39	(0.27)	0.001	1
Competitive overlap	24.9	(29.9)	0.001	257.9
Log competitor count	2.58	(1.58)	0	7.7
Law firm-client embeddedness	2.36	(2.58)	0	8.6
Law firm specialization	0.54	(0.10)	0.01	1.4
Law firm mean client size	6.38	(0.86)	0.7	10.4
Law firm size	6.68	(1.62)	0	9.5
Law firm founding year	1986.7	(17.1)	1903	2011
Law firm NBER experience (binary)	0.87	(0.34)	0	1
Law firm status	0.16	(0.37)	0	1
Law firm supply	0.94	(0.24)	0.5	1.9
Observations	1926178			

Table 3: Linear probability models with patent fixed effects predicting law firm choice

VARIABLES	(1) no interaction	(2) main model
Law firm expertise	0.207***	0.244***
	(0.002)	(0.002)
Competitive overlap	0.002***	0.004***
	(0.000)	(0.000)
Expertise X competitive overlap		-0.004***
		(0.000)
Law firm-client embeddedness	0.190***	0.190***
	(0.000)	(0.000)
Law firm status	0.020***	0.023***
	(0.001)	(0.001)
Law firm specialization	-0.353***	-0.335***
	(0.004)	(0.004)
Law firm mean client size	-0.025***	-0.025***
	(0.000)	(0.000)
Law firm size	0.003**	0.006***
	(0.001)	(0.001)
Law firm founding year	-0.000	-0.000*
	(0.000)	(0.000)
Constant	0.292***	0.330***
	(0.052)	(0.052)
Observations	1,926,178	1,926,178
R-squared	0.804	0.805
Patent FEs	included	included

Standard errors clustered by patent

***p<.001; **p<.01; *p<.05

Table 4: Linear probability models predicting law firm choice - triple interactions with embeddedness and status

	(1)	(2)
Law firm expertise	0.331***	0.283***
	(0.002)	(0.002)
Competitive overlap	0.005***	0.004***
	(0.000)	(0.000)
Law firm-client embeddedness	0.227***	0.190***
	(0.000)	(0.000)
Expertise X competitive overlap	-0.004***	-0.005***
	(0.000)	(0.000)
Expertise X embeddedness	-0.065***	
	(0.000)	
Comp. overlap X embeddedness	-0.001***	
	(0.000)	
Expert. X comp. X emb.	0.001***	
	(0.000)	
Law firm status	0.025***	0.105***
	(0.001)	(0.002)
Expertise X status		-0.238***
		(0.004)
Comp. overlap X status		-0.002***
		(0.000)
Exp. X comp. X status		0.005***
		(0.000)
Observations	1,926,178	1,926,178
R-squared	0.809	0.805
Patent FEs	included	included
Controls	included	included

Standard errors clustered by patent
***p<.001; **p<.01; *p<.05
Models also control for specialization, client size, size, and founding year

Table 5: Linear probability models predicting law firm choice - split sample analysis by law firm supply

	(1) Sample: Law firm supply below median	(2) Sample: Above median
Law firm expertise	0.288***	0.231***
	(0.003)	(0.002)
Competitive overlap	0.003***	0.005***
	(0.000)	(0.000)
Expertise X competitive overlap	-0.002***	-0.006***
	(0.000)	(0.000)
Observations	963,078	963,100
R-squared	0.790	0.824
Patent FEs	included	included
Controls	included	included

Standard errors clustered by patent
***p<.001; **p<.01; *p<.05
Models control for law firm-client embeddedness, law firm specialization, law firm mean client size, law firm size, law firm founding year, and law firm status

Table 6: Summary statistics for patent performance models

	(1) Mean	Std Dev	Min	Max
Patent days to acceptance	1266.6	(585.5)	89	5326
Patent citations	0.70	(0.99)	0	8.1
Law firm expertise	0.46	(0.27)	0.001	1
Competitive overlap	27.6	(31.8)	0.001	257.9
Law firm-client embeddedness	4.37	(1.95)	0	8.6
Law firm specialization	0.54	(0.098)	0.01	1.4
Law firm mean client size	6.44	(0.82)	0.7	10.4
Law firm size	6.71	(1.66)	0	9.5
Patent breadth	2.67	(0.59)	0.7	5.9
Twin patent	0.25	(0.43)	0	1
Observations	963089			

Table 7: Linear model (OLS) predicting days to patent acceptance and number of forward citations

	Patent days to acceptance			Log patent citations		
	(1)	(2)	(3)	(4)	(5)	(6)
Expertise	-98.38***	-91.13***	-50.06***	0.15***	0.12***	0.11***
	(15.67)	(14.44)	(11.33)	(0.02)	(0.01)	(0.01)
Competitive overlap	0.74***	0.48**	-0.59*	-0.00**	-0.00***	-0.00***
	(0.20)	(0.16)	(0.26)	(0.00)	(0.00)	(0.00)
Law firm-client embeddedness	-39.69***	-51.98***	-52.46***	0.01**	0.02***	0.02***
	(3.56)	(3.58)	(3.19)	(0.00)	(0.00)	(0.00)
Law firm specialization	120.63+	101.57	199.78*	-0.02	-0.10+	-0.07
	(68.21)	(82.69)	(86.42)	(0.06)	(0.06)	(0.07)
Law firm mean client size	19.21***	-5.43	-25.50+	-0.02***	-0.01+	0.00
	(5.57)	(6.96)	(13.08)	(0.01)	(0.01)	(0.01)
Law firm size	-5.58	-5.47	-15.70***	0.00	0.00	0.00*
	(3.46)	(3.50)	(4.29)	(0.00)	(0.00)	(0.00)
Patent breadth	55.24***	49.42***	46.12***	0.18***	0.11***	0.11***
	(6.66)	(6.57)	(6.16)	(0.01)	(0.00)	(0.00)
Twin patent	81.79***	39.56***	36.93***	0.08***	0.09***	0.09***
	(9.40)	(7.60)	(7.40)	(0.01)	(0.01)	(0.01)
Constant	1,030.76*	887.45+	999.44*	-0.18	0.26	0.25
	(501.40)	(466.61)	(448.94)	(0.47)	(0.39)	(0.41)
Observations	960,628	960,615	960,446	960,628	960,615	960,446
R-squared	0.18	0.28	0.32	0.11	0.16	0.17
Year FEs	included	included	included	included	included	included
Knowledge space dimensions	included	included	included	included	included	included
Patenting org FEs		included	included		included	included
Law firm FEs			included			included

Standard errors clustered by patenting organization and law firm
***p<.001; **p<.01; *p<.05

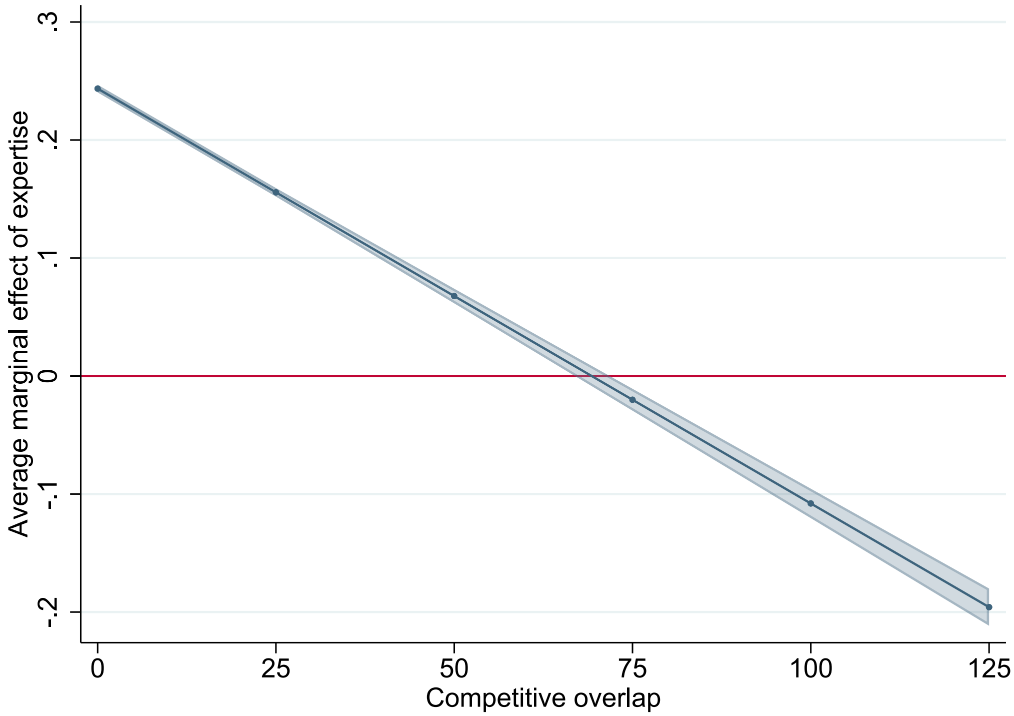

Figure 1: Competitive overlap and expertise interaction (based on model 2 of Table 3)

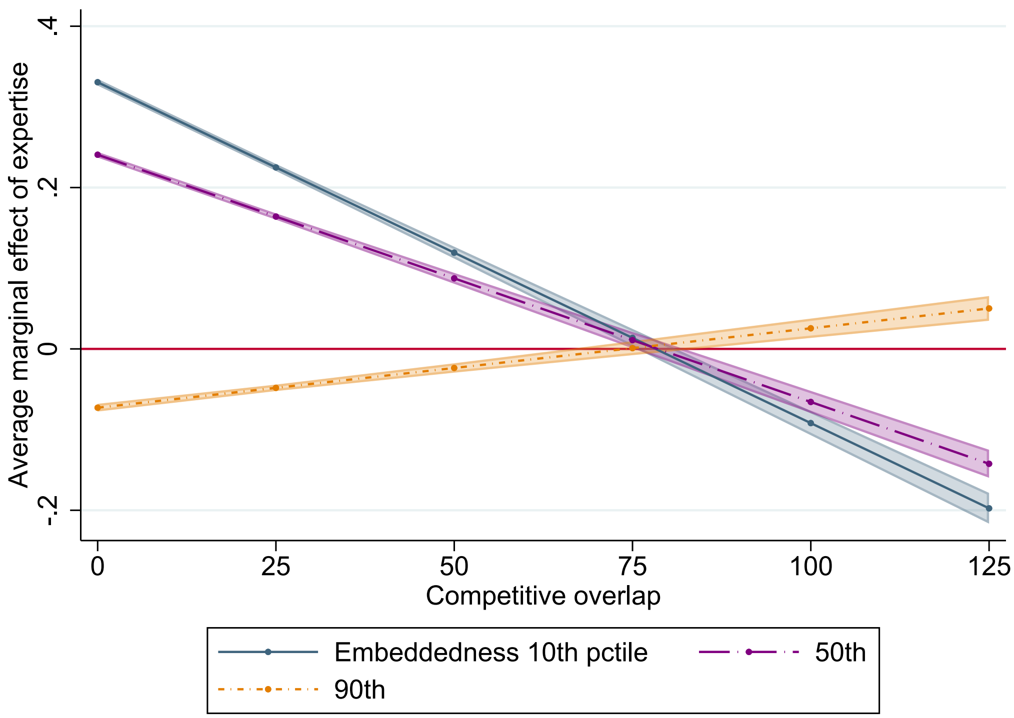

Figure 2: Competitive overlap and expertise trade-off by law firm-client embeddedness (based on model 1 of Table 4)

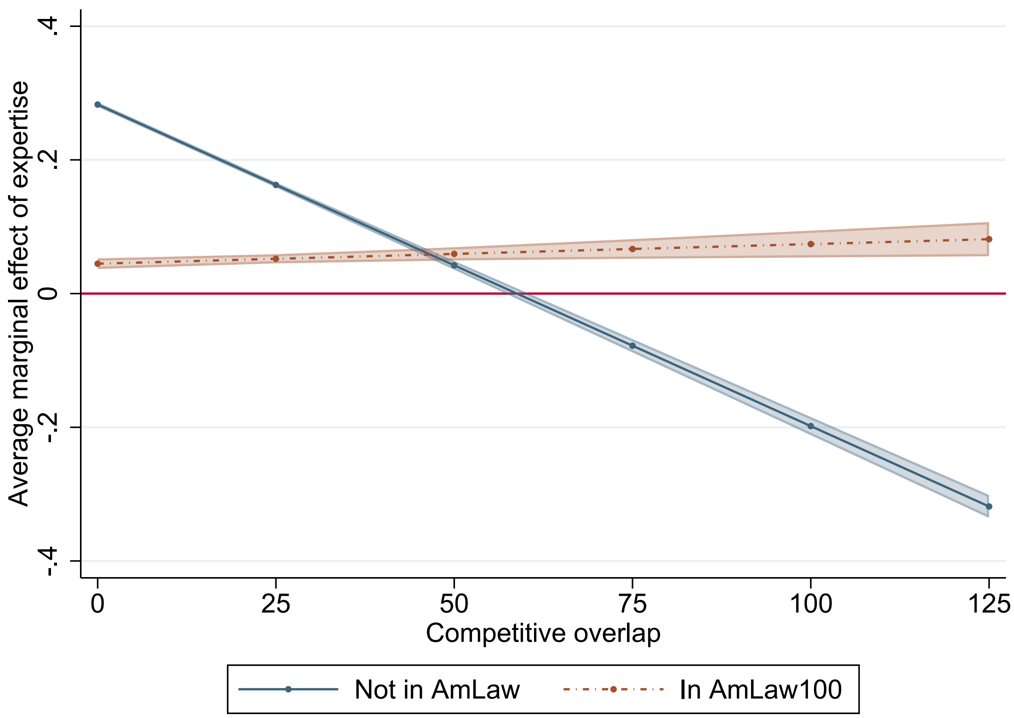

Figure 3: Competitive overlap and expertise trade-off by status (based on model 2 of Table 4)

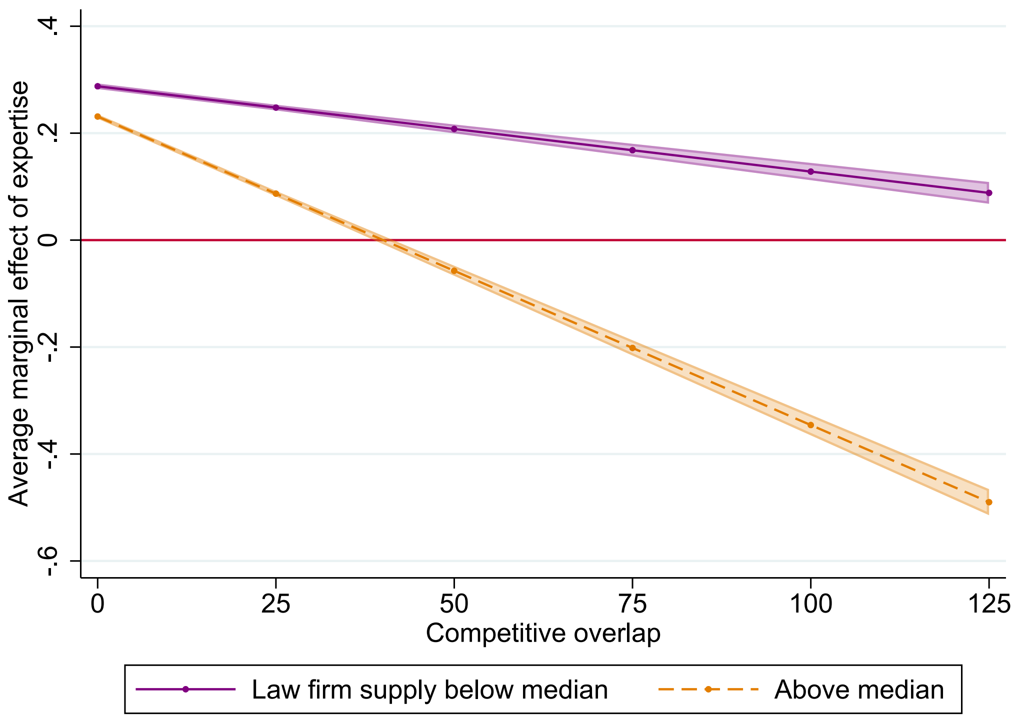

Figure 4: Competitive overlap and expertise trade-off by law firm supply (based on models 1 and 2 of Table 5)

A Appendix

US006542353B2

(12) **United States Patent**
Ardrey et al.

(10) **Patent No.:** **US 6,542,353 B2**
(45) **Date of Patent:** **Apr. 1, 2003**

(54) **CONTROL PANEL ASSEMBLY AND METHOD OF MAKING SAME**

(75) Inventors: **Kenneth J. Ardrey**, Canton, MI (US); **Mark R. Weston**, Brighton, MI (US)

(73) Assignee: **Key Plastics, LLC**, Northville, MI (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/972,335**

(22) Filed: **Oct. 5, 2001**

(65) **Prior Publication Data**

US 2002/0066660 A1 Jun. 6, 2002

Related U.S. Application Data

(62) Division of application No. 09/281,155, filed on Mar. 30, 1999, now Pat. No. 6,326,569.

(60) Provisional application No. 60/080,173, filed on Mar. 31, 1998.

(51) **Int. Cl.**^{7} **H02B 1/015**; H01H 9/00

(52) **U.S. Cl.** **361/660**; 200/310; 200/317; 361/622

(58) **Field of Search** 200/308–317, 200/329–345; 361/622, 600, 680

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,209,820	A	*	6/1980	Rundel et al.	361/622
4,430,288	A		2/1984	Bonis	264/510
4,862,499	A		8/1989	Jekot et al.	379/368
4,892,700	A		1/1990	Guerra et al.	264/510
4,937,408	A		6/1990	Hattori et al.	200/314
5,040,479	A		8/1991	Thrash	116/279
5,098,633	A		3/1992	Hausler	264/511
5,138,119	A		8/1992	Demeo	200/5 A
5,432,684	A		7/1995	Fye et al.	362/30
5,477,024	A		12/1995	Share et al.	219/121.69
5,536,543	A		7/1996	Papandreou	428/35.7
5,570,114	A		10/1996	Fowler	345/173
5,614,338	A		3/1997	Pyburn et al.	430/13
5,633,022	A		5/1997	Myers	425/384
5,736,233	A		4/1998	Fye	428/204
5,739,486	A		4/1998	Buckingham	200/5 A
5,744,210	A		4/1998	Hofmann et al.	428/106
5,747,556	A		5/1998	Boedecker	200/5 A
6,326,569	B1	*	12/2001	Ardrey et al.	200/314

* cited by examiner

Primary Examiner—J. R. Scott
(74) *Attorney, Agent, or Firm*—Rader, Fishman & Grauer PLLC

(57) **ABSTRACT**

A control panel assembly having a plurality of buttons located in a bezel and supported by a switch mat. A bezel and/or button is formed using a vacuum forming technique and has a transparent inner surface layer and a middle translucent color layer and an opaque outer surface layer. A portion of the opaque outer surface layer is removed to define a desired indicia on an outer surface of the bezel or button. In addition, the bezel or button can be backlit to allow visibility in low light conditions. A method of making a control panel component, such as a bezel or button, is also disclosed.

13 Claims, 3 Drawing Sheets

Figure A.1: Title page of a granted U.S. patent

Table A.1: Data structure illustration

Patent	Year	Patenting Organization	Law Firm	Realized patent-law firm dyad
1	2005	Microsoft	Foley and Lardner LLP	1
1	2005	Microsoft	Buchanan and Ingersoll PC	0
2	2006	Sony Corporation	Senniger Powers LLP	1
2	2006	Sony Corporation	Merchant and Gould PA	0
3	2006	IBM	White and Case LLP	1
3	2006	IBM	Buchanan and Ingersoll PC	0
4	2006	Microsoft	Foley and Lardner LLP	1
4	2006	Microsoft	Senniger Powers LLP	0
5	2007	Apple	Foley and Lardner LLP	1
5	2007	Apple	White and Case LLP	0

Table A.2: Correlation matrix for law firm selection models

	1	2	3	4	5	6	7	8	9	10
1 Law firm expertise	1.00									
2 Competitive overlap	0.20	1.00								
3 Log competitor count	0.37	0.63	1.00							
4 Law firm-client embedded.	0.28	0.19	0.48	1.00						
5 Law firm specialization	0.01	-0.37	-0.32	-0.07	1.00					
6 Law firm mean client size	0.08	-0.14	0.00	0.13	0.08	1.00				
7 Law firm size	0.01	0.60	0.61	0.21	-0.50	-0.03	1.00			
8 Law firm founding year	0.02	-0.38	-0.34	-0.10	0.30	0.09	-0.39	1.00		
9 Law firm NBER experience	0.20	0.07	0.02	0.05	-0.05	0.06	0.04	-0.01	1.00	
10 Law firm status	-0.07	0.00	-0.07	-0.06	-0.20	-0.10	0.11	0.00	0.02	1.00
Observations	1926178									

Table A.3: Correlation matrix for patent performance models

	1	2	3	4	5	6	7	8	9	10
1 Patent days to accept	1.00									
2 Patent citations	-0.11	1.00								
3 Law firm expertise	-0.03	0.04	1.00							
4 Competitive overlap	0.02	-0.04	0.08	1.00						
5 Law firm-client embedded.	-0.17	0.00	0.13	0.15	1.00					
6 Law firm specialization	0.04	0.03	0.17	-0.36	-0.04	1.00				
7 Law firm mean client size	-0.01	-0.00	0.09	-0.22	0.20	0.15	1.00			
8 Law firm size	-0.08	-0.04	-0.11	0.55	0.36	-0.45	-0.11	1.00		
9 Patent breadth	0.08	0.11	0.06	-0.09	-0.07	0.10	0.03	-0.12	1.00	
10 Twin patent	0.09	0.02	-0.04	0.12	-0.08	-0.04	-0.10	0.07	-0.05	1.00
Observations	960628									

Table A.4: Conditional logit predicting law firm choice (compare to Table 3)

	(1) no interaction	(2) main model
Law firm expertise	1.644***	1.854***
	(0.029)	(0.030)
Competitive overlap	0.009***	0.020***
	(0.000)	(0.001)
Expertise X competitive overlap		-0.019***
		(0.001)
Law firm-client embeddedness	1.854***	1.853***
	(0.008)	(0.008)
Law firm status	0.207***	0.229***
	(0.014)	(0.014)
Law firm specialization	-1.922***	-1.832***
	(0.049)	(0.050)
Law firm mean client size	-0.070***	-0.066***
	(0.005)	(0.005)
Law firm size	0.818***	0.830***
	(0.020)	(0.020)
Law firm founding year	-0.013***	-0.013***
	(0.000)	(0.000)
Observations	1,926,178	1,926,178
Patent FEs	included	included

Standard errors clustered by patent

***p<.001; **p<.01; *p<.05

Table A.5: Linear probability models predicting law firm choice - competitive overlap based on count (compare to Table 3)

	(1) no interaction	(2) main model
Law firm expertise	0.241***	0.436***
	(0.002)	(0.003)
Comp. overlap (count)	-0.001**	0.034***
	(0.000)	(0.001)
Exp X comp olap (count)		-0.092***
		(0.001)
Law firm-client embeddedness	0.191***	0.192***
	(0.000)	(0.000)
Law firm status	0.019***	0.027***
	(0.001)	(0.001)
Law firm specialization	-0.419***	-0.346***
	(0.004)	(0.004)
Law firm mean client size	-0.030***	-0.030***
	(0.000)	(0.000)
Law firm size	0.007***	0.014***
	(0.001)	(0.001)
Law firm founding year	-0.000***	-0.000***
	(0.000)	(0.000)
Constant	0.599***	0.697***
	(0.052)	(0.051)
Observations	1,926,178	1,926,178
R-squared	0.803	0.805
Patent FEs	included	included

Standard errors clustered by patent

***p<.001; **p<.01; *p<.05

Table A.6: Linear probability models predicting law firm choice - triple interaction with law firm supply

	(1)
Law firm expertise	0.512***
	(0.006)
Competitive overlap	0.002***
	(0.000)
Expertise X competitive overlap	0.003***
	(0.000)
Expertise X law firm supply	-0.256***
	(0.005)
Comp. overlap X supply	0.002***
	(0.000)
Exp. X comp. X supply	-0.006***
	(0.000)
Observations	1,926,178
R-squared	0.806
Patent FEs	included
Controls	included

Standard errors clustered by patent
***p<.001; **p<.01; *p<.05
Models also control for specialization, client size, size, founding year, law firm status, and law firm-client embeddedness

B NBER technology classification

NBER 1 digit category	Category name	NBER 2 digit category	Sub-category name
1	Chemical	11	Agriculture, food, textiles
		12	Coating
		13	Gas
		14	Organic compounds
		15	Resins
		19	Misc. (chem)
2	Computers & Communications	21	Communications
		22	Computer hardware and software
		23	Computer peripherals
		24	Information storage
		25	Electronic business methods and software
3	Drugs & Medical	31	Drugs
		32	Surgery, medical instruments
		33	Biotechnology
		38	Misc. (drugs&med)
4	Electrical & Electronics	41	Electrical devices
		42	Electrical lighting
		43	Measuring, testing
		44	Nuclear, X-rays
		45	Power systems
		46	Semiconductor devices
		49	Misc. (elec)
5	Mechanical	51	Materials processing & handling
		52	Metal working
		53	Motors, engines, parts
		54	Optics
		55	Transportation
		59	Misc. (mech)
6	Others	61	Agriculture, husbandry, food
		62	Amusement devices
		63	Apparel & textile
		64	Earth working & wells
		65	Furniture, house fixtures
		66	Heating
		67	Pipes & joints
		68	Receptacles
		69	Misc. (others)
7	Not Classified	70	Not classified

C Using deep learning to analyze the text of the patents and locate them in the knowledge space

In this Appendix, we explain how we use deep learning to locate patents in a 38 dimensional technology space (the 38 technological classes of the NBER classification system, see Appendix C). Deep learning refers to the process according to which the free parameters of a neural-network model are learned from data.

The neural network model we use takes a textual description of a patent as input and from the text it predicts the two digit NBER classification to which the patent belongs. The model has many free parameters and thus needs to be "trained" on data to learn the parameter values that lead to the best possible categorization performance. The training data consists of a large set of patent applications about which two pieces of information are used: the text of the abstract section of the patents, and the two digit NBER class to which the application has been assigned (by the US Patent Office). We decided to rely on the NBER two-digit classification because the USPC classification and the IPC classification systems have multiple hundreds of categories, which makes the deep learning task less reliable.

Deep learning involves the construction and fine-tuning of the 38-dimension technological space in which the patents are represented. This occurs via an updating algorithm that proceed by trial-and-error to maximize the out-of-sample categorization accuracy.

Next, we provide more details about how the text needs to be processed to be used as an input to the neural network. Then, we specify the structure of the neural networks we use in our empirical analyses.

Pre-processing the Text of the Patent Abstract

Text documents need to be represented in a numerical format to be used as inputs to the neural network. We follow general practice in the field of machine learning to pre-process our data (Chollet, 2017).

The first stage, called tokenization involves what could be informally described as text cleaning: we remove punctuation, make all letters lowercase, and remove 'stop words'. These are words that are important from a grammatical standpoint, but are not necessary to extract the kind of meaning we want to extract from the text. These include connectors ('and,' 'but,' ...) and articles ('the', 'a,' 'an,' ...). We used a standard stopword removing tool.[21] We also removed short words (shorter than 3 characters) and long words (longer than 14 characters) since they often correspond to meaningless strings of characters in input documents. The output of this first stage is a set of tokenized documents.

The second stage consists of encoding the words such that a numeric index is given to each unique word in our text corpus.

In the third stage, each document is represented as a sequence of word indices. Document length is set to be the same for all documents. The sequences of shorter documents are padded with 0s and longer documents are cut short such that all documents are represented as sequences of exactly L numeric indices.

The structure of the deep learning network

The neural network we use to predict NBER class probabilities from patent text consists of five layers. This model structure is the standard structure used in the deep learning literature for text categorization purposes

[21] Text processing was performed with Matlab's Text Analytics toolbox. The specification and training of the neural network model were done using Matlab's Deep Learning toolbox.

(Chollet, 2017).

1. *A word embedding layer.* It transforms the words and creates a representation of each word in the vocabulary (the set of unique words in the corpus) as D-dimensional vectors. This step places vectors representing words close together in the semantic space if the meanings of the words are similar.

   This layer has a number of parameters equal to the number of words times D. It represents the document as a L-long sequence of D-dimensional vectors:

$$w = \begin{pmatrix} w_{1,1} & \cdots & w_{1,l} & \cdots & w_{1,L} \\ \vdots & & \vdots & & \vdots \\ w_{d,1} & \cdots & w_{d,l} & \cdots & w_{d,L} \\ \vdots & & \vdots & & \vdots \\ w_{D,1} & \cdots & w_{D,l} & \cdots & w_{D,L} \end{pmatrix}$$

2. *A hidden biLSTM layer.* biLSTM stands for bi-directional Long Short Term Memory. This layer is the key representational layer: it learns dependencies in the text (and thus is sensitive to word ordering and sets of words, not just the presence or absence of individual words in the text) and represents each text as a point in a space with $2H$ dimensions, where H is a parameter that has to be chosen by the analyst:

$$x = \begin{pmatrix} x_1 & \cdots & x_h & \cdots & x_{2H} \end{pmatrix}.$$

   biLSTM layers are frequently used in sequence-classification applications. Predicting a category for a text is a special case of this type of application.

3. *A fully connected layer.* This layer takes the position of the text document in the technological space and outputs a 38 dimensional vector of real values:

$$\alpha = \begin{pmatrix} \alpha_1 & \cdots & \alpha_c & \cdots & \alpha_{38} \end{pmatrix}.$$

   For each c, α_c is obtained as a linear combination of the inputs $(x_1, \ldots, x_{2h})$. This layer represents the content of the patent as a position in 38-dimensional the technological space

4. *A softmax layer.* It applies a softmax function to the vector of category scores $(\alpha_1, \ldots, \alpha_K)$ and outputs a vector of categorization probabilities. Specifically, for $c \in \{1, \ldots, K\}$, we have:

$$P(c|x) = \frac{e^{\alpha_c}}{\sum_{j=1}^{K} e^{\alpha_j}} = \frac{e^{f_{\mathbb{H}}^{c}(x)}}{\sum_{j=1}^{K} e^{f_{\mathbb{H}}^{j}(x)}}. \tag{1}$$

5. *A classification output layer.* It returns a category based on the vector of categorization probabilities.

A trained neural network computes a vector of categorization probability for each patent.

Training parameters and details

We use a 10-fold training approach. The data is split into 10 randomly selected equal-sized sample. We first train the network using eight of these subsamples as a training sample and one of them as a test sample. We train the model for three epochs, with mini batch size of 64, as we found that this specification leads to the best average prediction accuracy (61 percent), without overfitting the data. After the network is trained on these nine subsamples, we use the trained network to do an out-of-sample prediction on the 10th subsample.

We repeat this approach nine more times, each time predicting a different subsample and using the other nine subsamples for training data and test data. Finally, we combine the 10 sets of predicted values and use them in the analyses.