To Complete a Puzzle, You Need to Put the Right Pieces in the Right Place

Kok, Holmer Jan

Chapter 3. Exploring Knowledge Recombination in R&D Alliances

Why Knowledge Pool Applicability Matters

Abstract: Knowledge applicability is a core driver of knowledge recombination activities. Extant research on the knowledge recombination implications of R&D alliances, however, tends to ignore the issue of knowledge applicability. Instead, it focuses on the size and diversity of the partner's knowledge pool. In this study, we address this gap, shifting attention to knowledge pool applicability – i.e. the extent to which components in the knowledge pool can be used in different application domains – and examining its implications for knowledge recombination activities in R&D alliances. We expect that both the partner's and the focal firm's knowledge pool applicability significantly impact firm's partner-specific knowledge recombination. Analysing 461 R&D alliance dyads of 88 firms in the fuel cell industry, our findings indicate that partner's knowledge pool applicability has an inverted U-shaped relationship with firms' partner-specific knowledge recombination. Surprisingly, we find that the knowledge pool applicability of the focal firm has a U-shaped relationship with its partner-specific recombination. Bringing forward the concept of knowledge pool applicability, this study contributes to a richer theoretical understanding of the knowledge recombination implications of R&D alliances.

This chapter was written together with Dries Faems and Pedro de Faria. Earlier versions of this chapter have been presented at the *Annual Meeting of the Academy of Management* in Atlanta (2017), *Strategic Management Society Annual International Conference* in Houston (2017), and at research seminars at the *University of Groningen* (2017), *École Polytechnique Fédérale De Lausanne* (2016), and *Stockholm School of Economics* (2017). A manuscript based on this chapter is currently under review for publication.

3.1. Introduction

R&D alliances are valuable mechanisms for firms to expand their knowledge pool, generating new opportunities for knowledge recombination across different knowledge components (e.g. Lahiri & Narayanan, 2013; Schilling & Phelps, 2007; Srivastava & Gnyawali, 2011). At the same time, not being able to effectively use partners' knowledge triggers the risk of ending up at the losing end of learning races, threatening the competitive position of firms (Hamel, 1991; Khanna, Gulati, & Nohria, 1998). Scholars have therefore started exploring the ability of firms to recombine knowledge from alliance partners. Existing research mainly focuses on the size and diversity of the partner's component knowledge pool as key drivers of knowledge recombination activities (e.g. Lahiri & Narayanan, 2013; Phelps, 2010; Schilling & Phelps, 2007). In this study, we complement this prior research by focusing on knowledge pool applicability as a crucial aspect to explain variance in knowledge recombination from alliance partners.

Knowledge recombination literature argues that components vary in terms of where they can be applied (Hargadon & Sutton, 1997; Wang *et al.,* 2014; Yayavaram & Ahuja, 2008). Some components are highly malleable and can be used in different application domains, whereas other components may be substantially constrained in their range of applications (Dibiaggio *et al.,* 2014; Hargadon & Sutton, 1997). For instance, in the fuel cell industry, which is the empirical setting of this study, some inventions can be used in multiple application domains, such as fuel cell stacks and fuel reformers, whereas the applicability of other inventions may be restricted to one single domain. Based on these insights, we argue that, next to size (i.e. number of components) and diversity (i.e. diversity of technology domains in which components are situated), knowledge pools can also be characterized in terms of their applicability or the extent to which single components within them have different application domains.

The notion that components vary in their applicability is largely neglected in existing research on alliances and their knowledge recombination implications. However, we theorize that knowledge pool applicability substantially influences the focal firm's partner-specific knowledge recombination or the extent to which the focal firm relies on knowledge from a particular partner when generating inventions. In particular, we hypothesize (i) an inverted U-shaped relationship between the partner's knowledge pool applicability and the focal firm's partner-specific recombination and (ii) a positive relationship between the focal firm's knowledge pool applicability and the focal firm's partner-specific recombination.

To empirically test our hypotheses, we collected unique data on 461 R&D alliance dyads of 88 focal firms in the fuel cell technological field and combined this with data on their worldwide fuel cell patenting activities. Our analyses show that, controlling for other characteristics of the partner's knowledge pool, partner's knowledge pool applicability has an inverted U-shaped relationship with firms' recombination of partner's components. Moreover, instead of the hypothesized linear positive relationship, we find a U-shaped relationship between the focal firm's internal knowledge pool applicability and partner-specific knowledge recombination.

These findings contribute to a richer theoretical understanding of the knowledge recombination implications of R&D alliances in two fundamental ways. First, we show the importance of applicability as a core characteristic of knowledge pools that influences knowledge recombination activities in the setting of R&D alliances. Existing studies have theoretically framed the knowledge pool of alliance partners on an aggregate level, arguing that components create value relative to other components present in the knowledge pool. We argue, however, that a closer examination of the components inside the knowledge pool may reveal that a lot of components do not actually have the assumed broad scope of applicability normally associated with large and diverse knowledge pools (e.g. Lahiri & Narayanan, 2013). Second, we develop novel theoretical arguments regarding firms' idiosyncratic abilities to engage in knowledge recombination in R&D alliances. Prior research tends to conceptually intertwine the focal firm's ability to identify and absorb component knowledge with its ability to actually recombine this component knowledge (e.g. Gilsing, Nooteboom, Vanhaverbeke, Duysters, & van den Oord, 2008; Rosenkopf & Almeida, 2003; Vasudeva, Zaheer, & Hernandez, 2013). We underline the fact that identification and transfer of component knowledge are necessary but not sufficient conditions for knowledge recombination to occur. Instead, firms need to be able to envision novel applications for components, in order to use them in knowledge recombination. In sum, this study underlines the importance of taking an in-depth knowledge recombination perspective to studying the performance implications of R&D alliances.

3.2. Theory

3.2.1 Knowledge recombination in the alliance context: state of the art

Knowledge recombination plays a central role in explaining performance differences across R&D alliances (e.g. Phelps, 2010; Schilling & Phelps, 2007; Wuyts & Dutta, 2014). In order to stimulate the creation of new inventions through knowledge recombination activities, firms have to enrich the contents of their own knowledge pool with novel components accessed from alliance partners (Fleming, 2001; Rosenkopf & Almeida, 2003; Savino *et al.,* 2017). Alliance scholars have prominently argued that, when the partners' knowledge pool contains more components, the focal firm is able to realize a larger set of new combinations (e.g. Lahiri & Narayanan, 2013; Schilling & Phelps, 2007).

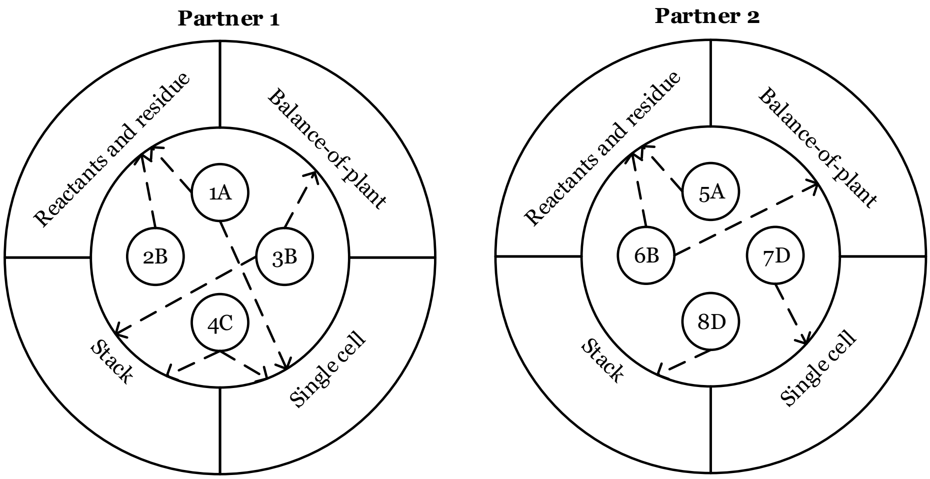
Scanning recent alliance literature, however, we observe a clear shift from focusing on the size of the partner's knowledge pool toward examining the diversity of the partner's knowledge pool (e.g. Phelps, 2010; Srivastava & Gnyawali, 2011; Wuyts & Dutta, 2014). Alliance scholars have argued that components accessed from alliance partners differ in terms of their usefulness and ease of retrievability (e.g. Gomes-Casseres, Jaffe, & Hagedoorn, 2006; Nooteboom, Vanhaverbeke, Duysters, Gilsing, & Van den Oord, 2007; Vasudeva & Anand, 2011). Accessing components from diverse technological domains allows generating more novel combinations (Phelps, 2010; Subramanian & Soh, 2017), that tend to be more valuable (Srivastava & Gnyawali, 2011; Rosenkopf & Nerkar, 2001; Wuyts & Dutta, 2014). At the same time, components from different technological domains tend to be more difficult to understand and apply in knowledge recombination (Phene *et al.,* 2006; Nooteboom *et al.,* 2007; Vasudeva & Anand, 2011). Hence, access to a diverse knowledge pool from a partner may involve substantial benefits and challenges, influencing the focal firm's ability to generate new technologies

(Nooteboom *et al.,* 2007; Phelps, 2010; Srivastava & Gnyawali, 2011; Subramanian & Soh, 2017; Wuyts & Dutta, 2014).

Thus, existing alliance literature has mostly focused on the size and diversity of partners' knowledge pool, arguing that they might influence the performance implications of R&D alliances. However, examining recent knowledge recombination literature, we observe that scholars have argued that more factors influence knowledge recombination than merely the diversity and number of available components (e.g. Dibiaggio *et al.,* 2014; Wang *et al.,* 2014; Yayavaram & Ahuja, 2008). In particular, what emerges from this literature is that components also vary in terms of where and how they can be applied in knowledge recombination. On the one hand, it has been argued that some components can be used in knowledge recombination across different settings, such as different industries (Hargadon & Sutton, 1997), countries (Petruzzelli & Savino, 2014) or technological generations (Furr & Snow, 2014), allowing firms to economize substantially on cognitive resources (Baker & Nelson, 2005; Wang *et al.,* 2014). Moreover, knowledge recombination scholars found that components that can easily be applied in different combinations also tend to enhance experimentation by the focal firm, since it becomes easier to mix-and-match components, enhancing exploratory innovation (Dibiaggio *et al.,* 2014; Guan & Liu, 2016; Wang *et al.,* 2014). On the other hand, however, when components can too easily be mixed-and-matched, it becomes costly to determine the optimal way in which to apply the components (Yayavaram & Ahuja, 2008). Jointly, these findings all point into the same direction: next to knowledge pool diversity and size, components in the knowledge pool may also vary in terms of where and how they may be applied in knowledge recombination.

To tackle this issue in the R&D alliance context, we introduce the concept of knowledge pool applicability, which denotes the extent to which components in the knowledge pool can be used in different application domains. In this way, we attempt to capture the extent to which components in the knowledge pool can flexibly be used in different recombination efforts. To make the concept of knowledge pool applicability more concrete, consider the example in Figure 3.1. In this figure, we show two partners from the fuel cell industry with the same knowledge pool size (i.e. four components, as denoted by the smaller circles) and diversity (i.e. same distribution of technological domains, where domain A may refer to hydrogen and domain D may refer to ceramic compounds). Moreover, we indicate that each component has one or more potential application domains (corresponding to the four main subsystems of a fuel cell system, in which new fuel cell inventions may be applied). Hence, although partners 1 and 2 both have a component pertaining to hydrogen (i.e. components 1 and 5), partner 1 is able to apply this component in two domains, whereas partner 2 is only able to do so in one domain. Partner 1 may, for example, have component knowledge of the composition of hydrogen that could be integrated into the design of the single cell, or to improve the performance of the fuel reformer. Thus, even when the knowledge pool size and diversity of two alternative alliance partners is the same, we show that their knowledge pool applicability may still differ substantially, potentially impacting recombination opportunities available to the focal firm.

Figure 3.1. Knowledge pool size, diversity, applicability



To summarize, knowledge recombination literature has shown that the extent of applicability of components is an important determinant of knowledge recombination activities. However, this has been largely neglected in alliance research, where the focus has instead been on the partners' knowledge pool size and diversity. In this study, we address this research gap by focusing on the knowledge pool applicability of the partner and focal firm, arguing that this dimension of the knowledge pool may substantially affect knowledge recombination activities in R&D alliances. In the following section, we developed two hypotheses connecting knowledge pool applicability to the focal firm's partner-specific recombination.

3.3. Hypotheses development

3.3.1. Partner's knowledge pool applicability

In the R&D alliance setting, we expect that a partner's knowledge pool applicability will positively impact the focal firm's partner-specific recombination. Specifically, partners with higher knowledge pool applicability are able to generate a larger number of combinations on the basis of a given set of components. Consequently, when collaborating with a partner with high knowledge pool applicability, the components that the focal firm learns to use from this partner tend to be more widely-applicable (Boh *et al.,* 2014; Wang *et al.,* 2014). Instead of being constrained to one single application, the focal firm gains flexibility in terms of where and how to apply the accessed components, greatly facilitating the generation of new combinations on the basis of a partner's components (Hargadon & Sutton, 1997; Wang *et al.,* 2014). The obtained flexibility can optimize the allocation of resources to the R&D alliances, ensuring that attempted efforts to learn how to recombine a partner's components bear fruit (Wang *et al.,* 2014). For example, in Figure 3.1, even if the focal firm is unable to learn how to apply partner 1's component 4 in the single cell domain, it may still be able to recombine this component in the stack domain. Such flexibility is, however, generally not present in partner 2's knowledge pool, severely restricting the focal firm's pursuit of recombination opportunities.

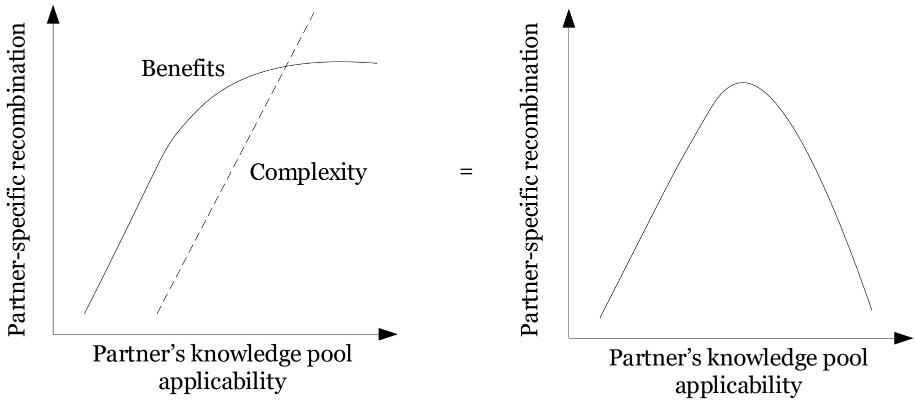
Beyond a certain point, we expect, however, that the benefits of a partner's knowledge pool applicability will taper off and diminish in magnitude. A key tenet in knowledge recombination literature is that, although in theory all components in the environment could be considered for recombination, firms' resource and cognitive constraints severely narrow down the number of combinations that can eventually be realized based on these components (Carnabuci & Bruggeman, 2009; Fleming & Sorenson, 2001; Olsson & Frey, 2002). As a consequence, the number of potential recombination opportunities generally outweighs the number of realized recombination opportunities (Strumsky & Lobo, 2015; Zahra & George, 2002). In line with this reasoning, alliance scholars found evidence to suggest that firms often face constraints in terms of fully realizing partner's recombination opportunities (Bos, Faems, & Noseleit, 2017; Cohen & Levinthal, 1990; Deeds & Hill, 1996; Vasudeva & Anand, 2011). Thus, we argue that, even if the partner's components can be applied in many different ways, the focal firm will only be able to identify and realize a fraction of the recombination opportunities associated with these components. In other words, the benefits accruing from higher partner knowledge pool applicability will become less outspoken beyond a certain threshold value.

In addition to diminishing marginal benefits of partner's knowledge pool applicability, we also expect certain challenges to emerge alongside with partner's knowledge pool applicability. The various applications for which a component can be used tend to be highly similar (Hargadon & Sutton, 1997), because they are strongly based on the same fundamental technological principles. For this reason, when the partner's components are very widely-applicable, firms may find it highly challenging to distinguish one component's application from another, rendering learning processes less effective and more cumbersome. Effectively, when partner's knowledge pool applicability is high, the focal firm will spend a lot of time trying to optimally allocate its attention and learning efforts in the partner's knowledge pool (Ghosh, Martin, Pennings, & Wezel, 2014; Katila & Ahuja, 2002; Yayavaram & Ahuja, 2008), as it becomes more difficult to identify which components' applications are most worthwhile to pursue. Thus, due to increased learning complexities (Fleming & Sorenson, 2001; Leiponen & Helfat, 2010), the focal firm's attention and resources may be detracted from actual knowledge recombination efforts towards determining the optimal way to recombine a partner's components, reducing the focal firm's partner-specific recombination.

Figure 3.2 summarizes our arguments. In this figure, we show that the benefits arising from a partner's knowledge pool applicability eventually level off, due to limited resource availability that constrain the full realization of a partner' potential recombination opportunities. At the same time, important challenges associated with identifying valuable recombination opportunities may emerge beyond a certain value of a partner's knowledge pool applicability. Hence, beyond a certain point, we expect the disadvantages of partner's knowledge pool applicability to rise faster than the benefits, shaping an inverted U-shaped relationship between partner's knowledge pool applicability and the focal firm's partner-specific recombination (Haans *et al.,* 2016). We hypothesize:

H1: The partner's knowledge pool applicability has an inverted U-shaped relationship with the focal firm's partner-specific recombination

Figure 3.2. Theoretical mechanisms underlying Hypothesis 1



3.3.2. Focal firm's knowledge pool applicability

In R&D alliances, firms often face difficulties accessing and recombining a partner's components (Faems, Janssens, & Van Looy, 2007; Gomes-Casseres *et al.,* 2006; Hamel, 1991). Many partners' components remain untapped in knowledge recombination, because firms perceive that the components' recombinant potential has already been exhausted or they are simply unable to use the components in any meaningful way (Ahuja & Katila, 2004; Fleming, 2001). Faced with such constraints, most firms are unable to fully utilize the component knowledge that may be accessed from an alliance partner, increasing the gap between the potential and realized recombination opportunities arising from the R&D alliance (Wuyts & Dutta, 2014; Zahra & George, 2002). We expect, however, that focal firms with higher knowledge pool applicability are able to bring the realization of combinatorial opportunities closer to its potential level through two principal mechanisms: (i) flexibility and (ii) effectiveness of knowledge recombination.

First, following the notion that “capabilities are built through experience” (Eggers, 2012: 318), we argue that by building widely-applicable component knowledge, firms develop a greater comprehension of how component knowledge, in general, can be flexibly applied in different ways. This provides them with the ability to better leverage available component knowledge (Lewin, Massini, & Peeters, 2011; Wuyts & Dutta, 2014). Hence, these firms experience more flexibility in terms of which of the partner’s components they may consider for knowledge recombination, being able to envision and realize more applications for a given component. Indeed, as argued by Henderson (1995), technological limits of components tend to be mostly present in the mind of inventors, rather than being grounded in actual technological limits. In other words, whereas one firm may quickly run into exhaustion of recombinant potential of a component, another firm may be able to elevate the recombinant potential of a component to a higher level. This means that, when provided with access to a component set of, for example, five components, a focal firm with higher knowledge pool applicability may be able to envision four different uses for these components, whereas a focal firm with low knowledge pool applicability may only be able to envision two of such uses.

Second, whereas firms with low knowledge pool applicability may pursue costly trial-and-error processes trying to generate new combinations on the basis of a particular set of components, firms with high knowledge pool applicability may have a better understanding of where to locate their attention. This is because the latter type of firm, having prior experience developing component knowledge with multiple applications, has a better understanding of where the limits of a component’s recombinant potential lie, avoiding ultimately fruitless recombination efforts in the partner’s knowledge pool (Nemet & Johnson, 2012). Hence, leveraging these unique capabilities, firms with higher knowledge pool applicability will be able to deploy resources towards the utilization of a partner’s components more effectively, making fuller use of a partner’s component knowledge pool, whilst using fewer resources in the process.

We also expect the benefits of higher internal knowledge pool applicability to consistently outweigh the costs of applying these capabilities in R&D alliances.

Capabilities emerging concomitantly with building internal knowledge pool applicability have already been assimilated and embedded into the firm's routines (Cohen & Levinthal, 1990; Lewin *et al.,* 2011; Wuyts & Dutta, 2014). Therefore, the various applications of these capabilities are well-understood, with a deep understanding of how to align available recombination opportunities with capabilities to realize these opportunities. Hence, the application of these recombination capabilities is more certain and less prone to mistakes (Katila & Ahuja, 2002). This makes it far more likely that sufficient resources are available to make full use of these recombination capabilities in order to better leverage a partner's component knowledge pool. Hence, we expect a strictly positive relationship between the firm's internal knowledge pool applicability and its partner component recombination because (i) higher internal knowledge pool applicability allows to more flexibly and effectively leverage the partner's knowledge pool and (ii) costs associated with applying capabilities emerging from internal knowledge pool applicability are negligible:

H2: The focal firm's knowledge pool applicability has a positive relationship with the focal firm's partner-specific recombination.

3.4. Methodology

3.4.1. Sample and data collection

Empirical context. We tested our hypotheses using data from the fuel cell technological field. Fuel cells are electrochemical devices that produce electricity through a chemical reaction between hydrogen and oxygen. We focused on the fuel cell R&D alliances of 88 firms in the period 1993-2007. These 88 firms were retrieved after compiling a list of the top 200 patent applicants in the fuel cell technological field and removing (i) firms with incomplete ownership data and (ii) firms that did not form any fuel cell R&D alliance during the time period of this study[1].

[1] We were able to collect complete ownership data for 139 parent firms. We aggregated all patents of subsidiaries in which these firms had a controlling interest to the parent-firm level. To collect ownership data of these parent firms, we used Bureau van Dijk's ORBIS database. Moreover, we complemented this with data on executed mergers and acquisitions, retrieved from the SDC Platinum Mergers and Acquisitions database. We also corrected for potential name changes and aliases of firms, using data

The fuel cell technological field is highly comparable to other technological fields in which R&D alliances have been studied, such as pharmaceuticals (e.g. Bos *et al.,* 2017; Wuyts & Dutta, 2014), semiconductors (e.g. Srivastava & Gnyawali, 2011), and telecommunications (e.g. Phelps, 2010). First, patenting propensities are elevated in the fuel cell technological field and rank among the highest in clean energy technologies (Albino *et al.,* 2014). Second, knowledge resources and capabilities are highly heterogeneously distributed in the fuel cell technological field (Hellman & van den Hoed, 2007; Vasudeva & Anand, 2011), creating a greater need to form R&D alliances (Harrison, Hitt, Hoskisson, & Ireland, 2001). Third, the environment in which fuel cell firms operate is highly uncertain and dynamic (Hellman & van den Hoed, 2007; Verbong, Geels, & Raven, 2008), increasing firms' tendencies to form strategic alliances (Schilling, 2015).

Patent data. To obtain a proxy for the knowledge recombination activities of firms, we collected data on worldwide patenting activities of the firms in our sample from the PATSTAT database (Autumn 2013 version). Consistent with prior studies, patents represent the knowledge components in a firm's knowledge pool (Ahuja & Katila, 2001). To retrieve fuel cell patents, we collected all patents filed in IPC class H01M8 which corresponds to fuel cell technology (Tanner, 2014). We aggregated these patent applications to the patent family level (following the European Patent Office's DOCDB definition). A DOCDB patent family captures all patent applications related to the same invention but filed at different patent offices (Albrecht *et al.,* 2010). Relying on patent families helps to overcome the home-country bias of single patent office applications (de Rassenfosse *et al.,* 2013). This bias arises because, for example, North-American firms have a much higher likelihood to file a patent at the USPTO rather than at, for example, the EPO or JPO. As a result, solely relying on USPTO patent applications considerably underestimates the knowledge recombination activities of firms outside North-America (de Rassenfosse *et al.,* 2013). This bias is especially problematic in the fuel cell technological field, as many prominent players in this field are Asian (e.g. Toyota, Honda, Nissan, Hitachi, Panasonic, Toshiba, Samsung Electronics, Asahi

---

from the ORBIS database. Subsequently, of these 139 firms, we retained 88 firms which had engaged in at least one R&D fuel cell alliance between 1993 and 2007. Harmonized patent applicant names were obtained through EEE-PPAT (ECOOM-EUROSTAT-EPO PATSTAT Person Augmented Table) from ECOOM.

Glass) or European (e.g. Daimler, Siemens, BASF, Shell, Renault) firms (Vasudeva, 2009). An additional important advantage of using patent families is that it captures a broader and more complete set of backward citations (Albrecht *et al.,* 2010).

Alliance data. To collect the alliance data, we identified R&D alliances in the LexisNexis database. We used this method because there is considerable evidence that other databases, such as Thompson Reuters' SDC Platinum Joint Venture and Strategic Alliances database (Schilling, 2009), severely underestimate the number of alliances formed (Lavie, 2007; Lavie & Rosenkopf, 2006)[2]. The LexisNexis database compiles press releases from different sources, including newspapers, trade journals, wire transcripts, etc. We employed a broad set of keywords to detect fuel cell R&D alliances[3,4], manually screening over 50,000 press releases. To give an example, the following press release extract identifies an R&D alliance between Nuvera Fuel Cells and TotalFinaElf:

> *Nuvera Fuel Cells, Inc., a leading global designer and developer of fuel cell and fuel processing technology today announced it has entered into an agreement with TotalFinaElf, one of the world's leading oil companies, to study the effects of gasoline on fuel processors and fuel cell stacks designed for the automotive industry* (PR Newswire, 2003)

We searched for all fuel cell R&D alliances formed before 2008.The first fuel cell R&D alliance that we detected, between Westinghouse Electric and Energy Research Corporation, started in 1978[5]. We included all fuel cell R&D alliances in which at least one firm was involved. Moreover, we also included fuel cell R&D alliances that were part of wider government-funded projects, such as the United

[2] To verify this, we conducted a broad search in the SDC database for all alliances in which the deal text mentioned the keyword "fuel cell" in the period 1978-2007. This search produced a total of 126 alliances, comprising not only R&D alliances, but also other types of alliances (e.g. marketing, supply). In contrast, during the same time period, we detected 849 R&D alliances in the fuel cell industry using the LexisNexis database. Similarly, Lavie and Rosenkopf (2006) reported that only 25% of alliances in their dataset were detected in the SDC database.

[3] We did not specifically search for non-R&D alliances because (i) the focus of our study is on technological activities and (ii) the language used to describe non-R&D alliances is highly idiosyncratic (Schilling, 2009), especially for supply and distribution alliances. Hence, our sample only contains alliances with an R&D element (Hagedoorn, 2002).

[4] The set of employed keywords is available from the authors on request.

[5] Although we focus on the 1993-2007 period, we also searched for R&D alliance data for the 1978-1992 period, in order to capture alliances which had started before 1993, but were still ongoing in 1993.

States Department of Energy's (DOE) *Hydrogen and Fuel Cells Program*. Finally, multi-partner R&D alliances were transformed into dyads, following earlier studies (e.g. Phelps, 2010).

Whereas numerous studies assume a fixed lifespan for alliances, ranging from one to five years (e.g. Schilling & Phelps, 2007; Srivastava & Gnyawali, 2011; Vasudeva & Anand, 2011), often within the same industry, we tracked alliances over time to approximate their starting and termination dates (Ahuja, 2000; Hashai, Kafouros, & Buckley, 2018; Lavie, 2007; Phelps, 2010). This is an important methodological step, given that there exists substantial heterogeneity in the lifespan of alliances (Deeds & Rothermael, 2003). Moreover, as emphasized by Wassmer (2010), alliance research should attempt to only focus on ongoing partnerships since, otherwise, statistical inferences are drawn about partnerships that do not actually exist anymore. When termination of the alliance was not formally announced, we followed Ahuja (2000) and utilized either (i) the expected tenure of the alliance or (ii) tracked the ongoing status of the alliance through subsequent press releases. In case a termination date could not be approximated, we followed Ahuja (2000) and assumed that the alliance was terminated in the year subsequent to the starting year.

3.4.2. Variables

Dependent variable. To measure *partner-specific recombination*, we followed earlier knowledge recombination (Katila, 2002; Phene *et al.,* 2006; Rosenkopf & Nerkar, 2001) and alliance (Frankort, 2016; Gomes-Casseres *et al.,* 2006; Mowery, Oxley, & Silverman, 1996; Rosenkopf & Almeida, 2003; Schildt, Keil, & Maula, 2012; Subramanian, Bo, & Kah-Hin, 2018; Vasudeva & Anand, 2011) literature and used backward citations of patents. Backward citations reflect the prior technological knowledge upon which an invention builds and can thus be used to denote the components that are recombined to generate a new invention (Katila, 2002; Phene *et al.,* 2006; Jaffe & de Rassenfosse, 2017). Although some studies argue that backward citations are a rather noisy indicator of knowledge flows (e.g. Alcacer & Gittelman, 2006), others provide reasonable evidence for the equivalence between patent citations and knowledge flows (e.g. Duguet & MacGarvie, 2005; Jaffe, Trajtenberg, & Fogarty, 2000). To compute the dependent variable, we counted the total number of fuel cell citations that focal firm *i* made to a partner *j* in a given year[6]. In order to correct for differences in firm scale, following Phelps (2010), we compute the share of citations made by focal firm *i* to a partner *j* in year *t*:

$$\text{Focal firm's partner-specific recombination}_{ijt} = \frac{\text{Fuel cell citations to partner}_{ijt}}{\text{Total fuel cell citations}_{it}}$$

Independent variable. In order to capture *knowledge pool applicability*, we inspected the International Patent Classification (IPC) codes listed on fuel cell patent applications. IPC codes are principally assigned to patents to facilitate patent examiner's search activities, and reflect the technological content of an invention (Strumsky & Lobo, 2015). Prior innovation studies have similarly employed IPC codes as a categorization tool to localize patents into specific technological domains (e.g. Kapoor & Adner, 2012; Phelps, 2010). IPC codes are classified in a hierarchical manner (i.e. lower levels represent subdivisions of higher levels), such that the first digit indicates the highest level of abstraction (e.g. H refers to 'Electricity' while G refers to 'Physics'), and subsequent digits increase the level of granularity (Benner & Waldfogel, 2008). To illustrate this hierarchical construction, we show an example of subgroup H01M8/24 in Figure 3.3.

Figure 3.3. Hierarchical construction of IPC code H01M8/24

Section	H	Electricity
Class	H01	Basic electric elements
Subclass	H01M	Processes or means, e.g. batteries, for the direct conversion of chemical into electrical energy
Main group	H01M8	Fuel cells; Manufacture thereof
Subgroup	H01M8/24	Grouping of fuel cells, e.g. stacking of fuel cells

[6] Following the methodology described by Bakker et al. (2016) and Nakamura et al. (2015), we aggregated all single patent office applications within a patent family to the patent family level to obtain a more precise account of the components that were recombined to generate the invention. For example, consider two patents A and B that belong to patent family 1. If patent A cites patents C and D, and patent B cites patents D and E, then patent family 1's backward citations are C, D, and E. Naturally, we also correct for patent family membership at the backward citation-level in such a way that, if patent family 1's cited patents D and E actually pertain to the same patent family, they are not counted twice.

Whereas higher IPC levels (such as the subclass, or, four-digit IPC code) can be used to gain an understanding of the broadness of technological fields that firms use to build inventions (Leten, Belderbos, & Van Looy, 2016), they tend to not be precise enough to denote the actual domains within which new inventions can be applied (Thompson & Fox-Kean, 2005). Instead, in order to determine the application domains of patents within a particular technological field (for fuel cell patents, main group H01M8), it is useful to look at the application domains nested within this field (for fuel cells, subgroups H01M8/02-H01M8/24). Here, an application domain is thus a subset of a technological field within which a particular invention can be applied. We identified four main application domains of fuel cell inventions[7,8]; namely: The single cell, stacking techniques, reactant production and residue treatment, and balance-of-plant aspects of fuel cell systems (EG&G Technical Services, 2004; Sharaf & Orhan, 2014). To capture these four applications domains, we used subgroups H01M8/02, H01M8/04, H01M8/06, and H01M8/24 (see Table 3.1 for a description of these subgroups)[9]. Effectively, from the IPC codes listed on a patent, we inferred that, when a firm generates a new invention that describes a technological solution to a problem that spans

[7] While our measure is somewhat conservative, in the sense that it focuses solely on the fuel cell technological domain as a potential location for application of inventions, we verified, through examination of fuel cell articles (e.g. EG&G Technical Services, 2004; Sharaf & Orhan, 2014; Steele & Heinzel, 2001), that these four application domains are indeed most relevant to fuel cell technology. Moreover, looking at the citation scope of fuel cell patents, we observe that this is a highly insular technological domain (George *et al.,* 2008), in which technologies draw, and are applied, mostly within the same technological domain (i.e. close to 75% of all backward citations of fuel cell patents go to other fuel cell patents). Finally, we only examine R&D alliances that are specifically targeted at fuel cell technology, meaning that the four application domains are likely to be the most relevant ones in this context.

[8] There are 12 subgroups within main group H01M8 that can be used to identify the aspects of fuel cell technology that are addressed by the fuel cell patent. Eight of these subgroups (i.e. H01M8/08-H01M8/22) are used to categorize the patent as pertaining to the design of a specific type of fuel cell. Fuel cell types can (principally) be distinguished on the basis of the electrolyte inside the cell (Steele & Heinzel, 2001). For example, H01M8/10 refers to the design of fuel cells with a solid electrolyte (e.g. polymer exchange membrane fuel cells) and H01M8/12 refers to the design of fuel cells with a solid oxide electrolyte (e.g. solid oxide fuel cells). The remaining four subgroups (i.e. H01M8/02, H01M8/04, H01M8/06, and H01M8/24) can be used to detect the application domains of fuel cell inventions. For each IPC subgroup, we examined hundreds of patent documents to verify their correspondence to a particular application domain. We further validated this correspondence by comparing the distribution of firms' patents across different application domains with press releases and firm documents describing their fuel cell technological activities.

[9] Out of all the patents in the sample, 90% of the patents had at least one of the four IPC subgroups listed on them (Table 3.1). For the remaining patents (10% of the sample), we examined patterns of co-occurrence of IPC subgroups (i.e. Breschi, Lissoni, & Malerba, 2003; Dibiaggio *et al.,* 2014) in order to categorize each subgroup into the subsystem with which it was most likely associated (e.g. subgroup C08J5/22, which refers to manufacturing films, membranes, or diaphragms made of macromolecular substances, co-occurred 99% of the time with subgroup H01M8/02, suggesting a strong association with the single cell).

multiple application domains, the firm demonstrates an understanding of how this invention can potentially be applied to each of these domains, in new knowledge recombination efforts (Boh *et al.,* 2014).

Having classified each patent into its corresponding application domain(s), we subsequently calculated the average number of application domains that were listed on each patent, and then aggregated this to the firm-level. For example, in Figure 3.1, partner 1 has 4 patents in its knowledge pool, three of which have two potential application domains, and one has one potential application domain. In this case, knowledge pool applicability takes a value of $\frac{(3\times2)+(1\times1)}{4} = 1.75$. Similarly, partner 2's knowledge pool applicability equals $\frac{(1\times2)+(3\times1)}{4} = 1.25$.

Table 3.1. Application domains in the fuel cell technological field

Application domain	Description	IPC subgroup
Single cell	This application domain relates to elements inside the fuel cell. This includes the design of gas diffusion layers, electrolytes, electrodes, etc.	H01M8/02
Balance-of-plant	This application domain relates auxiliary equipment of fuel cells. This includes the design of heat exchangers, air pumps, controlling systems etc.	H01M8/04
Producing reactants / treating residues	This application domain relates to producing reactants (e.g. hydrogen) and treating residues. This includes fuel reforming, hydrogen purification, hydrogen supply, etc.	H01M8/06
Stacking techniques	This application domain relates to stacking of fuel cells. This includes compression techniques of single cells, composition of stacks, etc.	H01M8/24

Control variables. We controlled for relevant attributes of the R&D alliance. Since trust and relational assets between partners are often developed over time (Dyer and Singh, 1998), we controlled for the age of the alliance, calculated as the time that elapsed since the (current) ongoing tie between the firm and the partner was initiated (*Age alliance*)[10]. Since equity arrangements in an R&D alliance may curb opportunistic behavior (Kogut, 1988), we included a control

[10] We looked at the age of the tie, rather than the alliance itself, because some dyads in our sample had multiple ongoing alliances at the same time (which we also control for). For example, Toshiba and UTC initiated a joint venture alliance in 1985, and a separate joint venture in 2001, both of which were terminated in 2004.

variable that takes a value of 1 when the alliance is a joint venture (*Joint venture*). We included a control variable that takes a value of 1 when the alliance is part of a government-funded program (*Government-funded*). Multi-partner alliances - i.e. alliances in which more than two partners are involved - can influence the behavior of any participating organization, inciting, in some cases, free-riding behavior (Das & Teng, 2002). Therefore, we included a control variable that takes a value of 1 when the alliance is a multi-partner alliance (*Multi-partner*). Knowledge recombination in R&D alliances with foreign partners may be strongly influenced by, for example, cultural differences between the firm and the partner (Lavie & Miller, 2008). Therefore, we included a control variable that takes a value of 1 when the partner is non-domestic (*International*). Finally, we controlled for the number of concurrent alliances ongoing between the firm and the partner (*Concurrent alliances*) (Gomes-Casseres *et al.,* 2006).

We also controlled for several alliance portfolio-level characteristics of the focal firm (Faems, Van Looy, & Debackere, 2005; Wassmer, 2010). Having a high number of technology collaborators may detract attention from the partner, potentially reducing the focal firm's partner-specific recombination. We therefore controlled for the number of technologically active fuel cell technology partners of the firm in the current year (*Inventive partners*) (Deeds & Hill, 1996; Wassmer, 2010). Technologically active fuel cell technology partners are firm partners which have filed at least one fuel cell patent in the past five years. In a similar way, we controlled for the number of non-technologically active fuel cell technology partners of the firm in the current year (*Non-inventive partners).* Moreover, we controlled for the number of upstream partners (i.e. universities, research institutes, and government laboratories) of the firm in the current year (*Upstream partners*) (Faems *et al.,* 2005).

We also controlled for several attributes of the knowledge pool of the firm and the partner. We controlled for the size of knowledge pool of the focal firm (*Firm knowledge pool size*) and the partner (*Partner knowledge pool size)* by computing a cumulative count of patents filed in the past five years [t-6, t-1]. These two variables were divided by 1000, to improve readability of the results. Furthermore, we controlled for the total number of backward citations made by the firm in a given year (*Total recombination*). We also divided this variable by 1000, to improve readability of the results. Importantly, we also controlled for the number of fuel cell citations that the firm made to the partner's patent stock in the past five years (*Past partner-specific recombination*) to control for any path-dependent effects (Gomes-Casseres *et al.,* 2006). We controlled for the focal firm's (*Firm knowledge pool diversity*) and partner's knowledge pool diversity (*Partner knowledge pool diversity*). Following earlier studies (e.g. Phelps, 2010; Sampson, 2007; Subramanian & Soh, 2017), we aggregated all IPC codes (at the main group level) of a firm's patents to the firm-level, and subsequently calculated the distribution among them relying on the widely-used Herfindahl index. Moreover, using IPC main group codes, we subsequently measured overlap in component knowledge (*Knowledge pool distance*), between the firm and the partner, by relying on the widely-used measure of technological distance introduced by Jaffe (1986) and used in numerous alliance studies (e.g. Sampson, 2007; Van de Vrande, 2013). The measure ranges between 0 and 1, where a value of 0 indicates full overlap and a value of 1 indicates no overlap in any technological domain. We controlled for potential coordination costs associated with conducting internal R&D and building new component knowledge and capabilities, by calculating the average number of inventors listed on the focal firm's patents (*Internal coordination costs*) (Grigoriou & Rothaermel, 2017). Moreover, we also controlled for the focal firm's experience with recombining internal components (*Internal component reliance*), dividing the number of internal backward citations by the total number of backward citations, and with older components (*Old component reliance)*, calculating the average age of backward citations of a focal firm's patents (Wuyts & Dutta, 2014). We also controlled for the focal firm's focus on particular patent offices, in order to control for any between-patent office heterogeneity that might affect patent citation behavior (Bakker *et al.,* 2016). Specifically, we included three control variables representing the share of patents in the focal firm's knowledge pool that were filed in each of the three main patent offices (i.e. EPO, JPO, USPTO).

3.4.3. Analytical method

The 88 focal firms were engaged in 461 R&D alliance dyads between 1993 and 2007. In the analyses, we focused on firm-partner dyads, where the firm and the partner need to be technologically active (i.e. they must have filed at least 1 fuel cell patent in the past five years) firms (i.e. we do not include upstream partners such as universities). Since most alliances lasted for longer than one year, and many of them were multi-partner alliances, this resulted in a total of 1691 firm-partner year observation. Each observation stands for a specific firm-partner dyad in a year in which it has an ongoing partnership. Following Gomes-Casseres et al. (2006), the citation output of a firm to an alliance partner in the next year is a function of the characteristics of the respective firm-partner dyad in the current year.

We relied on Generalized Estimating Equations (GEE) in order to test our hypotheses (Baum, 2008; Phelps, 2010). GEE models are especially fit for analyzing models in which the dependent variable appears as a fraction or proportion (i.e. the share of citations made to a partner by the firm) (Baum, 2008). Moreover, these specifications allow to substantially correct for non-independence across similar observations over a period of time (Hardin, Hilbe, & Hilbe, 2012). We applied an exchangeable correlation structure, to correct for correlation amongst observations from the same focal firm, and included robust standard errors in order to alleviate issues of heteroskedasticity (Hardin *et al.,* 2012). Furthermore, due to the distribution of our dependent variable, we used a binomial family and logit link function (Baum, 2008; Phelps, 2010). To control for variance over time, we included year dummies in each model. Finally, we lead the dependent variable by one year to reduce concerns of reverse causality.

3.4.4. Results

Descriptive statistics. Table 3.2 presents the descriptive statistics and correlation matrix. The average focal firm makes 1.9% of its citations to a partner in a given year. The average alliance tie has a lifespan of 2.88 years, in line with earlier studies (e.g. Phelps, 2010). The knowledge pool applicability of the focal firm and partner are well-distributed, with an average of 1.44 and 1.47 respectively, a minimum value of 1, and a maximum of 3. None of the pair-wise correlations are above 0.7. Moreover, VIF values, based on OLS regression, in all models are well-below the threshold value of 10, with an average value of 1.59 and a maximum value of 2.64 (Mason & Perreault, 1991). Therefore, we are confident that multicollinearity issues are not present in our models.

Table 3.2. Descriptive statistics and correlation matrix

	Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	Partner-specific recombination	1																								
2	Age alliance	0.16	1																							
3	Joint venture	0.08	0.23	1																						
4	Government-funded	-0.11	-0.11	-0.10	1																					
5	Multi-partner	-0.04	-0.10	-0.05	0.37	1																				
6	International	-0.09	0.23	0.15	-0.09	0.01	1																			
7	Concurrent alliances	0.05	0.19	0.26	0.08	0.21	0.01	1																		
8	Inventors	-0.16	0.01	0.02	0.02	0.09	0.18	0.02	1																	
9	Non-inventors	-0.12	0.01	-0.04	0.25	0.15	0.26	-0.01	0.30	1																
10	Upstream partners	-0.15	0.01	-0.06	0.27	0.12	0.02	0.00	0.38	0.24	1															
11	Firm knowledge pool size	0.00	0.11	-0.03	-0.12	-0.09	-0.07	0.08	0.02	-0.16	0.01	1														
12	Partner knowledge pool size	0.34	0.15	-0.01	-0.11	-0.06	-0.06	0.11	-0.06	-0.11	-0.06	0.09	1													
13	Total recombination	-0.03	0.06	0.00	-0.16	-0.15	-0.04	0.07	0.15	-0.12	0.06	0.66	0.04	1												
14	Past partner-specific recombination	0.29	0.34	0.06	-0.09	-0.06	-0.04	0.17	0.00	-0.11	-0.02	0.55	0.43	0.38	1											
15	Firm knowledge pool diversity	-0.10	-0.12	-0.06	0.09	0.00	0.03	-0.04	0.14	0.01	0.17	-0.07	-0.02	0.03	-0.05	1										
16	Partner knowledge pool diversity	0.02	-0.03	-0.04	0.01	-0.07	0.00	-0.05	0.03	-0.02	0.06	-0.03	0.02	0.00	0.01	0.14	1									
17	Knowledge pool distance	-0.22	-0.27	-0.10	0.20	0.03	-0.01	-0.15	-0.01	0.14	0.04	-0.21	-0.25	-0.16	-0.29	0.22	0.23	1								
18	Internal component reliance	-0.02	0.12	0.03	0.07	0.06	-0.06	-0.01	0.23	-0.04	0.12	0.15	0.04	0.13	0.10	0.09	-0.07	-0.12	1							
19	Old component reliance	-0.05	-0.06	0.04	0.18	0.04	0.12	-0.03	0.17	0.23	0.15	-0.26	-0.15	-0.17	-0.17	0.21	0.04	0.18	-0.04	1						
20	JPO patents	0.21	0.00	-0.02	-0.20	-0.09	-0.35	0.03	-0.35	-0.40	-0.33	0.32	0.18	0.18	0.18	-0.11	-0.07	-0.26	0.06	-0.33	1.00					
21	EPO patents	-0.09	-0.02	0.02	0.23	0.12	0.31	-0.07	0.10	0.32	0.18	-0.30	-0.21	-0.21	-0.20	0.32	0.08	0.40	-0.10	0.49	-0.48	1				
22	USPTO patents	-0.14	0.03	0.01	0.25	0.05	0.22	-0.03	0.31	0.42	0.28	-0.27	-0.13	-0.06	-0.13	0.12	0.08	0.32	0.08	0.36	-0.55	0.49	1			
23	Internal coordination costs	-0.04	-0.17	-0.02	0.08	0.08	-0.05	-0.05	0.12	0.03	0.07	-0.09	-0.04	-0.09	-0.07	0.29	0.06	0.12	-0.09	0.22	-0.06	0.28	-0.07	1		
24	Partner knowledge pool applicability	0.00	0.05	-0.03	0.04	-0.02	0.18	0.01	-0.06	0.00	0.03	-0.06	-0.04	-0.07	0.01	0.08	-0.19	-0.15	-0.03	0.03	-0.08	0.09	0.00	0.02	1	
25	Firm knowledge pool applicability	-0.09	0.06	-0.04	0.02	-0.05	0.17	-0.03	0.25	0.31	0.16	-0.09	-0.06	0.01	-0.04	-0.17	0.03	0.01	0.01	0.20	-0.32	0.19	0.41	0.08	0.00	1
	Mean	0.02	2.88	0.08	0.28	0.51	0.62	1.12	4.8	1.86	1.14	0.17	0.13	0.19	9.64	0.79	0.76	0.18	0.12	8.82	0.35	0.39	0.52	2.95	1.44	1.47
	SD	0.05	3.32	0.27	0.45	0.50	0.49	0.34	3.27	2.38	1.65	0.36	0.33	0.38	26.49	0.13	0.19	0.16	0.08	2.18	0.38	0.30	0.33	0.69	0.32	0.28
	Min	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	2	0	0	0	1	1	1
	Max	0.60	20	1	1	1	1	3	15	13	8	3.85	3.85	3.74	354	0.96	0.97	0.76	0.50	25	1	1	1	5.57	3	3

Regression results. In Table 3.3, we present the results of the GEE regressions. Among the dyad characteristics, and consistent with expectations (Sampson, 2007), we find evidence that alliance dyads that are part of a joint venture lead to more recombination of a partner's components by the focal firm (Model 1: $\beta_{\text{Joint venture}} = 0.440$, $p < .05$). Moreover, focal firms tend to recombine components of partners from different countries less often (Model 1: $\beta_{\text{International}} = -0.474$, $p < 0.01$) as expected (Gomes-Casseres *et al.,* 2006). In terms of alliance portfolio characteristics, we find that if the focal firm is engaged in R&D alliances with many other inventive partners, this has a negative and statistically significant influence on the intensity of partner-specific knowledge recombination (Model 1: $\beta_{\text{Inventors}} = -0.100$, $p < 0.01$), but the effect of non-inventive and upstream partners is statistically non-significant. Among the knowledge pool characteristics of the focal firm and the partner, we notice that partner's knowledge pool diversity positively influences partner- specific recombination (Model 1: $\beta_{\text{Partner knowledge pool diversity}} = 1.589$, $p <0.01$), whereas knowledge pool distance decreases it (Model 1: $\beta_{\text{Knowledge pool distance}} = -4.807$, $p < 0.001$).

In model 4, we test Hypothesis 1 following the procedure described by Haans et al. (2016) for testing curvilinear relationships. We find that the linear effect of a partner's knowledge pool applicability is positive and statistically significant (Model 4: $\beta_{\text{Partner knowledge pool applicability}} = 8.347$, $p < 0.001$), whereas the quadratic term is negative and statistically significant (Model 4: $\beta_{\text{Partner knowledge pool applicability squared}} = -2.574$, $p < 0.01$). We plot this relationship in Figure 3.4. The plot shows an inflection point at a value of 1.62, which is within one standard deviation of the mean, and therefore well within the range of observable points. Furthermore, the slope before the inflection point is positive and statistically significant ($p < 0.001$) and the slope after the inflection point is negative and statistically significant ($p < 0.01$). Moreover, the 95 percent Fieller confidence interval of the inflection point is within the range of observable points ([1.50,1.86]). We also find that the linear and quadratic coefficients of partner's knowledge pool applicability are jointly statistically significant ($Chi^2 = 13.82$, $p < 0.001$). Hence, we find support for Hypothesis 1.

In model 5 we test Hypothesis 2. We find that the linear effect of the firm's knowledge pool applicability has a statistically non-significant impact on the firm's partner-specific recombination. However, when we further assess whether this relationship is instead driven by curvilinear effects, we surprisingly find that the linear effect of firm's knowledge pool applicability is negative and statistically significant (Model 6: $\beta_{\text{Firm knowledge pool applicability}} = -5.547$, $p < 0.001$), whereas the quadratic term is positive and statistically significant (Model 6: $\beta_{\text{Firm knowledge pool applicability squared}} = 1.603$, $p < 0.001$). The curve, as displayed in Figure 3.5, has an inflection point at a value of 1.73, which is within one standard deviation of the mean, and therefore well within the range of observable points. Furthermore, the slope before the inflection point is negative and statistically significant ($p < 0.001$), whereas the slope after the inflection point is positive and statistically significant ($p < 0.001$). Moreover, the 95 percent Fieller confidence interval of the inflection point is within the range of observable points ([1.53,1.88]. We also find that the linear and quadratic coefficients of the focal firm's knowledge pool applicability are jointly statistically significant ($Chi^2 = 42.98$, $p < 0.001$). Hence, rather than a linear and positive relationship, we detect a U-shaped relationship between the focal firm's knowledge pool applicability and the firm's partner-specific knowledge recombination. In model 7, we jointly introduce the quadratic terms of knowledge pool applicability of the partner and focal firm, and find robust results.

Figure 3.4. Main effect of partner knowledge pool applicability

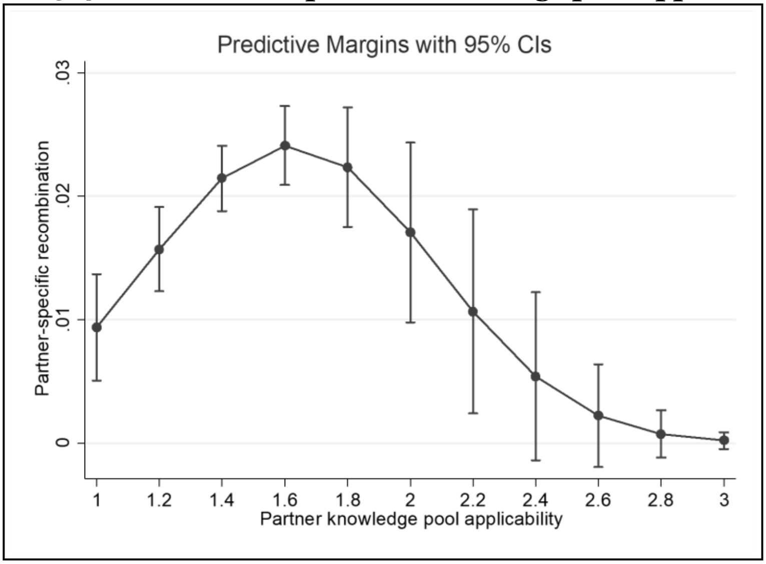


Figure 3.5. Main effect of focal firm knowledge pool applicability

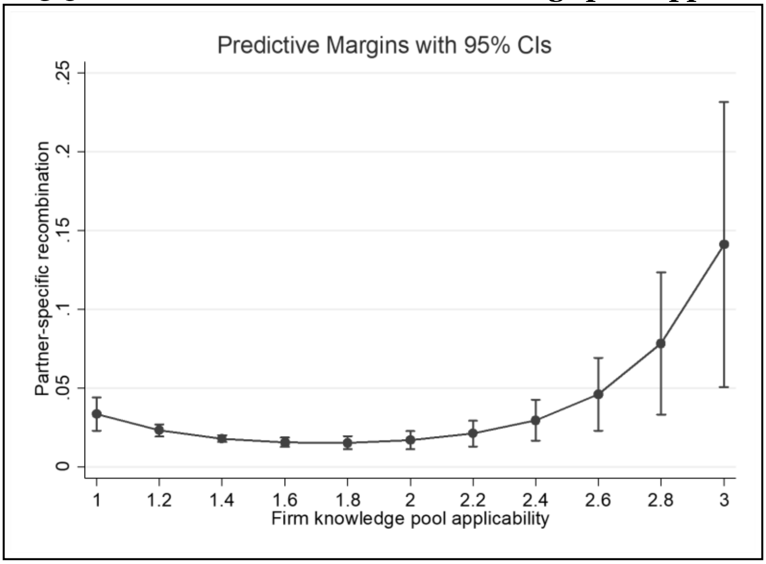


Table 3.3. GEE results

DV: Partner-specific recombination	1	2	3	4	5	6
Age alliance	0.03	0.03	0.03	0.03	0.03†	0.03†
	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]
Joint venture	0.44*	0.50*	0.44*	0.47*	0.43*	0.46*
	[0.21]	[0.22]	[0.22]	[0.23]	[0.22]	[0.23]
Government-funded	-0.23	-0.28	-0.29	-0.29	-0.32	-0.32
	[0.24]	[0.24]	[0.24]	[0.23]	[0.23]	[0.23]
Multi-partner	0.22	0.23	0.21	0.21	0.21	0.21
	[0.19]	[0.18]	[0.18]	[0.18]	[0.18]	[0.18]
International	-0.47**	-0.55**	-0.53**	-0.57***	-0.61***	-0.64***
	[0.17]	[0.17]	[0.17]	[0.16]	[0.17]	[0.16]
Concurrent alliances	-0.23	-0.24	-0.23	-0.24	-0.23	-0.24
	[0.20]	[0.20]	[0.20]	[0.20]	[0.20]	[0.20]
Inventors	-0.10**	-0.10**	-0.10**	-0.09*	-0.08*	-0.08*
	[0.03]	[0.04]	[0.04]	[0.04]	[0.04]	[0.04]
Non-inventors	0.02	0.03	0.03	0.03	0.04	0.04
	[0.03]	[0.03]	[0.03]	[0.03]	[0.03]	[0.03]
Upstream partners	-0.11	-0.11	-0.11	-0.10	-0.11†	-0.11
	[0.07]	[0.07]	[0.07]	[0.07]	[0.07]	[0.07]
Firm knowledge pool size	-1.64**	-1.65**	-1.65**	-1.64**	-1.62**	-1.61**
	[0.56]	[0.57]	[0.56]	[0.60]	[0.50]	[0.54]
Partner knowledge pool size	0.32†	0.34†	0.33†	0.36*	0.31†	0.34*
	[0.17]	[0.18]	[0.18]	[0.17]	[0.17]	[0.17]
Total recombination	0.07	0.09	0.10	0.15	0.16	0.20
	[0.19]	[0.19]	[0.19]	[0.20]	[0.19]	[0.20]
Past partner-specific recombination	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Firm knowledge pool diversity	-0.48	-0.49	-0.72	-0.75	-0.67	-0.72
	[0.54]	[0.55]	[0.59]	[0.61]	[0.61]	[0.62]
Partner knowledge pool diversity	1.59**	1.75***	1.74***	1.59**	1.72***	1.57**
	[0.49]	[0.52]	[0.51]	[0.54]	[0.49]	[0.52]
Knowledge pool distance	-4.81***	-4.75***	-4.85***	-4.05***	-5.12***	-4.37***
	[0.74]	[0.75]	[0.79]	[0.81]	[0.82]	[0.83]
Internal component reliance	-0.16	-0.10	0.04	-0.15	-0.09	-0.27
	[0.85]	[0.86]	[0.87]	[0.90]	[0.88]	[0.91]
Old component reliance	0.07	0.07	0.07	0.07	0.07	0.07
	[0.05]	[0.05]	[0.05]	[0.05]	[0.05]	[0.05]
JPO patents	0.82**	0.84**	0.83**	0.73**	0.69*	0.61*
	[0.26]	[0.25]	[0.26]	[0.27]	[0.27]	[0.28]
EPO patents	1.22***	1.22***	1.23***	1.25***	1.08**	1.12**
	[0.32]	[0.33]	[0.32]	[0.34]	[0.34]	[0.35]
USPTO patents	-0.37	-0.36	-0.24	-0.38	-0.29	-0.40
	[0.30]	[0.29]	[0.30]	[0.28]	[0.30]	[0.29]
Internal coordination costs	0.00	-0.01	0.02	-0.01	-0.06	-0.08
	[0.13]	[0.13]	[0.13]	[0.13]	[0.13]	[0.13]
Partner knowledge pool applicability		0.41†	0.40†	8.35***	0.42†	7.70***
		[0.24]	[0.24]	[2.41]	[0.24]	[2.33]
Firm knowledge pool applicability			-0.36	-0.33	-5.55***	-5.25***
			[0.42]	[0.41]	[1.06]	[1.02]
Partner knowledge pool applicability squared				-2.57**		-2.36**
				[0.79]		[0.77]
Firm knowledge pool applicability squared					1.60***	1.51***
					[0.27]	[0.25]
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1691	1691	1691	1691	1691	1691
Wald chi^2	1312.48	1297.41	1553.23	1612.86	1637.08	1895.41

† p < .10, * p < .05, ** p < .01, *** p < .001. Robust standard errors between brackets.

Robustness checks. We run several robustness checks to verify our main results[11]. First, we examine whether our results are driven by outliers in the dependent and independent variables. Winsorizing the dependent and two independent variables at the 1st and 99th percentile, our results remain highly stable[12].

Second, we apply different model specifications: (i) we use a count variable as the dependent variable (i.e. number of citations made by the focal firm to a partner) with GEE models (with a negative binomial family and link function, and an exchangeable correlation structure), (ii) we use logit models, where the dependent variable takes a value of 1 when the focal firm makes at least one citation to the partner, and (iii) following Phelps (2010), we log-odds transform the dependent variable, and run an OLS regression with focal firm and year dummies. In all three model specifications, the main results remain stable.

Third, we check for potential interaction effects between the knowledge pool applicability of the partner and focal firm (i.e. through a linear interaction and a quadratic interaction). Our results show no evidence of an interaction between the two variables, suggesting that the focal firm's knowledge pool applicability does not substitute or complement the partner's knowledge pool applicability.

Fourth, we exclude observations from all years before 1998. We chose this year as the cut-off point because, in 1998, DaimlerChrysler (now Daimler), Ford and Ballard Power Systems (a Canadian fuel cell manufacturer) formed a 1$bn joint venture to develop fuel cell systems for automotive vehicles. The formation of this joint venture is widely regarded as a turning point in the fuel cell technological field, as it was seen as a strong indicator of the commercial feasibility of fuel cell technology. Excluding these observations did not affect our main results.

Fifth, because we are dealing with a non-linear model (Hoetker, 2007; Williams, 2012), we compute the marginal effects (also referred to as partial effects) of the two quadratic relationships. Indeed, as stated by Wiersema and Bowen (2009, p. 682) "in an LDV [limited dependent variable] model, an explanatory variable's estimated coefficient can rarely be used to infer the true

[11] We do not report most of these analyses for the sake of brevity. They are, however, available from the authors upon request.
[12] Log-transforming the independent variables, consistent with Haans et al. (2016), did not influence our main results either.

nature of the relationship between the explanatory variable and the dependent variable." For partner's knowledge pool applicability, we detect that the marginal effects are positive and statistically significant ($p<0.05$) when the partner's knowledge pool applicability is in the range [1.00,1.50], statistically non-significant ($p>0.05$) in the range [1.51,1.80], negative and statistically significant ($p<0.05$) in the range [1.81,2.54], and statistically non-significant ($p>0.05$) in the range [2.55,3.00]. For the focal firm's knowledge pool applicability, we detect that the marginal effects are negative and statistically significant ($p<0.05$) when the focal firm's knowledge pool applicability is in the range [1.00,1.55], statistically non-significant ($p>0.05$) in the range [1.56,1.92], and positive and statistically significant ($p<0.05$) in the range [1.93, 3.00].

Finally, we exclude observations in which the firm made no citations in the next year. Using this alternative specification, the results remain highly stable. In sum, using these alternative models, our main results remain consistent: Hypothesis 1 finds strong support and Hypothesis 2 is not supported.

3.5. Discussion and conclusion

In this study, using a unique dataset of 88 focal firms engaged in 461 R&D alliance dyads in the fuel cell industry, we examined how the knowledge pool applicability of the partner and focal firm influence the focal firm's partner-specific recombination. We find that the partner's knowledge pool applicability has a robust inverted U-shaped relationship with the focal firm's recombination of partner's components. The results, however, do not provide support for Hypothesis 2. We find that the focal firm's knowledge pool applicability has a U-shaped relationship with its rate of partner-specific recombination, rather than the hypothesized linear and positive effect.

A potential explanation for this unexpected finding is that firms first need to develop a base-level of knowledge pool applicability, before they can benefit from the capabilities that emerge from it, in the R&D alliance context. Deploying these combinative capabilities in the R&D alliance, when understanding of how these capabilities should be aligned with available recombination opportunities is only minimal, can be costly, as the focal firm will often draw mistaken inferences about the applications of a partner's component knowledge, attempting to generate combinations that do not actually function well (Hargadon & Sutton, 1997; Nemet & Johnson, 2012). Hence, in the initial phase (i.e. from low to average internal knowledge pool applicability), firms may not yet enjoy the benefits of being able to apply these combinative capabilities to the recombination of partner's component knowledge (Darr, Argote, & Epple, 1995).

3.5.1. Theoretical implications

Alliance scholars have principally focused on the size and diversity of partner's knowledge pool as core drivers of knowledge recombination activities in R&D alliances (e.g. Lahiri & Narayanan, 2013; Schilling & Phelps, 2007; Wuyts & Dutta, 2014). In this study, we demonstrate that, even when the partner's knowledge pool size and diversity are held constant, numerous recombination opportunities may still emerge depending on the partner's knowledge pool applicability. Numerous alliance studies implicitly assume that diverse and large knowledge pools present numerous opportunities to engage in knowledge recombination (e.g. Lahiri & Narayanan, 2013; Gilsing *et al.,* 2008; Schilling & Phelps, 2007). However, we argue that numerous components can only be used in one very specific application domain, substantially reducing the range of combinations in which they can be applied. Hence, only considering the aggregate knowledge pool, in terms of how diverse or how large it is, in many cases obscures the fact that many components actually have a limited range of applications. We therefore encourage future research to account for knowledge pool applicability in their conceptual and empirical framework, when studying R&D alliances and their performance implications.

We also develop novel theoretical arguments regarding firms' idiosyncratic abilities to engage in knowledge recombination in the partner's knowledge pool. Existing alliance research tends to conceptually focus more on the absorption of component knowledge into the focal knowledge pool, often remaining relatively agnostic about the actual recombination of a partner's component knowledge (e.g. Rosenkopf & Almeida, 2003; Vasudeva & Anand, 2011; Gilsing *et al.,* 2008). However, we emphasize that even if a component is transferred into the knowledge pool, it is also important to look at whether the focal firm actually knows how to use this component in knowledge recombination (Zahra & George, 2002). Hence, it is important to look into factors that are likely to drive firm-specific abilities to engage in meaningful knowledge recombination (Wuyts & Dutta, 2014). To this end, we argued that combinative capabilities that emerge from building widely-applicable component knowledge allow the focal firm to have a better understanding of components' true recombinant potential, overcoming technological limits that other firms might have (Henderson, 1995). Moreover, these capabilities allow the focal firm to recognize when the recombinant potential of a component has reached its limit already, thus avoiding fruitless recombination efforts. We therefore point to the need to consider a more nuanced knowledge recombination perspective when examining the application of focal firms' internal capabilities in R&D alliances.

3.5.2. Limitations and future research

There are several limitations in this study, which can form a starting point for future research. First, in order to test our hypotheses, we chose an empirical setting which allowed us to easily operationalize the knowledge pool applicability of the focal firm and partner. Specifically, we examined the IPC subgroups that are listed on fuel cell patents, in order to determine a firm's knowledge pool applicability. In the fuel cell technological field, this approach was facilitated by the fact that there were only four major application domains for fuel cell inventions, and that there was a high correspondence between these four domains and the IPC subgroups of fuel cell patents. Although it is likely that this type of approach would work in other settings, we still encourage future studies to attempt to capture application domains in other settings.

Second, in this study we focused on technological application domains nested within a larger technological domain, capturing application domains which have immediate relevance to the context of this study (i.e. fuel cell-oriented R&D alliances). We encourage future studies to examine other types of application domains for knowledge components, such as other industries or countries (e.g. Hargadon & Sutton, 1997; Petruzzelli & Savino, 2014), developing potentially different theoretical insights than the ones discussed in this study.