Technological Uncertainty and Firm Boundaries: The Moderating Effect of Knowledge Modularity

Sebastian K. Fixson, Davit Khachatryan, and Wonhee Lee

***Abstract*—One fundamental challenge of technology and innovation management is for firms to decide which future technologies to develop in-house versus buying them outside. This challenge is particularly pertinent when industries emerge, typically a time of high levels of technological uncertainty. It has been long understood that technological uncertainty functions as important stimulus for firms to manage their boundaries. However, two to some extent competing strategy perspectives—governance and competence—predict that firms faced with uncertainty would either increase or decrease their scope of activities, respectively. To reconcile these conflicting positions, we propose a model in which knowledge modularity moderates the effect of technological uncertainty on firms' research and development (R&D) scope decisions. We develop new measures for R&D scope, knowledge modularity, and technological uncertainty, drawing on population ecology, network theory, and technology management. We test our model empirically using data on patenting activity and firm boundary location in the emerging automotive air bag industry. Our results generally support our model and show that in case of knowledge-generating activities such as R&D, scope decisions under technological uncertainty are more driven by concerns about the risk of obsolescence than the risk of opportunistic behavior. We discuss implications for managerial practice and future research.**

***Index Terms*—Complex technological products, firm boundary, industry emergence, innovation, modularity, technological uncertainty, technology management.**

I. Introduction

THE SCHOLARLY investigation of how firms respond to uncertainty in their environment stretches back for at least half a century [1], [2]. One aspect of major interest in this literature stream on the effects of uncertainty on firm behavior is the decision on where to locate the firm boundary, i.e., the decision on which activities to maintain inside the firm and which activities to access outside of the firm through market mechanisms. While the majority of this work has focused in the past on the world of production, the recently growing relevance of knowledge work has raised this question also in the arena of knowledge-creating activities such as research and development (R&D).

The transfer of the existing arguments from production to R&D has led to some conflicting explanations from two separate viewpoints [3]. One argument suggests that the appropriate response to an increase in technological uncertainty is to expand the firm's boundary to integrate more activities in-house to reduce the risk of opportunism [4], [5], i.e., the chance that one or more of the transacting parties "disguises attributes or preferences, distort data, obfuscate issues, and otherwise confuse transactions" [6, p. 554] to take advantage of an unforeseen situation. However, another argument suggests the opposite, i.e., that firms facing high levels of technological uncertainty decrease the range of their internal R&D activities to avoid the risk of obsolescence, the risk of a technology being made inferior relative to other technologies by technological change [7]–[9].

We propose to explain these conflicting arguments through the underlying—often only implicit—assumption on the structure of the relevant knowledge involved. More specifically, building on modularity theory, we develop a model in which the property *knowledge modularity* moderates the effect of exogenous technological uncertainty on firms' R&D scope decisions, i.e., the range of their knowledge development work. The hypotheses suggested in this paper conjecture that in regimes of low knowledge modularity, the risk of obsolescence concerns outweigh opportunism reduction considerations. Conversely, in high knowledge modularity regimes, the hypotheses propose that opportunism reduction considerations dominate the risk of obsolescence concerns. We empirically test those hypotheses with longitudinal data on patenting activity and firm boundary location from the emerging automotive air bag industry and find that the result from the data analyses generally support the conjectured moderating effect of knowledge modularity.

The paper is organized as follows. In the next section, we lay out the theoretical background, propose our model, and develop hypotheses. In Section III, we introduce the data and develop a set of new measures for R&D scope, technological uncertainty, and knowledge modularity, drawing on population ecology and network theory in addition to technology and innovation management. The discussion of our empirical results follows in Section IV, and Section V concludes the paper.

Manuscript received December 29, 2015; revised August 23, 2016 and November 14, 2016; accepted November 29, 2016. Date of publication January 9, 2017; date of current version January 18, 2017. Review of this manuscript as arranged by Department Editor P. E. D. Love. The authors acknowledge financial support for this research from the National Science Foundation (SES-0620487).

S. K. Fixson is with the Division of Technology, Operations, and Information Management, Babson College, Babson Park, MA 02457-0310 USA (e-mail: sfixson@babson.edu).

D. Khachatryan is with the Division of Mathematics and Science, Babson College, Babson Park, MA 02457-0310 USA (e-mail: dkhachatryan@babson.edu).

W. Lee is with the Samsung Economic Research Institute, Samsung Life Seocho Tower, Seocho-daero 74-gil 4, Seocho-gu, Seoul (e-mail: lee.woni@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TEM.2016.2638847

II. Background and Hypotheses

A. Technological Uncertainty and R&D Scope

Various scholars have identified uncertainty as an important factor explaining firms' decisions on R&D scope. Uncertainty has been identified as a key variable affecting organizational behavior [1], and coping with this uncertainty becomes a primary task for firms [2]. Uncertainty has also been described as one of several critical dimensions deciding the costs of transactions [6]. Note that among different types of uncertainty (e.g., demand uncertainty, customer preference uncertainty, supply uncertainty, fuzziness, etc.) [10], our focus here is on technological uncertainty as it is particularly relevant in technology-intensive industries for decisions on R&D scope. Moreover, we consider technological uncertainty as a type of primary uncertainty [11] which reflects the uncertainty arising from exogenous sources such as natural events, regulatory changes, or industry-level technological changes. Therefore, technological uncertainty influences an industry as a whole and cannot be controlled by any one firm [11], [12].

While there is agreement that uncertainty matters in determining effective R&D firm boundary locations, different conclusions seem to be drawn as to *how* it matters. One stream of studies, the governance perspective, building on the transaction cost economics (TCE) tradition [3], focuses on the efficiency of transactions in explaining firms' decisions on their R&D scope. In particular, uncertainty raises the risk of opportunism in market relationships, thereby raising their costs, and makes the fact that separate companies do not "speak the same dialect" more problematic [6], [13]. Uncertainty also increases the risk of failure of a technological development, and in the event of failure, it is often difficult to determine who is at fault among contracting parties [14]. Seen from the TCE perspective, therefore, firms expand their R&D scope as technological uncertainty increases to lower the risk of being exposed to opportunism.

Another stream of studies views the problem from a different angle. The resource-based view (RBV), later expanded into knowledge and, more generally, capabilities, and here labeled as the competence perspective, focuses on the rare, valuable, and inimitable aspect of firms that lead to interfirm performance variation.[1] Some studies in the RBV tradition have emphasized the competitive importance for firms to create resource bundles that can produce products customers want and are difficult to imitate by competitors. To achieve this idiosyncrasy of its resource bundle, a firm tends to deepen its expertise in a limited number of areas [16] and the development of combinative capabilities [17]. Firms can build this type of specialized expertise by focusing on familiar knowledge [16] as familiar knowledge tends to yield more reliable results [7], [18]. This effect is reinforced at high levels of technological uncertainty as they increase the risk of obsolescence of the system, especially if the system spans multiple knowledge areas. Seen from this competence perspective, firms tend to reduce their R&D scope when uncertainty increases to lower the risk of system obsolescence.

[1]Our description here occurs with broad strokes. The conceptualizations of resource, capability, and knowledge are, of course, not identical. While research has produced finer delineations—Makadok [15], for example, shows when resources and capabilities are substitutes and when they are complements—for the discussion here the major point is that this group is collectively relatively clearly distinguishable vis-à-vis the TCE perspective.

More recently, researchers have proposed to develop views and approaches that combine the two perspectives mentioned earlier, often by incorporating additional elements such as context, interaction, and innovation. For example, Madhok [19] calls for triangular alignment of transaction attributes, resource attributes, and governance skills and suggests that "both TCE and RBV theory need to pay more attention to the context within which the activity occurs" [19, p. 542]. Similarly, Jacobides and Winter [20] argue that TCE and RBV are in reality intertwined as their drivers interact with each other while they coevolve. In their view, not only do transaction costs moderate the effect of the capability distribution on vertical scope, the scope itself affects the capability development process which determines the capability distribution, which in turn also motivates a change of the transaction costs. Building on this idea, Argyres and Zenger [21] made the time horizon explicit as the element that links TCE and RBV. Finally, Wolter and Veloso [22] explore the question of vertical structure as response to an exogenous shock (i.e., technological innovation) to the industry. Building on Henderson and Clark's [23] typology of incremental, modular, architectural, and radical innovations, Wolter and Veloso propose that the degree of vertical integration in an industry will not change in response to an incremental innovation, will decrease in response to a modular innovation, and will increase in response to architectural and radical innovations.

Extending this stream of recent research, we turn our attention to another contextual variable: the structure that characterizes the knowledge created during the invention process. We propose that the degree to which knowledge components are modular helps explain which line of argument that underlie TCE and RBV has greater explanatory power in explaining the R&D scope.

B. Knowledge Modularity

One of the criticisms of TCE has been that with the individual transaction as its unit of analysis it has assumed overly simplified systems. As Williamson himself states "the practice of examining transactions 'as if' they were independent will not do if there are significant interaction effects between them" [3, p. 1102]. At the heart of this statement lies the question of whether the entire system is more than the sum of its parts. In other words, the degree to which the parts of a system are interconnected has implications for where firms draw their boundaries.

A powerful way to think about this aspect of interconnectedness is to apply modularity theory. Half a century ago, Simon [24] introduced the concept of near-decomposability as a solution to limit system complexity and avoid cognitive overload. Later, researchers added the insight of the critical role of task division [25], of how organizational structures come to mirror the structure of the products these organizations produce [23], and of how product structures and industry structures are related [26]. Following this work, Schilling [27] proposes a model that products and systems strive to find a degree of modularity that balances competing external forces, and Baldwin and Clark [28] develop a theory of operators that explains how systems evolve to become more modular. More recently, the concept of modularity has been explored across various contexts such as product and process development [29] and has been further investigated and interpreted as a composite construct [30].

On the product level, Ulrich [31] laid the foundation for our understanding of modularity as a key aspect of product architecture. Since then, product architectures and their multiple dimensions have been linked to various firm effects and industry outcomes, ranging from operational performance measures in processes and supply chains [32], to the appropriate relationship between buyers and suppliers [33]. Baldwin [34] clarifies the mechanisms through which modularity affects the underlying task and transfer network, and subsequently the location of firm boundaries.

Modularity has also been invoked on the knowledge level. Admittedly, more abstract than the product level, knowledge has been identified as having various attributes, for example, tacit versus explicit [35], competence-enhancing versus competence-destroying [36], firm-specific versus non-firm-specific [37], or easy to recombine versus not so easy to recombine [38]. It is this fourth characterization of knowledge, the degree to which it is "recombinable" that borrows from modularity theory. We propose that such a construct of modularity of the knowledge structure can help explain why TCE and RBV favor different predictions for a firm's R&D scope response when facing technological uncertainty. Note that we describe knowledge modularity as a phenomenon that exists on the industry level and is dynamic. While it is possible that the degree of modularity of the components of a product varies from design to design, i.e., multiple designs can exist in an industry, we focus our attention on the modularity of the underlying knowledge structure. This knowledge structure is an industry-level characteristic and not necessarily related to the product structure [39]. It is also not static but can change over time with technology changes as the aggregate outcome of the search activities of firms [18].

C. Hypotheses

Fundamentally, the TCE perspective presupposes one unit of analysis, a transaction, as more or less independent of the next transaction. This implicitly assumed very low level of interdependence between the "components" under consideration represents a high degree of modularity of the knowledge structure. In contrast, the RBV perspective focuses on resources that are difficult to imitate by competitors. One aspect that makes a set of "components" difficult to imitate is their level of interconnectedness, i.e., the complex system they form via their interdependencies [40]. In short, the two perspectives appear to build on different assumptions concerning the underlying knowledge structure. With modularity as an indicator of the knowledge structure, we use this difference in assumptions to formulate our first hypothesis.

Hypothesis 1: Modularity of the knowledge structure moderates the relationship between technological uncertainty and R&D scope.

In addition, we hypothesize how different degrees of knowledge modularity affect the relationship between technological uncertainty and R&D scope. To do so we compare the factors which reduce or increase R&D scope, once for a low knowledge modularity regime, and once for a high knowledge modularity regime. This hypotheses development builds on the discussion in the previous Section II-A and II-B.

In regimes of low knowledge modularity, many knowledge elements exhibit interconnections to various other knowledge elements. For a system of a given size, with the number of interconnections increasing, the system complexity grows nonlinearly, often exponentially. As a result, coordination becomes substantially more difficult. In other words, a large system with many interconnections between its components is difficult to understand, and by extension difficult to develop, due to its low degree of modularity [41].

High levels of technological uncertainty acerbate this challenge of low modularity systems for two reasons. First, technological uncertainty makes it harder to assess the future state of individual components, and consequently the system behavior as a whole more difficult to decipher. Second, high levels of technological uncertainty make it more likely that the development of individual knowledge components is delayed or fails. Low degrees of knowledge modularity then can render the whole system useless under conditions of rapidly changing technological uncertainty [9].

In summary, we hypothesize that high levels of technological uncertainty increase the failure and obsolescence risks in the context of knowledge structures with low degrees of modularity and that firms respond to this threat by reducing their scope, effectively limiting complexity.

Hypothesis 2A: In regimes of low knowledge modularity, high technological uncertainty is associated with a reduced R&D scope.

In regimes of high knowledge modularity, the interconnectedness between individual knowledge components is significantly lower than in regimes of low knowledge modularity. Consequently, components can be viewed as simply additive in the high modularity regime, and the risk of system obsolescence as a result of component failure is substantially reduced.

High degrees of modularity slow down the growth of complexity cost with growing system size. Instead, cost of opportunism dominates, because when modularity is high, a component could be easily removed and used or applied in another system. Under these circumstances, a developer could behave opportunistically and sell or apply the knowledge component also to competitors [5]. This risk of opportunism increases substantially with increasing technological uncertainty, as technological uncertainty makes it more difficult to predict the future use of technology, or to agree on the future value of specific knowledge components [5]. More generally, high levels of technological uncertainty make it more difficult for the participants to set R&D goals, to agree on technological approaches, and to determine who is at fault in the event of failure [14].

In summary, we hypothesize that high levels of technological uncertainty in the context of highly modular knowledge structures lead firms to increase their R&D scope in order to reduce opportunities for opportunistic behavior.

Hypothesis 2B: In regimes of high knowledge modularity, high technological uncertainty is associated with an increased R&D scope.

III. Data and Methods

A. Setting

We selected the emerging automotive air bag industry for our study for two reasons. First, automotive air bags are multi-technology products, incorporating diverse technologies such as mechanical, electrical and electronic, computing, chemical, and textile technologies, across the four main components: inflator, sensor, diagnostic and control unit, and bag. Hence, they represent a good technological example to study the phenomenon of modularity of the associated knowledge structure. Second, emerging industries with their inherent uncertainty provide an appropriate setting for studying the effects of research activity [42]. Consequently, we focus in the emerging automotive air bag industry on the phases in which serious technical work was conducted and product sales started, grew, and approached maturation, i.e., the years from 1983 to 2001, which allows us to study a phase with relevant levels of technological uncertainty.[2]

TABLE I
U.S. Patent Class/Subclasses Relevant for Automotive Air Bags

Level 1	Level 2	Level 3	Title
280/728.1			Inflatable passenger restraint or confinement (e.g., air bag) or attachment
	280/728.2		With specific mounting feature
	280/728.3		Deployment door
	280/729		Plural compartment confinement (e.g., “bag within a bag”)
	280/730.1		Inflated confinement specially positioned relative to occupant
		280/730.2	Mounted in vehicle and positioned laterally of occupant
	280/731		Deflated confinement located within or on steering column
	280/732		Deflated confinement located in or on instrument panel
	280/733		In the form of or used in conjunction with a belt or strap
	280/734		Responsive to vehicle condition
		280/735	Electric control and/or sensor means
	280/736		With source of inflation fluid and flow control means thereof
		280/737	With means to rupture or open fluid source
		280/738	With means to aspirate ambient air
		280/739	With confinement deflation means
		280/740	With means to diffuse inflation fluid
	280/741		Inflation fluid source
	280/742		Flow control means
	280/743.1		Specific confinement structure
		280/743.2	With confinement expansion regulating tether or strap

B. Data

We approximate firms’ knowledge production by using patent data. Although patent data does not capture all innovative activities in which firms engage, it does exhibit a close relationship with firms’ innovative activities and can be used as a reasonable indicator of knowledge held by those firms [44]–[46]. The United States Patent and Trademark Office (USPTO) assigns each patent one or more classification numbers based on the technologies the patented invention employs [46], i.e., the classification numbers of a patent correspond to knowledge components used in the invention [47], [48]. Following prior research, we use these classification numbers as the proxy for technological knowledge components used in a patented invention [49]–[51]. After reviewing the manuals of the USPTO classification system and the International Patent Classification (IPC), and corroborating through term searches, we identify the automotive air bag industry containing the following subclasses: USPC 280/728.1 to USPC 280/743.2 (see Table I).[3]

To construct our dataset, we extracted all patents classified into at least one of the aforementioned named air bag subclasses from Cassis 2 DVD-ROM [52] that contain all U.S. patents issued between 1969 and 2003. Since there is typically a time lag of two to three years from application to issuing of a patent we focus our analysis on patents applied in the timeframe between 1967 and 2000 to avoid censoring problems. The extracted data were then preprocessed, during which the occasional duplicates were detected and removed.[4,5]

Next, with our interest in the emergence phase of the automotive air bag industry, which occurred in the 1980s and 1990s (see Fig. 1), and in consideration of the aforementioned NHTSA ruling taking place in 1983, we focus in this paper on inventive activity starting from 1983. Subsequently, we identified firms which participated in the air bag industry during the focal timeframe. Since firm boundaries change over time

[2]After years of legal battles between the governmental National Highway Traffic Safety Administration (NHTSA) and the major automobile manufacturers over the best means to protect drivers and passengers in car crashes, the legal uncertainty was removed in the mid-1980s. NHTSA’s 1983 decision to require driver airbags, together with several Supreme Court rulings, changes in management in some of the car companies, and pressure from the insurance industry led to almost all new passenger cars being equipped first with driver and subsequently with passenger front airbags during the late 1980s and early 1990s [43].

[3]Note that starting from January 1, 2015 USPTO moved from the US patent classification (USPC) system to the Cooperative Patent Classification (CPC) system.

[4]After the data extraction, the duplicates were found and removed. In particular, out of all patents that shared the same application number and assignee code, only those were retained that had a nonmissing issue date. In the remaining set, for the vast majority of the patents the abstract was unique. However, there were certain patents that shared the exact same abstract, and in addition there were patents that did not have an abstract (i.e., their abstract was blank). In the case when the abstracts were duplicated, we retained only the patents having the latest date of issue. For the patents that did not have an abstract, we only retained the patents whose title and the assignee code was not shared among the rest of the patents that had missing abstracts.

[5]Since by construction of the data, all of the patents were extracted from the range of subclasses falling under the “umbrella” class 280 (“Land Vehicles”), it is obvious that all of the extracted patents had 280 as one of their classes. Consequently, for each patent we excluded from its set of USPTO classifications the common class 280 and, since our study concerns the role of knowledge modularity of knowledge components (USPTO classes) in moderating the relationship between uncertainty and inventive scope, the focus of the current study are patents that are classified into at least one class that is different from the common class 280.

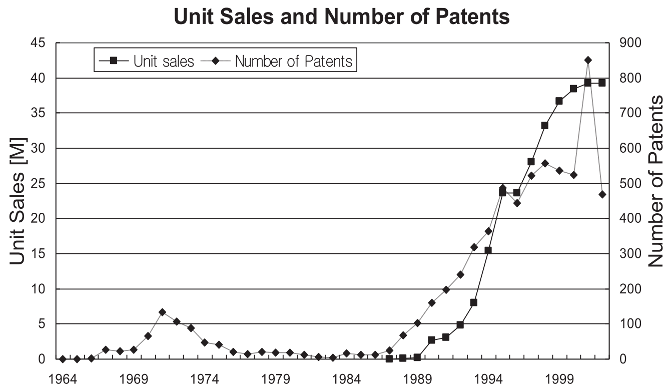

Fig. 1. R&D activity and sales in the automotive air bag industry.

through a variety of processes (e.g., mergers and acquisitions or divestments), we collected the firms' history data (e.g., name changes, subsidiaries, spin-offs, acquisitions, mergers, and joint ventures) from databases such as Lexis/Nexis, companies' annual reports (10-K's) and websites. Based on these data, we dynamically adjusted the patenting entity codes of the relevant patents. Some inventors (typically individuals or small firms) patent their developments but rarely commercialize them. Such inventors typically produce only a few patents sporadically and their impact on subsequent product development in the industry is often insignificant. The focal data set after the aforementioned preprocessing steps and considerations consists of 2283 unique patents. Since in our data set the top 20 patenting entities (in terms of their patenting activity) account for about 72% of the total patents under consideration, we focus our attention on these 20 patenting entities (Fig. 2 left) as the units of analysis when carrying out the scope calculations.[6] These 20 patenting entities include commercial firms, and no research centers whose risk behavior is known to be different [53]. Finally, since a number of patenting entities had years with no patenting activity, to reduce the sparseness of data without having to remove those patenting entities from the focal data set, we grouped the adjacent patent application years in two-year buckets starting from the 1983−1985 interval and ending at the 1999−2001 interval.

C. Variables

1) Dependent Variable: R&D Scope: As our dependent variable we use air bag-related R&D scope of patenting entity i in period t. Since patent classes have long been used as the proxy for technological components, we will base our measure of R&D scope on patent classes in which a patenting entity has patents during a given period of interest. However, to go beyond a simple count of patent classes and instead account for the degree of (dis)similarity between separate patent classes, we borrow a measure from population ecology. Specifically, we operationalize *R&D* scope using the quadratic diversity (also known as quadratic entropy) measure proposed by Rao [54]. The diversity coefficient has been widely used in ecological research for quantifying and measuring biodiversity of species or ecological communities [55]–[59]. In contrast to the more traditional biodiversity measures which measure biodiversity solely based on the number (or relative abundance) of functional groups, Rao's coefficient also accounts for dissimilarity among community members [59].

Rao's diversity coefficient is calculated using (1) as the expected dissimilarity between two members chosen randomly with replacement from a population of interest [54]. In (1) the non-negative, symmetric function $d(S_1, S_2)$ represents the dissimilarity between S_1 and S_2, while function P stands for the probability density function (pdf) characterizing the population of interest. As it can be noticed, Formula (1) simply is the expected value of the distance function $d\,(S_1,\, S_2)$.

$$H = \int d\,(S_1, S_2)\, P\,(dS_1)\, P\,(dS_2)\,. \tag{1}$$

To use (1) for measuring R&D scope of assignee i in period t, for each patent applied by patenting entity i in period t, we first note all the USPTO classes together with corresponding subcategory and category numbers [46] into which the patent has been classified. We then represent each patent by the following vector of length 3: <Category, Sub-category, Class> where the elements of the vector correspond to the class, the subcategory, and the category of the USPTO class into which the patent under consideration has been classified. Generally speaking, for a given patent there will be as many vectors of length 3 as there are references to classes made by the patent under consideration. As a result of this correspondence, the set of all patents applied by patenting entity i in period t is mapped onto a set of vectors of length 3. Furthermore, we define the dissimilarity between any two vectors using the Hamming distance [54], [60], which equals the number of positions at which the two vectors do not match. In our case, the Hamming distance between vectors S_1 and S_2 is

$$d\,(S_1, S_2) = 3 - \sum_{r=1}^{3} \delta_r \tag{2}$$

where the indicator δ_r takes the value 1 if the elements standing at the rth location of vectors S_1 and S_2 match, and 0 otherwise.

As it has been shown by Rao [54] for the vector mapping introduced earlier, the diversity coefficient simplifies to

$$H = \sum_{r=1}^{3} \left(1 - \sum_{s=1}^{k_r} p_{r,s}^2\right) \tag{3}$$

where $p_{r,s}$ $(s = 1,\ 2, \ldots, k_r)$ is the probability mass function (PMF) $(\sum_{s=1}^{k_r} p_{r,s} = 1)$ of the element standing at the rth position ($r = 1, 2, 3$) of the vector corresponding to a randomly selected patent. Consequently, we will use (3) for scope calculations[7,8]; see the Appendix for an example.

[6]All the individual inventors were grouped together and represented in the paper as one "assignee."

[7]Note that if a firm during a given period does not have patents, we define its scope to be 0.

[8]Since the R&D scope of an assignee may theoretically grow with the increase of available classes in the industry rather than as a response to changing uncertainty, we include the number of patent classes in the industry in period t to control for this effect.

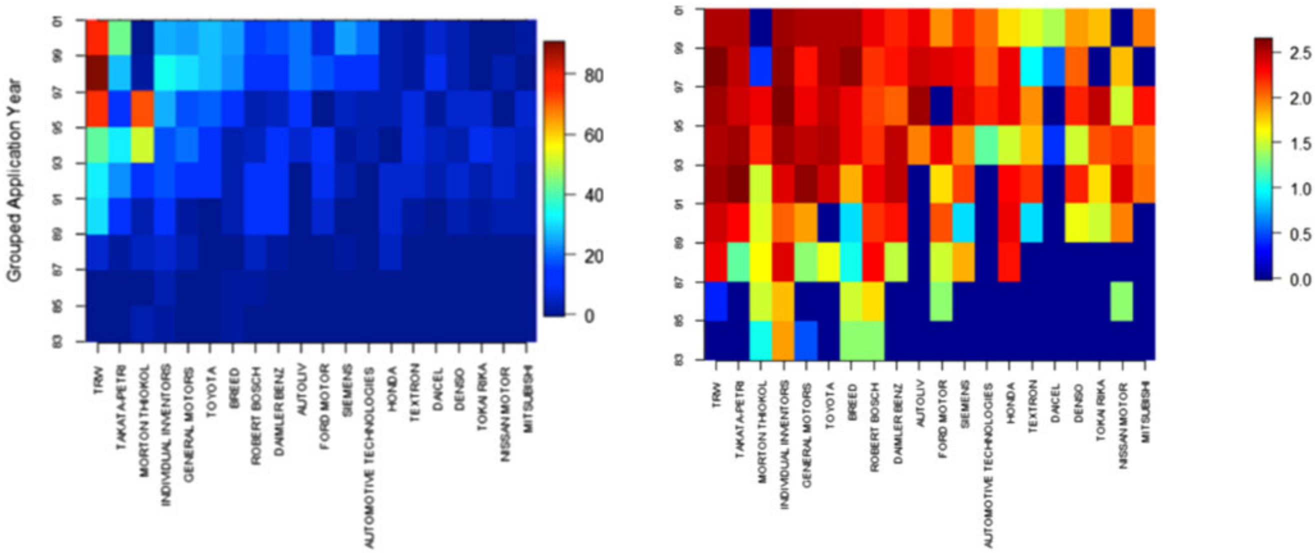

Fig. 2. Patenting volume (left) and R&D scope evolution (right) of top 20 patenting entities.

The dynamics of R&D scope for each of the top 20 patenting entities is provided as a heatmap in the right part of Fig. 2. Two observations deserve a special note. First, note that the heatmap gets first lighter and then more red as one moves vertically from the bottom towards the top, i.e., overall the scope increases over the years. Second, the heatmap also reveals the decrease of the scope as one moves horizontally from left to right across the graph. Since the heatmap in the right part of Fig. 2 is intentionally graphed in line with the one on the left, namely—in the descending order of patenting volume, it can be conjectured that the overall patenting volume of patenting entities may be correlated with their scope. Therefore, in light of these considerations we will include two additional control variables in our modeling: a variable for the time period to account for the first observation made earlier and the patenting volume (as a proxy for patenting entity "size") for each of the top 20 patenting entities to account for our second aforementioned note.

2) Independent Variable: Technological Uncertainty: Past research has measured technological uncertainty typically on the project level and often via self-assessment. For example, Raz *et al*. [61] asked managers to place their projects in one of four pre-defined categories, from low-tech to super high-tech. Similarly, Song and Montoya-Weiss [62] measured perceived technological uncertainty with a multi-item scale asking respondents using a Likert-type scale ranging from 0 to 10. In contrast, we develop a dynamic industry-level measure for *Technological Uncertainty,* similar to Luque [63] and Goerzen [64]. However, whereas Luque and Goerzen measure changes in patenting volume in an industry from year to year, we measure technological uncertainty through the changing pattern of the foci of inventive activity over time in the automotive air bag industry. Thus, we operationalize technological uncertainty in period t as the aggregate of the number of unique class references in period t to classes that saw no patenting activity in period $t-1$ and the unique class references to classes that had patents in period $t-1$ but did not see any patenting in period t, divided by the total number of unique class references in years t and $t-1$ across all classes used during those periods. More specifically, technological uncertainty in period t is presented as follows:

$$\frac{\sum_i (\text{Num. of references of class } i) + \sum_j (\text{Num. of references of class } j)}{\text{Total number of class references in periods } t \text{ and } t-1} \quad (4)$$

where $i \in$ classes, which are used in year t but were not used in period $t-1$, and analogously, $j \in$ classes that are not used in period t but were used in period $t-1$.

For example, in the [1993, 1995) period, 29 classes were newly referenced, i.e., they had not been used in the [1991, 1993) period and the total number of unique references in the [1993, 1995) period to those newly used classes equals 34. On the other hand there were 15 classes used in the [1991, 1993) period that were no longer referenced during the subsequent [1993, 1995) period, and the total number of unique references to those classes during the [1991, 1993) period was 19. Finally, the total number of unique class references during the two consequent periods under consideration was 1,140. Thus, according to (4), technological uncertainty during the [1993, 1995) period equals (34 + 19)/1,140 which is 0.046. Fig. 3 displays the evolution of technological uncertainty in the automotive air bag industry over our focal time span. Note that since patenting entity's R&D scope in a given period could be impacted by technological uncertainty in the previous period, in our modeling we include a lagged variable reflecting technological uncertainty in the previous period.

3) Moderating Variable: Knowledge Modularity: In the literature, modularity and related concepts such as interdependence have been measured through various metrics. For example, economists have used the construct of complementarity between business functions and practices [65], and management scholars have interpreted self-perceived complexity assessments of products and processes as metrics for interdependencies among a firm's productive activities [40]. Similar

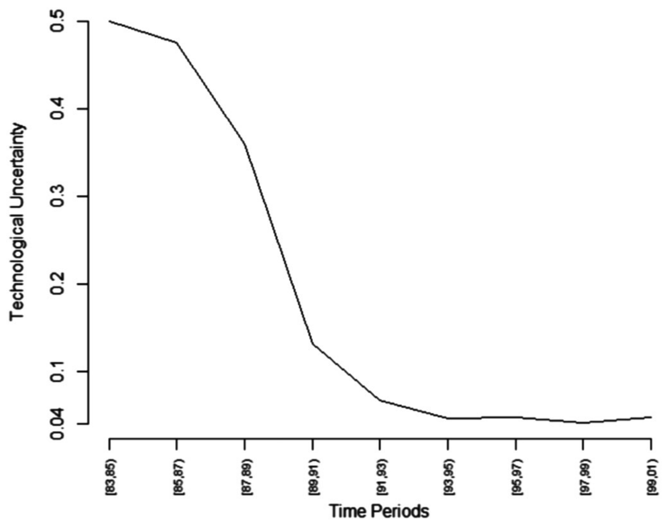

Fig. 3. Evolution of technological uncertainty over focal time span.

to Yayavaram and Ahuja [66], we use the co-occurrence of patent classes as an indicator of coupling between knowledge components (which may or may not coincide with physical components).

Further, to conceptualize modularity of the knowledge structure at the industry-level, we employ the co-occurrence of patent classes to create a graph presenting a knowledge network. In this knowledge network each patent class represents a node, and each patent referencing two or more classes is represented by undirected links between the classes into which it is classified [66]. The width of the edge connecting two nodes is based on the volume of patents referencing both classes that are represented as the nodes.[9] The structure of such a graph can shed light on how modular, or decomposable, the inventive knowledge base is.

To calculate modularity, we use the definition proposed by Newman and co-authors [67], [68]. Modularity is a network property that can be calculated for given (fixed) breakdowns of the network into communities. Based on the magnitude of the calculated modularity one can assess the "strength" of the proposed division, with strong divisions having a significant number of edges connecting nodes within the communities, with only sparse connections running across different communities [68].[10] The modularity is then defined as

$$Q = \frac{1}{2m} \sum_{v,w} \left(A_{vw} - \frac{k_v k_w}{2m} \right) \delta\left(c_v,\ c_w\right).$$

[9]Formulaically, the width of an edge connecting classes a and b during period k is $\mathrm{Width}_{a,b,k} = n_{a,b,k} / \sum_{i,j} n_{i,j,k}$, where $n_{i,j,k}$ is the number of patents applied for during period k that are classified in both classes i and j.

[10]More technically, in a network consisting of n vertices and m edges among them, denoted by A_{vw}, the indicator variable takes a value of 1 whenever there exists an edge between vertices v and w in the network, and 0, otherwise (i.e., A_{vw} are the elements of the adjacency matrix). It can be shown that $2m = \sum_{v,w} A_{vw}$ and that the degree (k_v) of a vertex v is given by $\sum_w A_{vw}$. Further, let us denote, respectively, by c_v and c_w the communities to which vertices v and w belong to, and denote by $\delta(c_v,\ c_w)$ the indicator function taking a value of 1 if and only if communities c_v and c_w are the same, and 0, otherwise.

TABLE II
MODULARITY LEVELS FOR EACH PERIOD

Period	Best Method	Modularity	Modularity Indicator
[1983, 1985)	Fast-greedy	0.695	0
[1985, 1987)	Walktrap	0.638	0
[1987, 1989)	Fast-greedy	0.372	1
[1989, 1991)	Fast-greedy	0.446	1
[1991, 1993)	Fast-greedy	0.467	1
[1993, 1995)	Fast-greedy	0.395	1
[1995, 1997)	Fast-greedy	0.503	0
[1997, 1999)	Fast-greedy	0.366	1
[1999, 2001)	Fast-greedy	0.468	0

The quantity Q measures the extent to which the proportion of edges that fall inside the contingent communities deviate from what that proportion would have been, had the edges among the vertices of the graph been drawn randomly. The higher the modularity the stronger the evidence of a detected community structure within a network.

For each period, we base our measure of modularity on how "communal" the network is. To find that modularity, we divide each knowledge network into communities employing algorithms that have been widely used in the literature on social networks, such as the "walktrap" [69], "fast-greedy" [68], and "edge-betweenness" [67] community detection methods and as modularity adopted the outcome of the method resulting in highest modularity. The modularity levels for each period are provided in Table II.

To investigate how the relationship between technological uncertainty and R&D scope changes in regimes of high and low knowledge modularity, we define an indicator (dummy) variable. In particular, the modularity indicator in period t is set to low (1) if the modularity in period t is less than or equal to the median value of modularity over our study period, and the modularity indicator is set to high (0), otherwise. To account for possible delays, our models also incorporates a lagged version of this variable.

D. Model Specification

For the testing of the hypotheses proposed in Section II-C, we will first model the relationship between technological uncertainty and R&D scope with the modularity serving as the moderating factor between the two. Subsequently, the hypotheses of interest will be formulated and tested as statistical significance tests regarding the parameters of the developed model. Since for each patenting entity under consideration a sequence of repeated observations measuring R&D scope are taken over time, and because by the very nature of our hypotheses we are interested in population-level analyses, we will adopt a marginal-model approach [70]. Repeated observations over time of the same subject require consideration of correlations between the observations [71]. To address this issue, we employ generalized estimating equations (GEE) [72], [73]. In addition to modeling the marginal expectation of the response and the variability of the response, GEEs allow to model the within-subject correlations separately [70]. The following are the three potential models

TABLE III
DESCRIPTIVE STATISTICS

Variable	Mean	Median	St. Dev	Minimum	Maximum
R&D Scope	1.422	1.780	0.993	0.000	2.627
Uncertainty (t)	0.191	0.067	0.196	0.041	0.500
Uncertainty ($t-1$)	0.229	0.132	0.198	0.041	0.500
Modularity (t)	0.483	0.467	0.115	0.366	0.695
Modularity ($t-1$)	0.491	0.467	0.115	0.366	0.695
Number of Classes	55.780	64.000	27.950	15.000	86.000
Number of Patents	82.300	60.500	72.000	19.000	352.000

tentatively entertained to model the mean (average) scope for patenting entity *i* during period *t:*

$$E(S_{i,\,t}) = \beta_0 + \beta_1 t + \beta_2 U_t + \beta_3 U_{t-1} + \beta_4 M_t + \beta_5 M_{t-1} + \beta_6 C_t + \beta_7 P_i + \beta_{24} U_t \times M_t \tag{5}$$

$$E(S_{i,\,t}) = \beta_0 + \beta_1 t + \beta_2 U_t + \beta_3 U_{t-1} + \beta_4 M_t + \beta_5 M_{t-1} + \beta_6 C_t + \beta_7 P_i + \beta_{35} U_{t-1} \times M_{t-1} \tag{6}$$

$$E(S_{i,\,t}) = \beta_0 + \beta_1 t + \beta_2 U_t + \beta_3 U_{t-1} + \beta_4 M_t + \beta_5 M_{t-1} + \beta_6 C_t + \beta_7 P_i + \beta_{25} U_t \times M_{t-1} \tag{7}$$

where t is the time period $(1, 2, \ldots, 9)$, $S_{i,t}$ is the scope of patenting entity *i* in period *t*, U_t is the technological uncertainty in period *t*, M_t is 1 if the modularity in period *t* is less than or equal to the median modularity during the focal span and 0 otherwise (modularity indicator), $C_t$is the number of classes (industry-wide) that were used by USPTO in period *t*, and P_i is the number of issued patents for patenting entity *i*.

E. Model Estimation and Empirical Results

The summary of descriptive statistics for the variables entering the models is provided in Table III. Models (5)–(7) were fitted using the "geepack" package of the R statistical software, leaving the within-subject correlation as unstructured. Among the three models (5)–(7) model selection was carried out using an extension of Akaike's Information Criterion (AIC) for GEE, represented by the QIC metric [74]. The parameter estimation summary for each of the three entertained models is presented in Table IV. Having the smallest QIC, model (7) is the most suitable candidate model for further exploration.

Subsequently, we entertained alternative within-subject correlation structures including first-order autoregressive, exchangeable, and independence forms, each time leaving the specification for the average response unchanged. The summary of those fits is presented in Table V. The unstructured within-subject configuration is still the one rendering the smallest QIC, making model (7) with unstructured within-subject configuration the final choice for hypothesis testing. Finally, after having dropped from model (7), all the variables corresponding to parameters that were not statistically significant at the 0.1 level, the final parameter estimates are provided in

TABLE IV
GEE REGRESSION COEFFICIENTS FOR MODELS (5)–(7)

Parameter	Estimates					
	Model (5)		Model (6)		Model (7)	
Intercept	2.244		1.596		1.658	
Period	0.079		−0.077		−0.062	
Number of classes	−0.011		0.008		0.007	
Number of patents	0.003	**	0.003	**	0.003	**
Uncertainty (*t*)	−2.436		−4.459		−4.567	†
Uncertainty ($t-1$)	−1.899	**	1.628	†	1.640	†
Modularity indicator (*t*)	−0.068		−0.030		−0.051	
Modularity indicator ($t-1$)	−0.012		0.298		0.559	*
Uncertainty (*t*)* modularity Indicator (*t*)	1.613	*				
Uncertainty ($t-1$)* Modularity indicator ($t-1$)			−2.510	***		
Uncertainty (*t*)* modularity indicator ($t-1$)					−8.726	***
QIC	195.370		194.885		194.863	

(Significance code: 0 '***' 0.001 '**' 0.01 '*' 0.05 '†' 0.1 ' ' 1)

TABLE V
GEE REGRESSION COEFFICIENTS TESTING THE INFLUENCE OF TECHNOLOGICAL UNCERTAINTY FOR MODEL (7) WITH DIFFERENT WITHIN-SUBJECT CORRELATION STRUCTURES

Parameter	Estimates							
	Independence		Exchangeable		Autoregressive		Unstructured	
Intercept	1.981		1.981		2.075		1.658	
Period	−0.051		−0.051		−0.049		−0.062	
Number of classes	0.002		0.002		0.000		0.007	
Number of patents	0.004	***	0.004	**	0.003	**	0.003	**
Uncertainty (*t*)	−5.191		−5.191		−5.218	†	−4.567	†
Uncertainty ($t-1$)	1.584		1.584		1.503		1.640	†
Modularity Indicator (*t*)	−0.073		−0.073		−0.060		−0.051	
Modularity Indicator ($t-1$)	0.532		0.532	†	0.550	*	0.559	*
Uncertainty (*t*)* modularity indicator ($t-1$)	−9.046	*	−9.046	*	−9.178	**	−8.726	***
QIC	197.219		197.219		196.970		194.863	

(Significance code: 0 '***' 0.001 '**' 0.01 '*' 0.05 '†' 0.1 ' ' 1)

Table VI.[11] We use these estimates for testing our hypotheses, which, formulated as statistical significance tests, are presented inTable VII.

In model 7, the interaction term of technological uncertainty and modularity exhibits a negative and statistically significant (*p*-value < 0.001) coefficient. This result suggests that the moderating effect of modularity is significant, which supports our Hypothesis 1. The sum of the coefficients corresponding to the

[11] We calculated both the "sandwich" and the model-based standard errors of estimated parameters from each of the entertained models. The standard errors and the resultant *p*-values that appear in the parameter estimation tables in the remainder of the paper are based on model-based standard error estimation since, given our relatively small sample size, the "sandwich" standard errors could be too liberal [70] and, thus, less accurate for testing the hypotheses of interest compared to the standard errors obtained under model-based estimation.

TABLE VI
GEE REGRESSION COEFFICIENTS FOR MODEL (7) WITH UNSTRUCTURED WITHIN-SUBJECT CORRELATION

Parameter	Estimate	Significance
Intercept	1.725	***
Number of patents	0.003	**
Uncertainty (t)	−4.391	***
Uncertainty ($t-1$)	1.299	*
Modularity indicator ($t-1$)	0.496	**
Uncertainty (t)* modularity indicator ($t-1$)	−7.797	***

(Signifiance code: 0 '***' 0.001 '**' 0.01 '*' 0.05 '†' 0.1 ' ' 1)

TABLE VII
FRAMEWORK FOR TESTING HYPOTHESES 1, 2A, AND 2B

Hypothesis 1	Hypothesis 2A	Hypothesis 2B
$\begin{cases} H_0 : \beta_{25} = 0 \\ H_A : \beta_{25} \neq 0 \end{cases}$	$\begin{cases} H_0 : \beta_2 + \beta_{25} \geq 0 \\ H_A : \beta_2 + \beta_{25} < 0 \end{cases}$	$\begin{cases} H_0 : \beta_2 \leq 0 \\ H_A : \beta_2 > 0 \end{cases}$

TABLE VIII
GEE REGRESSION FOR MODEL (6) WITH UNSTRUCTURED WITHIN-SUBJECT CORRELATION

Parameter	Estimate	Significance
Intercept	1.733	***
Number of patents	0.003	**
Uncertainty (t)	−4.355	***
Uncertainty ($t-1$)	1.222	†
Modularity indicator ($t-1$)	0.228	
Uncertainty ($t-1$)* modularity indicator ($t-1$)	−2.154	***

(Signifiance code: 0 '***' 0.001 '**' 0.01 '*' 0.05 '†' 0.1 ' ' 1)

technological uncertainty and the interaction term ($\beta_2 + \beta_{25}$) represents the change in the R&D scope as a result of an increase in technological uncertainty by a unit, assuming that the change is taking place under a low knowledge modularity regime. A one-sided Wald test for testing the Hypothesis 2A formulated in Table VII renders a p-value that is less than 0.001, which fully supports our Hypothesis 2A. On the other hand, the parameter estimate corresponding to the technological uncertainty (β_2) is negative and a one-sided Wald test designed for testing the third hypothesis of Table VII fails to support our Hypothesis 2B.[12]

Because the QIC values of model 6 and model 7 were very close (194.885 versus 194.863) it is logical to investigate how (if at all) our conclusions regarding the hypotheses of interest would be affected had we chosen model (6) to go with. Note that while in both models it is the last period's modularity that affects the magnitude of the relationship between uncertainty and current scope (i.e., moderates between scope and uncertainty), in model (7), it is the current period's uncertainty that is moderated by modularity, whereas in model (6) what is being moderated by modularity is the previous period's uncertainty. Thus, in essence the two models presuppose different time lags that it takes for the uncertainty to be reflected in the scope as moderated by modularity. Motivated by the absence of a formal relationship in reviewed extant research between technological uncertainty and inventive scope that would result in a choice between models (6) and (7) on rather theoretical grounds, we resort to empirical model building as compared to theoretical modeling [75] and thus repeated the earlier analyses with model (6).[13] The latter supports all the three hypotheses, including Hypothesis 2B and one-sided Wald test, designed for testing and renders a p-value less than 0.05. What the moderate support of Hypothesis 2B indicates based on this model is that in the regimes of relatively high knowledge modularity, increasing technological uncertainty is associated with firms broadening their inventive scopes in the following period. On the other hand, the lack of support for Hypothesis 2B based on model (7) shows the absence of evidence that following the period with high modularity firms' inventive scope broadens associated with increasing technological uncertainty in that period. Finally, note that the relatively moderate strength of support for Hypothesis 2B on the grounds of model (6) (p-value less than 0.05 compared to 0.001 for all other tests) suggests that the aforementioned support for Hypothesis 2B should be taken with a grain of salt.

IV. DISCUSSION

We began this paper with a discussion of the existing tension between the governance (TCE) and competence (RBV) perspectives on how firms adjust their R&D scope when facing technological uncertainty. In order to reduce the threat of opportunism, the governance perspective tends to predict an increase in R&D scope, whereas the competence perspective suggest a contraction of R&D scope in the face of technological uncertainty to increase a firm's focus and reduce the risk of obsolescence. To reconcile this apparent contradiction, we proposed a model in which the property knowledge modularity moderates the relationship between industry-level technological uncertainty and firm-level R&D scope. Testing our model with patent data from the U.S. automotive air bag industry we find that the interaction effect of technological uncertainty and knowledge modularity is significant, which confirms our overall hypothesis of knowledge modularity moderating between technological uncertainty and R&D scope. The results suggest that knowledge modularity contributes to explaining the different

[12]The model building procedure outlined in this section was repeated to evaluate the sensitivity of the results when the modularity indicator is defined based on a cutoff value 0.468 instead of the median 0.467 given how close the modularity of the last time period [1999; 2001) is to the median. In addition, the analyses were separately reran to assess whether controlling for the position of each firm in the supply chain would render different conclusions when testing the hypotheses of interest. For that purpose the set of all individual inventors was temporarily removed, and subsequently, a dummy variable was defined to denote whether or not a firm is an original equipment manufacturer (OEM) or a supplier. In both cases, the results were in line with those outlined in this section and the decision (support or lack of thereof) for each of the three hypotheses was unchanged.

[13]Following the same model identification-estimation procedure carried out for model (7), we estimated several within-subject correlation structures for model (6) and the results for the model with the most favorable QIC metric is summarized in Table VIII. Subsequently, the following three statistical hypothesis tests were considered for testing Hypotheses 1, 2A, and 2B, respectively: $H_0: \beta_{35} = 0$ versus $H_A: \beta_{35} \neq 0$; $H_0: \beta_3 + \beta_{35} \geq 0$ versus $H_A: \beta_3 + \beta_{35} < 0$; and $H_0: \beta_3 \leq 0$ versus $H_A: \beta_3 > 0$.

angles from which the relationship between technological uncertainty and R&D scope is seen by the two perspectives.

On a more detailed level, the air bag patent data also confirms our conjecture of a negative relationship between technological uncertainty and R&D scope in regimes of low knowledge modularity. This finding is consistent with earlier studies on cognitive limitations of organizations [18], [76]. In environments in which high levels of technological uncertainty exist, firms can build specialized competencies better by focusing on existing knowledge [16]. The data do not confirm, at least not consistently across our statistical models, however, our second hypothesis of a positive relationship between technological uncertainty and R&D scope in regimes of high knowledge modularity. Model (7) fails to confirm our Hypothesis 2B, whereas model (6) supports it. We consider it possible that our data—all taken from the domain of knowledge generation—do not show more consistently a positive relationship as conjectured because they do not really include a very high level of knowledge modularity such as those that can be found in some products in the physical domain. Another possible explanation could be our focus on the emerging phase of an industry, whereas very high levels of modularity tend to occur more in mature industries, perhaps preceding competence-destroying technological change.[14]

Taken together, our findings have both theoretical and managerial implications. On the theoretical level, the results suggest that a systemic aspect, such as knowledge modularity, has a moderating effect on how firms adjust their scope of R&D activities when facing technological uncertainty. In fact, our data suggest that this effect is so significant in the context of knowledge creating activities, such as R&D, that the competence perspective can explain our results better than the governance perspective. This supports the call on the TCE perspective to take better into account the possibility that the assumption of transaction unit independence is violated [3], especially when applied to knowledge-creating settings.

Although not the primary objective of the current study, the second theoretical contribution of the current work extends modularity theory. Existing theory suggests that high levels of standardization, often a key feature of modularity, help reduce uncertainty among market participants [77]. Our findings suggest a counterexample of this established rule. Indeed, our results show that, while knowledge modularity does moderate firms' scope decision in response to technological uncertainty, in settings that deal with abstract knowledge concepts, the relationship between technological uncertainty and modularity is not necessarily inversely related. Instead, in our study, the technological uncertainty decreases over our 15-year observation period, while the system-level knowledge modularity fluctuates. This suggests that at least in dynamic settings such as industry emergence, high levels of modularity are not necessarily seen as the cause for decreasing levels of technological uncertainty. More generally, there exists an industry-level knowledge modularity, which is not a choice of any individual firm yet is the outcome of all firms' technological decisions in the past. It, in turn, affects how firms react to technological uncertainty.

[14]We thank one of the anonymous reviewers for this relevant insight.

A third contribution of this paper is empirical. In contrast to extant research mostly focusing on firms' defensive moves in existing industries as a response to an external competitive shock via a technological innovation [22], [23], our study investigates knowledge creation activities and the corresponding scope decisions in an emerging industry, thus extending our knowledge by including the emergence of the innovations themselves.

Managers can use our findings to guide their strategic decisions about the R&D scope when facing different levels of technological uncertainty. By nature, R&D activities involve a significant risk for returns from investments. This risk is especially high under regimes of high technological uncertainty, as it increases the probability of failure of the development of technologies. In response, firms should focus initially on developing their deep expertise in areas close to what they already know, especially when the knowledge modularity is low.

V. Conclusion

The boundaries of the firm are a fundamental issue in the theory of the firm. How a firm manages its R&D scope is a special case of this question, one that is increasingly relevant for the field of technology and innovation management given the growing importance of knowledge work. Traditionally, the R&D boundary of a firm has been thought of as a "make-versus-buy" decision similar to decisions in the world of production. In contrast, we suggest that the effect of knowledge modularity on the relationship between technological uncertainty and R&D scope is significant, and because R&D is, by nature, less modular than production, the managing of R&D boundaries needs to take into account both external technological uncertainty and industry-level degrees of knowledge modularity. Consequently, the existing theories focusing on predicting the location of firm boundaries need to be adjusted when taken from the world of production and transferred to the world of R&D. Similarly, managers should consider the degree of knowledge modularity as part of their decision making about their organizations' R&D boundaries.

There are several possible limitations to our study. First, by its very nature, our analysis is descriptive, not normative. In other words, we cannot claim that our observations of firms' choices on R&D scope directly affected the firms' performance. This is because financial performance data are impossible to obtain if the unit of analysis (air bags) is only a fraction of the business of the firms under consideration. No firm in our sample reports separately financial results of its air bag-related business. For the same reason, we could not collect targeted input data such as R&D dollars that perhaps would allow imputing some measure of R&D efficiency. However, what we can say is that our focal firms are the most innovative firms in this industry in terms of the size of their knowledge pool, i.e., air bag patents, and probably in terms of the value of their knowledge pool. Consequently, our results may at least indicate the minimal conditions for successful R&D in the air bag industry. Second, patent data, in general, track only some inventions but not all of them. For example, process innovations that are not patented

TABLE A1
VECTOR MAPPING OF PATENTS BASED ON USPTO CLASSIFICATION

Class	Number of Class References
<5, 55, 180>	20
<5, 55, 244>	5
<4, 41, 200>	5
<4, 42, 313>	10

TABLE A2
HAMMING DISTANCES AMONG THE VECTOR MAPPINGS

Vector Mapping	<5, 55, 180>	<5, 55, 244>	<4, 41, 200>	<4, 42, 313>
<5, 55, 180>	0	1	3	3
<5, 55, 244>	1	0	3	3
<4, 41, 200>	3	3	0	2
<4, 42, 313>	3	3	2	0

are not accounted for in our data set, yet these innovations still require active attention of the managers of R&D.

A future direction of this research would be to test our model in other industries. We assume that knowledge modularity is an industry-level phenomenon; it is possible that the effect of knowledge modularity is more significant in some industries and weaker in others.

In addition to the theoretical contribution to the theory of the firm debate, we think that this research can help technology and innovation managers with this task in two ways. First, our research results point R&D managers to the value of understanding knowledge modularity for the decision on their firms' R&D scopes and explain how and why knowledge modularity affects firms' responses to technological uncertainty. Second, while knowledge modularity on the industry level is a theoretical construct, our newly constructed measure for knowledge modularity allows practitioners to actually gauge it through the use of patent data, which are publicly accessible.

APPENDIX

As a hypothetical example, assume a particular patenting entity in a given period has its patents classified into the following four USPTO classes: 180, 244, 200, and 313. The vector mapping with the corresponding patenting volumes and Hamming distances between the mappings are displayed in Tables A1 and A2 and relative abundances in Table A3. Note that in Table I, we are counting the number of class references, so if a given patent has multiple references to the same class, they all get included in the calculation.

ACKNOWLEDGMENT

The authors would like to thank Carliss Baldwin, Jeffrey Liker, Lawrence Seiford, and Jim Utterback for helpful comments on earlier versions of this paper. All errors and omissions remain our responsibility.

TABLE A3
PMF FOR THE ELEMENTS STANDING AT THE FIRST, SECOND, AND THIRD LOCATIONS OF VECTOR MAPPINGS

Category	Subcategory	Class	PMF
5			$25/40 = 62.5\%$
4			$15/40 = 37.5\%$
	55		$25/40 = 62.5\%$
	41		$5/40 = 12.5\%$
	42		$10/40 = 25\%$
		180	$20/40 = 50\%$
		244	$5/40 = 12.5\%$
		200	$5/40 = 12.5\%$
		313	$10/40 = 25\%$

Based on the aforementioned considerations and according to (3), the scope for the patenting entity during the given period would equal: $3 - ((0.625^2 + 0.375^2) + (0.625^2 + 0.125^2 + 0.25^2) + (0.5^2 + 0.125^2 + 0.125^2 + 0.25^2)) = 1.66$

REFERENCES

[1] J. G. March and H. A. Simon, *Organizations*. Hoboken, NJ, USA: Wiley, 1958.

[2] J. D. Thompson, *Organizations in Action*. New York, NY, USA: McGraw-Hill, 1967.

[3] O. E. Williamson, "Strategy research: Governance and competence perspectives," *Strategic Manage. J.*, vol. 20, pp. 1087–1108, 1999.

[4] O. E. Williamson, *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. New York, NY, USA: Free Press, 1985.

[5] G. P. Pisano, "The R&D boundaries of the firm: An empirical analysis," *Admin. Sci. Quart.*, vol. 35, pp. 153–176, 1990.

[6] O. E. Williamson, "The economics of organization - the transaction cost approach," *Amer. J. Sociol.*, vol. 87, pp. 548–577, 1981.

[7] D. A. Levinthal and J. G. March, "A model of adaptive organizational search," *J. Econ. Behavior Org.*, vol. 2, pp. 307–333, 1981.

[8] S. Balakrishnan and B. Wernerfelt, "Technical change, competition and vertical integration," *Strategic Manage. J.*, vol. 7, pp. 347–359, 1986.

[9] M. D. Santoro and J. P. McGill, "The effect of uncertainty and asset co-specialization on governance in biotechnology alliances," *Strategic Manage. J.*, vol. 26, pp. 1261–1269, 2005.

[10] H. Sicotte and M. Bourgault, "Dimensions of uncertainty and their moderating effect on new product development project performance," *R&D Manage.*, vol. 38, pp. 468–479, 2008.

[11] C. E. Helfat and D. J. Teece, "Vertical integration and risk reduction," *J. Law, Econ., Org.*, vol. 3, pp. 47–67, 1987.

[12] K. M. Sutcliffe and A. Zaheer, "Uncertainty in the transaction environment: An empirical test," *Strategic Manage. J.*, vol. 19, pp. 1–23, 1998.

[13] L. Poppo and T. Zenger, "Testing alternative thoeries of the firm: Transaction cost, knowledge-based, and measurement explanations for make-or-buy decisions in information services," *Strategic Manage. J.*, vol. 19, pp. 853–877, 1998.

[14] G. Hoetker, "How much you know versus how well I know you: Selecting a supplier for a technically innovative component," *Strategic Manage. J.*, vol. 26, pp. 75–96, 2005.

[15] R. Makadok, "Toward a synthesis of the resource-based and dynamic-capability views of rent creation," *Strategic Manage. J.*, vol. 22, pp. 387–401, 2001.

[16] C. K. Prahalad and G. Hamel, "The core competence of the corporation," *Harvard Bus. Rev.*, vol. 68, pp. 79–91, May-Jun. 1990.

[17] B. Kogut and U. Zander, "Knowledge of the firm, combinative capabilities, and the replication of technology," *Org. Sci.*, vol. 3, pp. 383–397, Aug. 1992.

[18] R. R. Nelson and S. G. Winter, *An Evolutionary Theory of Economic Change*. Cambridge, MA, USA: Harvard Univ. Press, 1982.

[19] A. Madhok, "Reassessing the fundamentals and beyond: Ronald Coase, the transaction cost and resource-based thoeries of the firm and the institutional structure of production," *Strategic Manage. J.*, vol. 23, pp. 535–550, 2002.

[20] M. G. Jacobides and S. G. Winter, "The co-evolution of capabilities and transaction costs: Explaining the institutional structure of production," *Strategic Manage. J.*, vol. 26, pp. 395–413, 2005.

[21] N. S. Argyres and T. R. Zenger, "Capabilities, transaction costs, and firm boundaries," *Org. Sci.*, vol. 23, pp. 1643–1657, 2012.
[22] C. Wolter and F. M. Veloso, "The effects of innovation on vertical structure: Perspectives on transaction costs and competences," *Acad. Manage. Rev.*, vol. 33, pp. 586–605, 2008.
[23] R. M. Henderson and K. B. Clark, "Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms," *Admin. Sci. Quart.*, vol. 35, pp. 9–30, 1990.
[24] H. A. Simon, "The architecture of complexity," in *Proc. Amer. Philosoph. Soc.*, 1962, vol. 106, pp. 467–482.
[25] E. von Hippel, "Task partitioning: An innovation process variable," *Res. Policy*, vol. 19, pp. 407–418, 1990.
[26] R. N. Langlois and P. L. Robertson, "Network and innovation in a modular system: Lessons from the microcomputer and stereo component industries," *Res. Policy*, vol. 21, pp. 297–313, 1992.
[27] M. A. Schilling, "Towards a general modular systems theory and its application to interfirm product modularity," *Acad. Manage. Rev.*, vol. 25, pp. 312–334, 2000.
[28] C. Y. Baldwin and K. B. Clark, *Design Rules: The Power of Modularity*, vol. 1. Cambridge, MA, USA: MIT Press, 2000.
[29] S. K. Fixson, "Modularity and commonality research: Past developments and future opportunities," *Concurrent Eng.: Res. Appl.*, vol. 15, pp. 85–111, 2007.
[30] F. Salvador, "Toward a product system modularity construct: Literature review and reconceptualization," *IEEE Trans. Eng. Manage.*, vol. 54, no. 2, pp. 219–240, 2007.
[31] K. T. Ulrich, "The role of product architecture in the manufacturing firm," *Res. Policy*, vol. 24, pp. 419–440, 1995.
[32] S. K. Fixson, "Product architecture assessment: A tool to link product, process, and supply chain design decisions," *J. Oper. Manage.*, vol. 23, pp. 345–369, 2005.
[33] J. Hsuan Mikkola, "Modularity, component outsourcing, and inter-firm learning," *R&D Manage.*, vol. 33, pp. 439–454, 2003.
[34] C. Y. Baldwin, "Where do transactions come from? Modularity, transactions, and the boundaries of firms," *Ind. Corp. Change*, vol. 17, pp. 155–195, 2008.
[35] I. Nonaka, "The knowledge-creating company," *Harvard Bus. Rev.*, vol. 69, pp. 96–104, 1991.
[36] M. L. Tushman and P. Anderson, "Technological discontinuities and organizational environments," *Admin. Sci. Quart.*, vol. 31, pp. 439–465, Sep. 1986.
[37] C.-J. Chen and B.-W. Lin, "The effects of environment, knowledge attribute, organizational climate, and firm characteristics on knowledge sourcing decisions," *R&D Manage.*, vol. 34, pp. 137–146, 2004.
[38] L. Fleming and O. Sorenson, "Technology as a complex adaptive system: Evidence from patent data," *Res. Policy*, vol. 30, pp. 1019–1039, 2001.
[39] S. Brusoni and A. Prencipe, "Unpacking the black box of modularity: Technologies, products and organisations," *Ind. Corp. Change*, vol. 10, pp. 179–205, 2001.
[40] M. J. Lenox, S. F. Rockart, and A. Y. Lewin, "Does interdependence affect firm and industry profitability? An empirical test," *Strategic Manage. J.*, vol. 31, pp. 121–139, 2010.
[41] D. A. Levinthal, "Adaptation on rugged landscapes," *Manage. Sci.*, vol. 43, pp. 934–950, 1997.
[42] L. F. Tegarden, W. B. Lamb, D. E. Hatfield, and F. X. Ji, "Bringing emerging technologies to market: Does academic research promote commercial exploration and exploitation?" *IEEE Trans. Eng. Manage.*, vol. 59, no. 4, pp. 598–608, 2012.
[43] J. D. Graham, *Auto Safety—Assessing America's Performance*. Dover, MA, USA: Auburn House, 1989.
[44] E. C. Engelsman and A. F. J. van Raan, "A patent-based cartography of technology," *Res. Policy*, vol. 23, pp. 1–26, 1994.
[45] Z. Griliches, "Patent statistics as economic indicators—A survey," *J. Econ. Lit.*, vol. 28, pp. 1661–1707, Dec. 1990.
[46] B. H. Hall, A. B. Jaffe, and M. Trajtenberg, "The NBER patent-citations data file: Lessons, insights, and methodological tools," in *Patents, Citations & Innovations—A Window on the Knowledge Economy*, A. B. Jaffe and M. Trajtenberg, Eds. Cambridge, MA, USA: MIT Press, 2002, pp. 403–459.
[47] L. Fleming, "Recombinant uncertainty in technological search," *Manage. Sci.*, vol. 47, pp. 117–132, 2001.
[48] L. Fleming and O. Sorenson, "Science as a map in technological search," *Strategic Manage. J.*, vol. 25, pp. 909–928, Aug.–Sep. 2004.
[49] M. S. Giarrantana, "The birth of a new industry: Entry by start-ups and the drivers of firm growth—The case of encryption software," *Res. Policy*, vol. 33, pp. 787–806, 2004.
[50] L. Rosenkopf and A. Nerkar, "Beyond local search: Boundary-spanning, exploration, and impact in the optical disk industry," *Strategic Manage. J.*, vol. 22, pp. 287–306, 2001.
[51] M. Stolpe, "Determinants of knowledge diffusion as evidenced in patent data: The case of liquid crystal display technology," *Res. Policy*, vol. 31, pp. 1181–1198, 2002.
[52] USPTO, Cassis2 DVD-ROM [DVD-ROM]. (2003). [Online]. Available: https://www.uspto.gov/web/offices/ac/ido/ptdl/pdf_files/corelist.pdf
[53] S. S. Erzurumlu, J. Davies, and N. R. Joglekar, "Managing highly innovative projects: The influence of design characteristics on project valuation," *IEEE Trans. Eng. Manage.*, vol. 61, no. 2, pp. 349–361, 2014.
[54] C. R. Rao, "Diversity and dissimilarity coefficients: A unified approach," *Theoretical Population Biol.*, vol. 21, pp. 24–43, 1982.
[55] C. Ricotta and M. Moretti, "CWM and Rao's quadratic diversity: A unified framework for functional ecology," *Oecologica*, vol. 167, pp. 181–188, 2011.
[56] C. Ricotta and L. Szeidl, "Towards a unifying approach to diversity measures: Bridging the gap between the Shannon entropy and Rao's quadratic index. Theor. pop. biology," *Theor. Pop. Biol.*, vol. 70, pp. 237–243, 2006.
[57] S. Pavoine and S. Dolédec, "The apportionment of quadratic entropy: A useful alternative for partitioning diversity in ecological data," *Environ. Ecol. Statist.*, vol. 12, pp. 125–138, 2005.
[58] J. Leps, F. De Bello, S. Lavorel, and S. Berman, "Quantifying and interpreting functional diversity of natural communities: Practical considerations matter," *Preslia*, vol. 784, 2006.
[59] Z. Botta-Dukát, "Rao's quadratic entropy as a measure of functional diversity based on multiple traits," *J. Veg. Sci.*, vol. 16, pp. 533–540, 2005.
[60] R. W. Hamming, "Error detecting and error correcting codes," *Bell Syst. Tech. J.*, vol. 29, pp. 147–160, 1950.
[61] T. Raz, A. J. Shenhar, and D. Dvir, "Risk management, project success, and technological uncertainty," *R&D Manage.*, vol. 32, pp. 101–109, 2002.
[62] M. Song and M. M. Montoya-Weiss, "The effect of perceived technological uncertainty in Japanese new product development," *Acad. Manage. J.*, vol. 44, pp. 61–80, 2001.
[63] A. Luque, "An option-value approach to technology adoption in U.S. manufacturing: Evidence from microdata," *Econ. Innov. New Tech.*, vol. 11, pp. 543–568, 2002.
[64] A. Goerzen, "Alliance networks and firm performance: The impact of repeated partnerships," *Strategic Manage. J.*, vol. 28, pp. 487–509, 2007.
[65] P. Milgrom and J. Roberts, "Complementarities and fit: Strategy, structure, and organizational change in manufacturing," *J. Acct. Econ.*, vol. 19, pp. 179–208, 1995.
[66] S. Yayavaram and G. Ahuja, "Decomposability in knowledge structures and its impact on the usefulness of inventions and knowledge-base malleability," *Admin. Sci. Quart.*, vol. 53, pp. 333–362, 2008.
[67] M. E. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Phys. Rev. E*, vol. 69, p. 026113, 2004.
[68] A. Clauset, M. E. Newman, and C. Moore, "Finding community structure in very large networks," *Phys. Rev. E*, vol. 70, p. 066111, 2004.
[69] P. Pons and M. Latapy, "Computing communities in large networks using random walks," in *Computer and Information Sciences*. Berlin, Germany: Springer-Verlag, 2005.
[70] G. M. Fitzmaurice, N. M. Laird, and J. H. Ware, *Applied Longitudinal Analysis*, 2nd ed. Hoboken, NJ, USA: Wiley, 2011.
[71] A. J. Dobson, *An Introduction to Generalized Linear Models*, 2nd ed. Boca Raton, FL, USA: CRC Press, 2002.
[72] G. Ahuja and C. M. Lampert, "Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions," *Strategic Manage. J.*, vol. 22, pp. 521–543, Jun.-Jul. 2001.
[73] R. Katila and G. Ahuja, "Something old, something new: A longitudinal study of search behavior and new product introduction," *Acad. Manage. J.*, vol. 45, pp. 1183–1194, Dec. 2002.
[74] W. Pan, "Akaike's information criterion in generalized estimating equations," *Biometrics*, vol. 57, pp. 120–125, 2001.
[75] G. E. P. Box, J. S. Hunter, and W. G. Hunter, *Statistics for Experimenters*. Hoboken, NJ, USA: Wiley, 1978.
[76] D. A. Levinthal and J. G. March, "The myopia of learning," *Strategic Manage. J.*, vol. 14, pp. 95–112, 1993.
[77] R. Sanchez and J. T. Mahoney, "Modularity, flexibility, and knowledge management in product and organization design," *Strategic Manage. J.*, vol. 17, pp. 63–76, 1996.

Sebastian K. Fixson received the Dipl.-Ing. (M.Sc.) degree in mechanical engineering from KIT (formerly Technical University of Karlsruhe, Germany), and the Ph.D. degree in technology, management, and policy from MIT.

He is an Associate Professor of technology and operations management at Babson College, Wellesley, MA, USA, where he teaches courses in operations, design, and innovation on various levels (undergraduate, M.B.A./M.Sc., and executive education programs). He is also the founding Faculty Director of Babson's new Master of Science in Management in Entrepreneurial Leadership (MSEL) program.

His work has appeared in books and journals such as *Concurrent Engineering, Creativity and Innovation Management, Design Management Review*, the IEEE TRANSACTIONS ON ENGINEERING MANAGEMENT, the *International Journal of Automotive Technology and Management, the Journal of Operations Management*, the *Journal of Product Innovation Management, MIT Sloan Management Review, Research Policy, Technology Analysis & Strategic Management*, and others. His research interest focuses on investigating various aspects of innovation and design, ranging from structural analyses concerning modularity and commonality, to process improvements in product development via better use of tools and design practices, in both established organizations and entrepreneurial contexts. He has also published on challenges and opportunities when teaching how to innovate.

Davit Khachatryan received the B.S. degree in applied mathematics and informatics from Yerevan State University, Yerevan, Armenia, the M.S. degree in statistics and the Ph.D. degree in management science from the University of Massachusetts, Amherst, MA, USA.

Prior to joining Babson College, he was a Senior Associate at the National Economics and Statistics practice of PricewaterhouseCoopers. In the latter role he consulted in the area of predictive modeling and advanced data analytics, helping clients from financial, healthcare, and government sectors with building automatic predictive models and enhancing business intelligence solutions. He is an applied statistician with research interests in analyzing intellectual property data to study the formation and diffusion of knowledge in emerging industries. He has also conducted research in the analysis of time series data producing publications in journals such as *Journal of the Royal Statistical Society (Series C), the American Statistician, and Quality and Reliability Engineering International*.

Wonhee Lee received the Ph.D. degree in industrial and operations engineering from the University of Michigan, Ann Arbor, MI, USA.

He has worked with Hyundai Motor Co. and Samsung Economic Research Institute. He is currently working for the Automotive Electronics Business Team at Samsung Electronics, Suwon, South Korea.