

University of Groningen

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

Kok, Holmer Jan

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*  
2018

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Kok, H. J. (2018). *To Complete a Puzzle, You Need to Put the Right Pieces in the Right Place: Exploring Knowledge Recombination and the Creation of New Inventions*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen, SOM research school.

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

---

## Chapter 2. Dusting off the Knowledge Shelves

### The Impact of Recombinant Lag on the Technological Value of Inventions

---

**Abstract:** Whereas knowledge recombination research tends to focus on original knowledge component attributes and their recombinant value implications, we contribute to an emerging literature stream on knowledge reuse trajectories, investigating the temporal dimension of reuse by introducing the concept of recombinant lag – i.e. the time that components have remained unused. Relying on organizational learning theory, we emphasize that it is not only important to consider the frequency of reuse, but also the recency of reuse. Our core argument is that recent reuse of knowledge components can trigger a rejuvenation effect that influences the value of resulting inventions. Analyzing 21,117 fuel cell patent families, we find an unexpected U-shaped relationship between recombinant lag and the value of inventions, which is moderated by frequency of reuse. Conducting post-hoc exploratory data analyses, we advance the concept of dormant components – i.e. valuable components that have remained unused prolongedly, as a potential explanation for this unexpected U-shaped pattern. Moreover, collecting and analyzing data on a second sample in the wind energy industry, we provide first indications for the generalizability of these unexpected findings. We contribute to a richer understanding of knowledge reuse trajectories, highlighting that, next to the magnitude of reuse information flows – i.e. information flows that are generated when components are reused, the timing of creation of these information flows shapes the value of subsequent recombination activities. We also contribute to extant research on the temporal dimension of knowledge recombination, pointing to recombinant lag as an important aspect next to component age.

---

This chapter was written together with Dries Faems and Pedro de Faria. Earlier versions of this chapter have been presented at the *Strategic Management Society Annual International Conference* in Denver (2015), *Annual Meeting of the Academy of Management* in Anaheim (2016), and in research seminars at *University of Groningen* (2015) and *Tilburg University* (2017). A manuscript based on this chapter is conditionally accepted for publication in the *Journal of Management*.

---

## 2.1. Introduction

Inventions originate from the recombination of existing components<sup>1</sup> (Fleming, 2001). The technological value of new inventions therefore hinges on attributes of recombined components, such as the technological field, geographical location, organizational context and temporal context from which they originate (e.g. Nerkar, 2003; Phene *et al.*, 2006; Rosenkopf & Nerkar, 2001). Whereas knowledge recombination research has focused on how the recombinant value of components is driven by their original attributes – i.e. attributes that were embedded into the component at the time of creation, this value is not necessarily pre-determined at creation (e.g. Fleming, 2001; Wang *et al.*, 2014). Instead, components go through a unique trajectory over time, which influences their recombinant value. Recently, some studies have started examining how components' recombinant value changes over time (e.g. Belenzon, 2012; Fleming, 2001; Yang *et al.*, 2010), focusing on how the frequency of reuse of components – i.e. the number of times a component was previously reused in a combination – shapes recombinant value (e.g. Boh *et al.*, 2014; Fleming, 2001; Katila & Ahuja, 2002). These scholars argue that each instance of component reuse produces new information flows about the component, which can improve subsequent recombination activities (Katila & Chen, 2008).

This emerging stream of literature on reuse trajectories, however, tends to ignore the temporal dimension of component reuse, neglecting that components differ in terms of when they were last reused<sup>2</sup>. This is surprising as two components created at the same time may go through different reuse trajectories over time, where one may have been last reused 10 years ago, and the other only 1 year ago. We therefore argue that, to increase our theoretical understanding of how knowledge reuse trajectories influence components' recombinant value, it is not only important to look at their frequency of reuse, but also essential to look at when this reuse occurred. To capture the temporal dimension of reuse, we introduce the

---

<sup>1</sup> In the context of this study, components refer to the “fundamental bits of knowledge or matter that inventors might use to build inventions” (Fleming & Sorenson, 2004: 910).

<sup>2</sup> An exception is Capaldo *et al.* (2017) who looked in a robustness check at the time elapsed since the last instance of reuse by the firm. However, they position the recency of component reuse as an alternative measure of component age. In contrast, we see the recency of component reuse as a distinct dimension of time that has a different effect from component age.

concept of recombinant lag – i.e. the time that recombined components have remained unused – and empirically test its impact on the technological value of resulting inventions<sup>3</sup>.

Using organizational learning theory insights (Argote & Miron-Spektor, 2011), we argue that recent reuse of knowledge components allows for the creation of information flows about the contemporary applications of the component in knowledge recombination. Such rejuvenation effect subsequently increases the value of resulting inventions. As components remain unused for longer periods, however, we expect this rejuvenation mechanism to reduce in strength in a non-linear way. Therefore, we hypothesize a non-linear negative relationship between recombinant lag and technological value of resulting inventions. We also predict that frequency of reuse moderates this relationship in such a way that the value-enhancing mechanism of rejuvenation becomes stronger when a component was frequently reused.

To test the hypotheses, we rely on a sample of 21,117 patent families in the fuel cell industry. Our analyses point to an unexpected U-shaped relationship between recombinant lag and the technological value of inventions. In addition, we observe that, for this fuel cell sample, this relationship mainly manifests itself when the frequency of reuse is low. Based on additional analyses – i.e. screening of raw data, exploration of fuel cell journals, an interview with a fuel cell expert, and additional tests in the wind energy industry, we explain this unexpected pattern by pointing to the existence of dormant components – i.e. valuable components that have remained unused for prolonged periods. Moreover, we provide first indications for the generalizability of this unexpected pattern, confirming the U-shaped relationship for an additional sample in the wind energy industry.

This study adds to the knowledge recombination literature in two important ways. First, we contribute to an emerging stream of literature on knowledge reuse trajectories and their impact on recombinant value of components. In particular, we theorize on the different mechanisms underlying frequency and recency of reuse and empirically demonstrate their impact on the technological value of

---

<sup>3</sup> Following earlier studies, we examine to what extent the recombination of particular components increases the technological value of resulting inventions, which we conceptualize as the number of times that these inventions serve as inputs for subsequent recombination efforts (Fleming, 2001; Rosenkopf & Nerkar, 2001).

inventions. Second, we contribute to a richer perspective on the temporal dimension of knowledge recombination. We show that it is not only important to consider when a component was created (i.e. component age), but also when it was last used to create new inventions. In terms of managerial implications, we highlight that the reevaluation of existing knowledge stocks may play an important role in the implementation of knowledge creation strategies.

## **2.2. Theoretical background**

In this section, we discuss how extant knowledge recombination literature relies on knowledge search concepts to study how original attributes of components shape their recombinant value. Subsequently, we discuss an emerging stream within knowledge recombination literature that shifts focus from original component attributes to knowledge reuse trajectories as drivers of recombinant value. Finally, we point to the need to explicitly consider the temporal dimension of component reuse, introducing the concept of recombinant lag.

### **2.2.1. Original component attributes and recombinant value**

In the early 1990's, inventors from Mitsubishi Electric Corporation and Kansai Electric Power Company recombined existing component knowledge on (i) fuel reformers and (ii) electrodes in order to generate highly efficient fuel cell systems in which the exothermic heat produced by the fuel cell could directly be used to fasten the endothermic reforming process (Ohtsuki, Seki, Miyazaki, & Sasaki, 1995). This example illustrates how new inventions originate from processes of knowledge recombination in which inventors seek out existing components and recombine them in novel ways (Fleming, 2001). Since recombined components largely determine how a new invention functions, the value and usefulness of a new invention hinges on the attributes of the recombined components (Capaldo *et al.*, 2017; Li *et al.*, 2008). Relying on knowledge search theory (Stuart & Podolny, 1996), existing research has mainly focused on original attributes of components, assuming that attributes that are embedded in components at the time of creation determine the value that can be realized from using them in recombination. Following this theoretical perspective, scholars have pointed to two important underlying mechanisms affecting the value of inventions that result from the

recombination of components: novelty and retrievability (Miller *et al.*, 2007; Phene *et al.*, 2006; Rosenkopf & McGrath, 2011). Whereas novelty refers to the extent to which the component is new to the focal inventor or context (Rosenkopf & McGrath, 2011), retrievability signifies the extent to which the component can be absorbed into the focal inventor's knowledge pool (Miller *et al.*, 2007; Phene *et al.*, 2006). Relying on these insights, scholars have examined how the origins of recombined components in terms of technological field (e.g. Nemet & Johnson, 2012; Rosenkopf & Nerkar, 2001), geographical region (e.g. Ahuja & Katila, 2004; Phene *et al.*, 2006), organizational context (e.g. Miller *et al.*, 2007), and temporal context (e.g. Capaldo *et al.*, 2017; Katila, 2002; Nerkar, 2003) influence the value of resulting inventions.

### **2.2.2. Component reuse, learning opportunities and knowledge recombination**

Whereas existing knowledge recombination research mainly investigates original attributes of components as drivers of recombinant value, some scholars have started shifting attention to how the recombinant value of components is also driven by their reuse over time (Belenzon, 2012; Fleming, 2001; Wang *et al.*, 2014). These scholars assume that components are highly malleable (Hargadon & Sutton, 1997; Wang *et al.*, 2014), and can be reused in numerous and diverse ways (Dibiaggio *et al.*, 2014; Fleming, 2001; Hargadon & Sutton, 1997; Yayavaram & Ahuja, 2008) by different inventors situated in different organizations (Belenzon, 2012; Yang *et al.*, 2010) at different points in time (Katila & Chen, 2008).

Relying on insights from organizational learning theory (Argote & Miron-Spektor, 2011), they frame component reuse as a learning process by which new information flows are generated that allow inventors to guide and improve their own recombination activities (Katila & Chen, 2008). Through each instance of reuse, new information is produced about how the component behaves in a new combination (Yang *et al.*, 2010). We label this release of new information through the reuse of a component as reuse information flows. Through these reuse information flows, inventors may obtain an improved understanding of the technological specificities underlying this component. To acquire such information flows, inventors may disassemble combinations in which components were reused,

gaining important information about the interconnections that exist between constituent components (Hargadon & Sutton, 1997; Sorenson, Rivkin, & Fleming, 2006; Zander & Kogut, 1995). Several studies provide evidence to support these learning dynamics (Katila & Chen, 2008), showing how inventors learn from prior recombination efforts by acquiring technologies for reverse-engineering (Zander & Kogut, 1995) or by closely inspecting patent documents and scientific publications (Murray & O'Mahony, 2007; Yang *et al.*, 2010). For example, inventors often acquire and subsequently test fuel cell stacks for prolonged periods of time, obtaining an understanding of how each individual component that comprises the fuel cell stack (such as electrolytes, electrodes, bipolar plates) contributes to the overall performance of the combination.

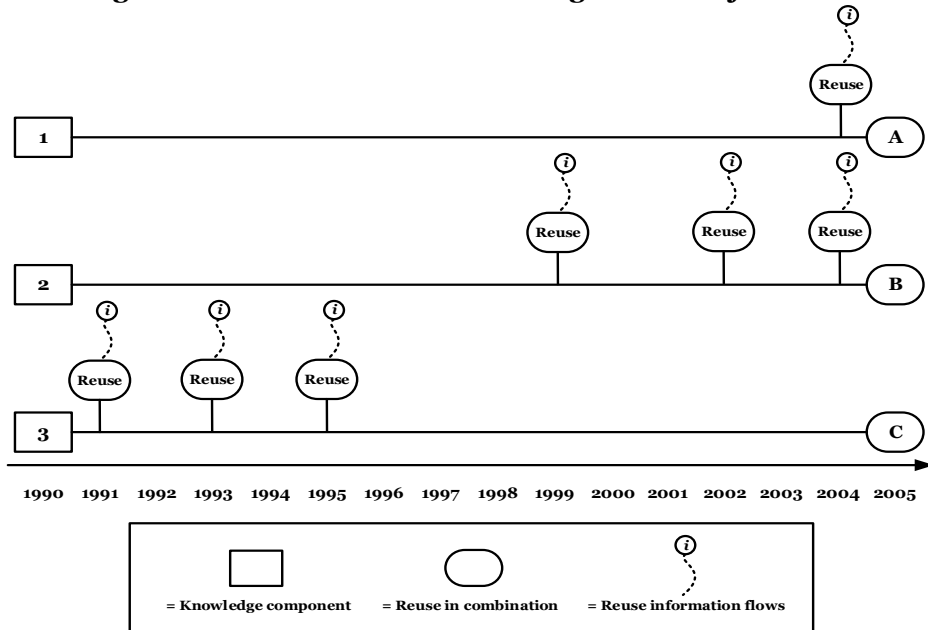
### **2.2.3. Frequency and recency of component reuse**

Pointing to the importance of knowledge reuse trajectories, scholars have primarily focused on the frequency of reuse. They have mainly argued that the frequency of component reuse is positively related to the value of inventions (Boh *et al.*, 2014; Dibiaggio *et al.*, 2014; Fleming, 2001). As instances of reuse provide important opportunities to obtain a richer understanding of components' specificities, frequently reused components tend to be more reliable and well-understood in knowledge recombination (Fleming, 2001; Katila & Ahuja, 2002; Wang *et al.*, 2014). Effectively, the higher the number of prior instances of reuse, the higher the number of reuse information flows and learning opportunities that are available (Yang *et al.*, 2010).

In this study, we argue that, next to its frequency, component reuse also varies in terms of its recombinant lag, which we define as the time that recombined components have remained unused. To illustrate the notion of recombinant lag, consider Figure 2.1, where we compare three components that were created in 1990 and which are recombined in an invention in 2005. Having as reference point the recombination that occurred in 2005, components 1, 2, and 3 have similar component age (i.e. 15 years). Moreover, we see that, before 2005, component 1 has been reused once, whereas component 2 and 3 have been reused three times. Although the frequency of reuse of component 2 and 3 is similar, the recombinant lag of component 2 is equal to one (i.e. component 2 was last reused in 2004),

whereas the recombinant lag of component 3 is ten. In the next section, we theorize on how these differences in recombinant lag are likely to influence the technological value of resulting inventions.

**Figure 2.1. Three different knowledge reuse trajectories**



## 2.3. Hypotheses

In this section, we hypothesize how recombinant lag influences the value of knowledge recombination. Our core argument is that a recent instance of reuse creates reuse information flows that can be used by potential inventors to learn how to apply the component in contemporary knowledge recombination and to create inventions with higher technological value. At the same time, we expect this rejuvenation mechanism to lose strength in a non-linear way. Moreover, we theorize that the strength of this rejuvenation mechanism is contingent upon the frequency of reuse of components.

### 2.3.1. Recombinant lag and the technological value of inventions

Organizational learning scholars have long acknowledged that the value of learning opportunities associated to information flows is dependent on the particular



## Chapter 2

temporal context during which they occur (e.g. Argote & Miron-Spektor, 2011; Eggers, 2012). Building on these insights, we argue that, next to considering the magnitude of reuse information flows (i.e. frequency of component reuse), it is also important to look at when reuse information flows are generated (i.e. recency of component reuse). Specifically, we argue that recent reuse of a component implies the generation of reuse information flows, which are embedded in the state-of-the-art of technology. Recent reuse of a component thus creates learning opportunities that allow inventors to infer how to apply the component in contemporary knowledge recombination activities, essentially ‘rejuvenating’ the component’s recombinant potential. Effectively, the way components are recombined into new inventions changes over time in line with the evolution of technological paradigms (Dosi, 1982). If we draw a parallel to cooking, we can see knowledge components as food ingredients which are combined and cooked in a particular way in order to prepare a meal (Petruzzelli & Savino, 2014). Culinary preferences change based on newly-acquired tastes and trends in the market, making it necessary to integrate certain ingredients into meals in different ways over time. In the same way, the more recent the last instance of reuse of a component, the more modern and up-to-date the ways in which the component was recombined. As a result, more valuable reuse information flows are generated. Tapping into these reuse information flows, recently reused components become more suitable to inventors for addressing present-day technological problems and opportunities.

To give an example of the importance of recent reuse, consider the case of fuel reformers in the fuel cell industry. Fuel reformers are typically used in fuel cell systems to extract hydrogen from a hydrocarbon (such as gasoline) or an alcohol fuel (such as methane), to be subsequently used as the reactant in the fuel cell. During the 1980’s and 1990’s, fuel reformer components were often used to design new fuel reformer systems for large-scale fuel cell power plants. In the early 2000’s, however, it was expected that the existing oil and gas infrastructure could be used for fuel cell vehicles (known as FCV). Inventors at firms such as Shell and ExxonMobil therefore started recombining existing component knowledge of fuel reformers in order to develop on-board fuel reformer systems that could be installed inside FCVs. A fuel cell expert we interviewed described these as “small chemical plants under the hood”. Consequently, fuel reformer components that

had been used in combinations for fuel cell plants decades before were now being reapplied in FCVs in radically different ways. By accessing these recently-produced reuse information flows, inventors were able to infer the most up-to-date applications of fuel reformer components, generating ultimately more useful inventions as a result.

However, given the generally rapid pace of technological change (Fabrizio, 2009; Stuart & Podolny, 1996), it is likely that the rejuvenation effect depreciates in a non-linear way. We expect that the difference in technological value between an invention with a recombinant lag of 1 year and an invention with a recombinant lag of 4 years is likely to be substantial as the learning opportunities of 1-year old reuse information flows are likely to be much higher than 4-year old reuse information flows. In contrast, the difference in technological value between an invention with a recombinant lag of 4 year and an invention with a recombinant lag of 7 years is likely to be less outspoken as the learning opportunities of 4-year old and 7-year old reuse information flows are likely to be more similar.

In sum, we expect that, for components with a high recombinant lag, reuse information flows will provide less useful opportunities to learn how to apply the component in contemporary knowledge recombination compared to components with a low recombinant lag. Consequently, we expect the technological value of inventions that result from the recombination of components with a high recombinant lag to be lower than the technological value of inventions resulting from components with a low recombinant lag. However, because we expect the most recent instances of reuse to provide substantially more useful reuse information flows to inventors than relatively less recent ones, we predict a non-linear relationship between recombinant lag and technological value. In particular, we expect the negative effect of moving from low to medium recombinant lag to be more outspoken than the negative effect of moving from medium to high recombinant lag. We therefore hypothesize:

*Hypothesis 1: The recombinant lag of components used in knowledge recombination has a negative and diminishing impact on the technological value of resulting inventions*

### **2.3.2. The moderating effect of the frequency of reuse**

Components do not only differ in terms of their recency of reuse, but also in terms of how frequently they have been reused. Jointly considering these two dimensions of component reuse, we expect that the frequency of reuse amplifies the rejuvenation effect associated with low recombinant lag.

A core tenet of organizational learning theory is that learning opportunities that are less ambiguous tend to be more useful (Argote & Miron-Spektor, 2011; Bohn, 1995; Lampel, Shamsie, & Shapira, 2009). Relying on these insights, we argue that reuse information flows from a recent instance of component reuse are less ambiguous (and, therefore, more useful) when numerous prior combinations are available in which the component was also reused. In particular, when frequency of reuse is higher, ambiguity regarding the unique features of the most recent and contemporary application of the component will be substantially reduced. When a component was reused more frequently, the inventor can access numerous reuse information flows regarding the component's prior instances of reuse, and use these to contrast how the component's most recent instance of reuse deviates from older ones (e.g. in Figure 2.1, component 2's most recent recombination in 2004 can be contrasted with its recombinations in 2002 and 1999).

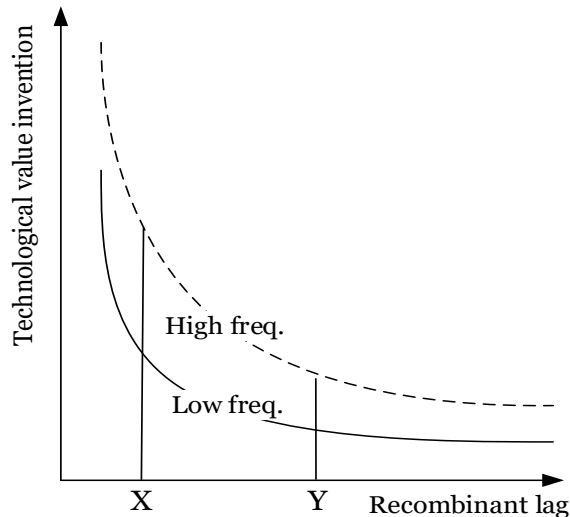
Going back to our previous example: before they were reused in FCVs, fuel reformer components had already been used extensively in combinations targeted at fuel cell power plants, producing sizeable reuse information flows. Later, when inventors started developing on-board fuel reformers, the recent reuse of fuel reformer components in FCVs could easily be contrasted with prior reuse in fuel cell plants. Consequently, this allowed inventors to better understand how recombination of fuel reformer components in FCVs differed from recombination in fuel cell plants.

Higher frequency of reuse is thus expected to enhance the rejuvenation effect of recombinant lag. However, we argue that the strength of this moderation effect will depend on the level of recombinant lag. In particular, as we argued in the previous section, the rejuvenation effect of recombinant lag is likely to reduce in strength in a non-linear way as recombinant lag increases. Therefore, at higher values of recombinant lag, higher frequency of reuse will only slightly raise the

impact of recombinant lag on the technological value of inventions. To clarify our reasoning, we depict this hypothesized moderating relationship in Figure 2.2. Here, we observe that, when recombinant lag has a value of X, the difference in impact between high and low frequency of reuse is considerably large. This is because, through higher frequency of reuse, the rejuvenation effect of recent reuse is amplified. However, moving towards a value of recombinant lag of Y, we observe that the difference between low and high frequency of reuse becomes smaller. At such high values of recombinant lag, the rejuvenation effect is nearly dissipated, making the moderating effect of higher frequency of reuse negligible. In other words, the convex relationship between recombinant lag and technological value of invention is expected to become steeper when the frequency of reuse is higher. Therefore, we hypothesize:

*Hypothesis 2: The frequency of reuse of components used in knowledge recombination moderates the relationship between components' recombinant lag and the technological value of resulting inventions in such a way that the relationship becomes steeper for higher frequency of reuse.*

**Figure 2.2. Frequency of reuse and recombinant lag**



## 2.4. Methodology

### 2.4.1. Empirical context

To test our hypotheses, we collected data on inventions related to fuel cell technology. We studied the patent family applications of the 200 firms with the highest number of patent applications in this industry. Invented in 1839 by William Grove, fuel cells produce electricity through a chemical reaction that combines a fuel (usually hydrogen) with an oxidizing agent (usually oxygen). This technology witnessed its first practical application in the 1960's when it was used by NASA in the Space Program to provide electricity (and drinking water) to spacecrafts (Perry & Fuller, 2002). In subsequent decades, the potential of this technology has been exploited in distributed energy generation, automobiles, and portable electronic devices (Sharaf & Orhan, 2014).

The fuel cell industry is suitable for testing our hypotheses for several reasons. First, given the long technological lineage of fuel cell technology, components used in fuel cell inventions vary substantially in terms of when they were created and when they were last used. Second, knowledge recombination as a means to generate new inventions is pervasive in the fuel cell industry. In fact, the successful integration of disparate components into coherent combinations is often heralded as the foundation of success of new fuel cell technologies (Sharaf & Orhan, 2014). Third, we use patent data to track inventions, and studies have shown that patenting propensities in fuel cell technology are among the highest in clean energy technologies (Albino *et al.*, 2014).

### 2.4.2. Data

**Patent data.** To study fuel cell inventions, we relied on patent data retrieved from the October 2013 version of the PATSTAT database<sup>4</sup>. In line with recent studies (e.g. Bakker, Verhoeven, Zhang, & Van Looy, 2016), we used patent families to identify inventions and knowledge recombination. To delineate patent families, we used the European Patent Office worldwide bibliographic database

---

<sup>4</sup> Some authors have expressed concerns about the use of patent data to study inventions, arguing that some firms tend to rely more on alternative appropriation mechanisms (Arundel & Kabla, 1998; de Faria & Sofka, 2010). Nevertheless, given the fact that we sampled from an industry in which patenting propensity rates are generally elevated, these concerns are alleviated (Arundel & Kabla, 1998).

(DOCDB) patent family definition. The DOCDB patent family captures all patent applications related to the same invention but filed at different patent offices (Albrecht, Bosma, Dinter, Ernst, Ginkel, & Versloot-Spoelstra, 2010). Effectively, a patent applicant seeking protection for an invention in more than one juridical region has to file a new patent application in each separate region (e.g. the USPTO for the U.S. and the JPO for Japan). These different patent applications from different patent offices collectively comprise more information about the invention than if only one single patent office is considered (Nakamura, Suzuki, Kajikawa, & Osawa, 2015). Therefore, we collected patent applications from all patent offices in the world, and aggregated these to the patent-family level. To capture the date that is closest to ideation of the invention, we looked at the priority date (i.e. the first time that the applicant sought patent protection for its invention at a patent office) of the patent family. The use of patent families to denote inventive activities has a number of advantages over the use of single patent office applications (Bakker *et al.*, 2016; de Rassenfosse, Dernis, Guellec, Picci, & van Pottelsberghe de la Potterie, 2013; Martínez, 2011). First, it captures a wider array of inventions since it does not limit itself to one patent office (Bakker *et al.*, 2016). Second, it overcomes the home-country bias of single patent office applications (Criscuolo, 2006; de Rassenfosse *et al.*, 2013). Third, studying patent families provides a more complete coverage of backward citations than single patent office applications (Albrecht *et al.*, 2010; Nakamura *et al.*, 2015).

Following earlier research, we studied patents' backward citations to examine the components that are recombined to create new inventions (Jaffe & de Rassenfosse, 2017; Phene *et al.*, 2006; Rosenkopf & Nerkar, 2001)<sup>5</sup>. Since we studied patent families, we aggregated all backward citations at the patent family-level (see for an example: Nakamura *et al.*, 2015). We collected all patent family applications filed by firms in IPC class H01M8 (titled 'Fuel Cells; Manufacture thereof') which corresponds to fuel cell technology (Tanner, 2014). Our data collection procedure allowed us to identify a total of 21,117 patent family

---

<sup>5</sup> We recognize that patent citations included by patent examiners may bias some of our results (Alcacer & Gittelman, 2006). However, we are confident that our findings are not driven by this data limitation. As Sorenson *et al.* (2006: 1001) note: "At worst, if examiners add citations that do not reflect true knowledge flows and do so in an unbiased way, this should only add noise, increasing the difficulty of finding statistical support for our hypothesis".

applications. These patent family applications were retrieved after removing (i) patent families that were not filed by the firms that we consolidated, (ii) patent families with incomplete backward citation information and (iii) patent families filed after 2007.

***Firm ownership data.*** To ensure that the examined patents captured the full extent of the firms' inventive activities, we aggregated the subsidiaries of the 200 firms with the highest number of patent applications in the fuel cell industry at the parent firm-level (Ahuja & Lampert, 2001; Nerkar, 2003). It was necessary to consolidate patenting activities at the parent firm-level in order to identify which patent citations were internal (i.e. citations between patents from the same applicant) and which were not.

We identified all subsidiaries in which each of these 200 firms had a controlling interest. In order to do so, we relied on the most recent ownership data available for these firms in Bureau van Dijk's Orbis Database. We subsequently matched the names of these subsidiaries to those available in the patent database<sup>6</sup>. Some of the firms in the top 200 were subsidiaries of other firms in the top 200, therefore their patent applications were aggregated at the parent firm-level. Other firms had incomplete ownership data due to, for example, bankruptcy, and therefore were not included in the analysis. As a result, our final group of firms included 139 firms.

### **2.4.3. Variables**

***Dependent variable.*** To measure the *Technological value* of inventions, we relied on forward citations (i.e. citations made to the patent family). Forward citations have often been used to capture the technological value of patented inventions (Ahuja & Lampert, 2001; Fleming, 2001; Jaffe & de Rassenfosse, 2017). Forward citations correlate positively with the economic value of patents (Hall, Jaffe, & Trajtenberg, 2005) and technology improvement rates (Benson & Magee, 2015). A high citation count indicates that a patented invention is frequently used

---

<sup>6</sup> We made efforts to connect the subsidiaries to their respective parent firms by inspecting potential name changes and other names that firms were known as (also available in the Orbis Database). Moreover, we collected data on mergers and acquisitions of the firms in our sample (retrieved from the SDC Platinum Mergers and Acquisitions Database). We also cross-checked ambiguous cases using the LexisNexis Academic Database. Finally, when necessary, we also inspected the address that was listed on the patent application of the applicant (provided that they were available). Harmonized applicant names were obtained through the EEE-PPAT, provided by ECOOM.

as an input for new patented inventions. Since older patents may receive more citations because they have been in existence for longer (Fleming, 2001; Nemet & Johnson, 2012), we applied a fixed four-year window to forward citations. In other words, irrespective of the year in which the patent was filed, we counted the number of forward citations that was made to this patent within the first four years after it was filed (e.g. for a patent filed in 2000, we counted the number of forward citations made to this patent up until 2004). In line with prior research (e.g. Miller *et al.*, 2007), we excluded internal forward citations.

***Independent variables.*** To measure *Recombinant lag*<sup>7</sup>, we considered the forward citations made to patents that were cited by the focal patent. Katila and Chen (2008: 606) already noted that “because one of the requirements for patenting is novelty, each time an existing patent is cited as an antecedent for a new patent, it is used in a different context than before. Thus each repeat use of a citation serves as a distinct source for learning.” Hence, the use of patent citations to track reuse and learning opportunities generated therefrom is highly suitable. For each patent cited by the focal patent, we calculated how many years elapsed between the priority year of the focal patent and the priority year of the last citation that was made to the cited patent. For example, in Figure 2.3 we have patent C that cites patent A, which was filed in 1995 and was last cited by patent B in 1997. The number of years elapsed between the creation of the focal patent (i.e. 2002) and the last citation that was made to patent A by patent B (i.e. 1997) is 5, which represents the recombinant lag of patent A for patent C. When a focal patent cites a patent, which had not been cited before, the recombinant lag equals the number of years elapsed between the priority year of the focal patent and the priority year of the cited patent. In Figure 2.3, the recombinant lag of patent A for patent B is therefore 2.

For each patent, we took the median value of recombinant lag of its backward citations (Nerkar, 2003). We took the median value of recombinant lag to more aptly capture the typical time that recombined components had remained

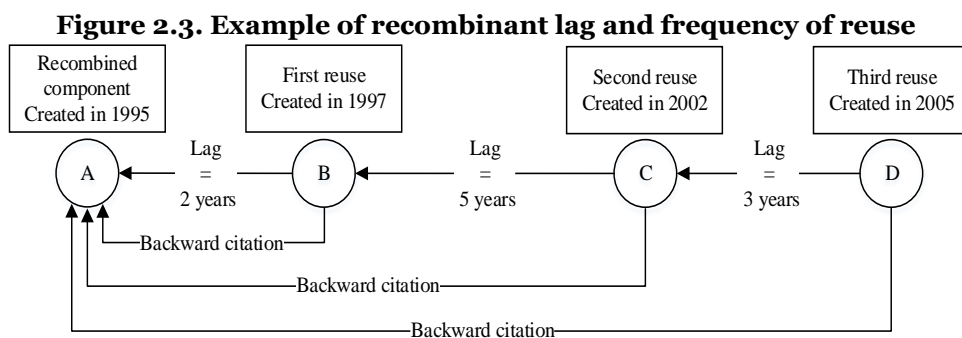
---

<sup>7</sup> Capaldo *et al.* (2017) tested the impact of a similar measure in one of their robustness checks. Whereas we examine the time that a patent has not been cited by *anyone*, they examined whether the time that elapsed since the last citation *by the firm* has a similar effect on the value of a patent as the age of a patent citation. We computed the same measure in a robustness check. We found that this measure had a very high correlation with age (0.85). Moreover, its impact on the value of a patent was similar to that of component age (i.e. strictly negative and linear).



unused. Relying on the median value of a variable is also warranted when the distribution of the variable is skewed. In our case, the distribution of recombinant lag was skewed to the right, indicating that most components had remained unused for short periods of time (i.e. 58 percent of backward citations had a recombinant lag of 1).

To measure *Frequency of reuse*, we examined how often the patents cited by the focal patent family were themselves cited by other patents (Hohberger, 2017; Miller *et al.*, 2007). In this way, we could assess to what extent an invention recombines components that were frequently used in other combinations. For example, in Figure 2.3, patent A was cited once before being cited by patent C. The frequency of reuse of patent A for patent C is therefore 1. Similarly, patent A was cited twice before being cited by patent D. The frequency of reuse of patent A for patent D is therefore 2. For each patent, we took the average value of frequency of reuse of its backward citations.



**Control variables.** Following prior research on the technological value of inventions, we included several control variables in the models. We controlled for several attributes of recombined components. It is expected that recombination of older components yields a negative impact on the technological value of inventions (Benson & Magee, 2015; Fabrizio, 2009; Nerkar, 2003; Schoenmakers & Duysters, 2010). We control for this by including the variable *Component age*, which is measured by the median number of years that elapsed between the priority year of the focal patent family and the priority years of the backward citations (Nerkar, 2003). Inventions that recombine a larger number of components tend to be more valuable (Kelley, Ali, & Zahra, 2013). To control for this fact, we included the

variable *Number of components*, which is measured by counting the number of backward citations of the patent family. Moreover, inventions that rely strongly on internal components tend to be less valuable (Kim, Song, & Nerkar, 2012; Rosenkopf & Nerkar, 2001). The variable *Internal components* controls for this fact and is calculated by dividing the number of internal backward citations by the total number of backward citations of the patent family. The technological diversity of recombined components may further influence the technological value of resulting inventions (Kelley *et al.*, 2013). We computed the variable *Technological breadth* using the measure developed by Gruber *et al.* (2013), which calculates technological breadth at the backward citation-level on the basis of IPC codes (we used the subclass level).

We also controlled for attributes of the focal patented invention. Single inventors tend to generate inventions with poorer outcomes than teams of inventors (Singh & Fleming, 2010). Hence, to control for these effects, we included the variable *Team size* which counts the number of inventors that are listed on the patent family application. Moreover, earlier research found that the number of patent authorities in which a patent was filed correlates with the value of the invention (Harhoff, Scherer, & Vopel, 2003). Hence, the number of patent offices in which a patent was filed may be indicative of the quality of the underlying invention. We included the variable *Patent offices* which counts the number of unique patent offices in which patents in the patent family were filed. Finally, since granted patents have passed patent examiners' evaluation of patentability, their technological value is generally higher. To control for this fact, we included the binary variable *Patent granted* which takes a value of 1 if at least one patent in the patent family had been granted.

#### **2.4.4. Analytical method**

Our unit of analysis is the patent family. In total, we analyzed 21117 patent families filed by 139 unique applicants over the time period 1959-2007. Each patent family was only observed once, in the year corresponding to its priority date. As our dependent variable is an overdispersed count variable (i.e. the standard deviation of the variable exceeds the mean), we used negative binomial regressions to test our hypotheses (Hausman, Hall, & Griliches, 1984). This method of analysis has

## Chapter 2

also been employed by prior research using patent data (e.g. Fleming, 2001; Nemet & Johnson, 2012; Rosenkopf & Nerkar, 2001). To control for variance associated with the year of creation of the patent family, we included year dummies in all models. Moreover, to hold constant characteristics of the applicant of the patent (i.e. the focal firm), we included firm dummies in all models. Finally, to control for heteroskedasticity, we included robust standard errors in all models.

### 2.4.5. Results

**Descriptive statistics.** Table 2.1 shows the descriptive statistics and correlation matrix. On average, the patents in our sample receive 2.48 citations in the first four years after creation. Moreover, we find that 38.5 percent of these patents receive no forward citations in the first four years after creation. The patents in our sample typically have a recombinant lag of 1.72 years, suggesting that most inventions rely on components that have remained unused for relatively short periods of time. Finally, our descriptive statistics indicate that the patents in our sample typically cite patents that were, on average, previously cited 6.96 times by other patents.

To check for potential multicollinearity problems, we consider the variance inflation factors (VIF) and the condition numbers of our models. The VIF analysis shows a maximum value of 1.45 and an average value of 1.20 for all variables, well below the threshold value of 10 (Mason & Perreault, 1991). Moreover, the condition numbers remain below the threshold value of 30 at 9.95 (Mason & Perrault, 1991). Consequently, we are confident that multicollinearity is not an issue in our models.

**Table 2.1. Descriptive statistics and correlation matrix**

Variables	1	2	3	4	5	6	7	8	9	10
1 Technological value	1									
2 Recombinant lag	-0.09	1								
3 Frequency of reuse	0.12	-0.20	1							
4 Component age	0.02	0.45	0.20	1						
5 Number of components	0.30	-0.13	0.22	0.08	1					
6 Internal components	-0.06	-0.07	-0.04	-0.16	-0.06	1				
7 Technological breadth	0.10	-0.05	0.12	0.04	0.22	-0.05	1			
8 Team size	0.07	-0.01	0.01	-0.01	0.06	-0.02	0.03	1		
9 Patent offices	0.30	-0.11	0.15	0.04	0.37	-0.03	0.14	0.04	1	
10 Patent granted	0.19	-0.06	0.09	0.04	0.22	0.02	0.09	0.03	0.26	1
Mean	2.48	1.72	6.96	5.42	9.05	0.16	0.65	2.99	2.41	0.72
SD	4.23	1.77	8.08	3.94	9.53	0.22	0.36	1.96	2.01	0.45
Min	0	0	0	0	1	0	0	1	1	0
Max	85	58	327.88	58	179	1	1	22	19	1

**Regression Results.** Table 2.2 presents the results of the negative binomial regressions. Model 1 is the baseline model which only includes the control variables. Overall, the control variables have the expected signs and have a statistically significant effect on the technological value of inventions. In line with prior research (e.g. Nerkar, 2003; Schoenmakers & Duysters, 2010), we find that the age of recombined components has the expected negative and statistically significant effect on the technological value of inventions (Model 1:  $\beta_{\text{Component age}} = -0.021$ ,  $p < 0.001$ ). The number of recombined components has a positive and statistically significant effect on the technological value of inventions (Model 1:  $\beta_{\text{Number of components}} = 0.017$ ,  $p < 0.001$ ), suggesting that inventions that recombine many different components are more technologically valuable (Kelley *et al.*, 2013). The invention's reliance on internally-generated components has a negative and statistically significant effect on the technological value of the invention (Model 1:  $\beta_{\text{Internal components}} = -0.340$   $p < 0.001$ ), indicating that strong reliance on internal components may inhibit the ability of others to build upon the newly-created invention (Kim *et al.*, 2012). The results also suggest that fuel cell inventions benefit from relying on technologically broad components (e.g. Kelley *et al.*, 2013), as indicated by the positive and statistically significant effect of technological breadth on the technological value of inventions (Model 1:  $\beta_{\text{Technological breadth}} = 0.146$ ,  $p < 0.001$ ).

The size of the team that contributed to the invention has a positive and statistically significant effect on the technological value of inventions (Model 1:  $\beta_{\text{Team size}} = 0.026$ ,  $p < 0.001$ ), supporting the notion that larger teams of inventors may be better able to resolve technological problems (Singh & Fleming, 2010). The number of unique patent authorities in which the patent was filed has a positive and statistically significant effect on the technological value of inventions (Model 1:  $\beta_{\text{Patent offices}} = 0.113$ ,  $p < 0.001$ ), providing evidence that a broader scope of patent protection may be indicative of the quality of an invention (Harhoff *et al.*, 2003). Finally, inventions which meet patent examiners' patentability evaluation tend to be more technologically valuable (Model 1:  $\beta_{\text{Patent granted}} = 0.220$ ,  $p < 0.001$ ).

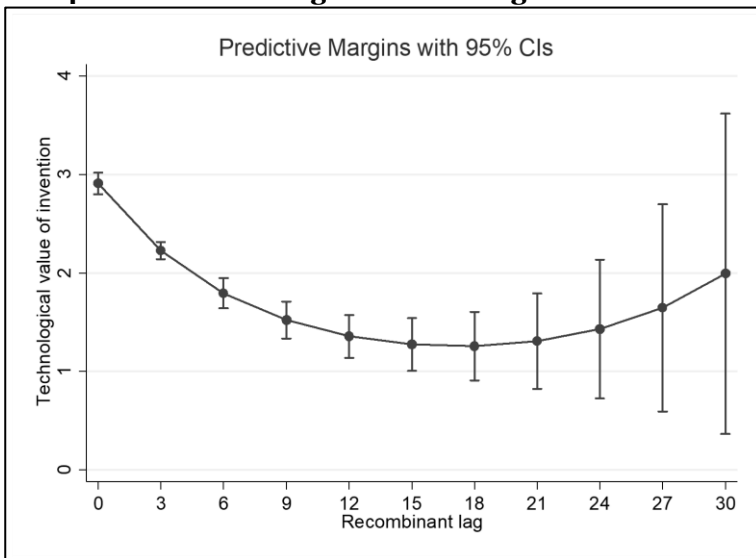
In model 3 we test Hypothesis 1. We find a negative and statistically significant effect of recombinant lag on the technological value of inventions (Model 3:  $\beta_{\text{Recombinant lag}} = -0.097$ ,  $p < 0.001$ ) and a positive and statistically

## Chapter 2

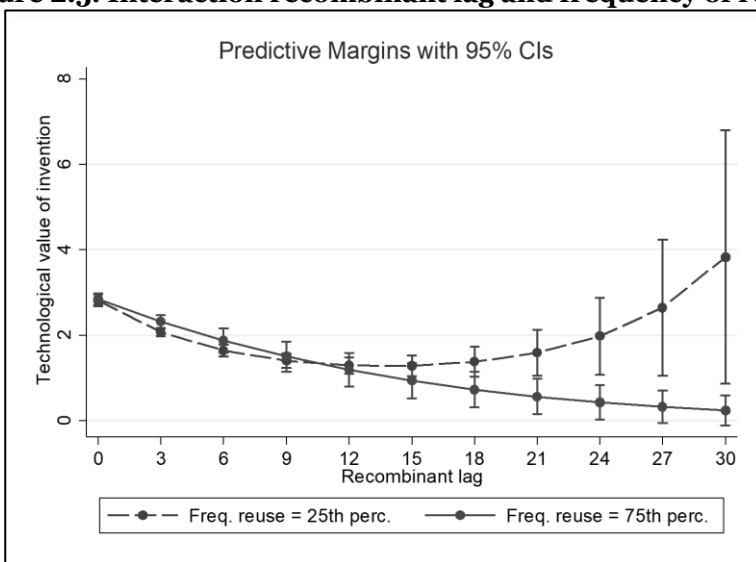
significant quadratic effect (Model 3:  $\beta_{\text{Recombinant lag squared}} = 0.003, p < 0.001$ ), indicating the existence of a non-linear relationship between recombinant lag and the technological value of inventions. We execute several tests to examine whether this is the relationship that we hypothesized (i.e. negative and with diminishing marginal effects) (Haans, Pieters, & He, 2016; Karim, 2009; Lind & Mehlum, 2010). We find that: (i) the linear coefficient is negative and statistically significant and the quadratic coefficient is positive and statistically significant, (ii) the 95 percent Fieller confidence interval of the inflection point is within the range of observable points ([13.19, 28.00]), (iii) the slope before the inflection point is negative and statistically significant at the minimum value of recombinant lag ( $p < 0.001$ ) and the slope after the inflection point is positive and statistically significant at the maximum value of recombinant lag ( $p < 0.001$ ), and (iv) the linear and quadratic coefficients of recombinant lag are jointly statistically significant ( $\text{Chi}^2 = 71.50, p < 0.001$ ). This means that, instead of the predicted negative relationship with diminishing marginal effects, we actually find a U-shaped relationship between recombinant lag and the technological value of inventions. Figure 2.4 plots this relationship and shows that the inflection point occurs at a value of recombinant lag of 17.2, implying an inflection point at relatively high levels of recombinant lag. Thus, although negative value implications of recombinant lag are clearly present for the initial range of values of recombinant lag, we do not find full support for Hypothesis 1.

In model 5 we test Hypothesis 2. We find a statistically significant interaction between the frequency of reuse of components and recombinant lag on the technological value of inventions (Model 5:  $\beta_{\text{Recombinant lag} \times \text{Frequency of reuse}} = 0.007, p < 0.05$ ; Model 5:  $\beta_{\text{Recombinant lag squared} \times \text{Frequency of reuse}} = -0.001, p < 0.01$ ). In Figure 2.5, we show that the upward slope of the U-shaped relationship between recombinant lag and technological value mainly emerges when frequency of reuse is low. Below, we first discuss our robustness checks. Subsequently, we present additional analyses and data to explain this unexpected U-shaped relationship.

**Figure 2.4. Recombinant lag and technological value of invention**



**Figure 2.5. Interaction recombinant lag and frequency of reuse**



**Table 2.2. Negative binomial regression results**

DV: Technological value of invention	1	2	3	4	5
Component age	-0.02*** [0.00]	-0.01*** [0.00]	-0.01** [0.00]	-0.01** [0.00]	-0.01*** [0.00]
Number of components	0.02*** [0.00]	0.02*** [0.00]	0.02*** [0.00]	0.02*** [0.00]	0.02*** [0.00]
Internal components	-0.34*** [0.05]	-0.35*** [0.05]	-0.35*** [0.05]	-0.35*** [0.05]	-0.35*** [0.05]
Technological breadth	0.15*** [0.03]	0.15*** [0.03]	0.15*** [0.03]	0.15*** [0.03]	0.15*** [0.03]
Team size	0.03*** [0.00]	0.03*** [0.00]	0.03*** [0.00]	0.03*** [0.00]	0.03*** [0.00]
Patent offices	0.11*** [0.00]	0.11*** [0.00]	0.11*** [0.00]	0.11*** [0.00]	0.11*** [0.00]
Patent grant	0.22*** [0.02]	0.22*** [0.02]	0.21*** [0.02]	0.21*** [0.02]	0.21*** [0.02]
Frequency of reuse	0.01*** [0.00]	0.01*** [0.00]	0.01*** [0.00]	0.01*** [0.00]	0.00 [0.00]
Recombinant lag		-0.06*** [0.01]	-0.10*** [0.01]	-0.10*** [0.01]	-0.13*** [0.02]
Recombinant lag squared			0.00*** [0.00]	0.00*** [0.00]	0.01*** [0.00]
Recombinant lag × Frequency of reuse				-0.00 [0.00]	0.01* [0.00]
Recombinant lag squared × Frequency of reuse					-0.00** [0.00]
Firm dummies	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes
Observations	21117	21117	21117	21117	21117
Pseudo R <sup>2</sup>	0.118	0.119	0.119	0.119	0.119
AIC	76248.10	76175.56	76147.40	76149.40	76140.12
BIC	77807.84	77743.26	77723.05	77733.00	77731.68
Log Likelihood	-37928.05	-37890.78	-37875.70	-37875.70	-37870.06
Wald chi <sup>2</sup>	11516.31***	11589.89***	11632.16***	11634.81***	11646.36***

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

**Robustness checks.** To assess the robustness of our findings, we run several additional model specifications (see Table 2.3). First, we test whether component age also has a non-linear relationship with the technological value of inventions (model 6). We find no statistical evidence of a non-linear relationship between the age of recombined components and the technological value of inventions. In contrast to recombinant lag, age appears to have a strictly negative linear relationship with the technological value of inventions, which is in line with the prior work of Nerkar (2003). Second, in models 7 and 8, we exclude patent families with a single backward citation (representing 4.58 percent of the sample), since these may not reflect knowledge recombination processes (i.e. only one component is used to build a new invention). The main results remain unchanged. Third, in models 9 and 10, we exclude all patent families created before 1990, since fuel cell technological development principally took off after this year (Perry & Fuller, 2002; Sharaf & Orhan, 2014). Results remain largely unaffected.

**Table 2.3. Robustness checks**

DV: Technological value of invention	6	7	8	9	10	11	12	13	14	15
Component age	-0.01 <sup>†</sup> [0.01]	-0.01 <sup>***</sup> [0.00]	-0.01 <sup>***</sup> [0.00]	-0.01 <sup>***</sup> [0.00]	-0.01 <sup>***</sup> [0.00]	-0.01 <sup>†</sup> [0.00]	-0.01 <sup>**</sup> [0.00]	-0.01 <sup>***</sup> [0.00]	-0.01 <sup>**</sup> [0.00]	-0.01 <sup>**</sup> [0.00]
Number of components	0.02 <sup>***</sup> [0.00]	0.02 <sup>***</sup> [0.00]	0.02 <sup>***</sup> [0.00]	0.02 <sup>***</sup> [0.00]	0.02 <sup>***</sup> [0.00]	0.01 <sup>***</sup> [0.00]	0.01 <sup>***</sup> [0.00]	0.02 <sup>***</sup> [0.00]	0.02 <sup>***</sup> [0.00]	0.01 <sup>***</sup> [0.00]
Internal components	-0.35 <sup>***</sup> [0.05]	-0.35 <sup>***</sup> [0.05]	-0.35 <sup>***</sup> [0.05]	-0.37 <sup>***</sup> [0.05]	-0.37 <sup>***</sup> [0.05]	-0.26 <sup>***</sup> [0.04]	-0.25 <sup>***</sup> [0.04]			
Technological breadth	0.15 <sup>***</sup> [0.03]	0.14 <sup>***</sup> [0.03]	0.14 <sup>***</sup> [0.03]	0.12 <sup>***</sup> [0.03]	0.12 <sup>***</sup> [0.03]	0.13 <sup>***</sup> [0.03]	0.13 <sup>***</sup> [0.03]	0.25 <sup>**</sup> [0.08]	0.25 <sup>***</sup> [0.08]	0.27 <sup>***</sup> [0.08]
Team size	0.03 <sup>***</sup> [0.00]	0.03 <sup>***</sup> [0.00]	0.03 <sup>***</sup> [0.00]	0.03 <sup>***</sup> [0.00]	0.03 <sup>***</sup> [0.00]	0.02 <sup>***</sup> [0.00]	0.02 <sup>***</sup> [0.00]	0.04 <sup>**</sup> [0.01]	0.03 <sup>*</sup> [0.01]	0.03 <sup>*</sup> [0.01]
Patent offices	0.11 <sup>***</sup> [0.00]	0.11 <sup>***</sup> [0.00]	0.11 <sup>***</sup> [0.00]	0.11 <sup>***</sup> [0.01]	0.11 <sup>***</sup> [0.01]	0.10 <sup>***</sup> [0.00]	0.10 <sup>***</sup> [0.00]	0.05 <sup>***</sup> [0.01]	0.05 <sup>***</sup> [0.01]	0.05 <sup>***</sup> [0.01]
Patent grant	0.22 <sup>***</sup> [0.02]	0.20 <sup>***</sup> [0.02]	0.20 <sup>***</sup> [0.02]	0.16 <sup>***</sup> [0.03]	0.16 <sup>***</sup> [0.03]	0.18 <sup>***</sup> [0.02]	0.18 <sup>***</sup> [0.02]	0.27 <sup>***</sup> [0.05]	0.27 <sup>***</sup> [0.05]	0.25 <sup>***</sup> [0.05]
Frequency of reuse	0.01 <sup>***</sup> [0.00]	0.01 <sup>***</sup> [0.00]	-0.00 [0.00]	0.01 <sup>***</sup> [0.00]	0.00 [0.00]	0.01 <sup>***</sup> [0.00]	-0.00 [0.00]	0.04 <sup>***</sup> [0.00]	0.05 <sup>***</sup> [0.01]	0.04 <sup>***</sup> [0.01]
Recombinant lag	-0.06 <sup>***</sup> [0.01]	-0.10 <sup>***</sup> [0.01]	-0.15 <sup>***</sup> [0.02]	-0.10 <sup>***</sup> [0.02]	-0.12 <sup>***</sup> [0.02]	-0.09 <sup>***</sup> [0.01]	-0.13 <sup>***</sup> [0.02]	-0.07 <sup>***</sup> [0.01]	-0.07 <sup>***</sup> [0.02]	-0.11 <sup>***</sup> [0.02]
Recombinant lag squared		0.00 <sup>***</sup> [0.00]	0.01 <sup>***</sup> [0.00]	0.00 <sup>**</sup> [0.00]	0.01 <sup>***</sup> [0.00]	0.00 <sup>***</sup> [0.00]	0.01 <sup>***</sup> [0.00]	0.00 <sup>***</sup> [0.00]	0.00 <sup>**</sup> [0.00]	0.00 <sup>**</sup> [0.00]
Recombinant lag × Frequency of reuse			0.01 <sup>**</sup> [0.00]		0.01 <sup>*</sup> [0.00]		0.01 <sup>***</sup> [0.00]		-0.01 <sup>*</sup> [0.00]	-0.00 [0.01]
Recombinant lag squared × Frequency of reuse			-0.00 <sup>***</sup> [0.00]		-0.00 <sup>*</sup> [0.00]		-0.00 <sup>***</sup> [0.00]		0.00 <sup>**</sup> [0.00]	0.00 [0.00]
Component age squared	0.00 [0.00]									
Firm dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21117	20150	20150	19164	19164	20119	20119	3674	3674	3554
Pseudo R <sup>2</sup>	0.119	0.118	0.119	0.119	0.119	0.097	0.098	0.073	0.073	0.073
AIC	76176.64	73696.49	73686.45	70662.48	70660.60	65969.82	65957.05	18213.05	18209.22	17687.52
BIC	77752.29	75262.86	75268.64	71967.37	71981.22	67535.89	67538.93	18523.50	18532.09	17934.55
Log Likelihood	-37890.32	-36650.24	-36643.22	-35165.24	-35162.30	-32786.91	-32778.52	-9056.53	-9052.61	-8803.76
Wald chi <sup>2</sup>	11594.05 <sup>***</sup>	11124.52 <sup>***</sup>	11142.48 <sup>***</sup>	10631.71 <sup>***</sup>	10638.41 <sup>***</sup>	8683.21 <sup>***</sup>	8714.80 <sup>***</sup>	2963.14 <sup>***</sup>	2984.11 <sup>***</sup>	1462.46 <sup>***</sup>

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



## Chapter 2

Fourth, in models 11 and 12, we exclude inventions with very high technological value (i.e. above 95<sup>th</sup> percentile) which corresponds to patents that receive more than 10 external forward citations within the fixed four-year window. Results remain highly stable.

We also execute several additional analyses, which we do not report in Table 2.3 for the sake of brevity, but which are available upon request: (i) we recalculate recombinant lag by taking the mean value of recombinant lag of the backward citations of the patent family, (ii) we run the analyses excluding component age as a control variable, (iii) we increase the fixed window of external forward citations to 5 years, and (iv) we rerun the analyses, excluding patents created before 1980. In all four cases, the main results remain highly stable.

### **2.4.6. Post-hoc exploratory data analysis**

Whereas we hypothesized a negative relationship with diminishing returns between recombinant lag and the technological value of inventions, we actually detected a robust U-shaped relationship. To make sense of this unexpected finding, we performed four steps. First, we reexamined our data, trying to identify inventions that drove this unexpected relationship. Subsequently, we delved into fuel cell technology literature and conducted an interview with a fuel cell technology expert to better understand why some inventions contributed to this unexpected relationship. Third, we conducted additional tests on a sample of inventions in the wind industry to explore the generalizability of our unexpected findings. Finally, we connected the additional information that emerged out of these analyses to existing knowledge recombination literature.

**Data examination.** As Figure 2.4 illustrates, the inflection point of the U-shaped curve is situated at relatively high levels of recombinant lag. In an attempt to understand what drives this U-shaped curve, we screened inventions beyond the inflection point of the curve, representing a small group of 49 inventions with a recombinant lag of at least 17 years. Screening these inventions, we observed that, in accordance with our hypothesis, 26 of them had received zero forward citations within the first four years of creation. However, we also identified a group of 23 inventions for which, following our hypothesis, we had expected very limited technological value, but which actually had considerable technological value (i.e.

on average these inventions received 1.83 external forward citations within the first four years of creation). These data suggest that the upward slope of our unexpected U-shaped relationship was driven by a limited number of observations with high levels of recombinant lag and higher-than-expected technological value.

**Sense making.** To understand why, in some exceptional cases, high recombinant lag is associated with considerable technological value, we reexamined the context of our study. In particular, we inspected issues of leading fuel cell technology journals (e.g. *Fuel Cells Bulletin*), and read several review articles on fuel cell technology (e.g. Sharaf & Orhan, 2014; Steele & Heinzl, 2001). In addition, we arranged an interview with a fuel cell expert who has published extensively in fuel cell-oriented journals and was responsible for coordinating several large Dutch- and European-level fuel cell programs.

Based on these data sources, we found strong indications that the cyclical nature of technology development can explain why some components with extreme recombinant lag are associated to inventions with considerable technological value. In the fuel cell literature, it is emphasized that, in line with other industries, fuel cell technology has experienced several cycles of technological development, where periods of revived interest and intense technological development are followed by periods of relative technological stability (Tushman & Anderson, 1986). These cycles of technological development are principally triggered by the emergence of new application fields for the technology, such as consumer electronic products for fuel cells in the early 2000's (Sharaf & Orhan, 2014). These new application domains often emerge following important technological breakthroughs in the primary technology, as similarly argued by Tushman and Anderson (1986). According to the interviewed fuel cell expert, one of the most notable technological cycles in fuel cell technology began in the 1960's when NASA placed fuel cells on board of their spacecrafts in order to generate electricity and provide drinking water. During this decade, polymer-electrolyte fuel cells (PEFC), developed by General Electric, competed against alkaline fuel cells (AFC), created by Pratt & Whitney. Due to the comparatively lower energy efficiency of PEFCs, NASA ended up selecting AFCs for most space missions, effectively putting PEFC development on the back burner. It was not until the early 1990's, following important improvements that increased the energy efficiency of PEFCs and reduced platinum

loading requirements for the catalyst (Prater, 1990), that interest in this technology was revived (Sharaf & Orhan, 2014). Following these technological improvements, automotive manufacturers recognized the potential of PEFC technology for the propulsion of automotive vehicles.

***Additional analyses in wind energy industry.*** We conducted an additional test to examine whether the U-shaped relationship between recombinant lag and technological value is generalizable to other industries. In particular, we collected additional data from the wind energy industry. This industry is an interesting setting for checking the generalizability of our unexpected findings. On the one hand, the fuel cell and wind energy industry are similar, as firms in these two industries principally focus their technological activities on improving the cost-efficiency of the technology (i.e. kWh/\$ rates) (Blanco, 2009; Sharaf & Orhan, 2014). On the other hand, an important difference between the two industries is that they experienced very different technology cycles in terms of duration, frequency, and intensity (Kaldellis & Zafirakis, 2011; Perry & Fuller, 2002).

The wind energy patent families were retrieved using IPC code F03D (titled ‘Wind motors’) (Popp, Hascic, & Medhi, 2011). To ensure comparability, we examine the same time period as the fuel cell industry analysis (1959-2007). This produced a sample of 3,674 patent families<sup>8</sup>. For this wind industry sample, we also find the unexpected U-shaped relationship between recombinant lag and the technological value of inventions (model 13 in Table 2.3)<sup>9</sup>. Notably, the inflection point of this relationship is situated at a higher level in the wind energy sample (recombinant lag of 26.6 years compared to 17.2 years in the fuel cell sample). Moreover, in the wind energy sample, we find a negative, rather than positive, interaction effect between frequency of reuse and recombinant lag (model 14). This interaction effect, however, was not robust to several model specifications. For

---

<sup>8</sup> Note that this data is unconsolidated, meaning that no distinction is made between internal and external citations.

<sup>9</sup> In model 13 of Table 2.3, The linear coefficient is negative and statistically significant ( $\beta_{\text{Recombinant lag}} = -0.075, p < 0.001$ ) and the quadratic coefficient is positive and statistically significant ( $\beta_{\text{Recombinant lag squared}} = 0.001, p < 0.001$ ), the 95 percent Fieller confidence interval of the inflection point is within the range of observable points ([20.34, 44.78]), the left part of the slope is negative and statistically significant at the minimum value of recombinant lag ( $p < 0.001$ ) and the right part of the slope is positive and statistically significant at the maximum value of recombinant lag ( $p < 0.01$ ), and the linear and quadratic coefficients of recombinant lag are jointly statistically significant ( $\text{Chi}^2 = 34.57, p < 0.001$ ).

instance, in model 15, we show the results for the sample in which patents created before 1980 are excluded, indicating a statistically non-significant interaction effect between frequency of reuse and recombinant lag.

**Connection to literature.** Based on our review of the fuel cell literature, interview with the fuel cell expert, and additional tests in the wind energy industry, we found strong indications to suggest that inventions with extremely high recombinant lag might rely upon components that have remained unused since a previous technological cycle. This suggests that components that have remained unused since a previous cycle may only become valuable when a new technology cycle emerges. Because they remain unused or ‘dormant’ for such extensive periods, we subsequently refer to these components as dormant components. The fact that components that have remained unused for a long time may still be valuable, is comparable to the concept of ‘shelved knowledge’ advanced by Garud and Nayyar (1994). They argue that, due to time lags in technological and market developments, some knowledge should be shelved, maintained, and then reactivated at a later point in time when complementary resources have emerged. Thus, they propose that some knowledge pieces remain unused for prolonged periods, not because they contain less value, but rather due to missing complementary resources or because they emerged ahead of their time (Garud & Nayyar, 1994). Applying these insights, we hence find strong indications that the upward slope of the U-shaped relationship between recombinant lag and the technological value of inventions can be explained by the existence of dormant components – i.e. valuable components that have remained unused for prolonged periods (Garud & Nayyar, 1994).

## **2.5. Discussion and conclusion**

This study explores the relation between recombinant lag – i.e. the time that components in knowledge recombination have remained unused – and the technological value of inventions. Whereas existing studies on knowledge reuse trajectories have mostly focused on the frequency of reuse of components, this study shows that the technological value of inventions is also substantially driven by the recency of component reuse. The core finding of this study is the U-shaped

## Chapter 2

relationship between recombinant lag and technological value, which we identified in two different industries. In this section, we first discuss the implications of our findings for the knowledge recombination literature. Subsequently, we discuss the practical implications of our findings. Finally, we discuss the core limitations of our study and suggest interesting avenues for future research.

### **2.5.1. Implications for knowledge recombination literature**

***Recency and frequency of reuse.*** In this study, we deviate from the majority of knowledge recombination research by shifting attention from the original attributes of knowledge components to how they are actually reused over time. We argue that knowledge components should not be considered as pieces of knowledge with a value that is fully determined at the origin. Instead, we follow an emerging stream of literature on reuse trajectories (Fleming, 2001; Katila & Chen, 2008; Yang *et al.*, 2010), emphasizing that components experience a history of reuse, which influences their recombinant value over time. At the same time, we contribute to this latter literature stream, illuminating that, next to the frequency of reuse, it is also important to consider the recency of reuse. Theoretically, we apply insights from organizational learning theory and emphasize that recency of reuse entails the generation of reuse information flows that are embedded in the state-of-the-art of technology, reflecting a rejuvenation effect. This is clearly different from frequency of reuse, which mostly captures the magnitude of available reuse information flows, and thus the amount of available learning opportunities. In our empirical analysis, we indeed find strong evidence of this rejuvenation effect, showing that the most recent instances of reuse yield inventions with the highest technological value.

At the same time, we unexpectedly observe that recombining components with extremely high recombinant lag may lead to considerable technological value. After conducting several additional analyses into this unexpected relationship, we found strong suggestions that these components reflect dormant component knowledge – i.e. valuable components that have remained inactive for prolonged periods. Thus, according to our findings, when recombinant lag is low, associated reuse information flows are valuable because they are embedded in the state-of-the-art of technology. But, according to our findings, value can also be associated

with high recombinant lag and, thus, a last instance of reuse that occurred a long time ago. This implies that reuse information flows associated with a temporally distant last instance of reuse may contain valuable information about how to apply a component in recombination. Based on the findings of our post-hoc exploratory data analyses, a potential explanation might be that, when generated, these reuse information flows represented information about the component's applications in recombination that was too far ahead of its time (Garud & Nayyar, 1994). Given inventors' inability to leverage these reuse information flows when they were generated, these particular reuse information flows were 'frozen' for the time being. Years later, during the emergence of a new technology cycle, inventors could 'defrost', interpret, and exploit these reuse information flows, leading to a value-enhancing recombination of the component.

To summarize: our findings make an important contribution to knowledge reuse trajectories literature, illuminating the importance of component rejuvenation or the generation of reuse information flows that are embedded within the state-of-the-art of technology. At the same time, we highlight that, when reuse information flows were generated a long time ago, they can represent information that was simply too far ahead of its time, implying opportunities for value-adding recombination when new technological cycles have emerged. Jointly, the findings provide a better understanding of how time and reuse information flows shape learning opportunities for knowledge recombination.

***Recombinant lag and component age.*** Emphasizing the temporal dimension of knowledge recombination, scholars have paid a lot of attention to component age (e.g. Katila, 2002; Nerkar, 2003). These studies tend to assume, explicitly or implicitly, that components inevitably become more widely-reused as they get older (e.g. Ahuja & Lampert, 2001; Heeley & Jacobs, 2008; Kelley *et al.*, 2013). As a result, they tend to ignore that there is actually much variation in terms of when and how frequently components are reused (Capaldo *et al.*, 2017). We contribute to our understanding of the temporal dimension of knowledge recombination by showing that component age and recombinant lag are distinct concepts, which influence components' recombinant value in different ways. From a theoretical point of view, recombinant lag implies a theoretical mechanism – i.e. the rejuvenation of recombinant potential of a component – that the age of a

## Chapter 2

component does not capture. From an empirical perspective, our results consistently show that, controlling for age, recombinant lag significantly influences the recombinant value of components. In sum, our results indicate that, despite the importance of component age in driving knowledge recombination value (as our empirical results confirm), it does not fully capture the impact of the temporal dimension of knowledge components on their recombinant value. Instead, we identify recombinant lag as an additional temporal dimension of knowledge recombination, emphasizing that it is not only important to account for when a component was created, but also to examine when a component has been reused.

### **2.5.2. Implications for practitioners**

Our findings carry important implications for practitioners. First, they imply that important learning opportunities become available when components were reused recently in knowledge recombination. When reuse occurs recently, the value of a component may be significantly enhanced because of access to additional up-to-date learning opportunities. For policy makers, these findings imply additional proof for the importance of transparency and information disclosure in new technology production. By allowing for information about new inventions to disseminate more quickly and accurately to other inventors, subsequent production of new inventions will be more valuable.

Second, our findings unexpectedly show that, beyond a certain tipping point, higher levels of recombinant lag can be positively associated with the technological value of inventions. Through several post-hoc exploratory data analyses, we found strong indications to suggest that this relationship was driven by dormant components – i.e. valuable components that have remained unused for prolonged periods. Thus, our findings suggest that the existing knowledge stock should be maintained and closely monitored over time. Concomitantly, the existing knowledge stock should be continually reevaluated – i.e. dusting off the shelves – in order to detect these potentially valuable knowledge components and their reuse information flows.

### **2.5.3. Limitations and future research**

***Generalizability and industry-specific conditions.*** Although we made substantial efforts in exploring the generalizability of our finding by

collecting data on an additional industry, we acknowledge that there are limitations to the generalizability of our findings. Comparing the results of the two samples (i.e. fuel cell sample and wind energy sample), we found a robust U-shaped relationship between recombinant lag and technological value. At the same time, we also observed that the exact nature of this relationship and the conditions under which it manifests most strongly differed across the two industries. We found that the inflection point of this U-shaped relationship was situated at a higher value of recombinant lag for the wind energy sample than the fuel cell sample. Moreover, whereas we found a robust interaction effect of frequency of reuse for the fuel cell sample, such a robust interaction effect remained absent in the wind energy sample. These interesting differences point to the need to further study the impact of industry-specific characteristics on the impact of reuse trajectories on technological value of inventions. We highlight two industry-specific conditions that could serve as an interesting starting point in this respect.

First, to explain the emergence of dormant components, we pointed to the nature of technology cycles in an industry. As prior research indicates, although most industries face various technology cycles during their evolution (Tushman & Anderson, 1986), the nature of these cycles, in terms of length and duration, differs across industries. This might explain why, in the wind energy industry, the inflection point beyond which recombinant lag positively influences technological value is situated at a higher level than in fuel cells. Second, when technological change in an industry occurs in a more punctuated, rather than incremental, manner (Tushman & Anderson, 1986), the value of recombinant lag might differ. For example, when technological change is punctuated, the recombinant value of a relatively less recent instance of reuse will be very minimal, as the reuse information flows contained therein will largely misrepresent the current state-of-the-art of technology. In contrast, in the absence of such technological punctuation, the state-of-the-art of technology is likely to be more stable, implying that rejuvenation effects might also emerge for components with relatively high recombinant lag. Jointly, we therefore encourage future studies to further explore the impact of and interaction between recombinant lag and frequency of reuse in a wide variety of different technological settings, which have different characteristics in terms of technological cycles and the absence/presence of technological shocks.



***Original component attributes and component reuse.*** In this study, we moved away attention from the original attributes of knowledge components to their reuse trajectories. In this way, we were able to demonstrate that, next to component age, also the recency of reuse is an important temporal dimension that influences the value of resulting inventions. A next step to expand our understanding of knowledge recombination could be to examine the interaction between original component attributes and knowledge reuse trajectories. Future studies could examine, for instance, whether certain components have original attributes that impede the disentanglement of combinations in which these components are integrated. Components with certain technological characteristics embedded in them at creation may, for example, be more difficult to physically detach from a combination in which they are reused. In fuel cells, for instance, the recombination of a particular fuel mixture, as opposed to a physical piece of hardware such as a seal, may result in combinations which are more difficult to take apart (Fleming & Sorenson, 2001), impeding the creation of useful reuse information flows. In other words, original attributes of the component might restrict the learning experience of inventors when such a component is reused. We therefore encourage future research to delve deeper into the interplay between original attributes of knowledge components and their reuse over time.

***Antecedents of recombining dormant components.*** Using patent data, we were able to examine the value of components on a large scale, using methods and measures validated by previous studies. However, these data did not allow fine-grained analyses of how and why particular components are utilized in knowledge recombination. Nevertheless, research on the antecedents of knowledge recombination is much-needed (Carnabuci & Operti, 2013). Therefore, we particularly urge future studies to conduct in-depth qualitative studies on the recombination and primary characteristics of valuable dormant components. These studies would carry important managerial implications, as they would help firms identify more precisely which components should be reconsidered for knowledge recombination.