Inventive capabilities in the division of innovative labor *

Ashish Arora[†], Colleen M. Cunningham[‡], and Wesley M. Cohen[§]

Abstract

We study how a firm's inventive capability conditions its participation in a division of innovative labor. Using a survey of US manufacturing firms, and treating inventive capability as unobserved, we estimate a finite-mixture model guided by simple theory linking inventive capability and product innovation outcomes. We find that firms' inventive capabilities condition how they benefit from different forms of external knowledge. High-capability firms' new-to-the-market product innovations benefit from externally available "raw" knowledge, which contributes to the internal generation of inventions. They benefit less, however, from externally generated inventions. In contrast, less capable firms are more likely to acquire and commercialize external inventions, and more typically introduce new-to-the-firm, not new-to-the-market products.

* **Acknowledgments:** This research was funded in part by the NSF (grant #0830349) and the Kauffman Foundation. We thank seminar attendees at Harvard Business School, INSEAD, IESE, and the Academy of Management and Strategic Management Society annual conferences for helpful comments. We also thank the U.S. Cluster Mapping Project for sharing cluster patent data.

† Duke University and NBER, ashish.arora@duke.edu

‡ *corresponding author*, University of Utah, colleen.cunningham@utah.edu

§ Duke University and NBER, wcohen@duke.edu

1 Introduction

The large, technologically self-sufficient firm is no longer (and never was) the sole driver of economic growth. Indeed, recent evidence suggests that nearly half of the product innovations introduced by American manufacturing firms originate from external sources (Arora *et al.*, 2016). If so, the innovative performance of the economy relies importantly upon a "division of innovative labor" (Jewkes *et al.*, 1958; Cassiman and Veugelers, 2006; Arora *et al.*, 2016). In this division of labor, one can distinguish between two important stages. First, *invention* yields an idea or artifact that may underpin a new or improved product. Second, there is *innovation*, that is, the commercialization of the invention. These two stages involve different capabilities: an *inventive capability* that encompasses the upstream technical expertise and functions that allow firms to come up with potentially novel artifacts and ideas, and a *commercialization capability* that enables the firm to develop the inventions and subsequently manufacture, market and sell them. In this paper, we assess the relationship between firms' inventive capabilities and external knowledge and, in turn, firms' decisions to compete by using either their own inventions or others'. In other words, we examine the role of inventive capability in the division of innovative labor.

Our contributions to the literature are twofold. First, we advance understanding of the role of firm capabilities in the innovation process by showing that, while firms' inventive capabilities complement externally available "raw" knowledge, they substitute for externally available inventions. Our findings rationalize seemingly conflicting findings in the strategy and innovation literatures on whether firms' capabilities complement or substitute for external knowledge in innovation and innovative performance. To draw the distinction sharply, while the literature on R&D spillovers, absorptive capacity, and the geography of innovation (Cohen and Levinthal, 1990; Volberda *et al.*, 2010; Audretsch and Feldman, 2004; Feldman, 1993; Rothaermel and Alexandre, 2009) views knowledge flows as complementing internal inventive capabilities, the literature on open innovation and markets for technology (Arora *et al.*, 2001; Cassiman and Veugelers, 2006; Veugelers and Cassiman, 1999; Chatterji and

Fabrizio, 2016; Laursen and Salter, 2006) has viewed external inventions as substituting for internal inventive capabilities. To reconcile these findings, we make two related distinctions: between inventive and commercialization capabilities, and between externally available "raw" knowledge that is the grist for invention (hereafter simply "knowledge") versus more fully formed, externally available inventions (hereafter "external inventions").

In addition to contributing to the strategy and innovation literatures, this reconciliation also promises to address the inconclusive findings, highlighted by David *et al.* (2000), on whether publicly supported R&D complements or substitutes for firms' own R&D efforts. We suggest that the relationship between industrial R&D and publicly supported R&D may turn on the form of the immediate outcome of this support—whether it is knowledge or an invention—as well as on the inventive capabilities possessed by the firms in question. For example, in biomedicine, does public support most immediately spawn knowledge of a cellular target implicated in a disease process, or does it suggest a drug that acts on the target to prevent or cure the disease? In this special issue honoring Paul David, we also note that Dasgupta and David (1994) partially anticipates our distinction between knowledge and invention in their consideration of the economics of science. They highlight the implications of the emergence of private, market-based incentives in driving academic research away from upstream, more basic science toward "technology"—closely related to our notion of invention.[1]

Our second contribution is methodological. We provide a novel way to "measure" capabilities, a core concept in strategy research (Helfat and Peteraf, 2003; Winter, 2003; Helfat and Winter, 2011). We apply a finite mixture model, a statistical approach commonly employed in marketing and health care to distinguish subgroups within a population by assessing systematic differences in the way a predetermined set of variables relates to outcomes of interest (e.g., McLachlan and Peel (2004); Kamakura and Russell (1989); Colombo and Morrison (1989)). In our application of this approach, we distinguish between firms of high-

[1] Dasgupta and David (1994) are, however, more concerned with the appropriability regime surrounding academic output—whether it is privatized or placed in the public domain. In contrast, our analysis focuses on whether the output is an invention (and therefore substitutes for internal inventive effort) or is knowledge that increases the efficiency of, and therefore complements, internal effort.

versus low-capability by examining how their product innovation behavior—that is, whether they introduce a new-to-the-market product, introduce a new-to-the-firm product, or do neither—differentially relates to externally available inventions and knowledge, as well as to their size. By treating inventive capability as an unobserved firm characteristic, we mitigate concerns over measurement-based endogeneity that have plagued the empirical research on firm capabilities (Arora and Nandkumar, 2012; Henderson and Cockburn, 1994; Wu, 2013; Franco *et al.*, 2009). We also go beyond prior uses of finite mixture models (Grimpe and Sofka, 2009) by both: 1) leveraging theory to match firms with the latent characteristic of high inventive capability; and 2) using our measure to test how firms' response to external inventions and knowledge differs depending on their inventive capability, and also conditions the relationship between size and innovative performance.

Our empirical analysis uses a survey of U.S. manufacturing firms from 2007 to 2009 that focuses on innovation and the division of innovative labor, including the sourcing of invention (Arora *et al.*, 2016). The survey samples all manufacturing firms, not just those with prior R&D spending or patents, allowing us to explore product innovation-related choices for a wide spectrum of firms. Our sample distinguishes our research from much of the literature that explores the use of external sources by innovators or R&D performing firms as surveyed by Vivas and Barge-Gil (2015), or only firms in industries where innovation is the basis of competition, e.g., pharmaceuticals or semiconductors (Ceccagnoli *et al.*, 2010; Fabrizio, 2009). For firms that earned revenue from a new or significantly improved product in 2007 to 2009, the survey focuses on whether this product was new-to-the-market or merely new-to-the-firm, as well as the source of new-to-the-market products.[2] In this study, we focus solely on product innovation, not process innovation. We link the survey data to other datasets, including patent data, as well as measures of the local knowledge and invention environment in order to exploit geographic variation in industry-specific measures of the external supply of inventions and knowledge (Jaffe, 1986; Feldman, 1993; Delgado *et al.*, 2014; Porter, 1998;

[2]The most significant new product is defined as the product innovation accounting for the most sales in 2009 of any of the new or significantly improved products introduced by the firm in the 2007 to 2009 period.

Carlino and Kerr, 2015).

To guide our empirical analysis, we develop a simple model in which firms use internally generated ideas or externally available inventions as the basis for one of three product innovation outcomes: a new-to-the-market product, a new-to-the-firm product, or no new product. In a prior stage, firms choose their R&D, which enhances the value of internally generated ideas. We distinguish firms by their inventive capability. High-capability firms are able to invest in R&D, whereas low-capability firms cannot.

Our model delineates two forms of external innovation inputs: 1) knowledge, which increases the productivity of internal R&D, and 2) externally inventions that firms can acquire and commercialize. The model implies that both high- and low-capability firms will introduce new products more when located in environments rich in knowledge and inventions.[3] In locations rich in external *inventions*, low-capability firms will benefit more by acquiring inventions. High-capability firms also benefit from external inventions. However, high-capability firms may also reduce R&D investment, suggesting that the net effect on new-to-the-market and new-to-the-firm products is ambiguous. In contrast, more external *knowledge* increases the R&D efficiency of high-capability firms, leading to greater internal invention, and therefore, an increase in new-to-the-market products, and perhaps new-to-the-firm products as well. We also examine more nuanced predictions from our model, including, for example, the implication that size and inventive capability are complements for product innovation.

Note that we do not estimate the causal impact of external inventions or knowledge on product innovation. Instead, we test whether the patterns of association between the external environment and the rate and the nature of innovative activity differ between high- and low-capability firms as predicted by our simple model. We do not directly measure

[3]We draw on the literature on geographic localization of knowledge flows(Audretsch and Feldman, 2004; Carlino and Kerr, 2015) to characterize the supply of external inventions and knowledge as, at least in part, geographically determined. While relevant local knowledge may not perfectly represent all available relevant knowledge, given the cross-sectional nature of our data, geographic differences provide a necessary source of variation.

inventive capability using empirical proxies such as prior patents, R&D, innovation, or sales. Instead, as noted above, we treat firm capability as an unobserved, latent variable by using a semi-parametric finite-mixture model approach.[4]

To prefigure our empirical results, we find that firms with greater inventive capabilities are more likely to introduce product innovations, both new-to-the-market and also new-to-the-firm. Consistent with prior findings, we find that location in an environment rich in either knowledge or inventions is positively associated with new-to-the-market and new-to-the-firm products.[5] But firms benefit in different ways depending on their inventive capability. Less capable firms use external invention, often to introduce new-to-the-firm products, whereas more capable firms use external knowledge to enhance their internal, new-to-the-market inventions.

Our paper proceeds as follows. The background section briefly discusses related literature. In the following two sections, we introduce a simple model that relates inventive capability to product innovation and how this is affected by the external knowledge environment. Next, in the FMM section, we detail our application of the finite mixture model approach to our particular setting, describing how we use this approach to estimate our measure of inventive capability. After describing our data, we examine the relationship between inventive capability and product innovation, and the role of size, and external inventions and knowledge. We further explore how inventive capability is related to firm performance. The discussion and conclusion section summarizes our results, discusses their implications for the related literature, and offers some managerial implications.

[4] FMM models are similar to some types of unsupervised machine learning. A multinomial logit relates characteristics to outcomes, and the coefficients can differ for firms of different capabilities (latent types). Each observation is assigned a starting probability, sometimes called the prior probability, of originating from a firm with a given capability level. This approach yields coefficient estimates relating observable characteristics such as age and size to outcomes such as a new product. Conditional on the actual outcome, there is a posterior probability of a given observation originating from a firm with a particular capability level. In turn, this posterior probability serves as the basis for the next round of estimates. The process continues until the prior and posterior converge.

[5] This is consistent with findings both from the markets for technology literature (West and Bogers, 2014; Arora *et al.*, 2001) that focuses on external sourcing of invention as the key driver, and findings from geography of innovation literature (Audretsch and Feldman, 2004) that focuses on knowledge spillovers and internal invention as the key driver.

2 Background

The literatures on R&D spillovers and on markets for technology largely posit opposite relationships between a firm's internal capabilities and external knowledge. When emphasizing the role of absorptive capacity, the literature on R&D spillovers highlights the complementarity between a firm's capability and external knowledge: To use external knowledge effectively, firms need to engage in internal inventive activity. In contrast, the literature on markets for technology argues that external inventions substitute for internal capability. These findings lead to an apparent puzzle: the firms with the most to gain from participating in a division of innovative labor (i.e., firms with limited capability) are also the least able to do so. They also offer very different predictions for competitive dynamics. If internal capability and external knowledge are complements, then more capable firms' use of external knowledge will reinforce their competitive advantage. If they are substitutes, then external knowledge will have a leveling effect.

While more apparent than real, this puzzle reflects a gap in our understanding of the relationship between capabilities and external knowledge. There are two reasons for this gap. One is conceptual, reflecting a lack of specificity around the characterization of external knowledge. The other is that measuring firm capabilities is very difficult in practice. We make progress on both fronts.

Scholars have typically not distinguished between the knowledge flows that provide inputs for invention versus inventions themselves (Griliches, 1992; Arora *et al.*, 2016), and have implicitly focused on only one or the other form of external inputs to innovation. For example, the literature on R&D spillovers and absorptive capacity (Cohen and Levinthal, 1990; Volberda *et al.*, 2010) has focused largely on the knowledge flows that are inputs to firm R&D. In contrast, those studying markets for technology (Arora *et al.*, 2001; Cassiman and Veugelers, 2006; Veugelers and Cassiman, 1999; Chatterji and Fabrizio, 2016; Laursen and Salter, 2006) have focused on external inventions which can substitute for internal R&D. By not distinguishing between knowledge flows and inventions, the literature also implicitly conflates different stages in the innovative process. Yet, the literature on R&D spillovers has implicitly focused on invention, and the literature on markets for technology has focused on how inventions are commercialized. We explicitly distinguish between knowledge and invention and empirically confirm that, whereas external knowledge complements inventive capability, external inventions substitute for it.

Second, prior related studies have typically employed observables such as R&D or patents as proxies for firms' technical capabilities (Arora and Nandkumar, 2012; Henderson and Cockburn, 1994; Wu, 2013; Franco *et al.*, 2009). This is a problem to the extent that innovative activity such as R&D, or indicators such as patents or past product innovations, are themselves endogenous to the external knowledge environment. Their use would typically bias regression estimates. In this paper, we explicitly treat capability as unobserved. Specifically, we treat it as a latent variable with the use of a finite-mixture model (or FMM).

Our work is closest to Mani and Nandkumar (2016), who use an FMM approach to distinguish firms, whose strategic position is based on technological competence, from firms relying on other capabilities. They find that the market value of firms competing on technology is lower when markets for technology are extensive, whereas that of firms whose strategic position is based on other complementary capabilities is unaffected (Mani and Nandkumar (2016)). Their analysis differs somewhat from ours since they analyze market value (a continuous variable) and thus use a mixture of Gaussian distributions. In contrast, we use a mixture of multinomial logit distributions since our dependent variable is categorical. More importantly, unlike Mani and Nandkumar (2016) we do not have observations over time. Instead, we use theoretically predicted differences in the responses of different firm types to firm characteristics (e.g., size) and their environments (e.g., supply of external invention) to categorize firm types distinguished by inventive capability.

Some prior work has estimated unobserved capability using stochastic frontier estimation, i.e., measuring how close a firm's output is to the maximum possible (Mahmood *et al.*, 2011; Dutta *et al.*, 2005). Such methods improve on the use of observable proxies. These approaches typically assume that unobserved ability is a draw from a random variable (that may persist over time), which mainly shifts the *level* of performance without changing the relationship between performance and characteristics such as size or environmental factors. That is, they rely entirely on observed outcomes to infer capability. Our approach allows the relationship between firm characteristics and product innovation outcomes to differ between high- and low-capability firms. We similarly allow product innovation outcomes of high- and low-capability firms to respond differently to environmental characteristics such as the supply of external inventions.

Our approach overlaps with random coefficient models (RCMs), which also allow for differences across firms in how outcomes are related to characteristics (Alcácer *et al.*, 2018; Hawk and Pacheco-de Almeida, 2018; Sampson and Shi, 2023). However, we leverage theory to predict how capabilities condition the relationship between size or external supply of inventions and the rate and type of innovation activity. In contrast, RCMs treat the relationship (i.e., the coefficient) as a random draw from a distribution. Thus, the only type of theorizing RCMs typically permit is the extent, rather than the sources, of heterogeneity across firms in the relationships (cf. Alcácer *et al.* (2018): 537-538).[6]

3 Conceptual framework

In this section, we provide an analytic framework that considers the relationship between firms' inventive capabilities and innovation outcomes.[7]

To fix ideas, suppose the firm will introduce at most one new product in a period. A firm will introduce a product if the payoff from commercialization, y, exceeds its commercialization cost. We assume that the payoff depends on the quality of the underlying invention. Further, larger businesses derive more value from an invention of a given quality than smaller ones (Cohen and Klepper, 1996). Larger size may reflect an appropriability advantage due

[6] By estimating firm specific coefficients one can get somewhat closer to the spirit of our approach. For instance, Alcacer et al. (2018) explain the estimated firm-specific coefficients as a function of firm characteristics. Formally, this is similar to estimating an OLS specification with interactions between the variable with the random coefficient and the firm characteristics that are supposed to condition the coefficient.

[7] The formal model behind the intuition here is presented in Appendix B

to the superior commercialization capabilities of larger firms associated with their manufacturing, marketing and sales capabilities (Arora *et al.*, 2023).

Inventions can be generated internally or obtained from the outside. Let x denote the quality of internal invention underpinning a new product, z the quality of external invention net of its cost of acquisition, q firm size, and c the cost of commercialization.[8] The firm will pick the highest quality invention available to it. More precisely, let $y = \max\{x, z\}$. We treat x and z as random variables, which implies that y is also a random variable. The firm introduces a new product if $qy - c \geq 0 \iff y \geq \frac{c}{q}$. New products can either be new-to-the-market or new-to-the-firm. We assume that inventions above a quality threshold, t, underpin new-to-the-market products, and those below are new-to-the-firm products. Figure 1 depicts the probability density function of y with these two key quality thresholds.

The quality of internal invention, x, depends in part on R&D investment, R, which in turn depends in part on the supply of external knowledge, k: an increase in R (for instance, in response to an increase in k) can be represented as a rightward shift in the distribution of x. The quality of external invention, z, depends in part on the supply of external inventions, θ: an increase in θ can be represented as a rightward shift in the distribution of z. The size, external invention, and external knowledge effects on new product introduction are summarized in Table 1.

Firms with high inventive capabilities can invest in R&D and improve the quality of their internal inventions. Firms with low inventive capabilities lack the ability to improve the quality of internal inventions via R&D and therefore do not invest in R&D. Therefore, even if both types of firms have access to the same quality of raw internal ideas, high-capability firms will have a higher quality distribution of internal inventions.

[8] Notice that we express invention quality in monetary units, to make it commensurable with commercialization cost. External invention, z, is therefore measured net of the cost of acquiring it.

3.1 Inventive capabilities and product innovation

It follows from the preceding setup that *higher capability firms will, all else equal, be more likely to innovate, and specifically, more likely to introduce a new-to-the-market product.* We can also generate several sets of comparative statics to help us differentiate high- and low-capability firms based on their product innovation outcomes. We outline the intuition behind these comparative static results below and summarize them in Table 2.

3.1.1 Effect of size

The direct effect of size is to lower the quality threshold for introducing new products. Because returns to introducing a new product are proportional to size, larger businesses can profitably introduce lower quality products for a given cost of commercialization. By itself, this would increase the probability of new-to-the-firm products, leaving the probability of new-to-the-market unchanged. However, there is an indirect effect as well, for high-capability (i.e., R&D performing) firms. Because size increases the returns to introducing a new product, it increases the marginal return to R&D. The effect of size on new-to-the-market and new-to-the-firm products thus differs by capability.

For low-capability firms, greater size simply lowers the threshold for new products, as in Figure 2. Thus, greater size is associated with a higher probability of new-to-the-firm products, a lower probability of no-product innovation, and an unchanged probability of a new-to-the-market product for low-capability firms.

For high-capability firms, greater size similarly lowers the threshold for new product introduction. But, per Figure 3, high-capability firms also will increase their R&D investments with an increase in size, leading to rightward shift in the distribution of y. In turn, this would increase the probability of a new-to-the-market product, in part at the expense of new-to-the-firm products. The effect on the probability of new-to-the-firm products for high-capability firms is ambiguous. On the one hand, greater size lowers the threshold, as it does for low-capability firms, increasing the probability of new-to-the-firm products. On the other hand, higher R&D investments increase the probability that the new product will be new-to-the-market. In sum, for high-capability firms, size is associated with a higher probability of new-to-the-market products, and a lower probability of no new product, with an ambiguous predicted relationship with the probability of new-to-the-firm product innovation.

3.1.2 Effect of external invention

A greater supply of external inventions will increase the probability of a new product, absent any change in the supply of internal inventions. However, the split between new-to-the-firm and new-to-the-market is less clear-cut, and again differs by capability.

For low-capability firms, as depicted in Figure 4, an increase in the quality of external inventions shifts the distribution of y to the right and thereby increases the probability of new-to-the-market product innovation. Regarding new-to-the-firm products, the effect is ambiguous. On the one hand, the newly available external inventions are more likely to be above the threshold for commercialization, but on the other hand, these higher quality inventions are also more likely to be commercialized as a new-to-the-market product.[9]

For high-capability firms, there is an additional offsetting effect because they will reduce their R&D, decreasing the quality of internal inventions. High-capability firms will respond to an increase in supply of external invention by reducing R&D, with a potentially offsetting effect on product innovation, both new-to-the-market and new-to-the-firm. If R&D is very sensitive to external inventions, its decline may be substantial.[10] As a result, for high-capability firms, an increase in external invention has an ambiguous effect across the board.

[9] As a simple numerical example, suppose internal idea quality can take values 0, 1, 2, & 3, with probabilities 0.7, 0.1, 0.1, 0.1. External quality is similarly distributed. Suppose the quality threshold for NTF is 1; inventions with quality threshold 2 or higher are NTM. The probability of no innovation is 0.49. The probability of NTM is 0.4 - 0.04 = 0.36, and the probability of NTF is 0.15. Now suppose the external invention quality distribution shifts right such that the probabilities are 0.68, 0.08, 0.12, 0.12. The probability of no innovation is 0.46 (lower), of NTM is 0.376 (higher), and NTF is 0.1456 (slightly lower). Note that the new distribution of external invention is such that the probability that the quality is equal to or greater than any given threshold is strictly higher.

[10] For instance, suppose R&D effort is binary. A substantial increase in external invention supply may cause the firm to abandon internal R&D. Its innovation may, as a result, decline. In this case the firm would effectively act as a low-capability firm.

3.1.3 Effect of external knowledge

External knowledge increases the effectiveness of internal R&D. By assumption, this does not affect low-capability firms. For high-ability firms, the resulting increase in R&D increases the overall probability of introducing a new product. The effects are depicted in Figure 5. The probability of a new-to-the-market innovation increases because, driven by increases in internal invention quality, inventions are more likely to be above the threshold t. However, there are opposing effects on the probability of a new-to-the-firm product. On the one hand, the quality of internal invention is more likely to be above the threshold for commercialization, but on the other hand, it may also be above the novelty threshold, making it a new-to-the-market product.

3.2 Inventive capability and the use of external inventions

For firms that introduced new-to-the-market products, we observe whether it was based on an internal or external invention. Given our simple model, we expect sourcing of invention to differ by capability. Because high-capability firms have better quality internal inventions, the share of external sources should be higher for low-capability firms than high-capability firms. Further, this difference should increase with size because, as discussed above, the quality of internal inventions increases with size for high-capability firms. Table 3 summarizes these points.

4 Finite mixture model

Our conceptual framework assumes two types of firms, high- and low inventive capability. These two types differ from one another in the relationships between introducing a new product and, respectively, size, external invention supply and external knowledge supply. For a moment, suppose firms only had a choice between a introducing a new-to-the-market product or not (i.e., new-to-the-firm products are ignored) and we could observe firm inventive capability, δ, and the payoff from innovating, v_n. Then, we could simply estimate two sets of regression coefficients, one for high-capability and one for low-capability with v_n as a dependent variable, as a function of size, external invention supply, knowledge, and controls. However, we do not know *ex ante* which firms are high-capability and which firms are low-capability. Many scholars have used proxies based on past outcomes and firm-choices, such as past invention, R&D intensity, patenting, exporting, and size, for capability (Arora and Nandkumar, 2012; Henderson and Cockburn, 1994; Wu, 2013; Franco *et al.*, 2009). As discussed earlier, the use of such proxies is problematic.

Instead, we treat capability as a latent characteristic. Specifically, we assume two types, high-capability and low-capability. Each type has a mean payoff from product innovation, which is a linear function of observed firm characteristics, such as size, and environmental characteristics, such as the supply of external inventions. A firm has an unknown probability of belonging to a high or a low type. We wish to jointly estimate these (unknown) probabilities, along with the coefficients linking payoffs to factors such as size and supply of external invention for each type. Effectively, the coefficients and the weights that provide the best fit (in the sense of maximizing the likelihood function) are reported. In our data, there are, however, two important aspects that complicate this simple description. First, each firm of a given type chooses between introducing a new-to-the-market product, introducing a new-to-the-firm product, or neither. Second, we do not observe the payoffs, only the firm's choice.[11] We address this by specifying mean payoffs from new-to-the-market and new-to-the-firm products, and using a multinomial logit framework (which allows us to infer expected payoffs based on the firm's choice) embedded in a finite-mixture model (FMM), which allows the observable outcomes associated with expected payoffs to differ between high and low types[12].

[11] A third complication is that we do not observe whether new-to-the-firm products are internal or externally sourced. We address this by initially not distinguishing between the source of the invention. We later analyze new-to-the-market product innovations based on whether they used an internal or an externally sourced invention.

[12] FMMs have been widely used in marketing research to categorize consumers by their revealed preferences (Bordley, 1989; Boxall and Adamowicz, 2002; Bucklin and Gupta, 1992; Colombo and Morrison, 1989; Kamakura and Russell, 1989; Kamakura *et al.*, 1996). These models allow for the attributes of various choices (e.g., price) and characteristics of the choosers (e.g., income) to relate differentially to choices across different subgroups or consumer types. For example, Kamakura *et al.* (1996) predict consumers' choice of peanut butter, segmenting those insensitive to price (i.e., brand loyalists) from those who make choices based

In our setting, we distinguish firms by their underlying inventive capability, allowing size, the external supply of inventions and the external supply of knowledge to differentially affect payoffs—and thus the behaviors—of high- and low-capability firms. Formally, we have three potential outcomes for each firm: (1) new-to-the-market product, (2) a new-to-the-firm product, or (3) neither, that is no new product innovation, here indexed by j. The probability of each outcome is: $P_i(j) = Prob(y_i = j)$, representing the class-specific estimates of the propensity to introduce a new-to-the-market product, a new-to-the-firm product, or do nothing. Expressed in the multinomial logit formulation:

$$P_i(j) = \frac{e^{x_i \beta_j}}{\sum_{j=1}^{3} e^{x_i \beta_j}} \tag{1}$$

where x_i are the observed features (i.e., size and external supply conditions) of firm i, and β_j are the coefficients relating those features to the j^{th} outcome. If we allow these probabilities to vary across latent classes (hereafter q), the class conditional probabilities are $P_{i|q}(j) = Prob(y_i = j|class = q)$ or, in the multinomial logit formulation:

$$P_{i|q}(j) = \frac{e^{x_i \beta_{qj}}}{\sum_{j=1}^{3} e^{x_i \beta_{qj}}} \tag{2}$$

We do not directly observe the firm type, and therefore we must estimate it (i.e., probability of being in a class q), along with class-specific outcome probabilities and corresponding class-specific coefficients (β_{qj}). FMM latent class models simultaneously estimate each of these elements. As comparison, a typical regression model would estimate a single set of coefficients (implicitly assuming a single latent class across all observations).

To estimate a FMM latent class model, one must first choose the number of classes q. Our choice of two classes followed both from our intended use of the analysis for latent capability measurement, and from diagnostic tests suggested by prior literature (Greene and Hensher, 2003; Grimpe and Sofka, 2009; Roeder *et al.*, 1999). We calculated R^2 values along with

on price.

the Akaike and Bayesian Information Criteria (AIC and BIC) for one (i.e. standard MNL regression) versus two, three or four classes. Table 6 highlights the various fit measures. While the diagnostics are not definitive, the additional fit from two classes versus one is substantial, whereas the additional fit from moving from two to three (or four) classes is minimal, supporting our choice of two classes.

After choosing the number of classes, the FMM estimation process involves several steps, including the estimation of the prior class probabilities, which are used to generate the class-specific probabilities and coefficients, and, in turn, the posterior class probabilities. (Appendix table A2 shows the posterior probabilities of being high-capability across various industries.) We outline this process more formally here. First, let H_{iq} denote the prior probability of firm i being in class q, z be a set of observable characteristics, and ϕ the corresponding coefficients. In our analyses, our predictors of the initial latent class probabilities include industry fixed effects (i.e., our z). H_{iq} has the multinomial logit structure,

$$H_{iq} = \frac{e^{z_i' \phi_q}}{\sum_{q=1}^{Q} e^{z_i' \phi_q}} \tag{3}$$

The probability of a particular outcome j (e.g., new-to-the-market) for a firm i is the sum of class-level probabilities ($P_{i|q}(j)$) weighted by class probabilities (H_{iq}), or:

$$P_i(j) = \sum_{q=1}^{Q} H_{iq} P_{i|q}(j) \tag{4}$$

Estimates of β_q and ϕ_q are generated via log likelihood maximization of equation 4. After obtaining estimates of ϕ_q from our two-class model, the latent class routine then computes choice probabilities and posterior estimates of the firm-specific class probabilities conditional on choice probabilities via Bayes theorem:

$$\hat{H}_{q|i} = \frac{\hat{P}_{i|q} \hat{H}_{iq}}{\sum_{q=1}^{Q} \hat{P}_{i|q} \hat{H}_{iq}} \tag{5}$$

This procedure is repeated until the posterior estimates of the firm-specific class probabilities, $\hat{H}_{q|i}$, converge, and hence also the class-specific coefficients, $\hat{\beta}_{jq}$. In summary, the analysis produces firm-specific estimates of latent class probabilities ($\hat{H}_{q|i}$), class-specific estimates of the propensity to innovate, introduce new-to-the-firm product, or do nothing ($\hat{P}_{i|q}$), and class-specific coefficients ($\hat{\beta}_{jq}$).[13] For comparison, as shown in Tables 4 and 7, we also estimated the standard Multinomial logit (MNL) model that implicitly assumes a single latent class, and therefore only includes a single set of coefficient estimates ($\hat{\beta}_j$).

5 Data

Our empirical analysis is based largely on the "division of innovative labor" (DoIL) survey of firms in U.S. manufacturing sector (Arora *et al.*, 2016). Administered in 2010, the DoIL survey collected data on new product introductions—focusing on the most significant single new product introduced—at the level of the business unit within firms, for 2007 through 2009. The sample frame for this survey was the Dun and Bradstreet Selectory database, the most complete publicly available frame for the United States at the time.

The survey sampled all American manufacturing firms, not just R&D performers, unlike prior innovation-related surveys (e.g., Cohen *et al.* (2002); Levin *et al.* (1987)). Such a sample is key for the present exercise because it both allows construction of a measure of inventive capability that is not conditioned on endogenous innovation inputs (e.g., R&D) or outcomes (e.g., patents), and enables observation of the relationship between external supply of inventions and raw knowledge, and innovation outcomes for all types of firms.

Sampling was stratified along multiple dimensions, including industry (at the 4 digit NAICS level), and size (categories: Fortune 500, over 1000 employees but not F500, 500 to 1000 employees, 100 to 499 employees, and 10 to 99 employees, and less than 10 employees). For Fortune 500 firms, the sampling unit was the firm's activity within a NAICS; for other firms it was based on primary NAICS. The initial sample was 28,709. Initial screening (for out-of-business or out-of-population) left a final sample of 22,034. The final respondent

[13] The above steps were performed in NLOGIT 5 using LCLOGIT routines.

count was 6,685, reflecting an adjusted response rate of 30.3%. A more detailed description of the sampling process and complete description of the phone survey procedures, along with tables of response rates across industries, detailed tests of response bias, and other related information are outlined in Arora *et al.* (2016).

The survey asked responding firms about whether they had introduced a new product in the previous three years, and if so, whether the product was new to the market or merely new to the firm.[14] We delineate firms that introduced a product that was new-to-the-market from firms that introduced a product that was only new-to-the-firm itself (i.e., not to the market). Across all firms, the average rate of new-to-the-market products is 17%, of new-to-the-firm products is 25%, and 58% of firms did not commercialize a new product at all over the three year sample period, 2007 to 2009. Firms introducing new-to-the-market products were also asked whether they had acquired the key invention underlying their product from an external source, such as a customer, a supplier, another firm in the industry, an independent inventor, an R&D contractor or a university.

The present study includes businesses operating across all manufacturing industries (NAICS 31-33) with 10 or more employees, or 5,175 respondents. Because of item non-response on key variables (e.g., business unit size), our final sample for our latent class analysis is 4,692, out of which there are 1,124 that introduced a new-to-the-market product.

In all analyses we use survey sample weights, constructed using Census data on the population of firms stratified by industry, size strata, and age to correct for non-response bias.[15] We link the survey data to other datasets at the level of respondent (using Duns number or other firm identifiers), industry (NAICS 3 or 4 digit level), and location (county or metropolitan statistical area). Specifically, we use the count of R&D specialist firms in the region as a measure of the external supply of inventions, and proximate, relevant university

[14]More precisely, of all the new products introduced, firms were asked to answer with respect to the most significant product, that which accounted for the largest share of revenues.

[15]We constructed a matrix of these three dimensions of stratification (industry, size, and age) from a custom report provided by the U.S. Bureau of the Census. We thank Ron Jarmin and his team at the U.S. Bureau of the Census for providing this report.

R&D spending as a measure of the supply of external knowledge. Descriptions of the relevant additional datasets are listed below along with variable descriptions.

6 Analysis

Our analysis proceeds in two steps. Both follow from our theoretical predictions about innovation choices and sourcing by firms introducing new-to-the-market product innovations. First, we analyze the likelihood that a firm introduced a new-to-the-market product, new-to-the-firm product, or did not commercialize a new product. Also in this step, we use FMM to simultaneously develop a measure of a firm's latent inventive capability and explore how inventive capability–along with measures of the external knowledge/invention environment, business unit size, and other firm and industry characteristics–shape the choice to introduce a new-to-the-market product, introduce a new-to-the-firm product, or not commercialize a new product. Second, we use our measure of latent inventive capability derived from our application of the FMM approach to examine if the patterns of association between firms' reliance on external invention, capability and business unit size confirm the predictions of our model.

6.1 New-to-the-market, new-to-the-firm, and inventive capability

Outcome variable

Our outcome variable reflects whether the focal product innovation of the business unit is *new-to-the-market*, *new-to-the-firm*, or neither (*none*).

Discriminating variables

Our FMM approach uses the relationships between a set of variables, based on our theory, and our outcomes to discriminate between high- and low-capability firms. Our theory emphasizes a distinction between two types of external inputs—external inventions and knowledge—which, along with business unit size, have different effects on new product commercialization depending on firm inventive capability. We use the (log) count of R&D specialist firms in the MSA (Metropolitan Statistical Area) of the respondent business unit weighted by use of such services in the industry of the respondent as a measure of *external invention* supply. The measure therefore varies by region and industry.[16] To measure external knowledge supply, we use geographically proximate, relevant university R&D spending. We used the NSF data on university R&D expenditures for 2004 through 2006. We count only R&D spending within a 100 mile radius of the focal firm and for research fields related to the industry of the respondent firm. We assessed relatedness based on whether R&D labs in the industry listed a field as relevant for research as reported in a prior survey of R&D lab managers (Cohen *et al.*, 2002).[17] Last, we measure *size* as the logged number of employees of the business unit.

Additionally, we include several variables to control for differences across firms, given our data are cross-sectional. First, we include an indicator for whether or not the business unit is *multiproduct* (where *multiproduct* = 1 if the business unit has more than one associated 6-digit NAICS) since the scope of related firm activities has been found to be positively associated with invention and innovative success (Cockburn and Henderson, 2001; Henderson and Cockburn, 1996). We also include an indicator for whether the business unit is part of a larger firm or is a *standalone* company. We also control for *firm age*.

At the industry level, we include a dummy for whether or not the respondent is in a *high tech industry*, which is defined as whether the share of firms in the industry of the business unit that perform R&D is above the median (high tech = 1). On average, we expect firms in high-tech industries to be more likely to introduce new products. We also control for whether

[16]Specialist R&D suppliers consist of suppliers of Architectural, Engineering, and Related Services (5413), suppliers of Specialized Design Services (5414), Computer Systems Design and Related Services (5415), Management, Scientific, and Technical Consulting Services (5416), and Scientific Research and Development Services (5417). For our measure, we take a log of the count the number of large establishments (>100 employees) in NAICS 5413-5417 in the relevant MSA according the the US Census County Business Patterns data for 2007. For the industry weights, we use Bureau of Economic Analysis input-output tables to get the share of inputs coming from R&D specialists in each industry. R&D specialists may also supply knowledge inputs to internal invention. To examine the robustness of our results that employ this measure, we also ran our analyses using counts of relevant patents in the region of the respondent as defined by the Cluster mapping project (Delgado *et al.*, 2014). Both sets of results are qualitatively similar.

[17]Proximate relevant university research activity may also proxy for university inventions. Inventions resulting from university research constitute, however, only a very small share of the external inventions employed by firms.(Arora *et al.*, 2016).

or not the business unit is in a *homogeneous market* (Sutton, 1998). Homogeneity is based on the share of total industry-level sales (4 digit NAICS) of the largest 7 digit NAICS category within the industry. We use total shipment values at the 4 and 7-digit NAICS level from the 2002 US Economic Census (homogeneous =1 for above median industries). Homogeneity should have competing effects on the choice between a new-to-the-market and a new-to-the-firm product.

Table 4 presents the results from simple multinomial logit (or single class) and FMM logit models. Table 5 includes the corresponding marginal effect estimates and average probability of each outcome for all three models. The best fit is provided by a model with two, rather than more or fewer latent classes. The fit statistics are outlined and described in Table 6.

We expect that one of the latent classes will be comprised of those with high inventive capability, characterized by higher probability of product innovation, i.e., new-to-the-market and new-to-the-firm products. Further, the effect of business unit size on the probability of new-to-the-market should be larger as compared to that of low-capability firms. Finally, the probability of new-to-the-market for less capable firms should increase more with external invention supply than is the case for more capable firms, and the reverse should be true for external knowledge.

We estimate the average probability of an observation belonging to latent class 1 is 35%, and to latent class 2 is 65%. Within those classes, we can also observe the relative incidence of whether a firm introduces a new-to-the-market product, introduces a new-to-the-firm product, or neither. For class 1, the predicted probability of a new-to-the-market product is 35%, of a new-to-the-firm product 40%, and of no new product 24%. For latent class 2, the corresponding probabilities are 8%, 16%, and 76%. Clearly, both the introduction of a new-to-the-market and a new-to-the-firm product are, on average, much more likely for those that have a higher probability of being in latent class 1.[18] Henceforth, we refer

[18]These outcome probabilities were generated using the full sample using the probability associated with being in each class as weights. If, instead, we assign each respondent to a single class based on whether its probability of belonging to that class exceeds 0.5, we get: class 1 new-to-the-market product is 45%, of new-to-the-firm product 53%, and of no new product 2%. For latent class 2, the breakdown is 4% new-to-

to latent class 1 as "high-capability", and use the likelihood that a respondent belongs to latent class 1 as a continuous measure of the inventive capability of the firm. Figure 6 highlights in more detail how the predicted probability of introducing a new-to-the-market product, introducing a new-to-the-firm product or neither of those changes as the firm's inventive capability increases. Consistent with our conceptual framework, the share of new-to-the-market product innovation increases monotonically with capability, while, as discussed above, the share new-to-the-firm products initially increases then declines with capability.

Table 4 (Columns 4-9) shows that business unit size increases the likelihood of new-to-the-market and new-to-the-firm products for all firms. Figure 7 shows that the probability of a new-to-the-firm product for high-capability firms is non-monotonic with size; it first increases and then decreases with size for high-capability firms. The likelihood of new-to-the-firm product is the highest for mid-sized business units of high-capability firms.[19]

Our conceptual framework implied that an increase the supply of external inventions would increase new-to-the-market and new-to-the-firm products for low-capability firms; the effect on high-capability firms is ambiguous. Conversely, greater external knowledge would increase new-to-the-market products by high-capability firms, and have no effect on low-capability firms. As shown in Table 4, new-to-the-market products and new-to-the-firms products indeed appear more responsive to the supply of inventions for low-capability firms (latent class 2) than high-capability firms. Figure 8 shows negligible difference in the probability of new-to-the-market or new-to-the-firm products across the range of external invention supply for high-capability firms and an increase of approximately 10% in new-to-the-firm products and 15% in new-to-the-market products across the range of invention supply for low-capability firms.

---

the-market product, 12% new-to-the-firm product, and 84% none.

[19]Other empirical patterns are consistent with latent class 1 representing "high-capability" firms (although they are not part of our theory): (1) for latent class 1, being in a high-tech industry is also positively associated with new-to-the-market products, having a large effect (+20%), compared to latent class 2, where being in a high tech industry increases the probability of new-to-the-firm product by 5%; (2) for latent class 1, being in a more homogeneous market increases the likelihood of new-to-the-market product innovation by 11%, and of doing nothing by 20%, compared to latent class 2, where being in a more homogeneous market increases the likelihood of new-to-the-firm product by 13%.

While we model external knowledge supply, k, and external invention supply, θ, as separate parameters, they are difficult to separate empirically. We use geographic variation to help identify these relationships, and areas rich in relevant knowledge tend also to be rich in related inventions. Indeed, our two measures are strongly correlated (r = 0.56). Given the collinearity between the two constructs, we now conduct a separate analysis using the external supply of knowledge. Accordingly, in Table 7, we duplicate our analysis from Table 4, using external knowledge supply (proximate, relevant university R&D spending) in place of invention supply. Here, we see that high-capability firms are more likely to introduce new-to-the-market products in the presence of external knowledge supply, consistent with the argument that high-capability firms are able to extract raw knowledge inputs from the outside. Somewhat unexpectedly, for low-capability firms, the probability of a new-to-the-firm product increase with external knowledge.[20] We do find that the probability of a new-to-the-market product does not change with external knowledge supply for low-capability firms, consistent with our framework.[21]

The divergent results between inventive capability and the external supply of, respectively, inventions and knowledge highlight a key point of our paper: firms of high- and low-capability use external resources differently. External inventions are a substitute for internal inventive efforts, and therefore, low-capability firms seek external inventions to fuel new-to-the-market product innovation. In contrast, the internal inventive efforts of high-capability firms benefit from more external upstream knowledge inputs. Put starkly, low-capability firms are less capable of product innovation without the help of external inventors than high-capability firms.

[20]One possible explanation for the increases in new-to-the-firm product for low-capability firms is that more knowledge, by allowing for greater new-to-the-market rates by more capable firms, may also yield more sources for low-capability firms to imitate (i.e., introduce a new-to-the-firm product).

[21]Tables 4 and 7 generate, for each observation, a measure of the probability that a firm belongs to the high-capability class. We correlate the two sets of estimates to examine the consistency of the classification. As Table A1 shows, both result in very similar assignments of firms to high- and low-capability classes, with a correlation of 0.89.

Validating our Capability Measure

Our analysis maintains that the capability we are measuring is inventive capability rather than, for example, the downstream capability to commercialize new products. Indeed, our finding that more capable firms are more likely to exploit external knowledge and not use external inventions, while less capable firms are more likely to exploit external inventions, is consistent with this interpretation. Nonetheless, we performed additional analyses to both confirm our view and investigate the robustness of our capability measure.

First, we examined the association between our measure of inventive capability and two key measures of inventive activity, namely R&D investment and past patenting (from 2002 to 2006). Controlling for industry, high-capability firms are more than three times more likely to have performed R&D (54% versus 15%) and almost twice as likely to have filed for a patent (9% versus 5%) over a three year period. We also examined the association between our measure and a measure drawn from the DoIL survey where respondents reported whether they internally developed a technology that they then licensed to another firm (Arora *et al.*, 2016). It is reasonable to presume that firms that possess inventive capability are more likely to have internal inventions available for sale or licensing. And, indeed, we find, controlling for industry, high-capability firms are two and a half times more likely to have been a technology supplier (23% versus 9%) over the three-year sample period.

Collectively, these patterns of association provide additional support for our interpretation of our latent variable measure as reflecting inventive capability.

6.2 Inventive Capability and External Sourcing

We next examine how the use of external inventions differs by capability and size. We first run a simple cross tab on our sample business units who introduced new-to-the-market products, where business units are distinguished by size (large and small) and inventive capability (high and low). For our measure of *inventive capability*, we use our estimate of the probability that the firm has high inventive capability in Table 4, and split firms into high- and low-capability on the basis of the median value. To distinguish small from large business units, we split at 500 employees.

We construct our measure of whether firms that introduced a new-to-the-market product acquired the underlying invention(s) externally or generated them internally from responses to the DoIL survey (Arora *et al.*, 2016), discussed above. We identify the source of the invention to be external if the respondent said one or more of the following was the main source of the overall concept, prototype or design underlying its new-to-the-market product: a supplier; a customer; another firm in the industry; a consultant, commercial lab, or engineering service provider; an independent inventor; or a university or government lab. The alternative was that the new-to-the-market product was based on an internal invention.

In our conceptual framework, less capable firms and smaller firms should be more likely to use external inventions. The cross-tab results presented in Table 8 accord with both predictions. The sharpest difference is observed if we compare small, less capable firms with larger, more capable firm. The former's reliance upon external invention, 49.3%, exceeds that of the latter by nine percentage points, representing a difference of almost one quarter.[22] These results affirm our main point: Whereas small, low-capability firms introduce new products mainly by commercializing inventions made by others, large, high-capability firms rely mainly upon internally generated inventions for their new products.

In our earlier characterization of inventive capability, we implied it is largely technical. Hence, firms with less inventive capability should be more likely to access inventions with less technical content. To probe this conjecture, we take advantage of the DoIL data that distinguishes externally acquired inventions by source. We conjecture that inventions originating from customers will have less technical content on average as compared to inventions originating from universities, R&D service contractors, and startups. As a consequence, the acquisition of inventions from customers should require less technical inventive capability as compared to acquisitions from other external sources.

[22] Consistent with our theoretical predictions, we also observe the percentage of smaller firms acquiring their inventions from an outside source exceeding that of larger firms by about four percent, and the percentage of less capable firms doing so exceeding that of more capable firms by just over a percent. Moreover we see a greater drop associated with greater capability among larger firms, where the drop is almost six percent.

Exploring this conjecture, Table 9 presents the results of a multinomial logit where the reference category is internal invention and column (1) presents the predicted likelihood of using an invention from a non-customer source versus an internal invention, and column (2) presents the predicted likelihood of using a customer-sourced invention relative to internal invention. Confirming our prior, the results in column (2) show that, as inventive capability increases, innovating firms are less likely to rely upon a customer-sourced invention as compared to internal invention. In addition to providing suggestive evidence that our capability measure reflects a technical capability, this result also highlights the broader point that external inventions may differ from one another in systematic ways, and that those differences have implications for the role of inventive capability in affecting their acquisition. The second result in Table 9 should, however, give us pause. The coefficient on inventive capability in column (1) is essentially zero. The interpretation is that the ratio of non-customer inventions to internal inventions does not increase with capability. This result is superficially inconsistent with our theory. But, in our view, what it highlights is the starkness of our model. To the degree that external inventions have greater technical content, internal inventive capability may enhance not only internal invention but also aid in the acquisition of more technical inventions.

6.3 Inventive Capability and Performance

In this section, we explore the relationship between inventive capability and business unit performance. Here our ambitions are modest—to simply consider whether the patterns of association between our measure of inventive capability and performance are consistent with our conception of the role and impact of inventive capability as presented in our conceptual framework.

Our measure of business unit performance is whether a business unit experienced market share growth between 2008 and 2009, the end of our sample period (Arora *et al.*, 2016). Our sample for this analysis is the full set of survey respondents regardless of product innovation outcome, and our measure of inventive capability here, is continuous, i.e. the likelihood of being high-capability. It is important to remember that firms with high inventive capability do not necessarily introduce a new-to-the-market product; they may introduce a new-to-the-firm product or do neither.

Table 10 reports the results of a linear probability model on the likelihood of market share growth. The column (1) result shows a positive significant relationship between inventive capability and market share increase, as expected. Per our model, inventive capability affects performance through the payoff firms realize from using their inventive capabilities to introduce new products. In column (2), along with inventive capability, we thus include whether the firm introduces a new-to-the-market or new-to-the-firm product. Given that our model would suggest that inventive capability should benefit the firms' performance through one of these two outcomes, one would expect the relationship between inventive capability and market share to weaken when the outcome measures are also included; performance should be related directly to outcomes, the inclusion of which should knock out the indirect effect of inventive capability. Indeed, this is exactly what we see in column (2).

In our conceptual framework, inventive capability increases R&D, which enhances the quality of internal inventions. The quality of new-to-the-firm products however remains bounded between $\frac{c}{q}$ and t. The main effect of an increase in the quality of internal invention should be reflected in the average quality of new-to-the-market products. Thus, if the firm actually innovates, one would expect capability to have a bigger impact on performance than if the firm introduces a new-to-the-firm product. To test this, column (3) includes interactions between inventive capability and the two outcomes new-to-the-firm product and new-to-the-market product. The results are again consistent with our theorizing: if firms have applied their inventive capability to a new-to-the-market product rather than commercializing a new-to-the-firm product, the likelihood of an increase in their market share increases.

As a side note, we observe that business unit size is associated with an increase in market share in column (1). However, once we control for whether the firm has introduced a new-to-the-market or new-to-the-firm product, the effect of business unit size dissipates, consistent with the theorized mechanism that the impact of business unit size on market share works through new product introduction.

7 Discussion and Conclusions

We investigate the links between firms' inventive capabilities, their product innovation outcomes, and the external supplies of inventions and knowledge. We develop a simple conceptual framework relating a firm's inventive capability to the three product innovation-related outcomes: introduction of a new-to-the-market product, introduction of a new-to-the-firm product, or neither. We further examine how that relationship is conditioned by the supplies, respectively, of external invention and knowledge, as well as firm size. Guided by the framework, we employ a finite-mixture model that simultaneously assigns firms to an "inventive capability" (i.e., high or low), and estimates how the relationships between, respectively, the supplies of, respectively, "raw" knowledge or inventions and firms' innovation-related outcomes are conditioned by that capability and size. From this first step, we develop and subsequently validate our estimated measure of firms' inventive capabilities. Using that measure, we confirm that firms' use of external inventions is conditioned by their inventive capabilities and size. In our final step, we examine the relationship between inventive capabilities and firm performance.

Perhaps the most striking finding is that increasing the external supply of inventions contributes more to new-to-the-market products for the less capable firms relative to the more capable firms. In this sense, a greater external supply of invention has a leveling effect across competitors. Introduction of new-to-the-market products by more capable firms does, however, increase with greater knowledge flows. External inventions can substitute—and thus compensate—for the inability of firms to invent. In contrast, external knowledge complements inventive capability, enabling greater new-to-the-market product innovation by more capable firms, reinforcing their advantage. Concretely, increasing the supply of external invention, for instance by thickening technology markets, strengthens the ability of less capable firms to innovate and compete. At the same time, however, more capable firms are able to capitalize on greater external supply of knowledge—such as that which may originate from universities or R&D spillovers from rivals—to introduce new-to-the-market products more.

To come to these conclusions, we have bridged several literatures that focus on the division of innovative labor, each of which emphasizes only one form of external knowledge. The literature on R&D spillovers, geography, and absorptive capacity focuses largely on raw knowledge, while the literature on markets for technology tends to focus only on the movement of inventions across organizations. As noted above, our finding that local knowledge flows have a positive relationship with new-to-the-market product innovation is consistent with prior findings on the geography of innovation, which typically argues that agglomeration allows firms to access knowledge spillovers that help them to generate inventions. Our findings, however, suggest that high-capability firms are more likely to use knowledge spillovers to innovate. In contrast, to the extent that less capable firms benefit from a propitious location, our findings suggest it is because of the availability of inventions rather than the flow of knowledge.

In addition to our substantive findings, our study contributes methodologically to the study of firm capabilities. We generate a measure of latent inventive capability that correlates with, but is not determined by, prior invention inputs like R&D, or prior invention outcomes like patenting. Demonstrating that our measure of inventive capability is also not defined in terms of outcomes of interest, we see that, within our sample, firms can have high inventive capability and not innovate or low-capability and innovate. Our application of the finite-mixture model to the study of firm capability is also notable in that, first, our application of the method is guided by theory, and, second, we use our our firm-specific estimate of inventive capability (i.e., the likelihood that a firm possesses high inventive capability) as a regressor in the subsequent empirical analyses (see Tables 7 and 8).

Our study also offers several implications for management and policy. For managers seeking growth, an important question is whether to invest substantial resources in R&D expenditure or pursue a different strategy. Assuming that capabilities are slow and costly to change, our findings suggest this will depend on both their existing capabilities and the external availability of inventions and knowledge. Specifically, for high-capability firms, investing in R&D will help them to invent and enables access to external knowledge inputs. High-capability firms in rich knowledge environments may also, however, be at greater risk of knowledge spilling over to rivals (Alcácer and Chung, 2007; Giarratana and Mariani, 2014). Our findings imply, moreover, that it is other high-capability firms who would benefit from such spillovers. In contrast, investing in R&D may not help low-capability firms for whom product innovation is more about sourcing external invention. For low-capability firms, investing in complementary commercialization resources (and perhaps technology sourcing capabilities) may be more advisable, especially if external inventions are plentiful. Moreover, an obvious implication for less capable firms is to locate where external inventions are in plentiful supply.

With respect to policy, our findings offer implications for the way we should think about the relationship between government support for R&D and firms' own innovative activities. As noted by David *et al.* (2000), the literature on whether publicly supported R&D complements or substitutes for firms' own R&D yields ambiguous results. Our findings suggest that to advance our understanding of that relationship, it would be productive to consider whether that public support generates raw knowledge or inventions and, at the same time, the capabilities of the firms that are the beneficiaries of that support.

Our conclusions require a number of qualifications. We only take one step in probing the relationship between firm capabilities and the division of innovative labor. Due largely to data limitations, we study the role of inventive capability in product innovation, not that of firms' commercialization capabilities such as manufacturing, marketing and sales. Our data are survey based, and therefore are subject to the caveats associated with such data. The cross-sectional nature of our data prevent us from exploring dynamic relationships between capability, product innovation, and performance. Our data cover innovations during 2007-2009, which includes the 2008 recession, and thus our results must be interpreted with caution. Finally, we have assumed that the supply of knowledge and inventions is exogenous. Concretely, this would imply, for example, that firm location in a rich or poor external knowledge environment is exogenous. Instead, firms may choose their location in part to access external inventions or knowledge.

Our theory is stark in its assumptions, including, for example, that inventive capability does not directly affect the quality of commercialized products originating from external invention. We suggest, however, that distinguishing between inventive capability and commercialization, and between knowledge and invention as inputs into product innovation, enables us to clarify whether external knowledge complements or substitutes for the firm's own capabilities. While we make an important methodological contribution by measuring a particular type of capability at a point in time, capabilities are clearly cumulative, shaped by prior decisions, experience, and learning. We hope future research continues to develop measures of latent capability.

References

Alcácer, J. and Chung, W. 2007. Location Strategies and Knowledge Spillovers. *Management Science*, 53(5):760–776.

Alcácer, J., Chung, W., Hawk, A., and Pacheco-de Almeida, G. 2018. Applying Random Coefficient Models to Strategy Research: Identifying and Exploring Firm Heterogeneous Effects. *Strategy Science*, 3(3):533–553.

Arora, A., Cohen, W., Lee, H., and Sebastian, D. 2023. Invention value, inventive capability and the large firm advantage. *Research Policy*, 52(1):104650.

Arora, A., Cohen, W. M., and Walsh, J. P. 2016. The acquisition and commercialization of invention in American manufacturing: Incidence and impact. *Research Policy*, 45(6):1113–1128.

Arora, A., Fosfuri, A., and Gambardella, A. 2001. *Markets for Technology.* The MIT Press, Cambridge, MA.

Arora, A. and Nandkumar, A. 2012. Insecure Advantage? Markets for Technology and the Value of Resources for Entrepreneurial Ventures. *Strategic Management Journal*, 33(3):231–251.

Audretsch, D. B. and Feldman, M. P. 2004. Chapter 61: Knowledge spillovers and the geography of innovation. In Thisse, J. V. H. a. J.-F., editor, *Handbook of Regional and Urban Economics*, volume Volume 4, pages 2713–2739. Elsevier.

Bordley, R. F. 1989. Note—Relaxing the Loyalty Condition in the Colombo/Morrison Model. *Marketing Science*, 8(1):100–103.

Boxall, P. C. and Adamowicz, W. L. 2002. Understanding Heterogeneous Preferences in Random Utility Models: A Latent Class Approach. *Environmental and Resource Economics*, 23(4):421–446.

Bucklin, R. E. and Gupta, S. 1992. Brand Choice, Purchase Incidence, and Segmentation: An Integrated Modeling Approach. *Journal of Marketing Research (JMR)*, 29(2):201–215.

Carlino, G. and Kerr, W. R. 2015. Chapter 6 - Agglomeration and Innovation. In Gilles Duranton, J. V. H. a. W. C. S., editor, *Handbook of Regional and Urban Economics*, volume 5 of *Handbook of Regional and Urban Economics*, pages 349–404. Elsevier.

Cassiman, B. and Veugelers, R. 2006. In Search of Complementarity in Innovation Strategy: Internal R&D and External Knowledge Acquisition. *Management Science*, 52(1):68–82.

Ceccagnoli, M., Graham, S. J. H., Higgins, M. J., and Lee, J. 2010. Productivity and the Role of Complementary Assets in Firms' Demand for Technology Innovations. *Industrial and Corporate Change*, 19(3):839–869.

Chatterji, A. K. and Fabrizio, K. R. 2016. Does the market for ideas influence the rate and direction of innovative activity? Evidence from the medical device industry. *Strategic Management Journal*, 37(3):447–465.

Cockburn, I. M. and Henderson, R. M. 2001. Scale and scope in drug development: unpacking the advantages of size in pharmaceutical research. *Journal of Health Economics*, 20(6):1033–1057.

Cohen, W. M. and Klepper, S. 1996. A Reprise of Size and R&D. *Economic Journal*, 106(437):925–951.

Cohen, W. M. and Levinthal, D. A. 1990. Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, 35(1):128–152.

Cohen, W. M., Nelson, R. R., and Walsh, J. P. 2002. Links and Impacts: The Influence of Public Research on Industrial R&D. *Management Science*, 48(1):1–23.

Colombo, R. A. and Morrison, D. G. 1989. Note—A Brand Switching Model with Implications for Marketing Strategies. *Marketing Science*, 8(1):89–99.

Dasgupta, P. and David, P. A. 1994. Toward a new economics of science. *Research Policy*, 23(5):487–521.

David, P. A., Hall, B. H., and Toole, A. A. 2000. Is public r&d a complement or substitute for private r&d? a review of the econometric evidence. *Research Policy*, 29(4):497–529.

Delgado, M., Porter, M. E., and Stern, S. 2014. Clusters, convergence, and economic performance. *Research Policy*, 43(10):1785–1799.

Dutta, S., Narasimhan, O., and Rajiv, S. 2005. Conceptualizing and measuring capabilities: methodology and empirical application. *Strategic Management Journal*, 26(3):277–285.

Fabrizio, K. R. 2009. Absorptive capacity and the search for innovation. *Research Policy*, 38(2):255–267.

Feldman, M. P. 1993. An Examination of the Geography of Innovation. *Industrial and Corporate Change*, 2(3):451–470.

Franco, A. M., Sarkar, M. B., Agarwal, R., and Echambadi, R. 2009. Swift and Smart: The Moderating Effects of Technological Capabilities on the Market Pioneering-Firm Survival Relationship. *Management Science*, 55(11):1842–1860.

Giarratana, M. S. and Mariani, M. 2014. The relationship between knowledge sourcing and fear of imitation. *Strategic Management Journal*, 35(8):1144–1163.

Greene, W. H. and Hensher, D. A. 2003. A latent class model for discrete choice analysis: contrasts with mixed logit. *Transportation Research Part B: Methodological*, 37(8):681–698.

Griliches, Z. 1992. The Search for R&D Spillovers. *The Scandinavian Journal of Economics*, 94:S29–S47.

Grimpe, C. and Sofka, W. 2009. Search patterns and absorptive capacity: Low- and high-technology sectors in European countries. *Research Policy*, 38(3):495–506.

Hawk, A. and Pacheco-de Almeida, G. 2018. Time compression (dis)economies: An empirical analysis. *Strategic Management Journal*, 39(9):2489–2516.

Helfat, C. E. and Peteraf, M. A. 2003. The dynamic resource-based view: capability lifecycles. *Strategic Management Journal*, 24(10):997–1010.

Helfat, C. E. and Winter, S. G. 2011. Untangling Dynamic and Operational Capabilities: Strategy for the (N)ever-Changing World. *Strategic Management Journal*, 32(11):1243–1250.

Henderson, R. and Cockburn, I. 1994. Measuring Competence? Exploring Firm Effects in Pharmaceutical Research. *Strategic Management Journal*, 15:63–84.

Henderson, R. and Cockburn, I. 1996. Scale, scope, and spillovers: The determinants of research productivity in drug discovery. *RAND Journal of Economics (RAND Journal of Economics)*, 27(1):32–59.

Jaffe, A. B. 1986. Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value. *American Economic Review*, 76(5):984–1001.

Jewkes, J., Sawkes, D., and Stillerman, R. 1958. *The sources of invention*. Macmillan.

Kamakura, W. A., Kim, B.-D., and Lee, J. 1996. Modeling Preference and Structural Heterogeneity in Consumer Choice. *Marketing Science*, 15(2):152–172.

Kamakura, W. A. and Russell, G. J. 1989. A Probabilistic Choice Model for Market Segmentation and Elasticity Structure. *Journal of Marketing Research (JMR)*, 26(4):379–390.

Laursen, K. and Salter, A. 2006. Open for innovation: the role of openness in explaining innovation performance among U.K. manufacturing firms. *Strategic Management Journal*, 27(2):131–150.

Levin, R. C., Klevorick, A. K., Nelson, R. R., and Winter, S. G. 1987. Appropriating the Returns from Industrial Research and Development. *Brookings Papers on Economic Activity*, 18(3):783.

Mahmood, I. P., Zhu, H., and Zajac, E. J. 2011. Where can capabilities come from? network ties and capability acquisition in business groups. *Strategic Management Journal*, 32(8):820–848.

Mani, D. and Nandkumar, A. 2016. The differential impacts of markets for technology on the value of technological resources: An application of group-based trajectory models. *Strategic Management*

Journal, 37(1):192–205.

McLachlan, G. and Peel, D. 2004. *Finite Mixture Models*. John Wiley & Sons.

Porter, M. E. 1998. Clusters and the New Economics of Competition.

Roeder, K., Lynch, K. G., and Nagin, D. S. 1999. Modeling Uncertainty in Latent Class Membership: A Case Study in Criminology. *Journal of the American Statistical Association*, 94(447):766–776.

Rothaermel, F. T. and Alexandre, M. T. 2009. Ambidexterity in Technology Sourcing: The Moderating Role of Absorptive Capacity. *Organization Science*, 20(4):759–780.

Sampson, R. C. and Shi, Y. 2023. Are U.S. firms becoming more short-term oriented? Evidence of shifting firm time horizons from implied discount rates, 1980–2013. *Strategic Management Journal*, 44(1):231–263.

Sutton, J. 1998. *Technology and Market Structure: Theory and History*. Mit Press.

Veugelers, R. and Cassiman, B. 1999. Make and buy in innovation strategies: evidence from Belgian manufacturing firms. *Research Policy*, 28(1):63–80.

Vivas, C. and Barge-Gil, A. 2015. Impact on Firms of the Use of Knowledge External Sources: A Systematic Review of the Literature. *Journal of Economic Surveys*, 29(5):943–964.

Volberda, H. W., Foss, N. J., and Lyles, M. A. 2010. PERSPECTIVE—Absorbing the Concept of Absorptive Capacity: How to Realize Its Potential in the Organization Field. *Organization Science*, 21(4):931–951.

West, J. and Bogers, M. 2014. Leveraging External Sources of Innovation: A Review of Research on Open Innovation. *Journal of Product Innovation Management*, 31(4):814–831.

Winter, S. G. 2003. Understanding dynamic capabilities. *Strategic Management Journal*, 24(10):991–995.

Wu, B. 2013. Opportunity costs, industry dynamics, and corporate diversification: Evidence from the cardiovascular medical device industry, 1976–2004. *Strategic Management Journal*, 34(11):1265–1287.

Table 1: Probability of New Product Introduction

Variable	**Prob(New Product)**
Size (q)	$\uparrow$ (threshold effect, T)
External Invention Supply (θ)	$\uparrow$ (external invention effect, Z)
External Knowledge (k)	$\uparrow$ (R&D effect, R)

Note: Firms introduce a new product when $y \geq c/q$. Increases in q, θ, or k increase the probability through distinct mechanisms: threshold (q), R&D (k), or external inventions (θ).

Table 2: Comparative Statics by Firm Capability

Firm Type	**Variable**	**New-to-the-market**	**New-to-the-firm**
High-capability	Size (q)	$\uparrow^R$	$?(\uparrow^T \leftrightarrow^R)$
	External Invention Supply (θ)	$?(\uparrow^Z \downarrow^R)$	$?(\uparrow^Z \downarrow^R)$
	External Knowledge (k)	$\uparrow^R$	$?(\uparrow^T \leftrightarrow^R)$
Low-Capability	Size (q)	0	$\uparrow^T$
	External Invention Supply (θ)	$\uparrow^Z$	$?(\uparrow^T \leftrightarrow^Z)$
	External Knowledge (k)	0	0

Note: Superscripts denote mechanisms: T = threshold effect, R = R&D effect, Z = external invention effect.

Table 3: Source of New-to-the-market product

Variable	**Share external**
High-Capability	$\downarrow^R$
High-Capability X Size	$\downarrow^R$

Note: R&D investment increases internal invention quality, and both capability and size increase R&D performance.

Table 4: New-to-the-market (NTM) product, new-to-the-firm (NTF) product or None: Multinomial and FMM Logits

	MNL			FMM Logit: 2 classes					
				Latent class 1			Latent class 2		
	NTM vs none (1)	NTF vs none (2)	NTM vs NTF (3)	NTM vs none (4)	NTF vs none (5)	NTM vs NTF (6)	NTM vs none (7)	NTF vs none (8)	NTM vs NTF (9)
Standalone	0.09	0.16	-0.07	-2.37	-3.43	1.06	-0.71	0.97	-1.69
	(0.15)	(0.14)	(0.16)	(1.99)	(1.90)	(0.88)	(0.46)	(0.56)	(0.88)
Multiproduct BU	0.21	-0.06	0.27	-0.25	-0.29	0.04	0.71	-0.07	0.78
	(0.10)	(0.08)	(0.11)	(0.43)	(0.39)	(0.32)	(0.34)	(0.22)	(0.43)
BU size (log)	1.01	0.66	0.35	4.25	3.29	0.96	0.80	0.80	-0.01
	(0.07)	(0.07)	(0.07)	(0.89)	(0.78)	(0.34)	(0.22)	(0.13)	(0.27)
Firm age (years)	-0.004	-0.003	-0.001	-0.05	-0.02	-0.02	0.02	-0.01	0.02
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
High tech industry	1.00	0.31	0.70	1.93	0.81	1.12	0.41	0.39	0.02
	(0.10)	(0.09)	(0.11)	(0.50)	(0.53)	(0.35)	(0.35)	(0.25)	(0.48)
Homogeneous industry	0.06	-0.01	0.07	-1.01	-2.31	1.30	-0.77	1.01	-1.79
	(0.10)	(0.09)	(0.11)	(0.54)	(0.79)	(0.49)	(0.50)	(0.32)	(0.64)
External Invention Supply	0.13	0.11	0.02	-0.23	-0.24	0.01	0.37	0.24	0.13
	(0.05)	(0.04)	(0.05)	(0.21)	(0.24)	(0.14)	(0.17)	(0.11)	(0.20)
Constant	-3.37	-2.04	-1.33	-1.4	1.37	-2.77	-4.29	-4.15	-0.14
	(0.24)	(0.21)	(0.25)	(2.05)	(1.74)	(1.22)	(1.17)	(0.81)	(1.59)
Avg class prob					0.35			0.65	
Observations		4692					4692		
ll		-3671.69					-3616.66		

Robust standard errors in parentheses. These regressions predict the likelihood of a firm introducing a new-to-the-market (NTM) product, a new-to-the-firm (NTF product, or not introducing a new product, using a multinomial logit (columns 1-3) and FMM logit (columns 4-9) specification. Regressors include: external invention supply, whether a firm is standalone (versus part of a multiunit firm), whether the firm operates in multiple submarkets within their industry (multiproduct BU), business unit size and age, and whether the firm is in a high tech industry and a homogeneous industry. Coefficients represent the change in log odds: marginal effects are listed in Table 5.

Table 5: New-to-the-market (NTM) product, new-to-the-firm (NTF) product or None: Average Marginal Effects

	MNL			Latent Class Logit: 2 classes					
				Latent class 1			Latent class 2		
	Pr(none) (1)	Pr(NTF) (2)	Pr(NTM) (3)	Pr(none) (4)	Pr(NTF) (5)	Pr(NTM) (6)	Pr(none) (7)	Pr(NTF) (8)	Pr(NTM) (9)
Standalone*	-0.03	0.03	0.00	0.34	-0.36	0.02	-0.07	0.12	-0.05
Multiproduct BU*	-0.01	-0.02	0.03	0.03	-0.02	-0.01	-0.03	-0.01	0.04
High tech industry*	-0.13	0.02	0.10	-0.14	-0.06	0.20	-0.07	0.05	0.02
Homogeneous industry*	0.00	-0.01	0.01	0.20	-0.31	0.11	-0.08	0.13	-0.05
External Invention Supply	-0.03	0.02	0.01	0.03	-0.01	-0.01	-0.05	0.03	0.02
BU size (log)	-0.18	0.08	0.09	-0.40	0.13	0.27	-0.13	0.09	0.04
Firm age (years)	0.001	-0.001	0.00	0.003	0.001	-0.004	0.00	-0.001	0.001

* Indicator variables

Marginal effects represent the increase in probability of an outcome (as compared to any other alternative). For indicator (0/1) variables, it is the increase in probability associated with a change from 0 to 1 (e.g., average increase in probability from being a standalone firm).

Table 6: Tests for fit: number of latent classes

Classes	LL	R2 (McF)	AIC	BIC	N	K
4	-3578.9	0.197	7291.8	7724.1	4692	67
3	-3599.0	0.192	7298.0	7620.7	4692	50
2	-3630.3	0.185	7326.5	7539.5	4692	33
1	-3671.7	0.058	7375.4	7478.6	4692	16

This table present the typically used fit statistics across the one, two, three, and four class models which we used to select the number of latent classes. The McFadden R^2 suggests models with more than one class provide much better fit (i.e., in one class model is 0.06 and is more than 0.19 in the two, three or four class models). Further, fit from McFadden R^2 is not substantially improved moving from 2 to 3 or 4 classes. The BIC is lowest in the single class model. Balancing these diagnostics, and aiming for ease of interpretability given our intended purpose to use the model to explore capabilities, we chose a model with two latent classes (i.e., high- and low-capability).
Note: BIC=2*LL+K*ln(N)

Table 7: New-to-the-market (NTM) product, new-to-the-firm (NTF) product or None: Multinomial and FMM Logits, External knowledge

	MNL			FMM Logit: 2 classes					
				Latent class 1			Latent class 2		
	NTM vs none (1)	NTF vs none (2)	NTM vs NTF (3)	NTM vs none (4)	NTF vs none (5)	NTM vs NTF (6)	NTM vs none (7)	NTF vs none (8)	NTM vs NTF (9)
Standalone	0.09	0.16	-0.07	-5.01	-3.59	-1.42	0.77	-0.18	0.95
	(0.15)	(0.14)	(0.16)	(4.03)	(3.71)	(1.07)	(0.46)	(0.35)	(0.69)
Multiproduct BU	0.22	-0.06	0.27	1.27	-0.14	1.40	-0.12	0.23	-0.35
	(0.10)	(0.08)	(0.11)	(0.68)	(0.41)	(0.75)	(0.22)	(0.25)	(0.36)
BU size (log)	1.01	0.65	0.35	4.11	2.87	1.24	0.74	0.71	0.03
	(0.07)	(0.07)	(0.07)	(1.12)	(0.78)	(0.56)	(0.15)	(0.16)	(0.24)
Firm age (years)	-0.003	-0.003	0.00	-0.01	-0.03	0.01	-0.01	0.01	-0.01
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)
High tech industry	1.04	0.34	0.70	0.35	0.34	0.02	1.38	0.05	1.33
	(0.10)	(0.09)	(0.11)	(0.62)	(0.44)	(0.63)	(0.22)	(0.34)	(0.46)
Homogeneous industry	0.05	-0.02	0.07	-2.56	-0.45	-2.11	0.68	-0.50	1.18
	(0.10)	(0.09)	(0.11)	(1.02)	(0.45)	(1.06)	(0.28)	(0.47)	(0.61)
External knowledge supply	0.02	0.01	0.01	0.17	-0.03	0.20	-0.01	0.12	-0.13
	(0.01)	(0.01)	(0.01)	(0.09)	(0.05)	(0.11)	(0.02)	(0.04)	(0.05)
Constant	-3.50	-2.09	-1.41	-2.69	1.84	-4.53	-3.78	-4.21	0.43
	(0.26)	(0.23)	(0.28)	(3.38)	(3.71)	(1.83)	(0.60)	(0.92)	(1.11)
Avg class prob					0.30			0.70	
Observations		4692				4692			
ll		-3674.6				-3622.56			

Robust standard errors in parentheses. These regressions predict the likelihood of a firm introducting a new-to-the-market (NTM) product, a new-to-the-firm product, or not introducing a new product, using a multinomial logit (columns 1-3) and FMM logit (columns 4-9) specification. Regressors include: external knowledge supply, whether a firm is standalone (versus part of a multi-unit firm), whether the firm operates in multiple submarkets within their industry (multi-product BU), business unit size and age, and whether the firm is in a high tech industry and a homogeneous industry. Coefficients represent the change in log odds.

Table 8: Share of new-to-the-market product innovators using external sources, by inventive capability and BU size

	Small BU	Large BU	All sizes
Low-Capability	0.493	0.457	0.487
	(0.040)	(0.049)	(0.034)
High-capability	0.479	0.403	0.475
	(0.029)	(0.065)	(0.028)
All capability	0.485	0.444	0.480
	(0.023)	(0.039)	(0.020)

This table presents a cross tabulation of the use of external sources in new-to-the-market product innovations across both inventive capability and business unit (BU) size. Standard errors in parentheses.

Table 9: Sources of invention: internal, customers, other external source, among innovating firms (Mlogit. Ref cat: Internal source)

	non-cust (1)	cust (2)
Inventive capability	0.11	-0.53
	(0.28)	(0.28)
BU size (log)	-0.04	-0.27
	(0.12)	(0.12)
Ext invention supply	0.32	-0.35
	(0.25)	(0.22)
Vertically integrated	0.56	0.52
	(0.24)	(0.24)
Multiproduct BU	-0.32	-0.43
	(0.24)	(0.24)
Industry FE	Yes(17)	Yes(17)
Constant	-0.69	-0.21
	(0.49)	(0.55)
Observations	1,124	
LL	-660.8	

Robust standard errors in parentheses. These regressions predict the likelihood of an firm introducing a new-to-the-market product using non-customer or customer sourced invention, as compared to internal (the reference category). The predictors are: firm inventive capability, business unit size, external invention supply, whether is vertically integrated (i.e., has supplier or customers inside the same firm), whether the firm operates in multiple submarkets within their industry (multiproduct BU), and 17 industry fixed effects.

Table 10: Market share increase, by product innovation outcome and capability (linear probability)

	Market Share Increase		
	(1)	(2)	(3)
Inventive capability	0.20	-0.00	-0.11
	(0.03)	(0.06)	(0.11)
New-to-the-market product		0.18	0.02
		(0.05)	(0.07)
New-to-the-firm product		0.13	0.17
		(0.03)	(0.06)
Capability * NTM product			0.32
			(0.13)
Capability * NTF product			0.01
			(0.15)
BU size (log)	0.03	-0.00	-0.01
	(0.01)	(0.01)	(0.02)
Start-up BU	0.20	0.20	0.19
	(0.06)	(0.05)	(0.06)
Constant	0.53	0.58	0.60
	(0.04)	(0.04)	(0.05)
Industry FE	Yes(45)	Yes(45)	Yes(45)
Observations	4,316	4,316	4,316
R-squared	0.04	0.05	0.05
LL	-3036	-3021	-3016

Robust standard errors in parentheses. These regressions predict the likelihood the focal respondent business unit experienced a market share increase from 2008 to 2009. The predictors are: firm inventive capability, product innovation outcome, business unit size, external invention supply, whether is vertically integrated (i.e., has supplier or customers inside the same firm), whether the firm operates in multiple submarkets within their industry (multiproduct BU), and 45 industry fixed effects.

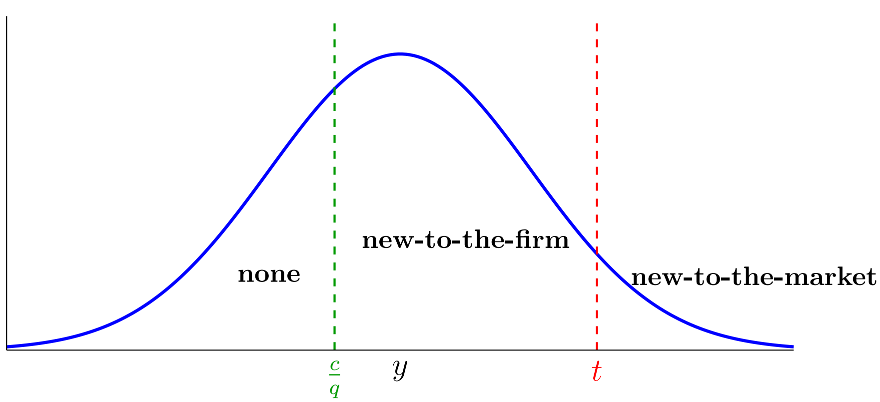


Figure 1: Product innovation categories based on (quality) thresholds of y: $y < \frac{c}{q}$ = no new product, $\frac{c}{q} < y < t$ = new-to-the-firm product, $y > t$ = new-to-the-market product.

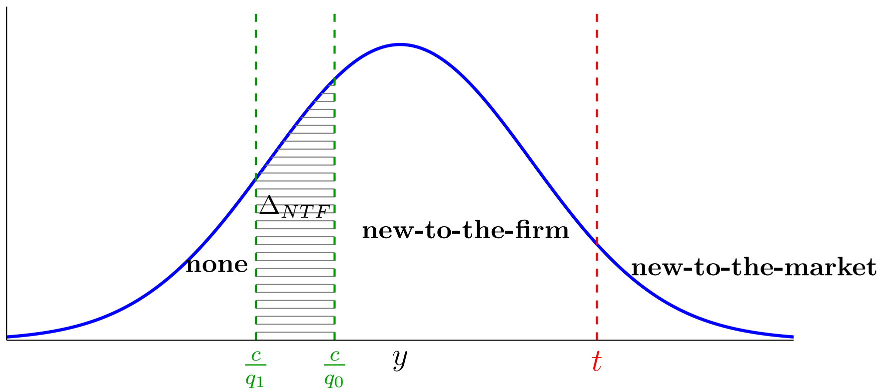


Figure 2: Effect of size (from q_0 to q_1) for *low-capability firms*. The effect is to increase in overall probability of a new product (by $y_0 = \frac{c}{q_0}$ to $y_1 = \frac{c}{q_1}$), as depicted by the area Δ_{NTF}. The increase is completely driven by new-to-the-firm with no increase in new-to-the-market.

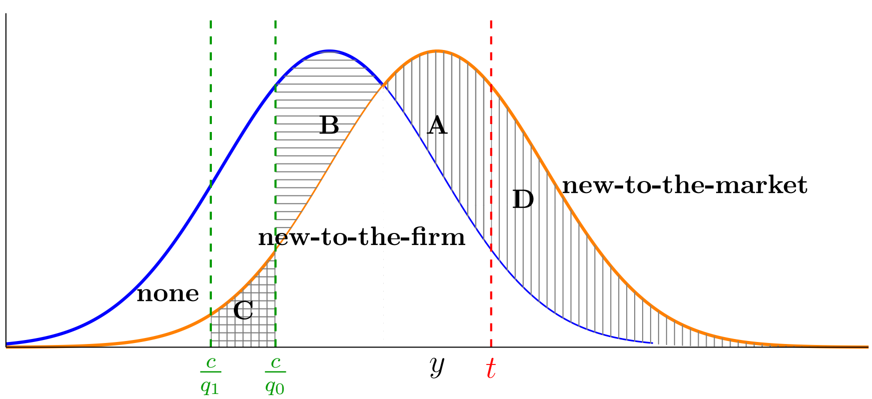


Figure 3: Effect of size (from q_0 to q_1) for *high-capability firms*. The direct effect is to lower the threshold for any new product from $y_0 = \frac{c}{q_0}$ to $y_1 = \frac{c}{q_1}$. The indirect effect, through increased R&D, is to shift the distribution of y to the right. Clearly the probability of any new product increases with size. The effects by type of product innovation are: $\Delta_{NTM} = D$, which implies that new-to-the-market product innovation increases because of the R&D effect. $\Delta_{NTF} = C + (A - B)$, which implies the effect on new-to-the-firm products is ambiguous. The threshold effect is positive (C), but the R&D effect is ambiguous ($A - B$).

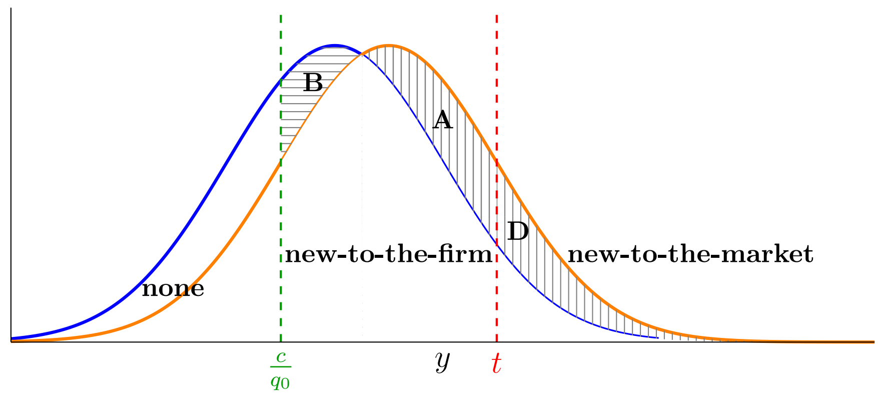


Figure 4: Effect of external supply of inventions for *low-capability firms*.
By increasing the quality of external inventions (z), the overall effect is to shift the distribution of y to the right, and thereby increase product innovation. The effects by type of product innovation are: $\Delta_{NTM} = D$, which implies that new-to-the-market product innovation increases because of the quality shift. $\Delta_{NTF} = A - B$, which implies the effect on new-to-the-firm products is ambiguous.

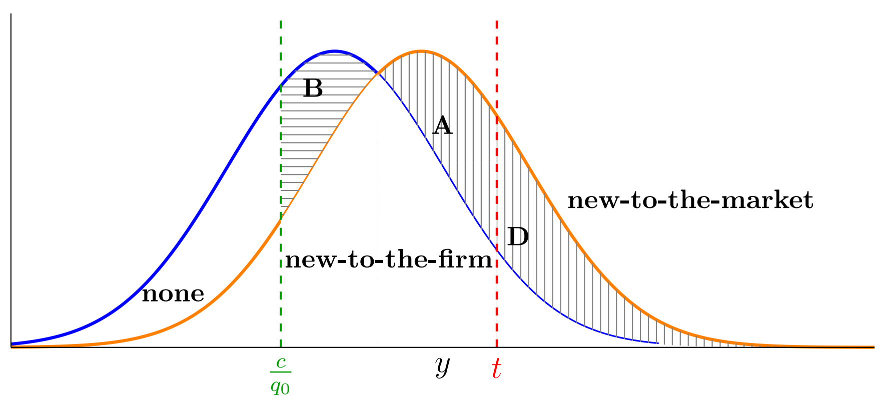


Figure 5: Effect of external knowledge for *high-capability firms*.
Increasing external knowledge reduces the costs of R&D and thereby increases the quality of internal innovation. The overall effect is to shift the distribution of y to the right, and thereby increase product innovation. The effects by type of product innovation are: $\Delta_{NTM} = D$, which implies that new-to-the-market product innovation increases because of the quality shift. $\Delta_{NTF} = A - B$, which implies the effect on new-to-the-firm products is ambiguous.

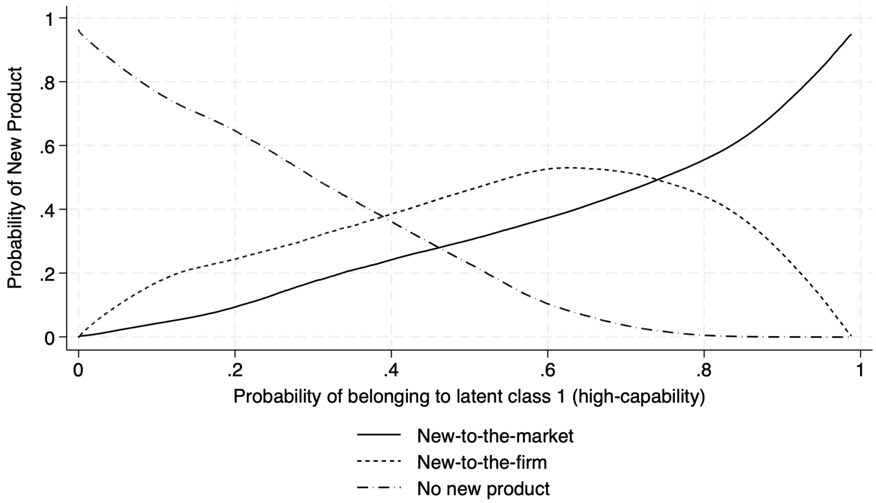


Figure 6: Likelihood of a New-to-the-market product, a New-to-the-firm product, or No new product, by probability of belonging to latent class 1 (high-capability)

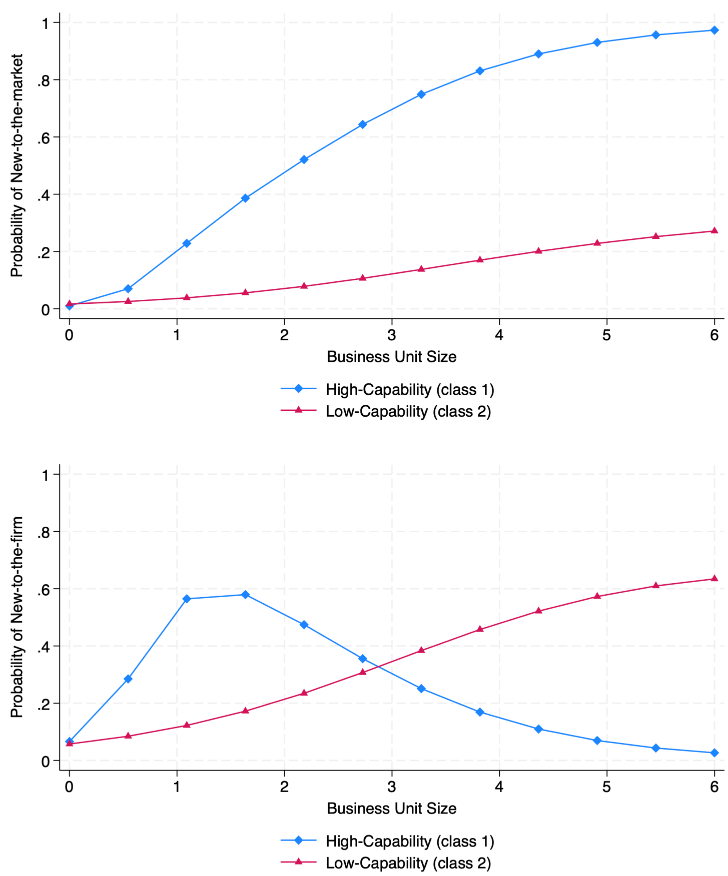
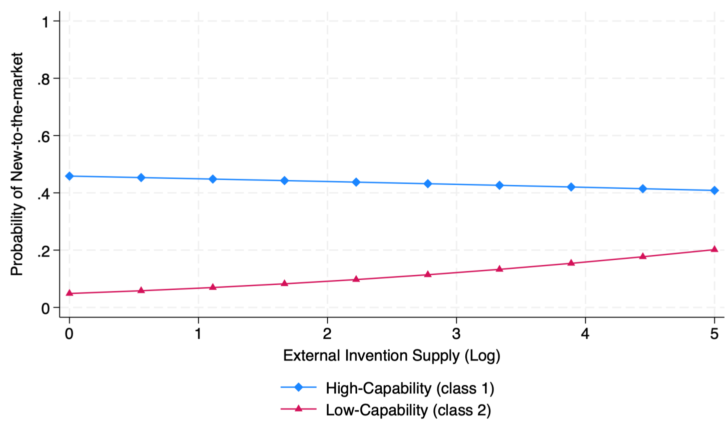


Figure 7: Likelihood of New-to-the-market product and New-to-the-firm product, by Business Unit Size for High- and Low-Capability firms



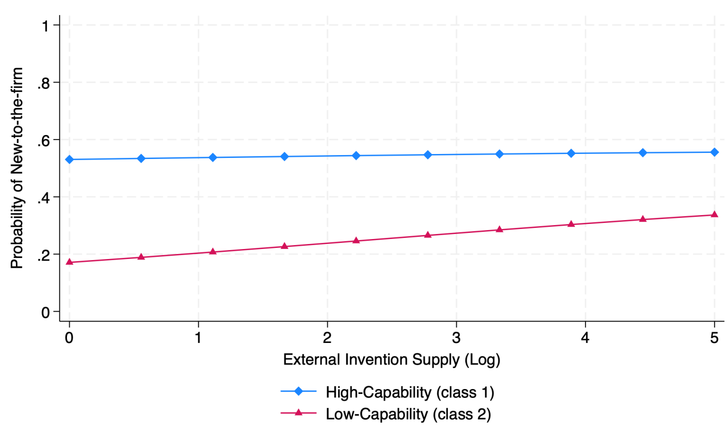


Figure 8: Likelihood of New-to-the-market and New-to-the-firm product, by External Supply of Invention (R&D specialists) for High- and Low-Capability firms

Appendix

A Appendix Tables

Table A1: Comparing latent class assignment: Invention (Table 4) vs. Knowledge (Table 7)

		Table 7 Low-Capability	Table 7 High-Capability
Table 4	Low-Capability	2865	282
	High-Capability	116	1711

This table presents a cross tabulation of the latent classes (capabilities) from Table 4 (external invention supply) and Table 7 (external knowledge supply). The off-diagonal values represent the number of firms differentially classified across the two analyses.

Table A2: Latent Class Probabilities

	class 1 (vs class 2)
Food & Textiles (NAICS 31)	-0.695
	(0.260)
Wood & Chemicals (NAICS 32)	-0.571
	(0.257)
Pharmaceuticals (NAICS 3254)	0.294
	(0.332)
Machinery & Transport (NAICS 331-3, 37)	-0.360
	(0.218)
Computers/Electronics (NAICS 334)	1.639
	(0.515)
Semiconductor (NAICS 3344)	0.611
	(0.318)
Instruments (NAICS 3345)	0.582
	(0.455)
Electrical Equipment (NAICS 335)	-0.106
	(0.258)
Transportation (NAICS 336)	0.055
	(0.316)
Medical Equipment (NAICS 3391)	0.683
	(0.427)
Constant (Ref: Misc Manu (NAICS 339))	0.241
	(0.457)

Robust standard errors in parentheses.

B Model Appendix

We first show that the optimal R increases with the size of the firm, q and with external knowledge, but decreases with an increase in the supply of external invention, θ. We use these results to show how size, external knowledge, and external invention affect the probabilities of new-to-the-market and new-to-the-firm products, and how this differs between high- and low-capability firms.

B.1 Inventive capability and R&D investment

Firms with high inventive capabilities can invest in R&D and improve the quality of their internal inventions. The cost of R&D is denoted by $\phi(R)$, where R is the amount of R&D investment. We assume that $\phi(R)$ is increasing and convex in R i.e., $\phi_R = \dfrac{\partial \phi}{\partial R} > 0$, and $\phi_{RR} = \dfrac{\partial^2 \phi}{\partial R^2} > 0$. We further assume that $\dfrac{\partial F}{\partial R} = F_R(x) < 0$. In words, a higher R&D investment shifts the distribution of internal inventions to the right. Firms with low inventive capabilities lack the ability to invest.

High-capability firms choose R to maximize expected profits given by

$$\max_R q \int_{\frac{c}{q}}^{A} (qy - c)h(y; R)dy - \phi(R) \tag{6}$$

Equation 6 can be written as

$$\max_R qY - c - q \int_{\frac{c}{q}}^{Y} H(y; R)dx - \phi(R) \tag{7}$$

The first order condition for an interior optimum is

$$-q \int_{\frac{c}{q}}^{Y} H_R(y)dy - \phi(R) \tag{8}$$

We assume that the sufficient condition for an interior maximum is satisfied, i.e.,

$$D = -q \int_{\frac{c}{q}}^{Y} H_{RR}(y)dy - \phi_{RR}(R) < 0 \tag{9}$$

B.2 Comparative statics for R&D (for high-capability firms)

Here we derive static comparisons for how the choice of R&D varies with three key factors: size, the supply of external invention, and the availability of external knowledge. We find the optimal R increases with the size of the firm, q, and decreases with an increase in the supply of external invention, θ, and increases with the availability of external knowledge, k.

B.2.1 Size

The comparative static for q is obtained by total differentiation of the first order condition

$$D\frac{\partial R}{\partial q} = \int_{\frac{c}{q}}^{Y} H_R(y)dy \tag{10}$$

It follows that $\dfrac{\partial R}{\partial q} \geq 0$ because $D < 0$ and $H_R = F_R G < 0$.

B.2.2 Supply of external invention

Let an increase in θ represent a rightward shift in the distribution of external invention. This implies that $G_\theta \leq 0$, and in turn, implies that $H_\theta = FG_\theta \leq 0$. The comparative static for θ is obtained by

$$D\frac{\partial R}{\partial \theta} = q\int_{\frac{c}{q}}^{Y} H_{R\theta}(y)dy \tag{11}$$

The result follows upon noting that $H_{R\theta} = F_R G_\theta > 0$

B.2.3 External knowledge

Let k represent external knowledge. We assume $\phi_{Rk} = \frac{\partial^2 \phi}{\partial R \partial k} \leq 0$ i.e., knowledge reduces the marginal cost of R&D.

$$\frac{\partial R}{\partial k} = -\frac{\phi_{Rk}}{D} \geq 0 \tag{12}$$

B.3 Effect of size on product innovation

Recall that inventions with quality greater than t are assumed to be *new-to-the-market*, and those below are *new-to-the-firm*. The probability of new-to-the-market, is $1 - H(t)$, and that of new-to-the-firm is $H(t) - H(\frac{c}{q})$. Let the probability of *new-to-the-market* be denoted by M and the probability of *new-to-the-firm*, be denoted by N.

B.3.1 Low-capability

$$\begin{aligned} \frac{\partial M}{\partial q} &= 0 \\ \frac{\partial N}{\partial q} &= h(c/q)\frac{c}{q^2} \geq 0 \end{aligned} \tag{13}$$

B.3.2 High-capability

$$\begin{aligned} \frac{\partial M}{\partial q} &= -H_R(t)\frac{\partial R}{\partial q} \geq 0 \\ \frac{\partial N}{\partial q} &= (H_R(t) - H_R(c/q))\frac{\partial R}{\partial q} + h(c/q)\frac{c}{q^2} \end{aligned} \tag{14}$$

The inequality for M follows by noting that $-H_R(t) = -F_R(t)G(t) \geq 0$ and $\frac{\partial R}{\partial q} \geq 0$. The second term in the expression for N is always positive. The first term in the expression for N is positive if $H_R(t) = F_R(t)G(t) > H_R(c/q) = F_R(c/q)G(c/q)$. If $F_R(t) \leq F_R(c/q)$, then the first term is negative because $G(t) > G(c/q)$ and because $F_R \leq 0$. In general, the expression for N cannot be signed.

B.4 Effect of increase in external invention

B.4.1 Low-capability

$$\begin{aligned} \frac{\partial M}{\partial \theta} &= -H_\theta(t) \geq 0 \\ \frac{\partial N}{\partial q} &= H_\theta(t) - H_\theta(c/q) \end{aligned} \tag{15}$$

The expression for N cannot be signed in general, unless $G_\theta(c/q) \leq G_\theta(c/q)$, in which case, the expression is negative.

B.4.2 High-capability

$$\begin{aligned}\frac{\partial M}{\partial \theta} &= -H_{\theta}(t) - H_R(t)\frac{\partial R}{\partial \theta} \\ \frac{\partial N}{\partial q} &= (H_{\theta}(t) - H_{\theta}(c/q)) + (H_R(t) - H_R(c/q))\frac{\partial R}{\partial \theta}\end{aligned} \tag{16}$$

The first term in the expression for M is positive but the second term is negative, reflecting the reduction in R. Neither term in the expression for N can be signed, so the expression as a whole cannot be signed.

B.5 Effect of external knowledge

By assumption, external knowledge cannot affect firms that do not invest in R&D. For high-capability firms

$$\begin{aligned}\frac{\partial M}{\partial k} &= -H_R(t)\frac{\partial R}{\partial k} \geq 0 \\ \frac{\partial N}{\partial k} &= (H_R(t) - H_R(c/q))\,\frac{\partial R}{\partial k}\end{aligned} \tag{17}$$

The expression for N is negative if $F_R(t)G(t) < F_R(c/q)G(c/q)$ and positive otherwise.

B.6 Share of internal inventions in new-to-the-market products

For new-to-the-market products, we also observe whether the source was external or not. It is easy to see that the share of external sources should be higher for low-capability firms than high-capability firms because high-capability firms have better quality internal inventions.