The Organizational Structure of Development and Firm Innovation

Jiho Yang
Imperial College Business School, London, UK
j.yang18@imperial.ac.uk

Paola Criscuolo
Imperial College Business School, London, UK
p.criscuolo@imperial.ac.uk

Brian S. Silverman
Rotman School of Management, University of Toronto, Toronto, Ontario, Canada
silverman@rotman.utoronto.ca

November 26, 2022

Keywords: organizational structure; innovation; development; duplication; commercialization

Abstract

Prior studies exploring the relationship between organization of corporate R&D and innovation have focused almost entirely on research activities. We examine the relationship between organizational structure and development activities. We hypothesize that organizational structure has an impact on development that is distinct from that on research. Specifically, centralization of development is associated with reduced duplication of development effort; however, the likelihood that a given invention will be commercialized is lower for centralized development than decentralized development. We test these hypotheses using a unique dataset that is composed of all invention disclosures submitted by R&D personnel at a global ICT company. Exploiting a shift in the organization of development activities, we examine subsequent changes in development outcomes. We find support for our hypotheses. (122 words)

We thank seminar participants at Bocconi University, Duke University, Grenoble Ecole de Management, HEC Paris, Hong Kong Business School, the University of Sussex (SPRU), and The Wharton School. and conference participants at the AOM conference and DRUID conference, for valuable comments on earlier versions of this paper. The usual disclaimer applies.

1. INTRODUCTION

The internal organization of corporate R&D has become an increasingly vibrant area of study over the last 20 years. As scholars have described, firms with more centralized R&D structures produce qualitatively different innovation outcomes than those with decentralized R&D (Hounshell & Smith, 1989; Kay, 1988). Centralized firms tend to produce "capabilities-broadening" innovation that exhibits broader search and impact (Argyres & Silverman, 2004), and generates more patents per R&D dollar (Arora, Belenzon, & Rios, 2014), than their decentralized counterparts. This is due at least in part to the ability of centralized structures to elicit a more connected set of researchers throughout the firm (Argyres, Rios, & Silverman, 2020), which facilitates the absorption and recombination of knowledge (Guler & Nerkar, 2012; Moreira, Markus, & Laursen, 2018; ter Wal, Criscuolo, McEvily, & Salter, 2020).

Although this literature is framed in terms of R&D – research and development – the implicit emphasis has been decidedly weighted toward research, to the near-exclusion of development. The literature's tacit assumption has been either that development is less economically important than research, or that the trade-offs that afflict research also afflict development in the same way. Yet the first of these assumptions is demonstrably untrue. In terms of investment by U.S. firms, R&D expenditure allocated to development has exceeded that allocated to research for at least six decades, increasing from 62% in 1958 to nearly 80% throughout most of the 21st century (National Science Foundation, 2020). In terms of bottom-line impact, commercialization – the purview of development – is crucial for appropriating returns to innovation, as is evidenced by vivid studies of successful and failed commercialization efforts (e.g., Cardinal, Turner, Fern, & Burton, 2011; Chesbrough & Rosenbloom, 2002).

In this paper, we take issue with the second of these assumptions. Specifically, we explore the impact of centralization or decentralization of development activities on a firm's development and commercialization outcomes. We hypothesize that although the centralization/decentralization of development and of research are afflicted by the same underlying trade-offs, these trade-offs are manifested in qualitatively distinct ways in development vs. research. Specifically, whereas centralization of research offers a primary benefit of enhancing the breadth of knowledge recombination (with a concomitant reduction in BU-specific knowledge creation), centralization of development offers a primary benefit of reducing duplication of development effort: Rather than each business unit investing in implementing its own manifestation of corporate knowledge, leading to excessive cost (Thomas, 2011), centralized development can produce a single version that satisfies the needs of multiple business units. The concomitant cost is that this satisficing version will not be ideal for any given business unit; on the margin, this leads to lower rates of commercialization of development's outputs.

To test these hypotheses, we exploit a unique dataset that is composed of all invention disclosures submitted by R&D personnel at a global ICT company that we shall call StarCo. StarCo's internal database contains detailed information on more than 60,000 inventions over a 30-year period, including those for which it chose not to seek patent protection. The dataset provides us with the identity of the R&D unit that produced a focal invention, various measures of invention quality, and (for the subset of inventions for which StarCo filed a patent application) whether the invention was incorporated into a commercialized product. Of particular note, while StarCo's R&D function had long been composed of a central research lab for exploratory research coupled with decentralized development units at each business unit, in the mid-2000s the firm centralized most of its development activities within a single development unit charged with developing technologies for all product lines and business units. This was purely a shift in organizational authority; few if any inventors changed physical location. Exploiting this shift in the organization of development activities, we examine subsequent changes in development outcomes. We find support for our hypotheses: The shift to centralized development is associated with 1) reduced duplication across inventions by development units, and 2) reduced likelihood that a given invention is commercialized.[1]

Our study makes three contributions to the literature. Phenomenologically, we draw attention to the distinct features of the development portion of R&D. Although scholars of product commercialization and R&D project management have explored factors that influence the success of development (e.g., Chiesa & Frattini, 2007), the phenomenon is relatively understudied in the strategic management literature. Theoretically, we extend theory to propose how the different nature of development from research leads to differences in the manifestations of the well-established trade-offs between centralization and decentralization. While these have their roots in the same overarching theories as the trade-offs in research, notably the incentive-based approach of organizational economics and the information-processing approach of the knowledge-based view, the manifestations are quite distinct. Empirically, this is the among the first studies to link formal organization structure of R&D to commercialization outcomes rather than purely innovation outcomes. Such assessment has been difficult to accomplish in prior research due to the challenge of linking specific inventions to specific products, with the exception of studies of the pharmaceutical industry where it is feasible to link patents to drugs (e.g., Cardinal, 2001; Eklund, in press). This also is among the first studies to assess the implications of organization on the duplication of development activities, thanks to the availability of data on all developed inventions rather than only those that have been patented.

The next section reviews key insights from prior literature and develops our hypotheses.

[1] A small set of StarCo's development activities remained decentralized after the reorganization. We do not find similar changes in the outcomes for those development activities. Although this does not give us the confidence of a difference-in-differences analysis, since the choice of centralizing or decentralizing particular development activities was not random, it does suggest that the outcomes we observe are not simply due to firm-wide trends in innovation.

Section 3 provides details of our empirical context, the firm we call StarCo. Section 4 describes our empirical methodology and presents results. In Section 5 we discuss the results and conclude.

2. THEORY AND HYPOTHESES

2.1. The organization of R&D, with an emphasis on D

A small but influential stream of work has proposed that a firm's formal R&D organization structure should affect the type of innovation that it generates (Hounshell & Smith, 1989; Kay, 1988). The literature has typically distinguished between centralized R&D, where decisions and budget come from corporate headquarters, and decentralized R&D, where business-unit executives have decision and budget authority (Argyres et al., 2020; Lerner & Wulf, 2007; Teece, 1996).[2] This research has yielded several stylized facts that are consistent with theory. Perhaps the most consistent finding is that corporations with centralized R&D activities tend to produce innovations that exhibit broader search and broader impact than those with decentralized R&D, as proxied by forward and backward patent citations (e.g., Argyres & Silverman, 2004). This occurs at least in part because a centralized R&D function in a corporation characterized by related diversification provides incentives to researchers to pursue innovations that are applicable firm-wide, and hence to seek out opportunities to recombine knowledge with inventors in distant parts of the firm; in contrast, divisional managers do not internalize research spillovers that they may provide to other divisions, and hence have weaker incentives to invest in R&D that might have firm-wide benefits (Argyres, 1995; Hoskisson, Hitt, & Hill, 1993; Kay, 1988). Although in principle it might be possible to establish a set of payments from other divisions to compensate a focal division for such spillover benefits, such payments are notoriously challenging to structure due to classic transaction cost problems (Williamson, 1985).

[2] Some firms have "hybrid" structures in which a firm has both centralized and decentralized R&D units (see, e.g., Eggers, 2016).

Centralized R&D also tends to elicit different patterns of communication and information processing than decentralized R&D. To the extent that innovation requires the recombination of knowledge (Fleming, 2001), and to the extent that a given inventor's knowledge is necessarily limited (Guler & Nerkar, 2012), communication among inventors is typically seen as a crucial ingredient in the innovation process (Singh, 2005; Toh, 2014). Centralized R&D structures are associated with more connectedness among a firm's inventors, which in turn generates innovation with broader impact (Argyres et al., 2020). Such structures also tend to reduce communication between inventors and downstream employees such as marketing personnel, which may "shield" inventors from pressures to narrow their focus. In contrast, decentralized R&D structures are characterized by more isolated clusters of inventors, who typically interact more readily with within-division staff from other functions than with inventors from other divisions (Karim & Kaul, 2015; Katz & Allen, 1982). This yields communication patterns that favor narrower innovation.

In sum, the extant literature proposes that decentralization of R&D may be associated with underinvestment in broad research, because no business unit has sufficient incentive to invest in knowledge recombination that has firm-wide implications (Nelson, 1959) and because inventors in silos may forgo access to useful "distant" knowledge. Centralized R&D can overcome this underinvestment in broad knowledge recombination, although perhaps at the expense of strong incentives for within-BU innovation (Argyres, 1995; Argyres & Silverman, 2004; Eklund, in press).

Although this literature is couched in terms of R&D – research and development – it is decidedly focused on research rather than development activities, as evidenced both by its empirical focus on patent statistics and by its theoretical emphasis on creation of new knowledge rather than on implementation of knowledge in products or processes. To the extent that development exhibits similar characteristics to research, then this would not matter.[3] But research and development comprise two qualitatively distinct activities, characterized by different challenges and different manifestations of the above-described trade-offs. Whereas research aims to produce new scientific knowledge that does not necessarily culminate in specific products/processes and that might potentially be applied to many different businesses, development consists of a set of activities designed to create new or refine existing products/processes that will be commercialized by specific business units (Zahra & Nielson, 2002). Stylized conceptions of research frequently tout the benefits of wide-ranging connections among researchers in order to gain access to scientific and technical knowledge (e.g., Guler & Nerkar, 2012; Moreira et al., 2018; Schilling & Phelps, 2007); in contrast, stylized conceptions of development frequently tout the benefits of connections to other functions within the firm such as manufacturing and marketing, as well as to customers (Cardinal et al., 2011; Chiesa, 2001). To this end, research activities and development activities are frequently performed by different departments within a firm (Du, Leten, Vanhaverbeke, & Lopez-Vega, 2014; Leifer & Triscari, 1987), with different incentive and measurement systems (Barge-Gil & López, 2015), even if the two departments ultimately report to the same executive.

What does this mean for centralization or decentralization of development? In principle, decentralization will be characterized by the same fundamental incentive and information-processing issues as decentralization of research: business unit executives will strive to maximize the performance of their divisions rather than the performance of the firm overall,[4] and

[3] Alternatively, if development exhibited different characteristics than research, but accounted for only a tiny portion of R&D effort, then one might ignore these differences. However, nearly 80% of total R&D spending by firms went into development activities in 2018, compared to 14%, 6% for applied and basic research, respectively (National Science Foundation, 2020; Arora, Belenzon, and Patacconi (2018) discuss the decline in science production by corporate R&D labs).
[4] This is a fundamental characteristic of decentralization of authority for nearly all organizational decisions (see, e.g., Dessein, Garicano & Gertner, 2010; Dessein, Lo, & Minami, in press; Rivkin & Siggelkow, 2003.)

decentralized development decisions will be based more heavily on input from people closer to the "front lines" of production and sales, and perhaps from customers as well. However, development will exhibit a distinct manifestation of these issues. Whereas business unit-level incentives for research lead to an underinvestment in generation of knowledge with firm-wide applicability, business unit-level incentives for development should lead to excessive investment in developing duplicative business unit-specific products/processes rather than drawing on prototypes available elsewhere in the firm. And whereas decentralization may lead to less communication among far-flung inventors within the research function – as well as less cross-divisional information reaching decision-makers who might "harmonize" development decisions across the firm (Guadalupe, Li, & Wulf, 2014) – thus leading to narrower innovation, its incorporation of input from production, sales, and customers should lead to development of products that are more closely tailored to market needs.

Consider a firm in which development is decentralized. Theoretical models demonstrate that division executives will make "locally optimal" decisions, which optimize division performance, rather than "globally optimal" decisions that optimize firm performance (Rantakari, 2008). This results in greater division-specific customization than is optimal for the firm because division managers do not internalize the benefits of coordination and hence "put excessive weight on adapting their decisions to the local conditions" (Alonso, Dessein & Matouschek, 2008: 161).[5] This excessive customization of development outcomes to each business is particularly problematic if business units are technologically related, or their products share a similar technological base. Under such circumstances, development activities by individual business units may end up

[5] Even if divisional managers were interested in supporting globally optimal development decisions, information flows in a decentralized corporate structure hinder the sharing of relevant information necessary for the manager to recognize and assess potential synergies (Guadalupe et al., 2014).

producing inventions that deal with common technical problems and have similar functionalities. Consequently, as individual business units each invest in customized development of a similar corporate innovation, there occurs firm-level inefficiency due to duplicative effort. Thomas (2011) demonstrates this empirically in a study of the global detergent industry, finding that decentralized firms produce a greater-than-optimal number of detergents. According to her analysis, Procter & Gamble and Unilever could increase profits while reducing their product variety by as much as 20% and 30%, respectively; put differently, these firms (which have decentralized brand-product authority) are producing 20% and 30% more SKUs than they should. In contrast, a firm with centralized decision-making will produce a smaller line of products; this occurs because the decisionmakers internalize the resulting cost savings from coordination across product lines. Extending this insight to development – which is the mechanism by which these detergents are created – a firm whose development units report to business unit executives will tend to produce an overabundance of duplicative innovation, as each business unit reinvents the wheel to generate its ideal product.[6] Thus, we hypothesize:

Hypothesis (H1). *Centralization of development activities reduces duplication in development outcomes.*

To the extent that centralization reduces duplicative development because it precludes business units from engaging in efforts to make locally ideal products, there is a concomitant trade-off. By virtue of developing products that satisfy multiple divisions while not optimizing for any, the firm will develop innovations that meet the needs of a given division's customers less precisely. On the margin, then, each commercialized product stemming from the development effort will

[6] A recent example of this, noted in the business press, was the decision of luxury Swiss watch producer Richemont to centralize the development of watch movements, as opposed to having each of its independent brands develop its own movement (Foulkes, 2018). The stated goal was to reduce duplicative costs and tap scale economies in R&D.

generate asymptotically lower revenue.[7] This is illustrated empirically in Thomas's (2011) analysis, where reductions in locally optimal detergent variety are associated with reduced revenue (which is offset by centralized cost savings). Relatedly, a recent *McKinsey Report* describes the decision by the CEO of an unnamed European equipment producer to centralize the product development function in an otherwise largely decentralized firm: Although this achieved its aggregate goal of increasing profitability, "with product managers reporting to the [centralized] technical function rather than to business units, some new products have been technically strong but less tailored to market needs…" (Campbell, Kunisch, & Müller-Stewens, 2011: 4). Given the lower divisional profit associated with satisficing vs. optimal development, we expect that, on the margin, divisions will be less likely to commercialize the development outcomes of centralized development units as compared with the outcomes of decentralized development units. Thus, we predict:

Hypothesis (H2). *Centralization of development activities negatively affects the likelihood that a given invention is commercialized.*

To be clear, we do not assert that either centralized or decentralized development is better than the other. As described above, each structure has advantages in overcoming one aspect of the trade-offs in incentives and information processing, while exhibiting disadvantages related to another aspect.[8]

---

[7] One exception: Revenue could increase if there are complementarities across the divisions' products such that centralized development allows for greater coordination across these products. In such a case, satisficing development may yield products that work more harmoniously together, thus yielding complementarities that increase customers' willingness to pay even as each division's component is individually less well suited to customer needs.

[8] Development activities in the automotive industry illustrate this. In the late 1990s, several automakers including Ford, GM, Subaru, and Volvo adopted "platform-sharing" – the global use of a centrally developed automobile platform across numerous makes and models – as an effective mechanism for cost reduction and faster design of new vehicles (Treece & Sherefkin, 2001). At the same time, other auto producers such as Honda eschewed the "industry gospel that fewer platforms equal higher profits" (Doi, 2001: 1), on the grounds that "the use of common platforms would not allow the automaker to respond effectively to the diverse requirements of its customers" (Office of Industries, U.S. International Trade Commission, 2002: 15).

3. RESEARCH CONTEXT: “STARCO” CORPORATION

3.1. Description of the setting and qualitative evidence

We tested our hypotheses using data from a large, multi-business firm in the ICT sector which we call StarCo for confidentiality purposes. Before presenting our quantitative analyses, we first describe qualitative information that we collected from multiple sources. First, we conducted six exploratory interviews with five former and current executives of the firm. The interviewees included a former director of the firm's Intellectual Property Rights (IPR) department (interviewed twice), a former senior engineering manager, a former vice president of technology portfolio management, a current head of a research group who was vice president of industry collaboration at the time of the reorganization, and a former CTO of one of the business units. We also consulted numerous texts on StarCo and its reorganization of development activities including annual reports, a book detailing the history of StarCo, and popular-press articles published about the firm. For confidentiality reasons, we cannot reveal exact sources; however we mark with acronyms whether cited information is from interviews (IN), annual reports (AR), the book (B), or articles (A).

StarCo is a large multinational in the ICT sector. It grew rapidly by diversifying into related industries and providing a diverse array of related products for different geographic and customer segments (A). During the observation period, the firm was a global market leader in terms of both sales and profits. Regarding R&D, it spent more than 10% of revenues annually in R&D and at its height employed more than 30,000 people in 10+ R&D sites, making it one of the most R&D-intensive and innovative firms in the industry (AR; A). StarCo was also widely known for continuously introducing a variety of new innovative products to existing markets as well as venturing into emerging markets with radically new technologies.

For many years prior to the beginning of the study period, the R&D units of StarCo had been composed of a central research lab for exploratory and long-term research and decentralized development units in charge of developing new products and related technologies for business units (AR; IN). In the mid-2000's, however, StarCo decided to centralize many of its development activities, introducing a new central development unit that coordinated technical requirements of different product lines and developed common technologies that could potentially be used in multiple products (A; AR; B; IN). Consequently, StarCo came to have two central units, each focusing on a different phase of R&D: the central research lab conducting basic, exploratory, and long-term research, and the new central development unit developing technologies to be commonly used by multiple product lines. The qualitative data all indicate that the new central development unit was distinguished from the existing central research lab in terms of technological and commercial focus ("advanced development" vs. "fundamental research" in the words of one interviewee), timeline from project inception until commercialization (1-year-plus vs. 5 to 10 years), and people employed ("engineers" vs. "researchers"). In addition, although many development activities were moved to the central development unit after the reorganization, there still remained some development activities conducted by the business units in a decentralized manner. These decentralized activities mostly focused on product-specific development and modification of the research unit's technologies to meet product-specific requirements (AR). Also, development activities for product lines that were based on different, incompatible core technologies stayed in a business unit (IN). After the reorganization, therefore, StarCo's R&D was separated into three different organizational units: the central research unit, the central development unit, and a small handful of decentralized development units. Figure 1 graphically demonstrates the restructuring of R&D organization at StarCo. This organization structure lasted for four years, after which StarCo introduced a new structure for the entire corporation that also involved restructuring of research and development activities (AR).

[Insert Figure 1 Here]

In the annual report the reorganization of development activities was justified as a response to macro trends in which different technologies and product markets were converging, as well as to benefits from economies of scale. The interview data also give preliminary evidence consistent with our predictions that the centralization of development helped StarCo reduce duplication in development efforts. For instance, one interviewee said, "The reason why [the reorganization] was done is that R&D was very inefficient before this. Every project was asking for more and more resource. Each project was developing its own building blocks, its own Lego blocks so to speak (Director of IPR department)." Another interviewee stated that the objective of the reorganization "was wrestling to, in a way, reduce the amount of duplications and redundancies in the organization. There were duplicate developments [before the reorganization] (VP of Technology Portfolio Management)." On the flip side, however, the reorganization also impeded coordination between development and downstream functions. According to the director of IPR department, "one of the downsides… with the [centralized development] structure was it made R&D very efficient… but it disconnected us from the client/the customer." Similarly, the former CTO of a business unit told "[After the reorganization], we had separated the product developed from the technology." As a result, the central development unit ended up developing technologies that were not implemented into products by the business units. As the head of research pointed out, "[in the central development unit] people developed features that no product really at the end of the day used."

3.2. Sample

We gained access to StarCo's internal database containing information on all inventions (> 60,000) submitted by all StarCo employees (> 15,000 inventors) spanning 30 years. Although the single-firm setting may limit the study's generalizability, this dataset has distinct advantages over large public datasets that enable us to examine our research questions. One such advantage is that it not only includes patented inventions but also those for which StarCo did not seek patent protection.

Therefore, the data represent a complete history of the firm's inventive outcomes and minimize a potential sample selection problem (Criscuolo, Alexy, Sharapov, & Salter, 2019). Moreover, the dataset provides information that is rarely available in other datasets such as the R&D-unit affiliation of inventors, whether an invention is the subject of a patent application filing, and, for the subset of filed inventions, whether the invention is implemented into a commercialized product and/or promoted as an industry standard, in addition to basic information such as descriptive titles of inventions, names of inventors, year of submission to the database, and relevant technological area. The comprehensiveness and richness of the dataset enable us to test the hypotheses using novel measures that better reflect the theoretical concepts of our interest.

The most important aspect of the dataset for our study is that it provides information on the R&D-unit affiliation of inventors. This allows us to focus on outcomes of development and exclude those of research, and to distinguish between development units that were centralized during StarCo's reorganization and those that remained decentralized. Specifically, an invention is considered as a development outcome if any of its inventors is affiliated with a development unit.[9] If an inventor's affiliation is missing for a particular invention, we looked at her other inventions in the dataset to see if information on his/her unit affiliation is available in any of these inventions. We then replaced the missing affiliation with the affiliation information from that inventor's most temporally proximate submission. We do not observe many changes in inventors' affiliation over time (e.g., only 5% of inventors in the centralized development units were from the research units).

Our sample consists of all inventions submitted by development-unit personnel at StarCo

[9] If an invention is a collaborative work between a development unit and a research unit, we assigned this invention as a development outcome. This is because it is less likely for basic/applied research, which requires high level of scientific knowledge, to involve development professionals who usually possess less deep expertise than researchers. On the contrary, development activities can often benefit from scientific knowledge of researchers in solving complex technological problems. Therefore, we assume that collaboration between engineers and researchers commonly occurs for development activities rather than for research activities. Our results are robust to excluding these collaborative inventions, which comprise 7% of all development inventions

from 4 years before the reorganization to 4 years after the mid-2000s reorganization. For confidentiality reasons we cannot disclose the precise number of inventions in the data set, but we can provide rough counts. During the eight-year sample period, StarCo's development inventors submitted more than 12,000 invention disclosures to the firm's IPR department. Roughly two-thirds were produced by development units that were centralized during the reorganization, with the remaining one-third from units that remained decentralized.

3.3. Variables

3.3.1. Dependent variables.

Duplication. We measure the extent of duplication for each focal invention by calculating the similarity of its title to those of all inventions submitted prior to the focal invention. Since an invention's title summarizes core information about the invention's technological features, high level of similarity in titles between an invention and prior inventions can indicate duplication of development outcomes. We measured similarity through a commonly accepted metric of cosine similarity (Magerman, Van Looey, & Debackere, 2015). Specifically, we first normalized titles of inventions by removing non-alphanumerical characters and lemmatizing all texts. We then used the scikit-learn data science package in Python to calculate the cosine similarity between a focal invention and each prior invention made by development personnel at StarCo (see Figure 2 for an example of this process). We then averaged these cosine similarity values across all focal invention-prior invention dyads to derive our dependent variable *Avg Title Cosine Similarity*.[10]

Commercialization. We are able to measure commercialization for the subset of inventions for which StarCo filed a patent application. As noted above, Starcom's IP management system records

[10] We normalized the variable by dividing its values by the observed maximum value such that it ranges from 0 to 1. In an unreported analysis, we constructed alternative dependent variables based on cosine similarities of a focal invention to inventions submitted only in the previous 1, 2, or 3 years from the submission of a focal invention – thus excluding prior inventions submitted in the more distant past – and obtained qualitatively consistent results.

additional information about such inventions, including whether the filed invention was implemented into a product. The IP system also notes whether a filed invention becomes part of an “essential industry standard,” which by definition means that it has been implemented in at least one artifact. We coded *Implemented* as 1 if the focal invention’s record indicates that it was implemented into at least one product or that it was incorporated into an essential industry standard, and 0 otherwise. In principle, inventions can be implemented into products several years after being filed, which could create a right-truncation concern for this variable. However, the StarCo system including information regarding implementation has been updated until nine years after our sample period. This significantly reduces the risk of right-truncation, since it is unlikely that an invention will experience its first implementation more than nine years after its development.[11]

3.3.2. Independent variables.

To estimate the inter-period differences in invention characteristics, we constructed a *Post Reorganization* dummy which is equal to 1 for all inventions submitted after StarCo centralized development activities and 0 for those submitted before this reorganization. We also distinguished between inventions by development units that were centralized during the reorganization from those by development units that remained decentralized. We constructed a *Centralized Group* dummy which equals 1 for the inventions by the ultimately-centralized development units and 0 for those by the development units that remain decentralized. Inventions are coded this way regardless of whether they were submitted before or after the reorganization. The main independent variable of interest is the interaction between *Post Reorganization X Centralized Group*, the coefficient of which will show the effect of centralization of development on innovation *net of* the

[11] We looked at all implemented inventions that were submitted at the beginning of the study period to see how many years passed between submission and implementation. 93% of these inventions were implemented within nine years after submission.

effects of other firm- or industry-wide events (e.g., changes in IPR policies). H1 predicts that centralization of development reduces duplication; we therefore expect this interaction term to be negatively associated with *Avg. Title Cosine Similarity*. H2 predicts that centralization of development decreases the likelihood of commercialization; we therefore expect this interaction term to be negatively associated with *Implemented*.

3.3.3. Controls.

We control for various invention-level characteristics. First, it is possible that more novel or broader inventions are less likely to be similar to previous inventions. To address this, we include *New Technology Combinations,* which measures the share of novel pairwise combinations of technology classes that are present in a focal invention.[12] We also include *Number of Tech Classes*, which is a count of the number of three-digit technology classes an invention encompasses. Second, some inventions stemmed from formal projects while others stemmed from inventors' engagement in self-directed projects; this is recorded on the invention submission. It is possible that inventions arising from formal projects are more likely to be implemented. Thus, we controlled for *Official Project Invention*, which equals 1 if an invention stemmed from an official project and 0 otherwise. Third, it is possible that inventions closer to the technological frontier may have more opportunity to be integrated into industry standards, which can have implications for implementation. We measure this by looking at all ultimately patented inventions in each StarCo technology class that were submitted in a given year, and calculating the average age of the prior arts cited on those patents. We include *Frontier Technology*, which equals 1 if an invention is in a technology class-year where the average age of cited prior art is less than three years old, and 0 otherwise. (Note that

[12] StarCo developed its own technology classification scheme to identify an invention's technological area. The scheme is hierarchically organized into three levels and consists of 18 one-digit classes, roughly 120 two-digit classes, and more than 300 subclasses. We used the three-digit classes to calculate our control variables.

this variable, unlike the other controls, is based on technology class characteristics and not on the characteristics of the specific invention.) Fourth, the size of inventor teams may affect similarity of invention if a larger team incorporates more first-person knowledge of other inventions at StarCo. Therefore, we included ln(*Team Size*), measured as the natural log of (1 plus the number of inventors) to control for this effect. Finally, we included *Year* and *Technology Class* fixed effects to control for time-based fluctuations and technology-field-based heterogeneity.

3.4. Descriptive statistics

Table 1, Panel A shows the descriptive statistics and correlations among the key variables. Overall, StarCo filed patent applications for roughly 29% of all inventions. Roughly 23% of all filed inventions are implemented. The mean number of co-inventors on an invention is 1.6; although some inventions have as many as 14 inventors, the vast majority have either 1 or 2 inventors. The Table indicates only modest correlation among the variables.[13]

Panel B of Table 1 shows separate summary statistics for two subsets of StarCo inventions pre- and post-reorganization. "Centralized Group" consists of inventions from development activities that were ultimately moved to the newly established centralized unit in the newly established centralized unit, while "Reference Group" consists of those from development activities that remained in decentralized units. As this Panel indicates, the Centralized and Reference Groups are not dramatically different in terms of title cosine similarity or implementation rate during the pre-reorganization timeframe. However, post-reorganization, the Centralized Group exhibits a sharp drop in the rate of invention implementation, consistent with the prediction of H2, while the Reference Group exhibits an increase. The Centralized Group

[13] There is a moderate level of correlation between *New Technology Combinations* and *Number of Tech Classes* (.237). The Variance Inflation Factor (VIF) for one *Technology Class* dummy is 10.32, while VIFs for all other variables are below the suggested threshold of 10 with the maximum value being 7.40 (*Post Reorganization*) and the mean VIF being 2.90. Overall, multicollinearity problems are not likely to exist in our models.

demonstrates comparable title cosine similarity post vs. pre-reorganization, while the Reference Group exhibits an increase in similarity, indicating a small relative decrease in post-reorganization similarity for the Centralized Group. This is not inconsistent with H1, although it is not as compelling as the evidence for H2.

[Insert Table 1 Here]

Of course, these univariate comparisons do not take into account the many other factors that could affect invention. Since univariate comparisons obscure the impact of such factors, we therefore turn to multivariate econometric analysis.

4. ECONOMETRIC ANALYSIS

We estimated the relationship between centralization of development activities and duplication using the following model:

$$\begin{aligned} Outcome_{ijt} = \; & \alpha + \beta_1 Centralized\ Group_j + \beta_2 Post\ Reorganization_t \\ & + \beta_3 Centralized\ Group * Post\ Reorganization_{jt} \\ & + \gamma Invention\ Characteristics_{ijt} + \theta Tech + \phi Year + \varepsilon_{ijt} \qquad (1) \end{aligned}$$

where *Outcome* is either *Average Title Cosine Similarity* or *Implemented* depending on the model, *Invention Characteristics* is a vector of the control variables defined above, *Tech* is a vector of technology fixed effects, and *Year* is a vector of year fixed effects. The main coefficient of interest is β_3, which reflects the change in relationship between invention and outcome for inventions in the centralized development units after centralization occurs.

By estimating Equation (1), we leverage the longitudinal nature of the sample, as well as the fact that some development units remained decentralized after StarCo's reorganization, to compare the before-and-after changes of two groups of inventions. Econometrically, this approach is akin to a difference-in-differences regression. However, since StarCo's decision to centralize or decentralize particular development activities was not random, we are cautious about drawing strong causal inference from our findings.[14]

To mitigate this non-random assignment issue, we constructed two subsamples that are designed to encompass inventions from the Centralized and Reference Groups that are more comparable. First, we matched inventions in the Centralized Group with those in the Reference Group using a Coarsened Exact Matching approach (Iacus, King, & Porro, 2012). Specifically, for each invention in the Centralized Group we identified all inventions in the Reference Group that were in the same two-digit internal technology class, were submitted to StarCo's IPR office in the same calendar year, had the same number of inventors (1, 2, 3, and 4 or more), and had the same status of being either a formal project or not. This subsample totaled more than 9,000 inventions, with more than 5,000 from the Centralized Group and almost 4,000 from the Reference Group.

For our second subsample, we leveraged insight from our interviews with the former director of StarCo's IPR department. This individual noted that one subset of inventions developed by a particular Centralized-group team – hardware components – was technologically similar to the inventions developed in the Reference Group. This was true for two reasons. First, the majority of development effort by business units was focused on developing customized hardware features that would appeal to customer. Second, hardware development effort at both the decentralized and centralized units drew on similar technological expertise.[15] We therefore restricted the sample of

---

[14] For instance, our interviews indicate that development activities which were expected to produce more generic technologies were moved to the new central development unit, while those related to business-/product-specific technologies were more likely to remain in the decentralized development units. Similarly, development efforts that relied on technology standards that were used by multiple business units tended to be centralized, while development for business unit-specific standards tended to remain decentralized.

[15] "Yes, they [i.e., hardware component development efforts at centralized and decentralized units] would be very similar, because… From a skills and competencies and what they're trying to work on and the problems they're solving, yes, there's a lot of similarity between these special technology projects within business groups and the [hardware component centralized development teams]. The only reason they're inside the business group as opposed to being in [centralized development unit] is because the business group set them up because there was something missing from the [centralized development unit]." (Former director, IPR Department, StarCo.)

Centralized-group inventions to only those developed by the team responsible for generating hardware components, while continuing to include all Reference-group inventions.[16] This subsample totaled more than 6,000 inventions, with almost 1,500 from the Centralized Group and more than 4,500 from the Reference Group.

Appendix Tables A2 and A3 present the summary statistics for the matched-invention subsample and the hardware components-only subsample, respectively. As is evident in these tables, the Centralized Group and the Reference Group exhibit much more similar means in these subsamples. Appendix Figures A1-A3 show the evolution of our dependent variables over time. Although noisy and clearly not definitive, Figure A1 indicates that the gap between the Centralized and the Reference Groups for *Avg. Title Cosine Similarity* tended to increase after the reorganization, while exhibiting relatively little difference in time trend beforehand, especially for the hardware components subsample. Figures A2 and A3 provide comparable suggestive evidence for *Implemented*.

There are two possible methodological concerns regarding estimation of Equation (1) when *Implemented* is the dependent variable. First, since *Implemented* is a binary variable, OLS does not offer unbiased point estimates. One alternative is to use probit or logit estimation; however, our core explanatory variable is an interaction term, and it can be difficult to interpret interaction terms in non-linear models (Ai & Norton, 2003). To address this, we follow recent practice by presenting both probit and OLS estimates.

Second, since *Implemented* is observable only for those inventions for which patent applications were filed, there is the possibility of selection bias or distortion if different StarCo divisions have different criteria for seeking patent protection. For example, if the Centralized

[16] Our data also indicate that before the reorganization, more than 15% of the inventions from the hardware-component development group were co-developed with inventors in the Reference Group.

Group used a lower threshold of perceived quality for deciding to pursue patent protection than did the Reference Group, then the filed inventions of the Centralized Group might be of lower quality on average; if the likelihood of implementation is correlated with invention quality, then the Centralized Group would have a lower rate of Implementation simply due to the lower threshold for pursuing patent protection. Throughout the sample time period, StarCo relied on a single centralized IPR department to make patent-application decisions regarding inventions from all units in the firm, which may mitigate this concern to some degree. To address this concern further, we estimate both the OLS and probit models via a Heckman two-stage approach in which the first stage addresses the likelihood that a patent application will be filed for a given invention. (We provide the single-stage non-corrected "naïve" models in the Appendix for the interested reader.) As exclusion restrictions we used two variables which affect the likelihood that a patent application is filed for an invention but are not related to other characteristics of the invention. StarCo's IPR department employs a set of "patent experts" whose job is to determine which submitted inventions warrant a patent application. Each expert is assigned a varying number of invention disclosures for which she needs to formulate a recommendation regarding pursuit of patent protection. The allocation of an invention disclosure to a patent expert is not a function of the quality or complexity of the invention, but is primarily driven by the technical profile of the patent expert (and secondarily her workload). Our first exclusion restriction measures the number of other inventions assigned to the patent expert at the time she is assigned the focal invention (ln*Patent Expert Workload*). If the workload of a patent expert is high, she has less time and fewer cognitive resources for evaluating a focal invention and is less likely to pursue a patent application for the invention (a similar instrument was used in Criscuolo et al., 2019 and Frakes & Wasserman, 2017). The other exclusion restriction (*Diligent*) is related to how meticulous patent experts are in performing their job. The StarCo IP system contains several free-text boxes where information evaluating a submitted invention should be recorded, but where text entry is optional. We identify the proportion of times that a given patent expert has filled in one of the free-text boxes across all of the inventions that she evaluates as a measure of her meticulousness. This variable captures an individual trait of a patent expert and therefore it does not vary with the quality of invention disclosure having been assessed.

4.1. Results

Table 2 shows the results of the analyses testing H1, which predicts that centralization of development reduces duplication in development outcomes. All three models in Table 2 feature *Avg. Title Cosine Similarity* as a dependent variable, but use different samples as described above. We expect that centralization of development negatively influences the average similarity between the title of a focal invention and those of all inventions submitted prior to the focal invention. In Model 1, we test this relationship using all inventions in the Centralized and the Reference Groups. The coefficient on the main independent variable (*Post Reorganization X Centralized Group*) is negative (b=-0.015, p=.011), lending support to H1. Model 2 replicates this estimation on the sample of matched inventions. The coefficient on *Post Reorganization X Centralized Group* retains the same sign and magnitude (b=-0.015, p=.051). Model 3 replicates the estimation on the "Hardware Components" sample; the coefficient on our chief explanatory variable remains negative and more than doubles in magnitude (b=-0.039, p=.001). Overall, the models in Table 2 provide consistent support for H1. The results are economically meaningful, as the coefficients account for roughly 9% to 19% of the sample means of *Avg. Title Cosine Similarity*.

[Insert Table 2 Here]

H2 predicted that inventions generated by a centralized development unit would be less likely to be implemented in a commercialized product than those of decentralized units. As described above, we have implementation information only for inventions for which patent applications are filed, and we use a two-stage approach to this estimation in order to address potential sample selection bias concerns around "selection" into being filed. Table 3 reports the results of the second-stage regressions, while Appendix B1 reports the results of the first stage. As shown in Appendix B1, the Wald test of independent equations indicates that ρ, the correlation between error terms of the first-stage and the second-stage regressions, is different from zero. This is a precondition for sample selection bias to exist (Certo, Busenbark, Woo, & Semadeni, 2016), thus indicating the presence of sample selection bias and justifying the use of Heckman models. Relatedly, the coefficients on our two instruments are in the predicted directions and of substantial magnitude (p-values<.001 in all models), suggesting that these satisfy the exclusion restriction.

The models reported in Table 3 examine the effect of centralization of development on *Implemented* with the different samples as described in the previous section. We present three probit estimations and three OLS estimations. In all three probit models, the coefficient on *Post Reorganization X Centralized Group* is negative and of substantial magnitude, ranging from -0.269 to -0.566 (with p-values ranging from .000 to .004); as in Table 2, the magnitude is greatest for the "Hardware Components" sample. To evaluate the magnitude of the effects on implementation, we computed predicted probabilities of implementation for inventions in each of the four subsamples (i.e. *Centralized Group* = 1 & *Post Reorganization* = 0; *Centralized Group* = 1 & *Post Reorganization* = 1; *Centralized Group* = 0 & *Post Reorganization* = 0; *Centralized Group* = 0 & *Post Reorganization* = 1). The details of this procedure are provided alongside Table B3 in the Appendix. As that Table shows, the post-reorganization probability that a filed invention is implemented fell for the Centralized Group by roughly between 2 and 13 percentage points, depending on the sample. This probability stayed almost the same or even increased for the Reference Group, making the net effect range between -6 and -11 percentage points, an effect of substantial economic significance. The OLS models provide similar results, with the coefficient on

Post Reorganization X Centralized Group ranging from -0.071 to -0.141 (with *p*-values ranging from .002 to .016), again with the magnitude greatest for the "Hardware Components" sample. The coefficients indicate that as compared to the Reference Group, inventions from the Centralized Group are 7 to 14 percentage points less likely to be implemented post-reorganization. These results thus provide consistent evidence that is supportive of H2.[17]

[Insert Table 3 Here]

4.2. Additional Analyses

Our analyses for H1 used the similarity in titles as a dependent variable (DV). However, this variable may not fully square with the construct of interest (Duplication) since the similarity in titles may be driven by, for example, the use of common words. To mitigate this concern, we conducted additional analyses using alternative operationalizations of duplication.

An alternative measure of duplication leverages additional information stored in StarCo's IP data management system. Specifically, when a patent expert evaluates a submitted invention to determine whether it warrants a patent application, she typically searches the system for previous inventions submitted to StarCo that preclude claiming novelty for the focal invention as well as conducting a more extensive search for existing patents owned by other companies. These searches are recorded in the IP system. From the texts of these searches, we created a dummy variable (*Previous Invention*) equal to 1 if a patent expert found at least one such internal invention.[18] The existence of these previous inventions that preclude claiming novelty indicates that there was duplication in development effort. Indeed, 85% of inventions in our sample where a previous

[17] Appendix Table B2 presents results of a "naïve" single-stage estimation of *Implemented*. The coefficients on *Post Reorganization X Centralized Group* are consistently of the same sign but with smaller magnitude than those in the Heckman-corrected Table 3 estimations (while the standard errors retain their magnitude). Hence, the two-stage regression generates more precise point estimates, while yielding the same general qualitative insights as the naïve model.

[18] An example from the data: "*The general concept of [redacted for confidentiality reasons] is a known concept e.g. from [Invention XXXX] and [Invention YYYY].*"

invention was found were not put forward for patent filing. Patent experts do not always search for prior arts or record their search results in the database, and such information is missing especially for inventions submitted in earlier years. Therefore, when *Previous Invention* is used as a DV, we limit our analysis to the sample of inventions submitted one year before and after the reorganization.

Relatedly, another way that reduced duplication in development influences innovative outcomes is through an increase in the patentability of inventions.[19] Indeed, one of the reasons that patent experts often cite for not filing for patent protection is because the firm already possesses similar patented inventions. Looking at the feedback provided by patent experts to inventors to communicate the outcome of their assessment, we found that roughly 28% of comments for non-patenting decisions mentioned other inventions previously submitted to StarCo. Conversely, the fact that a patent application is filed for an invention shows that no previous invention was detected and thus no duplication occurred, as distinctiveness from prior inventions is a necessary condition for granting a patent to an invention. Therefore, less duplication in development should lead to a greater likelihood that patent applications are filed for new inventions. Hence, our second measure of duplication is a dummy variable (*Recommended Patent Application*) equal to 1 if the IPR department recommended that a focal invention be filed for patent protection and 0 otherwise.

Third, centralization of development not only reduces duplications, but it should also lead to development of standardized technologies that can be used in multiple business units. To measure whether an invention is related to standardized technologies, we leverage an item in the invention disclosures which asks inventors to specify if their inventions can be proposed to industry standards. We constructed a dummy variable (*Industry-Standard Invention*) as 1 if its inventors

[19] Put differently, although inventions from centralized and decentralized development face the same threshold for warranting patent application, a greater proportion of centrally developed inventions are likely to warrant patent application because their novelty is less likely to be preempted by a previous StarCo invention.

checked 'Yes' on this item, and 0 otherwise. Although industry standards are conceptually different from internal standards (i.e. technology that can be used by multiple business units of the firm), inventions related to industry standards are a subset of inventions related to internal standards. In other words, industry standards are used by all business units of the firm (thus, they are also internal standards), but not all internal standards will become industry standards. Therefore, this measure underrepresents inventions related to internal standards in the sample. As our hypothesis predicts a positive association between centralization of development and internal standards, the use of this variable as a proxy for internal standards makes it more difficult to find the predicted relationship.

We tested the relationship between centralization of development (*Post Reorganization X Centralized Group*) and these alternative dependent variables using probit regressions. The same set of control variables as in the main analyses are included in these models. Also included in the models examining the effect on *Recommended Patent Application* are *Patent Expert Workload (log)* and *Diligent* as these variables would influence the likelihood of patent application filing. The results of the additional analyses are shown in Tables C1-C3 in the Appendix. Table C1 reports the results for the analyses where *Previous Invention* is used as a dependent variable. For Models 1, 3, and 5, the samples included inventions without prior-art-search texts and imputed 0 to *Previous Invention* for these inventions. This imputation assumes that texts for prior-art searches are missing for these inventions because patent experts have not found any relevant prior art. Supporting this assumption, the proportion of inventions given patenting decisions is more than 10% points higher for inventions without prior-art search texts (32.88%) than those having texts (21.92%). The coefficients of the main independent variable are negative in all three models (Model 1: b=-0.307, p=.009; Model 3: b=-0.316, p=.059; Model 5: b=-0.431, p=.042). In Models 2, 4, and 6, we excluded the inventions without prior-art-search texts from the sample and only used inventions that have texts for prior-art searches. The coefficients of the main independent variable remain negative and of similar magnitude as in models 1, 3, and 5, although the standard errors are substantially larger presumably due to the smaller sample size (Model 2: b=-0.259, p=0.077; Model 4: b=-0.327, p=0.109; Model 6: b=-0.371, p=0.151). Overall, these results suggest that the likelihood that internal prior arts are found for a focal invention decreases after centralization.

Table C2 reports the results for the analyses where *Recommended Patent Application* is the dependent variable. The table shows positive coefficients of the main IV in all three models (Model 1: b=0.253, p=.016; Model 2: b=0.424, p<.001; Model 3: b=0.468, p<.001), indicating that centralization of development has a positive impact on the likelihood that inventions are filed for patent protection thanks to the reduced duplication. Table C3 reports the results for the analyses where *Industry-Standard Invention* is the dependent variable. The coefficients of the main IV are positive in all models (Model 1: b=0.600, p=0.001; Model 2: b=0.402, p=.011; Model 3: b=0.575, p =.001). These results lend support to the prediction that centralization of development positively influences the likelihood that inventions are related to internal standards.

We also tested the robustness of H2 with additional analyses. First, we excluded essential patents from *Implemented* and reran the regressions. The results are consistent with the main analyses: in Table D1 in the Appendix, the coefficients of *Post Reorganization X Centralized Group* are negative across the samples (for 2nd stage probit models, b=-0.298, p=.006 for All Inventions; b=-0.450, p=.007 for "Matched Inventions"; b=-0.522, p=.006 for "Hardware Components"; for 2nd-stage OLS models, b=-0.067, p=.016 for All Inventions; b=-0.107, p=.011 for "Matched Inventions"; b=-0.113, p=.007 for "Hardware Components").

Second, we repeated the analyses for subsamples for which the theory predicts the effect of centralization on implementation would be greater. Specifically, it is expected that development activities that need to be tightly aligned with market needs would be most disrupted by centralization. We chose two such subsamples. The first subsample consists of inventions from two of StarCo's internal technology classes: one is related to "visible to end-users" technologies and the other to "user interface, usability, user experience." Inventions in these technology classes are directly visible to end-users and thus inputs from downstream functions and customers are crucial. The second subsample is composed of inventions that are submitted to User-interface (UI) internal patent board of StarCo. The subsample analyses revealed evidence in line with our expectations: in Table D2 in the Appendix, the absolute values of the coefficients of *Post Reorganization X Centralized Group* are substantially larger for both End-user technology inventions and UI patent-board inventions (b=-0.602, p=.001 for End-user technology inventions; b=-1.093, p=.002 for UI patent-board inventions) than their counterparts (b=-0.142, p=.206 for Non end-user technology inventions; b=-0.172, p=.086 for Non-UI patent-board inventions) and Chow tests indicate that the coefficients are different between the groups (χ^2=4.63, p=.031 for End-user vs. Non end-user; χ^2=6.57, p=.010 for UI patent board vs. Non-UI patent boards).

Third, we also looked at whether inventions in the Centralized Group took more time to be implemented, which would indicate the difficulties associated with commercializing development outcomes from the centralized unit. Our dataset contains information on dates when inventions were submitted, when decisions regarding patent filings were made, and (for inventions filedf for patent protection) when inventions were implemented. Table D3 in the Appendix shows the average time to implementation for those Centralized vs. Reference Group inventions that were ultimately implemented during our sample frame, both before and after the reorganization. Pre-reorganization, implementation of innovations produced by the Centralized Group took 109-119 days longer than those produced by the Reference Group. This pre-reorganization difference may reflect that the re-organization effort did not randomly select development activities for centralization. Of more direct interest here, after the reorganization, implementation of innovations produced by the Centralized Group took 122-139 days longer than those produced by the Reference

Group, reflecting an increase of 12%-17% in the excess time difference.

5. DISCUSSION AND CONCLUSION

As noted earlier, our study makes three contributions: phenomenological, theoretical, and empirical. Whereas the bulk of prior work on organization of R&D focuses on *research* activities, we draw attention to *development* activities. By explicating the distinct features of development, we extend theory to illuminate centralization/decentralization of development is associated with different outcomes than the convention outcomes for research. And, by exploiting a novel dataset with unusually rich data on a multi-business corporation's inventive activities, we are able to link formal organization structure of R&D to commercialization outcomes rather than purely innovation outcomes. Through these contributions, our study informs the literature on R&D organization structure and innovation. Whereas previous studies have found that centralization of research enables corporate headquarters to overcome business units' reluctance to invest in projects with internal spillovers and insulates research personnel from immediate market pressure – thus facilitating broader innovative search and innovative outcomes (albeit at a cost of less BU-specific knowledge generation) – our study reveals that centralization of development enables corporate headquarters to reduce business units' tendency to invest in excessive duplicative development, albeit at the cost of forgoing inventions that are optimized for specific business units in favor of inventions that "satisfice" across multiple BUs.

We also inform the broader literature on corporate strategy by revisiting and extending the longstanding debate on the fit between corporate strategy and structure. Early studies suggested that different types of diversification require different organization structures, and in particular that related diversification requires centralized structures to realize latent synergies (e.g., Hoskisson et al., 1993). But recent research that explores private information and managerial incentives suggests that, under certain circumstances, centralization can be detrimental to effective coordination of interdependent tasks among complementary functional activities (Chen, Kaul, & Wu, 2019; Qian, Roland, & Xu, 2006). Our findings support and extend both views. The fact that centralization of development helps a multi-business firm reduce duplication of development effort reinforces the synergy-enabling effect of centralization: it is not relatedness among a firm's businesses *per se*, but rather the firm's use of an appropriate organization structure, that facilitates the realization of related diversification's synergistic potential. By centralizing development and coordinating technical requirements of different businesses, a firm can realize economies of scope in development. That said, this study demonstrates that there are costs to centralization of development for such a firm. Specifically, centralization decreases the likelihood that a given invention is implemented by business units into their products. Our study thus suggests that there is no universally optimal organization structure for relatedly diversified firms, thus extending recent work that has identified relevant tradeoffs.

This study also generates implications for two other streams of research. The first relates to innovation around product platforms. Prior studies of product platforms have demonstrated numerous benefits of common architectures and components shared across multiple product lines (Gawer & Cusumano, 2014). However, it has devoted less attention to the factors that enable firms to create such platforms. Although a handful of conceptual studies have proposed that centralized decisionmaking is useful (and perhaps even necessary) to fully realize the benefits of product platforms (Magnusson & Pasche, 2014; Robertson & Ulrich, 1998), there are no extant large-scale empirical tests of this proposition. Our subsidiary analyses, presented in Appendix Table C3, provide initial support for this proposition: StarCo's central development unit tended to produce standardized technologies that could be shared across multiple product lines. At the same time, the fact that these inventive outputs were less likely to be included in commercialized products highlights potential downsides of centralizing to support platform-based product development.

This study may also be relevant to the literature on overinvestment in R&D. A recent paper by Ahuja and Novelli (2017) draws attention to the possibility of such overinvestment, and strives to lay out a range of factors that may serve as triggers for such behavior. Notably, the factors discussed in that study relate to technology or market features, such as uncertainty of R&D outcomes or market demand, or to sociological/behavioral features, such as the "legitimacy" of R&D activity or adherence to particular status-quo-like patterns of resource allocation. This study highlights an organizational design reason why firms may overinvest in the "D" part of R&D: incentives of BU managers to invest in duplicative development activities to develop locally-optimal inventions. More generally, this result suggests that there may be a range of organizational features that encourage overinvestment. Future research might explore such features in detail.

5.1. Limitations and Future Research

Although our study benefited greatly from access to detailed data of a firm's invention disclosures and their implementations into products, the single-firm setting invites questions about the generalizability of our findings. Given that StarCo was a leading innovator in the ICT industry where most firms invest a substantial portion of revenue into R&D, our results are likely applicable to R&D-intensive companies that have a portfolio of related product lines. Nonetheless, future research can benefit from testing the predictions suggested in this study with a larger sample of firms from a diverse range of industries. Perhaps the effect of organization structure on development outcomes varies by industry. For instance, in stable industries where the rate of change in technological and market environments is low, a centralized development unit may be able to produce technologies that business units are willing to implement into their products, thus ameliorating the potential negative effect of centralization. More generally, extending this research to other settings can help delineate boundary conditions for the hypothesized relationships.

Another limitation of our study is that we could not directly document the mechanisms through which organization structure of development influences innovative outcomes. For instance, centralization may reduce duplication because it facilitates coordination of development activities for different business units by (a) restricting the autonomy of business units which prefer proprietary development projects and (b) aggregating and specializing in processing information on development requirements. Although these mechanisms appear plausible, are well established in the relevant literature, and are consistent with the qualitative information that we compiled, we could not empirically test them with the available data. Future studies might extend this line of research by conducting moderation analyses using variables that are likely to correlate with the corporate headquarters' ability to exert control over development activities and transfer or process dispersed information and knowledge. For instance, as a firm's official development projects are more likely to be subject to formal control mechanisms, the effect of organization structure governing development projects would appear more strongly in the official projects. In addition, researchers may conduct deep qualitative studies to understand precisely how a central development unit makes decisions on which development projects to pursue and how it transfers and processes information from business units.

Finally, and more speculatively, we envision two additional avenues for future research. First, one intriguing result from our analysis is that, while StarCo's reorganization was associated with a decrease in the rate of implementation of inventions for the centralized development unit (as predicted), the implementation rate for those development units that remained decentralized actually increased.[20] This positive effect of reorganization on non-centralized development activities could be an artifact of changes in the external environment or in other R&D policies of

[20] This is visible in Table 3 and Appendix Table B3. In Models 4-6 in Table 3, there are positive direct effects of *Post Reorganization* on implementation. In Table B3, while the likelihood of implementation decreased by 2.6% for inventions in the Centralized Group (as predicted), it increased slightly for Reference-Group inventions.

the firm. However, it could stem from better division and allocation of tasks post-reorganization: if the remaining non-centralized development tasks are truly BU-specific and are no longer conflated with more general development, then perhaps inventors in these units may have become more effective in utilizing local information and adapting to market needs. Prior academic research has found that relocating some activities outside a firm's boundaries may improve the performance of the remaining in-house activities (e.g., Azoulay, 2004); our result might extend this to considering relocation of activities within the organization. Managerially, this may suggest that managers should consider centralization or decentralization of specific tasks rather than entire functions, in order to reap the benefits of appropriate organization on a task level. We thus speculate that future research on these issues would be beneficial.

A second speculative thought relates to interactions between research and development activities. As work in innovation management has noted, the handoff from research to development can be challenging, given the different incentives and goals of each (Chiesa & Frattini, 2007). Much of that work invokes an assumption that research tends to be run corporately while development tends to report to business units, implying that such different reporting arrangements hamper relations between research and development. Given this, one reasonable question arising from our study is: what happens to research-development relations when development is centralized instead of decentralized? Would this facilitate tighter integration between the two activities? We hope to see, or conduct, future research that can illuminate this issue.

Regardless of the direction of future research in this area, we are confident that the organization structure of development activities, as distinct from research activities, will prove to be a fruitful avenue of research for years to come.

REFERENCES

Ahuja, G., & Novelli, E. (2017). Activity overinvestment: The case of R&D. *Journal of Management* 43(8): 2456–2468.

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters* 80(1): 123–129.

Alonso, R., Dessein, W., & Matouschek, N. (2008). When does coordination require centralization? *American Economic Review* 98(1): 145–179.

Argyres, N. S. (1995). Technology strategy, governance structure and interdivisional coordination. *Journal of Economic Behavior & Organization* 28(3): 337–358.

Argyres, N. S., & Silverman, B. S. (2004). R&D, organization structure, and the development of corporate technological knowledge. *Strategic Management Journal* 25(8-9): 929–958.

Argyres, N. S., Rios, L. A., & Silverman, B. S. (2020). Organizational change and the dynamics of innovation: Formal R&D structure and intrafirm inventor networks. *Strategic Management Journal* 41(11): 2015–2049.

Arora, A., Belenzon, S., & Patacconi, A. (2018). The decline of science in corporate R&D. *Strategic Management Journal* 39(1): 3–32.

Arora, A., Belenzon, S., & Rios, L. A. (2014). Make, buy, organize: The interplay between research, external knowledge, and firm structure. *Strategic Management Journal* 35(3): 317–337.

Azoulay, P. (2004). Capturing knowledge within and across firm boundaries: Evidence from clinical development. *American Economic Review* 94(5): 1591–1612.

Barge-Gil, A., & López, A. (2015). R versus D: estimating the differentiated effect of research and development on innovation results. *Industrial and Corporate Change* 24(1): 93–129.

Campbell, A., Kunisch, S., & Müller-Stewens, G. (2011, June 1). To centralize or not to centralize? *McKinsey Quarterly* 97–102.

Cardinal, L. B. (2001). Technological innovation in the pharmaceutical industry: The use of organizational control in managing research and development. *Organization Science* 12(1): 19–36.

Cardinal, L. B., Turner, S. F., Fern, M.J., & Burton, R.M. (2011). Organizing for product development across technological environments: Performance trade-offs and priorities. *Organization Science* 22(4): 1000–1025.

Certo, S. T., Busenbark, J. R., Woo, H. S., & Semadeni, M. (2016). Sample selection bias and Heckman models in strategic management research. *Strategic Management Journal 37*(13) 2639–2657.

Chen, M., Kaul, A., & Wu, B. (2019). Adaptation across multiple landscapes: Relatedness, complexity, and the long run effects of coordination in diversified firms. *Strategic Management Journal* 40(11): 1791–1821.

Chesbrough, H., & Rosenbloom, R. S. (2002). The role of the business model in capturing value from innovation: evidence from Xerox Corporation's technology spin-off companies. *Industrial and Corporate Change* 11(3): 529–555.

Chiesa, V. (2001). *R&D strategy & organisation: Managing technical change in dynamic contexts*. London: U.K.: Imperial College Press.

Chiesa, V., & Frattini, F. (2007). Exploring the differences in performance measurement between research and development: evidence from a multiple case study. *R&D Management* 37(4): 283–301

Criscuolo, P., Alexy, O., Sharapov, D., & Salter, A. (2019). Lifting the veil: Using a quasi-replication approach to assess sample selection bias in patent-based studies. *Strategic Management Journal* 40(2): 230–252.

Dessein, W., Garicano, L., & Gertner, R. (2010). Organizing for synergies. *American Economic Journal: Microeconomics* 2(4): 77–114.

Dessein, W., Lo, D. & Minami, C. (in press). Coordination and organization design: Theory and micro-evidence. *American Economic Journal: Microeconomics*.

Doi, A. (2001). Honda turns its back on industry gospel that fewer platforms equals higher profits. *The Japan Automotive Digest*. VII(15): 1.

Du, J., Leten, B., Vanhaverbeke, W., & Lopez-Vega, H. (2014). When research meets development: antecedents and implications of transfer speed. *Journal of Product Innovation Management* 31(6), 1181–1198.

Eggers, J. P. (2016). Reversing course: Competing technologies, mistakes, and renewal in flat panel displays. *Strategic Management Journal* 37(8): 1578–1596.

Eklund, J. C. (in press). The knowledge-incentive trade-off: Understanding the relationship between organization design and innovation. *Strategic Management Journal*.

Fleming, L. (2001). Recombinant uncertainty in technological search. *Management Science* 47(1): 117–132.

Foulkes, N. (2018). Richemont attacks R&D overlap with centralising drive. *Financial Times*. January 18. Retrieved March 15, 2022 from https://www.ft.com/content/3707be66-daab-11e7-9504-59efdb70e12f?accessToken=zwAAAX-d9iG5kc83B75m2qsR59OVBFnv23DhLw.MEQCIC4MlbNSPivpHSVtIhB0vKRc3pPyOw7IHEPewQKg2tq_AiANuCvLkG3Skniu4IjUp8zKoXjdlg_JjQbS2Dwco79mww&sharetype=gift?token=4311cf7b-8d55-4497-a3de-b0bcbf6baee7

Frakes, M. D., & Wasserman, M. F. (2017). Is the time allocated to review patent applications inducing examiners to grant invalid patents? Evidence from micro-level application data. *Review of Economics and Statistics* 99(3): 550–563.

Gawer, A., & Cusumano, M. A. (2014). Industry platforms and ecosystem innovation. *Journal of Product Innovation Management* 31(3): 417–433.

Guadalupe, M., Li, H., & Wulf, J. (2014). Who lives in the C-suite? Organizational structure and the division of labor in top management. *Management Science* 60(4): 824–844.

Guler, I. & A. Nerkar. (2012). The impact of global and local cohesion on innovation in the pharmaceutical industry. *Strategic Management Journal* 33(5): 535–549.

Hoskisson, R. E., Hitt, M. A. & Hill, C. W. L. (1993). Managerial incentives and investment in R&D in large multiproduct firms. *Organization Science* 4(2): 325–341.

Hounshell, D. A., & Smith, J. K. (1989). *Science and corporate strategy: DuPont R&D, 1901–1980*. New York: Cambridge University Press.

Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis* 20(1), 1-24.

Karim, S., & Kaul, A. (2015). Structural recombination and innovation: Unlocking intraorganizational knowledge synergy through structural change. *Organization Science* 26(2): 439–455.

Katz, R., & Allen, T. J. (1982). Investigating the Not Invented Here (NIH) syndrome: A look at the performance, tenure, and communication patterns of 50 R&D project groups. *R&D Management* 12(1): 7–20.

Kay, N. (1988). The R&D function: Corporate strategy and structure. In G. Dosi, C. Freeman, R. Nelson, G. Silverberg, & C. Soete (Eds.), *Technical change and economic theory* (pp. 282–294). London, U.K.: Pinter.

Leifer, R., & Triscari, T. (1987). Research versus Development: Differences and similarities. *IEEE Transactions on Engineering Management* 34(2): 71–78

Lerner, J. & Wulf, J. (2007). Innovation and incentives: Evidence from corporate R&D. *Review of Economics and Statistics* 89(4): 634–644.

Magermann, T., Van Looey, B., & Debackere, K. (2015). Does involvement in patenting jeopardize one's academic footprint? An analysis of patent-paper pairs in biotechnology. *Research Policy* 44(9): 1702–1713.

Magnusson, M., & Pasche, M. (2014). A contingency-based approach to the use of product platforms and modules in new product development. *Journal of Product Innovation Management* 31(3): 434–450.

Moreira, S., Markus, A., & Laursen, K. (2018). Knowledge diversity and coordination: The effect of intrafirm inventor task networks on absorption speed. *Strategic Management Journal* 39(9): 2517–2546.

National Science Foundation. (2020). *National Patterns of R&D Resources: 2017–18 Data Update*. NSF 20-307. Alexandria, VA. Available at https://ncses.nsf.gov/pubs/nsf20307.

Nelson, R. R. (1959). The simple economics of basic scientific research. *Journal of Political Economy* 67(3): 297–306.

Office of Industries, U.S. International Trade Commission. (2002). *Industry and Trade Summary: Motor Vehicles*. USITC Publication 3545, September.

Qian, Y., Roland, G., & Xu, C. (2006). Coordination and experimentation in M-form and U-form organizations. *Journal of Political Economy* 114(2): 366–402.

Rantakari, H. (2008). Governing adaptation. *The Review of Economic Studies* 75(4): 1257–1285.

Rivkin, J. W., & Siggelkow, N. (2003). Balancing search and stability: Interdependencies among elements of organizational design. *Management Science* 49(3): 290–311.

Robertson, D., & Ulrich, K. (1998). Planning for product platforms. *Sloan Management Review* 39(4): 19–31.

Schilling, M. A., & Phelps, C. C. (2007). Interfirm collaboration networks: The impact of large-scale network structure on firm innovation. *Management Science* 53(7): 1113–1126.

Singh, J. (2005). Collaborative networks as determinants of knowledge diffusion patterns. *Management Science* 51(5): 756–770.

ter Wal, A.L.J., Criscuolo, P., McEvily, B. & Salter, A. (2020). Dual networking: How collaborators network in their quest for innovation. *Administrative Science Quarterly* 65(4): 887–930.

Teece, D. J. (1996). Firm organization, industrial structure, and technological innovation. *Journal of Economic Behavior and Organization* 31(2): 193–224.

Thomas, C. (2011). Too many products: Decentralized decision making in multinational firms. *American Economic Journal: Microeconomics* 3(1): 280–306.

Toh, P. K. (2014). Chicken, or the egg, or both? The interrelationship between a firm's inventor specialization and scope of technologies. *Strategic Management Journal* 35: 723–738.

Treece, J., & Serefkin, R. (2001, March 5). Platform sharing key to profits. *Automotive News*. p. 53.

Williamson, O. E. (1985). *Economic institutions of capitalism*. New York: Free Press.

Zahra, S. A., & Neilson, A. P. (2002). Sources of capabilities, integration and technology commercialization. *Strategic Management Journal* 23(5): 377–398.

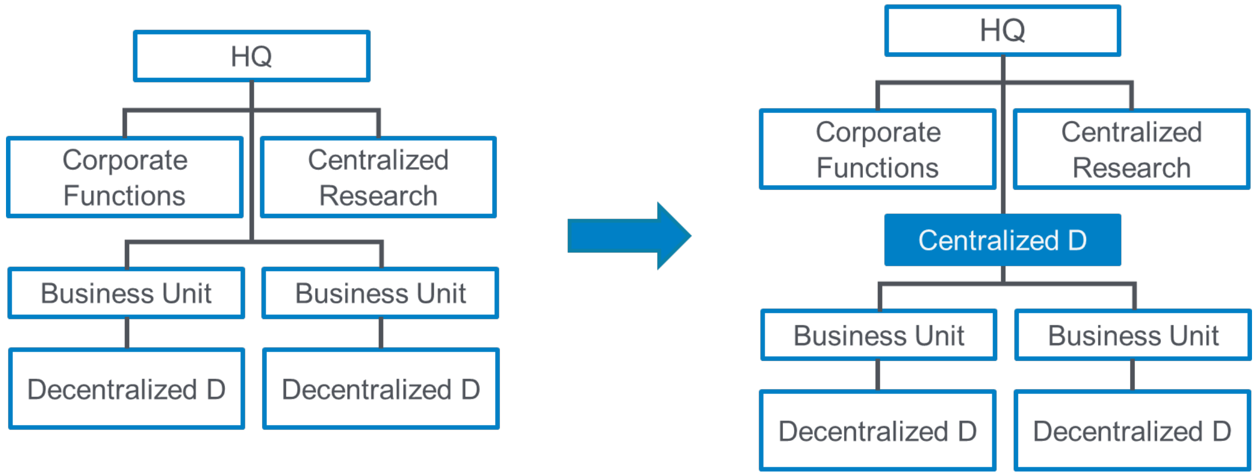


FIGURE 1 Graphical illustration of StarCo's restructuring of R&D organization

Step 1: Lemmatizing the title

Invention	Patent Number	Assignee	Title	Lemmatized
A	US1029725482	Google LLC	Task initiation using long-tail voice commands by weighting strength of association of the tasks and their respective commands based on user feedback	task initiation use long-tail voice command weight strength association task respective command base user feedback
B	US1049019082	Google LLC	Task initiation use sensor dependent context long-tail voice command	task initiation use sensor dependent context long-tail voice command

Step 2: Calculating cosine similarity (note: words are arranged alphabetically here)

Invention	association	base	command	context	dependent	feedback	initiation	long	respective	sensor	strength	tail	task	use	user	voice	weight
A	1	1	2	0	0	1	1	1	1	0	1	1	2	1	1	1	1
B	0	0	1	1	1	0	1	1	0	1	0	1	1	1	0	1	0

Cosine similarity = $\frac{A \cdot B}{\|A\|\|B\|}$ = 0.636

FIGURE 2 Example of process to lemmatize titles and calculate title cosine similarity

TABLE 1 Descriptive Statistics and Correlations

Panel A: Summary statistics and correlations for entire sample

	Variables	Obs.	Mean	Min	Max	1	2	3	4	5	6	7
1	*Avg. Title Cosine Similarity*	+12,000	.162	0	1							
2	*Implemented*	+3,000	.229	0	1	-.004						
3	*New Technology Combinations*	+12,000	.025	0	1	-.002	.011					
4	*Number of Tech Classes*	+12,000	1.502	1	7	.006	.034	.238				
5	*Official Project Invention*	+12,000	.567	0	1	-.048	.102	.003	.100			
6	*Frontier Technology*	+12,000	.261	0	1	-.014	.059	-.014	.127	.031		
7	*Team Size (log)*	+12,000	0.959	0.693	2.708	-.023	.125	-.007	.103	.166	.051	
8	*Recommended Pat. Appl.*	+12,000	.288	0	1	-.079	-.235	.063	.174	.132	.128	.201

Panel B: Summary statistics for the four distinct subsamples

	Centralized group, pre-reorganization		Centralized group, post-reorganization		Reference group, pre-reorganization		Reference group, post-reorganization	
Variables	Obs.	Mean	Obs.	Mean	Obs.	Mean	Obs.	Mean
Avg. Title Cosine Similarity	+2,000	.152	+4,000	.156	+1,000	.162	+2,000	.180
Implemented	+1,000	.247	+1,000	.220	+500	.219	+500	.226
New Technology Combinations	+2,000	.034	+4,000	.019	+1,000	.035	+2,000	.019
Number of Tech Classes	+2,000	1.712	+4,000	1.419	+1,000	1.588	+2,000	1.374
Official Project Invention	+2,000	.584	+4,000	.612	+1,000	.444	+2,000	.559
Frontier Technology	+2,000	.472	+4,000	.135	+1,000	.476	+2,000	.119
Team Size (log)	+2,000	0.972	+4,000	0.989	+1,000	0.945	+2,000	0.908

TABLE 2 Duplication (*Avg. Title Cosine Similarity*) as a function of centralization of development

	Sample: All Inventions	**Sample: Matched Inventions**	**Sample: Hardware Component**
	(1)	**(2)**	**(3)**
Post Reorganization x Centralized Group	-0.015	-0.015	-0.039
	(0.006)	(0.008)	(0.012)
Post Reorganization	0.028	0.025	0.035
	(0.008)	(0.011)	(0.013)
Centralized Group	-0.006	-0.001	0.000
	(0.005)	(0.006)	(0.010)
New Technology Combinations	-0.004	-0.016	-0.009
	(0.010)	(0.013)	(0.017)
Number of Technology Classes	0.005	0.007	0.002
	(0.002)	(0.003)	(0.003)
Official Project Invention	-0.013	-0.015	-0.005
	(0.003)	(0.004)	(0.005)
Frontier Technology	0.002	-0.005	0.008
	(0.004)	(0.005)	(0.007)
Team Size (log)	-0.003	0.003	0.007
	(0.004)	(0.006)	(0.007)
Technology Class fixed effects	Included	Included	Included
Year fixed effects	Included	Included	Included
Constant	0.127	0.137	0.163
	(0.011)	(0.017)	(0.022)
N	12,000+	9,000+	6,000+
R-squared	.013	.012	.013
Adjusted R-squared	.011	.009	.009

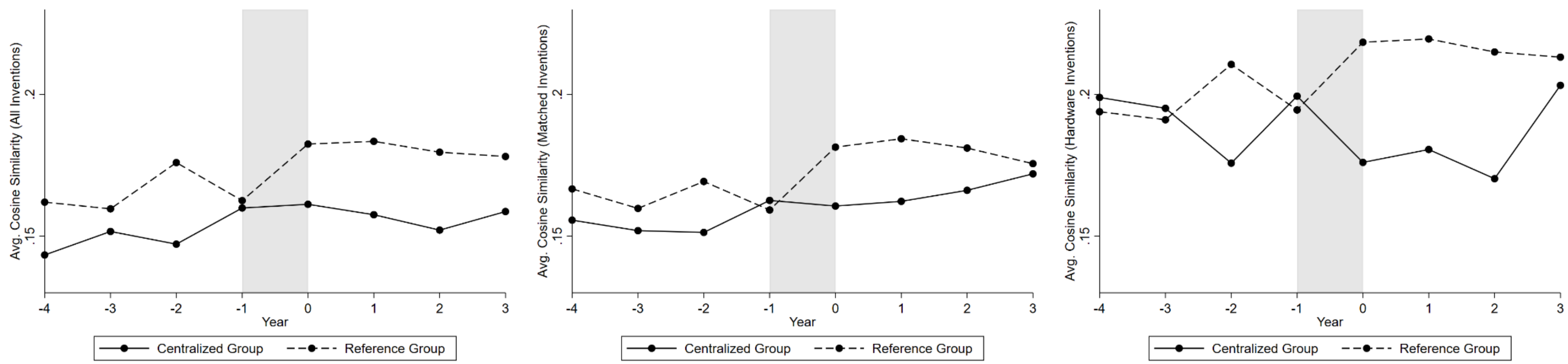
Notes: OLS estimation. Robust standard errors in parentheses.

TABLE 3 Commercialization of invention (*Implemented*) as a function of centralization of development

	Probit estimation: 2nd stage			OLS estimation: 2nd stage		
	All Inventions	**Matched Inventions**	**Hardware Component**	**All Inventions**	**Matched Inventions**	**Hardware Component**
	(4)	**(5)**	**(6)**	**(4')**	**(5')**	**(6')**
Post Reorganization x Centralized Group	-0.269 (0.094)	-0.479 (0.147)	-0.566 (0.163)	-0.071 (0.030)	-0.135 (0.046)	-0.141 (0.045)
Post Reorganization	0.330 (0.144)	0.585 (0.228)	0.413 (0.225)	0.070 (0.040)	0.146 (0.064)	0.087 (0.058)
Centralized Group	0.065 (0.063)	0.255 (0.099)	0.058 (0.119)	0.019 (0.020)	0.077 (0.030)	0.015 (0.035)
New Technology Combinations	0.154 (0.133)	0.309 (0.232)	0.230 (0.198)	0.053 (0.045)	0.105 (0.076)	0.062 (0.062)
Number of Tech Classes	-0.086 (0.025)	-0.097 (0.038)	-0.036 (0.040)	-0.025 (0.008)	-0.029 (0.012)	-0.007 (0.012)
Official Project Invention	0.122 (0.051)	0.139 (0.076)	0.180 (0.081)	0.042 (0.015)	0.044 (0.021)	0.055 (0.021)
Frontier Technology	-0.013 (0.053)	-0.079 (0.081)	-0.028 (0.084)	0.000 (0.017)	-0.022 (0.026)	-0.009 (0.025)
Team Size (log)	0.134 (0.069)	0.227 (0.114)	0.402 (0.115)	0.063 (0.020)	0.098 (0.034)	0.139 (0.029)
Technology Class	Included	Included	Included	Included	Included	Included
Year	Included	Included	Included	Included	Included	Included
Constant	-0.005 (0.204)	-0.238 (0.301)	-0.666 (0.365)	0.421 (0.064)	0.343 (0.092)	0.214 (0.100)
N	12,000+	9,000+	6,000+	12,000+	9,000+	6,000+
N (selected)	3,000+	2,000+	1,000+	3,000+	2,000+	1,000+
χ^2	110.529	315.337	82.979	115.927	65.459	113.972
ρ	-.549 (.062)	-.554 (.094)	-.479 (.114)	-.333 (.033)	-.336 (.051)	-.288 (.058)
atanh(ρ)	-.617 (.089)	-.625 (.136)	-.522 (.148)	-.347 (.038)	-.350 (.057)	-.296 (.063)

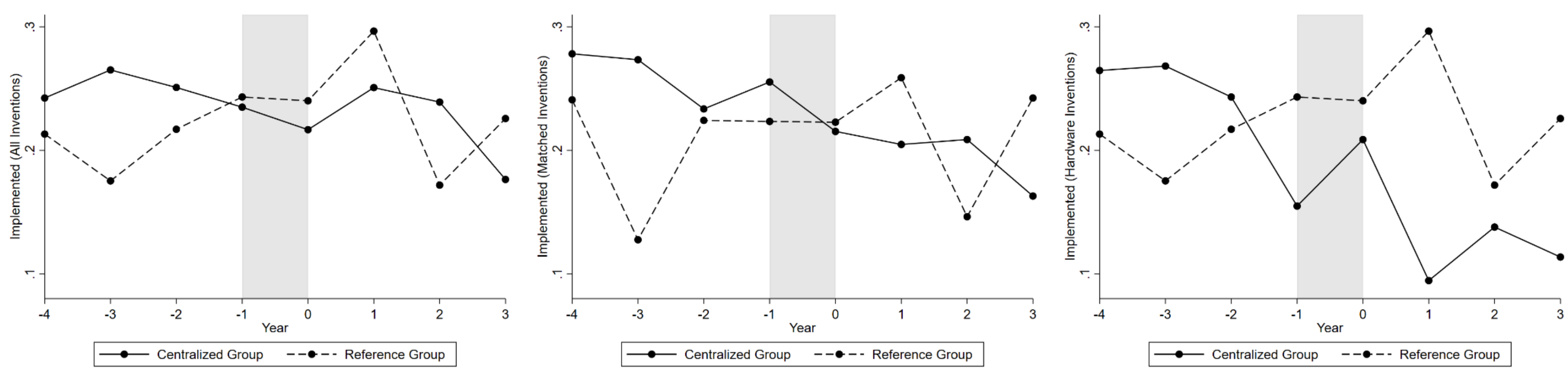
Notes: Second-stage models of two-stage estimation. Appendix Table B2 presents results of the first-stage model estimating selection into an invention's being filed for patent protection. Models 4-6 present results of probit estimation. Models 4'-6' present results of OLS estimation; Robust standard errors in parentheses.

APPENDIX



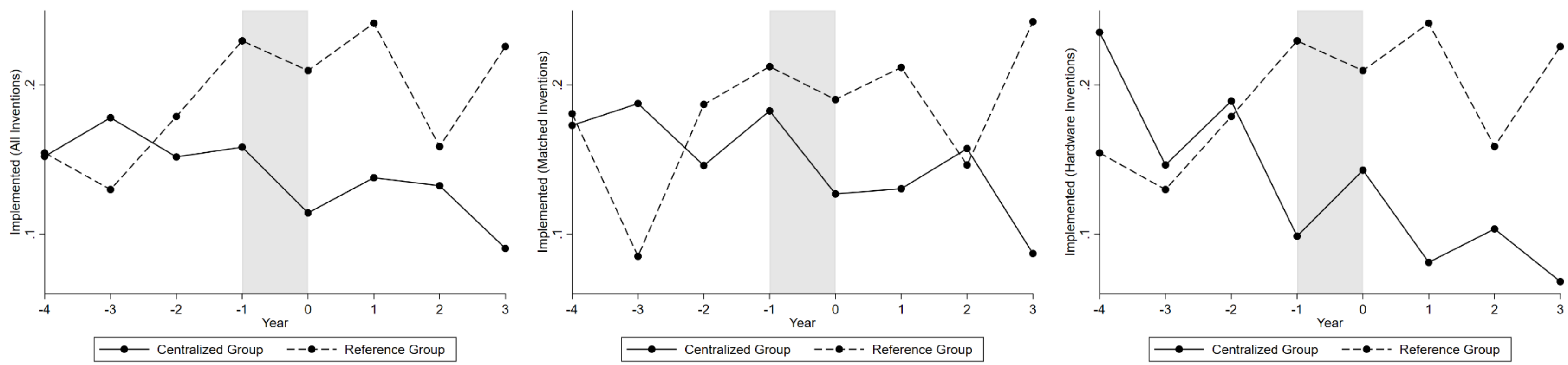
All Inventions	“Matched Inventions” subsample	“Hardware Components” subsample

FIGURE A1 *Avg. Title Cosine Similarity* by year of invention submission, Centralized Group vs. Reference Group



All Inventions	“Matched Inventions” subsample	“Hardware Components” subsample

FIGURE A2 *Implemented* by year of invention submission, Centralized Group vs. Reference Group



All Inventions	“Matched Inventions” subsample	“Hardware Components” subsample

FIGURE A3 *Implemented* (excluding essential patents) by year of invention submission, Centralized Group vs. Reference Group

TABLE A1 Differences in means between Centralized and Reference Groups (All Inventions)

Pre-Reorganization	**N (Reference)**	**Mean (Reference)**	**N (Centralized)**	**Mean (Centralized)**	**Difference in means (SE)**	**t**
Avg. Title Cosine Similarity	+1,000	0.162	+2,000	0.152	0.010 (0.004)	2.268
Implemented	+500	0.219	+1,000	0.247	-0.028 (0.020)	-1.373
Industry-Standard Invention	+1,000	0.061	+2,000	0.109	-0.048 (0.008)	-5.715
Previous Invention	+1,000	0.022	+2,000	0.032	-0.010 (0.005)	-2.004
Recommended Pat. Appl.	+1,000	0.359	+2,000	0.405	-0.046 (0.014)	-3.225
Post-Reorganization	**N (Reference)**	**Mean (Reference)**	**N (Centralized)**	**Mean (Centralized)**	**Difference in means (SE)**	**t**
Avg. Title Cosine Similarity	+2,000	0.180	+4,000	0.156	0.023 (0.004)	6.554
Implemented	+500	0.226	+1,000	0.220	0.006 (0.022)	0.299
Industry-Standard Invention	+2,000	0.025	+4,000	0.121	-0.096 (0.006)	-15.099
Previous Invention	+2,000	0.201	+4,000	0.169	0.032 (0.009)	3.591
Recommended Pat. Appl.	+2,000	0.175	+4,000	0.259	-0.084 (0.010)	-8.620

TABLE A2 Differences in means between Centralized and Reference Groups (Matched Inventions)

Pre-Reorganization	**N (Reference)**	**Mean (Reference)**	**N (Centralized)**	**Mean (Centralized)**	**Difference in means (SE)**	**t**
Avg. Title Cosine Similarity	+1,000	0.163	+1,000	0.157	0.007 (0.005)	1.239
Implemented	<500	0.207	+500	0.259	-0.052 (0.026)	-1.990
Industry-Standard Invention	+1,000	0.059	+1,000	0.093	-0.034 (0.009)	-3.622
Previous Invention	+1,000	0.027	+1,000	0.040	-0.014 (0.006)	-2.170
Recommended Pat. Appl.	+1,000	0.325	+1,000	0.342	-0.017 (0.017)	-1.000
Post-Reorganization	**N (Reference)**	**Mean (Reference)**	**N (Centralized)**	**Mean (Centralized)**	**Difference in means (SE)**	**t**
Avg. Title Cosine Similarity	+2,000	0.181	+3,000	0.165	0.016 (0.004)	3.927
Implemented	<500	0.216	+500	0.204	0.012 (0.026)	0.459
Industry-Standard Invention	+2,000	0.025	+3,000	0.077	-0.052 (0.006)	-8.619
Previous Invention	+2,000	0.207	+3,000	0.197	0.010 (0.011)	0.933
Recommended Pat. Appl.	+2,000	0.154	+3,000	0.213	-0.059 (0.010)	-5.680

TABLE A3 Differences in means between Centralized and Reference Groups (Hardware Components)

Pre-Reorganization	N (Reference)	Mean (Reference)	N (Centralized)	Mean (Centralized)	Difference in means (SE)	t
Avg. Title Cosine Similarity	+1,000	0.194	<500	0.192	0.002 (0.010)	0.254
Implemented	+500	0.219	+100	0.221	-0.002 (0.035)	-0.053
Industry-Standard Invention	+1,000	0.061	<500	0.062	-0.001 (0.012)	-0.113
Previous Invention	+1,000	0.022	<500	0.024	-0.002 (0.008)	-0.232
Recommended Pat. Appl.	+1,000	0.359	<500	0.387	-0.028 (0.025)	-1.128
Post-Reorganization	**N (Reference)**	**Mean (Reference)**	**N (Centralized)**	**Mean (Centralized)**	**Difference in means (SE)**	**t**
Avg. Title Cosine Similarity	+2,000	0.215	+500	0.178	0.037 (0.007)	5.070
Implemented	+500	0.226	+100	0.150	0.076 (0.030)	2.508
Industry-Standard Invention	+2,000	0.025	+500	0.052	-0.028 (0.007)	-4.222
Previous Invention	+2,000	0.201	+500	0.141	0.060 (0.015)	4.019
Recommended Pat. Appl.	+2,000	0.175	+500	0.286	-0.111 (0.015)	-7.329

TABLE B1 First-stage selection regressions for *Implemented*

2nd stage estimation	Probit			OLS		
Sample	**All Inventions**	**Matched Inventions**	**Hardware Component**	**All Inventions**	**Matched Inventions**	**Hardware Component**
Corresponding 2nd-stage models in Table 3	**(4)**	**(5)**	**(6)**	**(4')**	**(5')**	**(6')**
lnPatent Expert Workload	-0.388	-0.376	-0.366	-0.394	-0.381	-0.368
	(0.017)	(0.025)	(0.025)	(0.017)	(0.024)	(0.025)
Patent Expert Diligence	1.475	1.452	1.337	1.481	1.458	1.336
	(0.098)	(0.140)	(0.119)	(0.099)	(0.141)	(0.119)
Post Reorganization x	0.277	0.436	0.469	0.277	0.436	0.469
Centralized Group	(0.055)	(0.084)	(0.091)	(0.055)	(0.084)	(0.091)
Post Reorganization	-0.750	-0.925	-0.784	-0.745	-0.920	-0.781
	(0.078)	(0.126)	(0.107)	(0.078)	(0.126)	(0.107)
Centralized Group	-0.067	-0.137	0.079	-0.067	-0.136	0.079
	(0.041)	(0.065)	(0.070)	(0.041)	(0.065)	(0.070)
New Technology	0.221	0.137	0.168	0.221	0.141	0.166
Combinations	(0.091)	(0.155)	(0.132)	(0.091)	(0.155)	(0.132)
Number of Tech Classes	0.109	0.089	0.126	0.109	0.089	0.126
	(0.017)	(0.025)	(0.025)	(0.017)	(0.025)	(0.025)
Official Project Invention	0.152	0.143	0.155	0.152	0.143	0.155
	(0.028)	(0.041)	(0.040)	(0.028)	(0.041)	(0.040)
Frontier Technology	0.132	0.109	0.091	0.132	0.108	0.091
	(0.034)	(0.051)	(0.050)	(0.034)	(0.051)	(0.050)
Team Size (log)	0.621	0.637	0.586	0.623	0.640	0.588
	(0.036)	(0.058)	(0.052)	(0.036)	(0.058)	(0.052)
Technology Class	Included	Included	Included	Included	Included	Included
Year	Included	Included	Included	Included	Included	Included
Constant	-1.460	-1.293	-1.456	-1.456	-1.290	-1.455
	(0.127)	(0.176)	(0.187)	(0.127)	(0.177)	(0.188)
N	12,000+	9,000+	6,000+	12,000+	9,000+	6,000+
N (selected)	3,000+	2,000+	1,000+	3,000+	2,000+	1,000+
ρ	-.549	-.554	-.479	-.333	-.336	-.288
	(.062)	(.094)	(.114)	(.033)	(.051)	(.058)
atanh(ρ)	-.617	-.625	-.522	-.347	-.350	-.296
	(.089)	(.136)	(.148)	(.038)	(.057)	(.063)
Wald test of independent equations (ρ = 0)						
χ^2	48.503	21.151	12.496	85.081	37.459	21.935
p-value	< .001	< .001	< .001	< .001	< .001	< .001

Notes: The coefficients on our two instruments, *lnPatent Expert Workload* and *Patent Expert Diligence*, are negative and positive, respectively (p < 0.001 in all models), suggesting that these satisfy the exclusion restriction. Wald tests of independent equations indicate the presence of sample selection bias in all models, thus justifying the use of Heckman models. Robust standard errors in parantheses.

TABLE B2 “Naïve” single-stage estimations for *Implemented* without controlling for sample selection

	Probit			OLS		
	All Inventions	**Matched Inventions**	**Hardware Component**	**All Inventions**	**Matched Inventions**	**Hardware Component**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Post Reorganization x Centralized Group	-0.134 (0.101)	-0.327 (0.156)	-0.407 (0.170)	-0.036 (0.029)	-0.091 (0.045)	-0.099 (0.044)
Post Reorganization	-0.167 (0.142)	0.030 (0.219)	-0.046 (0.208)	-0.046 (0.038)	0.008 (0.059)	-0.018 (0.054)
Centralized Group	0.058 (0.069)	0.252 (0.107)	0.092 (0.127)	0.017 (0.020)	0.070 (0.030)	0.024 (0.035)
New Technology Combinations	0.226 (0.143)	0.409 (0.232)	0.261 (0.209)	0.066 (0.044)	0.119 (0.071)	0.067 (0.062)
Number of Tech Classes	-0.053 (0.026)	-0.073 (0.040)	0.004 (0.040)	-0.015 (0.008)	-0.021 (0.012)	0.003 (0.012)
Official Project Invention	0.233 (0.053)	0.246 (0.079)	0.277 (0.081)	0.063 (0.014)	0.063 (0.021)	0.071 (0.021)
Frontier Technology	0.037 (0.057)	-0.039 (0.087)	-0.007 (0.090)	0.011 (0.017)	-0.012 (0.025)	-0.004 (0.025)
Team Size (log)	0.409 (0.062)	0.518 (0.101)	0.638 (0.093)	0.120 (0.019)	0.155 (0.033)	0.183 (0.028)
Technology Class	Included	Included	Included	Included	Included	Included
Year	Included	Included	Included	Included	Included	Included
Constant	-0.895 (0.182)	-1.102 (0.267)	-1.477 (0.296)	0.217 (0.060)	0.155 (0.085)	0.044 (0.092)
N	3,000+	2,000+	1,000+	3,000+	2,000+	1,000+
Pseudo R-squared	.045	.053	.077			
R-squared				.048	.056	.078

Notes: Robust standard errors in parentheses.

TABLE B3 Predicted probability for *Implemented*

Sample: All Inventions	**Centralized Group**	**Reference Group**	**Diff [Cent. – Ref.]**
Pre-Reorganization	.219	.172	-.047
Post-Reorganization	.197	.207	-.010
Diff [Post – Pre]	-.022	.035	-.057

Sample: Matched	**Centralized Group**	**Reference Group**	**Diff [Cent. – Ref.]**
Pre-Reorganization	.244	.153	.091
Post-Reorganization	.199	.215	-.016
Diff [Post – Pre]	-.045	.062	-.107

Sample: Hardware	**Centralized Group**	**Reference Group**	**Diff [Cent. – Ref.]**
Pre-Reorganization	.241	.200	.041
Post-Reorganization	.111	.183	-.072
Diff [Post – Pre]	-.130	-.017	-.113

Notes on process: As STATA's -heckprobit command is not compatible with the command for estimating predicted values (-prvalue), we first manually applied Heckman's 2-step procedure for correcting selection bias to our data and used the results of the 2^{nd}-stage probit regression to estimate predicted probabilities. The results are comparable to those obtained by using -heckprobit command (*b* = -0.303, SE = 0.105 for 'All Inventions' sample; *b* = -0.539, SE = 0.164 for 'Matched Inventions' sample; *b* = -0.622, SE = 0.182 for 'Hardware Components' sample). We then computed probabilities for inventions in each subsample to be implemented by setting the values of continuous control variables at their sample means (*New Technology Combinations, Team Size*). For discrete or binary control variables, *Number of Tech Classes* is set to 1, *Official Project Invention* to 1, and *Frontier Technology* to 0. For *Technology Class* dummies, the most frequent technology class is set to 1 and the others to 0. For *Year* dummies, the third year is set to 1 for the samples of inventions submitted before the reorganization and the seventh year to 1 for those submitted after the reorganization.

TABLE C1 *Previous Invention* as a function of centralization of development

	All Inventions		**Matched Inventions**		**Hardware Component**	
	Previous invention	**Previous invention (non-missing)**	**Previous invention**	**Previous invention (non-missing)**	**Previous invention**	**Previous invention (non-missing)**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Post Reorganization x Centralized Group	-0.307	-0.259	-0.316	-0.327	-0.431	-0.371
	(0.118)	(0.146)	(0.167)	(0.204)	(0.212)	(0.258)
Post Reorganization	0.813	0.629	0.846	0.689	0.806	0.626
	(0.097)	(0.120)	(0.157)	(0.184)	(0.101)	(0.125)
Centralized Group	0.192	0.124	0.212	0.189	0.065	-0.087
	(0.099)	(0.126)	(0.148)	(0.182)	(0.181)	(0.222)
New Technology Combinations	-0.628	-0.673	-1.234	-1.311	-0.514	-0.530
	(0.253)	(0.283)	(0.468)	(0.462)	(0.319)	(0.363)
Number of Tech Classes	0.038	0.086	0.081	0.119	0.054	0.075
	(0.034)	(0.042)	(0.052)	(0.057)	(0.049)	(0.060)
Official Project Invention	0.006	-0.053	-0.059	-0.107	-0.061	-0.133
	(0.055)	(0.064)	(0.076)	(0.085)	(0.080)	(0.092)
Frontier Technology	-0.083	0.038	-0.105	0.025	-0.126	-0.014
	(0.066)	(0.0830)	(0.089)	(0.103)	(0.094)	(0.120)
Team Size (log)	-0.358	-0.286	-0.375	-0.333	-0.355	-0.309
	(0.089)	(0.104)	(0.158)	(0.184)	(0.133)	(0.154)
Technology Class	Included	Included	Included	Included	Included	Included
Year	Included	Included	Included	Included	Included	Included
Constant	-1.546	-0.897	-1.593	-0.956	-1.628	-0.791
	(0.227)	(0.300)	(0.340)	(0.417)	(0.362)	(0.519)
N	3,000+	1,000+	2,000+	1,000+	1,000+	<1,000
χ^2	275.093	135.678	208.512	94.042	158.254	76.878
Pseudo R-squared	.097	.063	.104	.064	.115	.075

Notes: Probit estimation. Robust standard errors in parentheses. Only inventions submitted one year before and after the reorganization are included in the analyses. Models 1, 3, and 5 treat inventions without prior-art-search texts as having 0 previous invention. Models 2, 4, and 6 treat such inventions as missing the data on previous invention.

TABLE C2 *Recommended Patent Application* as a function of centralization of development

	All Inventions	**Matched Inventions**	**Hardware Component**
	(1)	**(2)**	**(3)**
Post Reorganization x Centralized Group	0.253	0.424	0.468
	(0.105)	(0.115)	(0.119)
Post Reorganization	-0.723	-0.907	-0.771
	(0.163)	(0.198)	(0.212)
Centralized Group	-0.060	-0.133	0.077
	(0.099)	(0.106)	(0.113)
New Technology Combinations	0.226	0.159	0.199
	(0.092)	(0.153)	(0.131)
Number of Tech Classes	0.112	0.090	0.129
	(0.039)	(0.047)	(0.040)
Official Project Invention	0.162	0.145	0.161
	(0.043)	(0.051)	(0.046)
Frontier Technology	0.151	0.129	0.097
	(0.052)	(0.070)	(0.080)
Team Size (log)	0.605	0.634	0.577
	(0.067)	(0.087)	(0.091)
Patent Expert workload (log)	-0.389	-0.383	-0.360
	(0.052)	(0.055)	(0.059)
Diligent	1.400	1.412	1.326
	(0.270)	(0.367)	(0.264)
Technology Class	Included	Included	Included
Year	Included	Included	Included
Constant	-1.431	-1.259	-1.467
	(0.351)	(0.401)	(0.390)
N	12,000+	8,000+	6,000+
χ^2	940.645	531.444	689.442
Pseudo R-squared	.181	.176	.189

Notes: Probit estimation. Standard errors clustered on the patent expert in parentheses.

TABLE C3 *Industry-Standard Invention* as a function of centralization of development

	All Inventions	**Matched Inventions**	**Hardware Component**
	(1)	**(2)**	**(3)**
Post Reorganization x Centralized Group	0.600	0.402	0.575
	(0.096)	(0.159)	(0.167)
Post Reorganization	-0.993	-0.538	-0.745
	(0.129)	(0.234)	(0.194)
Centralized Group	0.231	0.053	0.061
	(0.066)	(0.112)	(0.135)
New Technology Combinations	-0.262	-0.396	-0.448
	(0.161)	(0.265)	(0.316)
Number of Tech Classes	0.096	0.149	0.191
	(0.022)	(0.034)	(0.037)
Official Project Invention	0.397	0.514	0.258
	(0.047)	(0.073)	(0.078)
Technology	0.255	0.345	0.179
	(0.046)	(0.079)	(0.082)
Team Size (log)	0.519	0.326	0.307
	(0.051)	(0.092)	(0.090)
Technology Class	Included	Included	Included
Year	Included	Included	Included
Constant	-1.580	-1.791	-1.251
	(0.139)	(0.210)	(0.222)
N	12,000+	8,000+	6,000+
χ^2	1018.250	503.669	368.149
Pseudo R-squared	.318	.339	.304

Notes: Probit estimation. Robust standard errors in parentheses.

TABLE D1 *Implemented* as a function of centralization of development, excluding essential patents

	Probit estimation: 2^{nd} stage			OLS estimation: 2^{nd} stage		
	All Inventions	**Matched Inventions**	**Hardware Component**	**All Inventions**	**Matched Inventions**	**Hardware Component**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Post Reorganization x Centralized Group	-0.298 (0.110)	-0.450 (0.165)	-0.522 (0.188)	-0.067 (0.028)	-0.107 (0.042)	-0.113 (0.042)
Post Reorganization	-0.162 (0.173)	0.010 (0.278)	0.263 (0.253)	-0.023 (0.042)	0.013 (0.063)	0.051 (0.055)
Centralized Group	-0.077 (0.072)	0.114 (0.106)	-0.045 (0.134)	-0.021 (0.018)	0.023 (0.025)	-0.014 (0.032)
New Technology Combinations	0.260 (0.149)	0.383 (0.239)	0.341 (0.207)	0.066 (0.041)	0.095 (0.066)	0.083 (0.059)
Number of Tech Classes	-0.040 (0.030)	-0.079 (0.042)	-0.057 (0.044)	-0.009 (0.007)	-0.018 (0.010)	-0.012 (0.011)
Official Project Invention	0.159 (0.060)	0.157 (0.086)	0.207 (0.088)	0.037 (0.014)	0.036 (0.020)	0.050 (0.020)
Frontier Technology	-0.020 (0.063)	-0.077 (0.095)	-0.035 (0.093)	-0.004 (0.015)	-0.019 (0.023)	-0.010 (0.023)
Team Size (log)	0.291 (0.078)	0.427 (0.123)	0.506 (0.121)	0.071 (0.022)	0.109 (0.034)	0.141 (0.028)
Technology Class	Included	Included	Included	Included	Included	Included
Year	Included	Included	Included	Included	Included	Included
Constant	-1.176 (0.255)	-1.340 (0.358)	-1.429 (0.436)	0.120 (0.072)	0.075 (0.088)	0.024 (0.094)
N	12,000+	9,000+	6,000+	12,000+	9,000+	6,000+
N (selected)	3,000+	2,000+	1,000+	3,000+	2,000+	1,000+
χ^2	100.140	2262.058	91.270	101.637	137.318	119.639
ρ	0.038 (0.101)	-0.045 (0.151)	-0.262 (0.143)	0.024 (0.090)	-0.025 (0.101)	-0.149 (0.070)
atanh(ρ)	0.038 (0.101)	-0.045 (0.152)	-0.268 (0.154)	0.024 (0.090)	-0.025 (0.101)	-0.150 (0.072)

Notes: The models use implementation, but not essential patents, as a dependent variable. Robust standard errors in parentheses.

TABLE D2 Sub-sample analyses for *Implemented*

	End-user technologies[a)]	Other technologies[a)]	χ^2 for Difference[b)]	UI patent board[a)]	Other patent boards[a)]	χ^2 for Difference[b)]
	(1)	(2)		(3)	(4)	
Post Reorganization x Centralized Group	-0.602 (0.182)	-0.142 (0.112)	4.633 [0.031]	-1.093 (0.345)	-0.172 (0.100)	6.568 [0.010]
Post Reorganization	0.850 (0.283)	0.168 (0.162)		0.922 (0.462)	0.141 (0.132)	
Centralized Group	0.056 (0.137)	0.076 (0.071)		0.526 (0.291)	0.033 (0.064)	
New Technology Combinations	-0.482 (0.270)	0.278 (0.158)		-0.501 (0.361)	0.200 (0.149)	
Number of Tech Classes	-0.038 (0.074)	-0.070 (0.026)		-0.078 (0.134)	-0.065 (0.024)	
Official Project Invention	0.159 (0.112)	0.117 (0.059)		0.455 (0.186)	0.088 (0.053)	
Frontier Technology	0.158 (0.119)	0.042 (0.054)		-0.240 (0.474)	0.046 (0.049)	
Team Size (log)	0.163 (0.147)	0.105 (0.081)		0.195 (0.283)	0.118 (0.073)	
Technology Class	Excluded	Excluded		Excluded	Excluded	
Year	Included	Included		Included	Included	
Constant	-0.828 (0.445)	-0.407 (0.183)		-1.085 (0.739)	-0.300 (0.163)	
N	4,000+	7,000+		2,000+	10,000+	
N (selected)	<1,000	2,000+		<500	3,000+	

Notes: All results from 2nd-stage probit regressions. End-user technologies refer to StarCo's internal technology classes that are visible to end users and/or related to user-interface/usability features. UI patent board means StarCo's patent board where all inventions pertaining to user interface are submitted to and evaluated.

a) Robust standard errors reported in parentheses.

b) p-values reported in brackets.

TABLE D3 T-tests of differences in average days to implementation (*Implemented*) between Centralized and Reference Groups

Pre-Reorganization	**N (Reference)**	**Mean (Reference)**	**N (Centralized)**	**Mean (Centralized)**	**Difference in means (SE)**	**t**
Days passed until Implementation (from Submission)	100+	1607.5	100+	1716.8	109.3 (128.5)	-0.850
Days passed until Implementation (from Filing Decision)	100+	1549.8	100+	1669.0	119.2 (128.6)	-0.927
Post-Reorganization	**N (Reference)**	**Mean (Reference)**	**N (Centralized)**	**Mean (Centralized)**	**Difference in means (SE)**	**t**
Days passed until Implementation (from Submission)	<100	1081.0	200+	1202.6	121.5 (87.8)	-1.384
Days passed until Implementation (from Filing Decision)	<100	990.7	200+	1129.8	139.1 (87.7)	-1.586

Notes: Conditional on a filed invention being implemented. Data exhibits right-censoring given that some filed inventions that ultimately will be implemented were not yet implemented by the end of our sample period. This presumably explains the lower mean time to implementation for post-reorganization inventions.