What are We Really Good at?

Product Strategy with Uncertain Capabilities*

Jeanine Miklós-Thal† Michael Raith‡ Matthew Selove§

June 15, 2016

Abstract

Firms often learn about their own capabilities through their products’ successes and failures. This paper explores the interaction between such learning from experience and product strategy in a formal model. We consider a firm that can launch a sequence of products, where each product’s performance depends on the fit between the firm’s capabilities and the product. A successful new product always causes the firm to become more optimistic about the capability most relevant for that product; however, it can also cause the firm to become less optimistic about some of its other capabilities, including capabilities the new product does not use. A product launch generates useful information for future decisions if it leads to learning about capabilities used by potential future products. We find that a product sharing few or even no capabilities with potential future products may generate more useful information than a product with greater overlap.

*We are grateful to Birger Wernerfelt, Juanjuan Zhang and seminar participants at Cambridge University, Chapman University, the Marketing Science Conference, the Utah PSI Conference, the Marketing Dynamics Conference, and the Strategic Management Society Conference for helpful comments.

†Simon Business School, University of Rochester; e-mail: jeanine.miklos-thal@simon.rochester.edu.

‡Simon Business School, University of Rochester; e-mail: raith@simon.rochester.edu.

§Argyros School of Business and Economics, Chapman University; e-mail: selove@chapman.edu.
1 Introduction

It is widely accepted that a firm’s product strategy should exploit its existing capabilities and help to build new ones (Wernerfelt 1984; Teece, Pisano, and Shuen 1997; Dutta, Narasimhan, and Rajiv 1999; Slotegraaf, Moorman, and Inman 2003). Implementing this advice can be challenging, however, because firms often face considerable uncertainty about the nature of their capabilities. Capabilities are hard to identify directly, e.g., through “capabilities audits” (Ulrich and Smallwood 2004), because they are deeply embedded in a firm’s people, processes, and culture (Day 1994; Grewal and Slotegraaf 2007). This implies that managers must infer what their firm is good at largely based on what has, and has not, succeeded in the past, but such inference tends to be imperfect.

Eastman Kodak is a case in point. The company’s long-time dominance of the camera-film business arguably rested on some combination of expertise in two main areas: imaging and chemical engineering. In the early 1980s, when Kodak’s management foresaw the eventual decline of film and began to look for new opportunities, the prevailing view was that Kodak was primarily a chemicals company, which prompted several costly investments to diversify into pharmaceuticals (Feder 1988). A failure to develop any profitable new drugs, however, caused a shift in the leadership’s beliefs about the company’s strengths (along with turnover of the leadership itself), and by the early 1990s, the company’s leadership “decided that its expertise lay not in chemicals but in imaging” (Economist 2012). This change in beliefs led to a major shift in strategic focus toward digital photography (Gavetti, Henderson, and Giorgi 2005), which, however, was ultimately ill-fated too because of competition from low-cost electronics companies.\(^1\)

Another example involves the website eHarmony, which matches men and women interested in long-term relationships based on their responses to a detailed personal questionnaire. Its success arguably resulted from a combination of (i) expertise

\(^1\)In addition to the citations listed above, our interpretation of this example is also based on our personal interview with former Kodak executive Larry Matteson.
in romantic relationships, and (ii) the capability to use data analytics to create matches. In an attempt to leverage its expertise in romantic relationships, in 2010 the company introduced Jazzed.com, a dating website for people interested in short-term relationships. (For short-term relationships, high levels of compatibility are less important than physical attraction (Harwell 2015), and so the company’s complex matching expertise was less relevant in this new market.) The Jazzed.com website failed to attract many customers and was closed down two years after its launch. In a remarkable shift in strategy, the company is now set to launch a website that will match potential employees with jobs (Nash 2015), and is contemplating other matching services in the future, such as helping people find friends or find financial advisors (Jacobs 2014). It appears that the company has concluded that its core competence is matching, not romantic relationships, and has adjusted its strategy accordingly.\footnote{In the words of founder and CEO Neil Clark Warren, “We have a new approach to what we’re trying to do. We want to be a relationship site. We don’t want to just match people for marriage, we also want to try to help them find the right job.” (Wells 2015).}

This paper takes a first step towards exploring the implications of uncertainty about capabilities for product strategy in a formal model. We analyze how firms can learn about their capabilities from past product successes and failures, and then use these results to provide guidance for optimal product launch decisions. As in the examples above, there is a two-way link between product strategy and beliefs about capabilities (or “skills”). Naturally, the firm would like to introduce new products that rely on skills the firm believes it has, based on its past product experiences. Each new product introduction, in turn, causes the firm to update its beliefs about its skills, which affects the expected profitability of future product introductions.

In our formal model, a firm is “born” with prior beliefs about its vector of skills, and updates its beliefs upon launching new products and observing their success or failure. The likelihood of success in any market depends on how well the firm’s skills
match the skills that are relevant for that market.\textsuperscript{3} Skills can differ in importance across and within markets. For skills never utilized before, we can think of the firm as either already possessing a latent skill, or equivalently, as acquiring the skill when it first enters a market that uses that skill. The firm’s learning process is imperfect because success or failure in any market depends on multiple skills, and on luck. To illustrate our point as clearly as possible, we abstract away from uncertainty about customer preferences and uncertainty about competitive offerings.\textsuperscript{4}

We show that success in a market always leads to an upward revision of beliefs about the skill most relevant for that market. At the same time, however, a success may also lead to a downward revision of beliefs about some of the firm’s other skills. Importantly, the multi-dimensional nature of skills implies that a (successful or failed) product introduction can generate information about skills that the product itself does not use. This occurs, for example, if the success or failure of the current product helps the firm determine which skill is likely to have caused the success of a previously launched product. Interpreting Kodak’s story in the context of our model, the failure of the company’s pharmaceutical products cast doubt on its general chemical expertise, which, given its earlier success in the camera film business, reinforced management’s belief that the company had strong imaging skills. Hence, failure in one market (pharmaceuticals) led to an upward revision of beliefs about a skill (imaging) not used in that market.

An intuitive implication of this learning process is that if a company repeatedly launches successful products that are similar in one respect but different in others, the firm (and outsiders) will come to view the capability supporting the intersection of those products as a core capability (or “core competence”, Prahalad and Hamel

\textsuperscript{3}We are concerned with skills that are fungible rather than specific to particular products. In the words of Teece (1982), “... a firm’s capability lies upstream from the end products – it lies in a generalizable capability which might well find a variety of final product applications.”

\textsuperscript{4}McGrath (2011) advocates entering markets adjacent to a firm’s current areas of success, so that failure, if it does occur, will produce useful learning. Specifically, the article advocates entering markets where the firm faces uncertainty either over market variables (such as customer preferences) or over capability issues (such as availability of talent), but not both. In terms of this framework, our paper focuses on cases in which new products involve uncertainty over the firm’s capabilities.
For instance, a company like Honda that successfully launches motorcycles, then cars, and eventually a broad line of products including garden power tools and outboard motors and most recently small jets (Ostrower 2015) would be inferred to possess a core competence in engine design, even without detailed direct knowledge of organizational processes.

Building on Wernerfelt’s (1984) argument that new products may be “stepping stones” for future products, we proceed to analyze the information value of a new product for potential future product launch decisions. As one would expect, a product has a positive information value for a future product if the two products are related. Less obviously, a product also has a positive information value if it is unrelated to the future product, as long as it is related to a past product and the past product is related to the future product. For instance, Kodak’s negative experience with pharmaceuticals led management to believe that the company had strong imaging skills, which would be useful information for a potential future product unrelated to pharmaceuticals but relying on the imaging skill.

Not only is the information value of an indirectly related product positive; we also show that it may exceed the information value of a directly related product. More generally, we show that the information value of suitably comparable products for the same future product is not necessarily increasing in their relatedness with the new product. A product that fully overlaps with the future product in terms of the required skills may have a lower information value than a product that has only partial (or no) direct overlap with the future product, and a product with partial overlap may have a lower information value than a product without any direct overlap.

At a prescriptive level, when choosing which projects to undertake, the firm should consider both the profits directly at stake and the value of the resulting information for assessing future opportunities. Consistent with recent advice to practitioners that firms should follow “Strategies for Learning from Failure” (Edmondson 2011) or embrace “Failing by Design” (McGrath 2011), our analysis implies that firms should, in some cases, undertake projects that have a high probability of failure and/or low direct
returns, provided such projects generate sufficiently valuable information.

It is important to note that the firm may not need to conduct a full-scale product launch in order to acquire such information. For example, one can interpret the product “launch” decision in our model as a decision to incur the cost of product development and a test market, or of creating a “minimum viable product” (Eisenmann, Ries, and Dillard 2013), assuming such actions are sufficient to reveal whether the product will succeed, and thus, sufficient to provide useful information about the firm’s skills.

2 Related literature

A sizeable literature in marketing uses game-theoretic models to study optimal product line strategy (e.g., Desai 2001; Kuksov and Villas-Boas 2009; Guo and Zhang 2012; Thomadsen 2012; Amaldoss and Shin 2015), analyzing how a firm’s strategic interactions with competitors and consumers affect the number as well as (horizontal and vertical) positioning of a firm’s products. This literature, however, largely ignores the role played by a firm’s own capabilities in shaping its optimal product line decisions. Our work, in contrast, does not explicitly model consumer choice or competition; in our model, these factors are implicitly reflected in the profits that arise from a product’s success or failure. Abstracting away from strategic interactions with consumers and competitors, we focus instead on how the company’s capabilities drive its product line decisions.

In strategy, a large literature has studied how a firm’s resources, competencies, and capabilities determine which products it should launch (e.g., Wernerfelt 1984; Barney 1991; Teece, Pisano, and Shuen 1997). Our paper differs from this literature in two main respects. First, the vast majority of this literature assumes that firms are
fully aware of their existing capabilities. By contrast, we explore the implications for product strategy when a firm cannot observe its capabilities directly but learns about them based on its products’ successes and failures. Second, few existing papers have used a formal modeling approach to study capabilities. By developing a formal model of how capabilities drive product success outcomes and of the associated learning about capabilities, we are able to derive explicit conditions for each of our results to hold, thus providing boundary conditions for the key insights of our paper.

Our work also speaks to the trade-off between exploration and exploitation central to bandit problems (e.g., Gittins 1979; Lin, Zhang, and Hauser 2015), as the firm’s optimal product line strategy balances its objective to earn high profits in the short-term with its goal to learn more about its own capabilities. In some respects, our model is simpler than standard bandit problems; whereas bandit models typically assume a player repeatedly chooses from a finite number of options over an infinite number of periods, we assume a firm makes a launch versus no-launch decision for three products in an exogenously specified order. Simplifying the decision problem in these respects allows us to develop a rich model of the correlation between outcomes in different markets, and to show how the capability overlap among different markets affects belief-updating and the information value of new products.

Finally, although our paper is not concerned with consumer beliefs or brands, there are noteworthy parallels between a firm’s learning about its skills in our model

---

5Some papers in the strategy literature have argued that a firm might be uncertain about which capabilities caused its past successes (Lippman and Rumelt 1982; Reed and DeFillippi 1990; Barney 1991). Whereas these earlier papers focus on how such “causal ambiguity” might serve as a barrier to resource imitation, we explore how a firm can resolve such ambiguity by launching additional new products.

6Matsusaka (2001) develops a formal model in which a firm sequentially searches for a market that is a good fit for its capabilities. One key difference is this previous paper assumes new markets are independent and identically distributed in terms of organizational fit, whereas our model implies that the outcome in one market provides information about success probability in other related markets. Other papers have developed formal models of capabilities that focus on different questions than our paper. For example, Sakhartov and Folta (2014) develop a model in which a high degree of resource relatedness across a firm’s products reduces the cost of redeploying resources from one product market to another.

7Most of the bandit literature assumes the payoffs to various options are uncorrelated. Recent models that allow for correlated payoffs (e.g., Rusmevichientong and Tsitsiklis 2010) have focused on deriving lower bounds for the performance of various heuristic strategies.
and consumer evaluations of brand extensions. In particular, lab experiments and survey data have shown that customers evaluate brand extensions more favorably if the new product is a good fit with the firm's previous products (Aaker and Keller 1990), and that new product success or failure can have feedback effects onto how customers evaluate brands (Luo, Chen, Han, and Park 2010). More strikingly, similar to the intuitive and counter-intuitive belief-updating in our model, lab experiments have shown that a successful brand extension can strengthen some associations with a brand while weakening other associations (Dacin and Smith 1994). These parallels with our results suggest that consumer evaluations of brand extensions resemble multidimensional learning about what attributes a brand stands for.

3 Model

There are $N$ skills, indexed by $n = 1, 2, ..., N$. We focus on the decisions of a single firm whose skill set is represented by the vector $\theta$, with each element $\theta_n \in \{0, 1\}$ indicating whether the firm has skill $n$. The firm does not know which skills it has, but it can make inferences about its skills based on the performances (success or failure) of the products it launches. Prior to any product introductions, the firm's beliefs about its skills are independently distributed with probabilities $P(\theta_n = 1) = \alpha_n \in (0, 1)$ for all $n$.

There are many ways to model the firm's opportunities to launch products over time. The simplest setting rich enough to convey our results is a three-period model where in each period, the firm can either launch, or not launch, an exogenously available product. With a slight abuse of notation, we use $t \in \{1, 2, 3\}$ to denote both time period $t$ and the product that becomes available in period $t$.

For each product $t$, exactly two skills (out of $N$) are relevant and denote by $I(t)$ the set of those skills (that is, their skill indices). The two skills in $I(t)$ can differ in their importance for product $t$, and we use $H(t) \in I(t)$ and $L(t) \in I(t)$ to denote the indices of the high- and the low-importance skill. In case of equal importance,
$H(t)$ is taken to be the skill with the the lower index.\textsuperscript{8} For example, if skills 1 and 2 are both relevant for product 3, but skill 1 matters more for product 3’s performance than skill 2, then $I(3) = \{1, 2\}$, $H(3) = 1$, and $L(3) = 2$.

Each product $t$ is characterized by $H(t)$, $L(t)$, and a function $R_t(\theta_{H(t)}, \theta_{L(t)}) \in [0, 1)$ that specifies the probability that product $t$ succeeds ($q_t = S$) or fails ($q_t = F$), depending on the firm’s skills.\textsuperscript{9} The following assumptions on this function ensure that each of the skills in $I(t)$ increases product $t$’s success probability, and that skill $H(t)$ matters (weakly) more than $L(t)$ for product $t$.

**Assumption 1.** For each $t$,

$$0 \leq R_t(0, 0) < R_t(0, 1) \leq R_t(1, 0) < R_t(1, 1) < 1.$$  

These assumptions are flexible enough to allow skills to have equal or unequal importance, and to allow either a positive or negative interaction between the skills in terms of their impact on the success probability for any particular product.

The “fit” between the firm’s skill set and a product $t$ is better if the firm possesses a larger number of the skills in $I(t)$. Moreover, if the firm has exactly one of the skills in $I(t)$, its fit with product $t$ is better if the firm possesses skill $H(t)$ than if it possesses skill $L(t)$. Better fit thus increases success probability.

As mentioned, we consider a three-period game. For each $t \in \{1, 2, 3\}$, product $t$’s success results in a gross profit of $\pi_t$, while failure results in a gross profit of zero. At the beginning of each time period $t$, the firm observes the cost $c_t$ of launching product $t$. Then, it decides whether or not to launch product $t$.\textsuperscript{10} If the firm launches product $t$, its performance is realized and the firm earns $\pi_t - c_t$ if $q_t = S$ and $-c_t$ if $q_t = F$.

\textsuperscript{8}This assumption has no implications for any of our results.
\textsuperscript{9}Ruling out success probabilities of 1 ensures that the firm can never perfectly learn its skills.
\textsuperscript{10}The assumption that the firm knows with certainty what products will be available for launch in future periods is made to simplify the exposition of our analysis. As discussed later, the key insights of our analysis would remain unchanged in the presence of uncertainty about future extension opportunities.
\( q_t = F,^{11} \)

The fixed cost \( c_t \) of launching product \( t \in \{1, 2, 3\} \) is drawn from a uniform distribution with support \([0, \pi_t]^{12}\). The support of the distribution ensures that product launch is always profitable ex post if it succeeds, but unprofitable if it fails.

After observing each success or failure, the firm updates its beliefs about its skills using Bayes’ rule. The information set at the beginning of period \( t \) is denoted by \( \Omega_t \). At time \( t = 1 \), it is empty: \( \Omega_1 = \emptyset \). At time \( t = 2 \), \( \Omega_2 = \{q_1\} \) if the firm chose to launch product 1 in period 1 and \( \Omega_2 = \emptyset \) otherwise. At time \( t = 3 \), \( \Omega_3 = \{q_1, q_2\} \) if the firm chose to launch both product 1 and product 2, \( \Omega_3 = \{q_1\} \) if the firm chose to launch product 1 but not product 2, \( \Omega_3 = \{q_2\} \) if the firm chose to launch product 2 but not product 1, and \( \Omega_3 = \emptyset \) otherwise.

The firm has a discount factor \( \delta \in (0, 1] \) and maximizes expected discounted profits. Note this is a finite-period dynamic programming problem.

---

\(^{11}\)When a product is launched, it stays on the market for just one time period. This assumption could easily be relaxed. If the profits from a successful product occur over multiple time periods, our assumptions are equivalent to defining \( \pi_t \) as the total discounted profits from a successful product \( t \).

\(^{12}\)We assume that \( \pi_t \) and \( R_t(\cdot, \cdot) \) are known in order to simplify the formal analysis. Allowing \( \pi_t \) to be uncertain in place of (or in addition to) \( c_t \) would not affect our qualitative insights. The assumption that the probability distribution of \( c_t \) is uniform is inconsequential and merely made to simplify the exposition.
4 Learning About Skills

Our analysis proceeds in two steps. First, in this section, we analyze how the firm learns about its skills from the observed performances of introduced products. Second, in Section 5, we use these findings to explore the firm’s optimal forward-looking product-introduction decisions.

In learning models with a single skill dimension, a standard finding is that product success provides favorable news, that is, it leads to positive belief updating about the firm’s skill level. Conversely, product failure provides unfavorable news and leads to negative belief updating. In what follows, we will refer to such a pattern as “intuitive belief-updating.”

When each product uses multiple skills, on the other hand, then, as we will show, belief updating about some skills can be “counter-intuitive.” In that case, a success provides unfavorable news about a particular skill, while a failure provides favorable news. The main goal of this section is to provide conditions under which belief updating about a particular skill is either “intuitive” or “counter-intuitive” in our

<table>
<thead>
<tr>
<th>Table 1. Overview of notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
</tr>
<tr>
<td>$\theta_n \in {0, 1}$</td>
</tr>
<tr>
<td>$\alpha_n \in (0, 1)$</td>
</tr>
<tr>
<td>$t \in {1, 2, 3}$</td>
</tr>
<tr>
<td>$I(t)$</td>
</tr>
<tr>
<td>$H(t)$</td>
</tr>
<tr>
<td>$L(t)$</td>
</tr>
<tr>
<td>$R_t(\theta_{H(t)}, \theta_{L(t)})$</td>
</tr>
<tr>
<td>$q_t \in {S, F}$</td>
</tr>
<tr>
<td>$\Omega_t$</td>
</tr>
</tbody>
</table>
framework.

Formally, we define the following terms:

**Definition 1** (intuitive versus counter-intuitive belief-updating). “Intuitive belief-updating” occurs at time \( t \) for skill \( n \) if:

\[
E[\theta_n|q_t = S, \Omega_t] > E[\theta_n|\Omega_t] > E[\theta_n|q_t = F, \Omega_t].
\] (1)

“Counter-intuitive belief-updating” occurs at time \( t \) for skill \( n \) if conditions (1) are reversed.

The simplest case to analyze is where the current product uses only skills that have not been used by any previously launched product. Because the beliefs about such skills are independently distributed at the beginning of the current period and each of these skills increases the current product’s success probability, success is favorable news about both of them. The following lemma summarizes the implications of this reasoning for belief updating in periods 1 and 2. (Learning in the third and last period is neglected throughout, because it has no impact on future product introduction decisions in our simple three-period model.)

**Lemma 1.** In period \( t = 1 \), intuitive belief-updating occurs for any skill that is used by product 1 (and no belief updating occurs for all other skills). In period \( t = 2 \), if there is no skill overlap between products 1 and 2 or product 1 was not launched, then intuitive belief-updating occurs for any skill that is used by product 2 (and no belief-updating occurs for all other skills).

Lemma 1 states that belief updating in the first period is intuitive for all skills used in that period. The remainder of this section will focus on learning in the second period, where Lemma 1 leaves open the possibility of counter-intuitive belief updating in case of skill overlap between products 1 and 2.
4.1 Belief-updating with partial skill overlap

Suppose now that products 1 and 2 have exactly one skill in common (“partial skill overlap”). Let \( n \) denote the skill used only by product 1, \( n' \) denote the skill used by both products, and \( n'' \) denote the skill used only by product 2. That is, \( I(1) = \{n, n'\} \) and \( I(2) = \{n', n''\} \). The following lemma states that intuitive belief-updating must occur in period 2 for both of the skills that product 2 uses.

**Lemma 2.** If there is partial skill overlap between products 1 and 2, then in period \( t = 2 \), intuitive belief-updating occurs for the skill that is used by both products (skill \( n' \)) and for the skill that is used only by product 2 (skill \( n'' \)).

In contrast, our next result shows that, under a fairly weak condition on the success probability function \( R_1(\cdot) \) of product 1, the performance of product 2 leads to counterintuitive belief-updating for the skill used only by product 1 (skill \( n \)). The general condition for this to happen can be written as follows:

**Condition 1.**

\[
\frac{\Pr(q_1|\theta_n = 0, \theta_{n'} = 1)}{\Pr(q_1|\theta_n = 1, \theta_{n'} = 1)} > \frac{\Pr(q_1|\theta_n = 0, \theta_{n'} = 0)}{\Pr(q_1|\theta_n = 1, \theta_{n'} = 0)}
\]

Condition 1 is a statement about the relative likelihood by which different skill profiles generate the performance that was observed in the first period. It captures the idea that skills may be substitutes or complements in the sense that having one skill may increase or decrease the role played by the other skill in explaining the observed history.

To fix ideas, suppose that \( q_1 = S \). Condition 1 then always holds if \( R_1(0,0) = 0 \) and thus \( \Pr(q_1|\theta_n = 0, \theta_{n'} = 0) = 0 \), that is, if product 1 cannot succeed unless the firm has at least one of the relevant skills. In this case, the firm knows that if it does not have skill \( n' \) (the skill shared by products 1 and 2), this means that it must have skill \( n \) (the skill used only by product 1). It is therefore intuitive that unfavorable news about skill \( n' \), e.g., due to a failure of product 2, leads to positive belief updating about skill \( n \).
If $R_1(0, 0) > 0$ and skill $n$ is product 1’s high importance skill (still assuming that $q_1 = S$), Condition 1 simplifies to

$$\frac{R_1(1, 0)}{R_1(0, 0)} > \frac{R_1(1, 1)}{R_1(0, 1)}.$$  

(2)

Condition (2) implies that product 1’s success is more indicative of having skill $n$ (the first argument in $R_1$) if the firm does not have skill $n'$ (the skill used by both products 1 and 2) than if it does. Intuitively, this means that, once product 1 has been successful, any unfavorable news about skill $n'$ again leads to positive belief updating about skill $n$. By Lemma 2, a failure (success) of product 2 thus leads to positive (negative) belief updating about skill $n$.

The following lemma formalizes this intuition (a formal proof can be found in the Appendix):

**Lemma 3.** Suppose there is partial skill overlap between products 1 and 2. Then for the skill that is used only by product 1 (skill $n$), counter-intuitive belief-updating occurs in period $t = 2$ if and only if Condition 1 holds given $q_1$; if Condition 1 is reversed, intuitive belief-updating occurs.

For illustration, consider the following numerical example: $q_1 = S$, $\alpha_1 = \alpha_2 = \alpha_3 = 0.5$, and

$$I(1) = \{1, 2\} \text{ and } (R_1(0, 0), R_1(0, 1), R_1(1, 0), R_1(1, 1)) = (0.1, x, x, 0.9),$$

$$I(2) = \{2, 3\} \text{ and } (R_2(0, 0), R_2(0, 1), R_2(1, 0), R_2(1, 1)) = (0.1, 0.5, 0.5, 0.9).$$

Under these assumptions, Condition 1 holds if and only if $x > 0.3$. For example, if $x = 0.5$, then $E[\theta_1|q_1 = S] = E[\theta_2|q_1 = S] = 0.7$, and a successful launch of product

$^{13}$Details on the numerical computations are available upon request.
2 leads to the following belief updating:

\[ E[\theta_1 | q_1 = q_2 = S] \approx 0.67 < E[\theta_1 | q_1 = S], \]
\[ E[\theta_2 | q_1 = q_2 = S] \approx 0.84 > E[\theta_2 | q_1 = S], \]
\[ E[\theta_3 | q_1 = q_2 = S] \approx 0.67 > \alpha_3 = 0.5. \]

Hence, belief updating is intuitive for skills 2 and 3, but counter-intuitive for skill 1. This illustrates the results of Lemmas 2 and 3.

It is worth highlighting that Condition 1 is a very weak form of substitutability in that it can hold even if the skills are complements in the additive sense. For instance, in the numerical example just discussed, skills are complements in the additive sense, i.e.,

\[ R_1(1, 1) - R_1(0, 1) > R_1(1, 0) - R_1(0, 0), \]

for any \( x < 0.5 \). There thus exists a range of value for \( x \), namely \( x \in (0.3, 0.5) \), such that Condition 1 holds (for \( q_1 = S \)) although the skills that product 1 uses are complements in the standard additive sense.

### 4.2 Belief-updating with full skill overlap

Finally, consider cases in which products 1 and 2 use both of the same skills (“full skill overlap”); i.e., \( I(1) = I(2) = \{ n', n'' \} \). Suppose with loss of generality that \( n' \) is the more important skill for product 2: \( H(2) = n' \) and \( L(2) = n'' \).

**Lemma 4.** Suppose there is full skill overlap between products 1 and 2. Then in \( t = 2 \), the firm updates its beliefs as follows:

(a) For product 2’s high importance skill (skill \( n' \)), intuitive belief-updating occurs.

(b) For product 2’s low importance skill (skill \( n'' \)), intuitive belief-updating occurs unless (i) Condition 1 holds and (ii) skill \( n'' \) is sufficiently unimportant for the
performance of product 2, i.e. $R_2(0,1) - R_2(0,0)$ and $R_2(1,1) - R_2(1,0)$ are sufficiently small relative to $R_2(1,1) - R_2(0,1)$.

Lemma 4 implies that if products 1 and 2 fully overlap and both skills are equally important for the success of product 2, then the belief updating about both skills is intuitive. If skills differ in their importance for product 3, belief updating about the more important skill continues to be intuitive. In contrast, if product 2’s less important skill is sufficiently unimportant and Condition 1 holds, then counter-intuitive belief-updating occurs for this skill.

Intuitively, as skill $n''$ becomes less important to product 2’s success, the full-overlap case resembles the partial-overlap case. In the limit, if products 1 and 2 only overlap on a single skill that has a significant impact on the success probability of both products (skill $n'$), then counter-intuitive belief-updating can occur for the skill that has a significant impact on product 1 but not on product 2 (skill $n''$).

To illustrate these results, consider the following example: $\alpha_1 = \alpha_2 = 0.5$, $x \in (0,0.5)$, and

$$I(1) = \{1,2\} \text{ and } (R_1(0,0), R_1(0,1), R_1(1,0), R_1(1,1)) = (0,0.4,0.4,0.8),$$

$$I(2) = \{1,2\} \text{ with } H(2) = 2, \text{ and } (R_2(0,0), R_2(0,1), R_2(1,0), R_2(1,1)) = (0,x,0.5,0.5 + x).$$

It is easy to check that Condition 1 holds for both $q_1 = S$ and $q_1 = F$ in this case. Now consider updated beliefs. At the end of the first period, assuming that product 1 was introduced successfully, we have $E[\theta_1 | q_1 = S] = E[\theta_2 | q_1 = S] = 0.75$. If product 2 is launched and succeeds, then $E[\theta_2 | q_1 = q_2 = S] = \frac{3 + 4x}{3 + 6x} > 0.75$ for any $x \in (0,0.5)$, which implies that intuitive updating occurs for the more important skill 2. In contrast, the belief about the less important skill becomes $E[\theta_1 | q_1 = q_2 = S] = \frac{2 + 6x}{3 + 6x}$, which exceeds $E[\theta_1 | q_1 = S] = 0.75$ if and only if $x > \frac{1}{6} \approx 0.17$. That is, if skill 1 is relatively unimportant for the success of product 2 (low $x$), the success of product 2 makes it more likely that the firm attributes the success of product 1 to

---

$^{15}$Details on the numerical computations are available upon request.
having skill 2 than skill 1.

5 Product strategy

We now examine the interaction between the firm’s product introduction decisions and learning about its own capabilities. The results and examples in this section illustrate two key points of our paper. First, when a firm is uncertain about its skills, new products can generate valuable information about the firm’s likelihood of having various skills, which helps the firm make better product launch decisions in the future. Second, information value can arise either through direct skill overlap with a potential future product, or through skill overlap with a past product that, in turn, has skill overlap with a potential future product. The firm should consider both possible channels of learning when it assesses the total information value that a new product will generate.

5.1 Myopic vs. forward-looking product strategy

Let us define a myopic product strategy as the strategy of introducing in each period the product that maximizes the firm’s expected profit in that period. In our simplified setting with only one product available per period, a myopic firm introduces the available product if and only if its expected profit is positive. Our main concern in this section will be to determine when and why the firm’s optimal strategy will be forward-looking instead of myopic. A forward-looking product strategy takes into account how learning about one’s skills from the success or failure of the next product will affect optimal product launches in the future.

We solve for the optimal product launch decision rules backwards. Consider the product launch decision in the final period. Given the observed cost $c_3$ and past product performances, as summarized in the information set $\Omega_3$, launching product 3 is optimal if and only if

$$\Pr(q_3 = S|\Omega_3) \pi_3 \geq c_3.$$
Using this decision rule, the firm’s expected profit in period 3 given the history of past products, but prior to learning the entry cost $c_3$, is

$$E [\Pi_3 | \Omega_3] = \Pr (c_3 \leq \Pr (q_3 = S|\Omega_3) \pi_3) (\Pr (q_3 = S|\Omega_3) \pi_3 - E [c_3 | c_3 \leq \Pr (q_3 = S|\Omega_3) \pi_3])$$

$$= \frac{\Pr (q_3 = S|\Omega_3)^2 \pi_3}{2}.$$  \hspace{1cm} (3)

Now consider the firm’s decision in period 2. If the firm launches product 2, its expected discounted profit over periods 2 and 3, given product 1’s performance $q_1$ (assuming that product 1 was launched), is

$$\Pr (q_2 = S|q_1) \pi_2 - c_2$$

$$+ \delta [\Pr (q_2 = S|q_1) E [\Pi_3|q_1, S] + (1 - \Pr (q_2 = S|q_1)) E [\Pi_3|q_1, F]],$$

where $\delta > 0$ is the discount factor.

If the firm does not launch product 2, it earns no profit and receives no additional information about its skills in period 2. In this case, the firm’s total expected discounted profit from period 2 onwards is simply $\delta E [\Pi_3|q_1]$.

It follows that launching product 2 is optimal for the firm if and only if

$$\Pr (q_2 = S|q_1) \pi_2 - c_2$$

$$+ \delta [\Pr (q_2 = S|q_1) E [\Pi_3|q_1, S] + (1 - \Pr (q_2 = S|q_1)) E [\Pi_3|q_1, F] - E [\Pi_3|q_1]]$$

$$\geq 0. \hspace{1cm} (4)$$

A firm that follows a myopic product strategy launches product 2 if and only if the expected profit from that product alone, in the first line of (4), is positive. In other words, the myopic strategy always prescribes entering the market that best “fits” the firm’s skill set given current beliefs. A forward-looking firm, however, should also take into account the difference in expected third-period profits in the second line of condition (4), which captures the information value of launching product 2.
Intuitively, beyond its direct profitability, launching product 2 can be valuable for the firm because the resulting learning about about its capabilities can lead to better product introduction decisions in the future.

Formally, let us denote this information value of launching product 2 by

\[ \Delta_2 = \Pr (q_2 = 1|q_1) E [\Pi_3|q_1, S] + (1 - \Pr (q_2 = 1|q_1)) E [\Pi_3|q_1, F] - E [\Pi_3|q_1] . \]

By revealed preference, \( \Delta_2 \) is non-negative: the optimal product launch choices based on information sets \( \Omega = \{q_1, q_2 = S\} \) and \( \Omega = \{q_1, q_2 = F\} \), respectively, are weakly better than the optimal choice based on \( \Omega = \{q_1\} \) only (a formal statement appears in the proof of the next result).

To understand when and to what extent the optimal strategy deviates from the myopic strategy, we need to study how product and firm characteristics affect \( \Delta_2 \). Our first result follows from our earlier findings on belief updating:

**Proposition 1.** Suppose that the firm has launched product 1 in period \( t = 1 \). Generically\(^{16} \), launching product 2 has a positive information value (\( \Delta_2 > 0 \)) if and only if at least one of the following two conditions holds:

(i) Product 2 has (partial or full) skill overlap with product 3.

(ii) Product 1 has skill overlap with both product 2 and product 3.

Proposition 1 provides necessary and sufficient conditions for product 2’s information value to be positive. Importantly, product 2’s performance can yield useful information about the success probability of product 3 even in the absence of any skill overlap between products 2 and 3 (see condition (ii) in the proposition). In our three-period setting, this happens if products 1 and 2 have partial skill overlap, and products 1 and 3 overlap in the skill used by product 1 but not product 2. Table 2

\(^{16}\)Throughout the paper, “generically” means “for almost all parameter values.” Defining the Lebesgue measure over the feasible set of parameters, the set of parameters such that \( \Delta_2 = 0 \) if the conditions in Proposition 1 hold would have a Lebesgue measure of zero.
illustrates such a pattern of skill overlaps in a “resource-product matrix” (Wernerfelt, 1984). By Lemma 3, because of its partial skill overlap with product 1, product 2’s performance is informative about both of product 1’s skills. This in turn implies that product 2’s performance leads to learning about one of the skills that product 3 uses (skill 1 in Table 2), even if there is no skill overlap between products 2 and 3.

Table 2: Product 2 has a positive information value in the absence of any skill overlap with product 3

<table>
<thead>
<tr>
<th></th>
<th>Product 1</th>
<th>Product 2</th>
<th>Product 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skill 1</td>
<td>×</td>
<td></td>
<td>×</td>
</tr>
<tr>
<td>Skill 2</td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>Skill 3</td>
<td></td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>Skill 4</td>
<td></td>
<td></td>
<td>×</td>
</tr>
</tbody>
</table>

The optimal second-period decision rule of a forward-looking firm differs from that of a myopic firm if and only if $\Delta_2 > 0$. Proposition 1 therefore implies the next result.

The optimal second-period decision rule of a forward-looking firm differs from that of a myopic firm if and only if product 2 has (partial or full) skill overlap with product 3 or/and product 1 has skill overlap with both product 2 and product 3.

5.2 Information value and product characteristics

In this section, we will discuss how the characteristics of products 2 and 3 affect the size of the information value $\Delta_2$. Our first result states that the information value of product 2 increases with the scale of product 3, provided the information value is positive.

While this claim holds for the optimal decision rule, the outcome of a forward-looking strategy in any given period $t$ may of course be the same as under a myopic rule, depending on the characteristic of the market(s) available at $t$ and beyond.
Proposition 2. If $\Delta_2 > 0$, then the information value $\Delta_2$ of product 2 is strictly increasing in $\pi_3$.

This finding is intuitive. When the information gained from launching product 2 permits the firm to make a better decision in period 3, then the value of this information is increasing in the importance of the period 3 decision, as captured by $\pi_3$.\textsuperscript{18}

The next lemma will prove useful for examining how the degree of skill overlap between products 2 and 3 affects the size of product 2’s information value.

Lemma 5. Keeping $\Pr (q_2 = S | q_1)$ and $\Pr (q_3 = S | q_1)$ constant, the information value $\Delta_2$ of product 2 is strictly increasing in $|\Pr (q_3 = S | q_1, q_2 = S) - \Pr (q_3 = S | q_1)|$.

Lemma 5 states that the information value of product 2 is greater when its performance leads the firm to revise its belief about the success probability of product 3 more strongly. It is the absolute value of this impact that matters: If the success of product 2 means bad news for 3 (or conversely failure of 2 means good news for 3), it is just as informative for the decision whether to launch product 3 as when success of 2 means good news for 3.

In what follows, we examine how the degree of skill overlap between products 2 and 3 affect the information value of product 2. Intuitively, one might expect that, all else equal, product 2’s information value is greater the more product 2 overlaps with product 3 in the skills used, because greater overlap would seem to lead to more learning about the skills that are relevant in period 3.

To investigate whether greater overlap indeed implies a higher information value, we vary one of the skills that product 2 uses while holding the characteristics of products 1 and 3 constant. Comparing two different period 2 products is generally challenging, however, because the products can differ not only in the skills they use but also in their success probabilities. To make comparisons meaningful, we therefore rely on the following definitions:

\textsuperscript{18}Recall that $c_t \sim U [0, \pi_t]$, so an increase in $\pi_3$ implies that the expected value of $c_3$ also increases.
**Definition 2** (weak and strong comparability). Consider a product 2' that uses skills \( n \) and \( n' \) and a product 2'' that uses skills \( n \) and \( n'' \neq n' \). We will say that products 2' and 2'' are “weakly comparable” (given \( \Omega_2 \)) if

(i) the shared skill \( n \) is either the high-importance skill for both or the low-importance skill for both (\( H(2') = H(2'') = n \), or \( L(2') = L(2'') = n \)),

(ii) \( R_{2'}(\cdot, \cdot) = R_{2''}(\cdot, \cdot) \), and

(iii) \( \Pr(q_{2'} = S|\Omega_2) = \Pr(q_{2''} = S|\Omega_2) \).

The products 2' and 2'' are “strongly comparable” if instead of (iii), the following condition holds (which implies (iii)):

\[ (iii') \Pr[\theta_n = x, \theta_{n'} = y|\Omega_2] = \Pr[\theta_n = x, \theta_{n''} = y|\Omega_2] \text{ for all } \{x, y\} \in \{0, 1\} \times \{0, 1\}.\]

Both weak and strong comparability require that products 2' and 2'' use the same production function and are equally likely to succeed. Strong comparability, however, requires not only equal overall success probabilities of 2' and 2'', but requires identical marginal distributions over the respective skills used by each product, conditional on \( q_1 \) (requirement (iii')). It is not always possible to satisfy requirement (iii'), depending on how product 1 overlaps with products 2' and 2''.

**Proposition 3.** Let \( n \) and \( n' \) denote the skills used by product 3. Consider a product 2'' that uses skills \( n \) and \( n'' \neq n' \) (partial skill overlap with product 3) and a product 2' that uses skills \( n \) and \( n' \) (full skill overlap with product 3).

(a) If products 2' and 2'' are strongly comparable, then \( \Delta_{2'} > \Delta_{2''} \); that is, the product that has full skill overlap with product 3 has a greater information value than the product with only partial skill overlap.

(b) If products 2' and 2'' are weakly but not strongly comparable, then the product that has full skill overlap with product 3 can have a smaller or greater information value than the product with only partial skill overlap.
Recall the intuition that product 2 has a higher information value the more its skills overlap with those of product 3. Proposition 3 shows that the conditions for this to be true are actually quite restrictive. The relation is unambiguous only if $2'$ and $2''$ are strongly comparable (part a). Even with weak comparability, which is still a strong requirement, the product with less overlap may have a greater information value. The following example provides an illustration.

**Example A** To see how a less-overlapping product can have a higher information value, suppose that product 1 was introduced successfully ($q_1 = S$) and that the products have the following skill requirements:

<table>
<thead>
<tr>
<th></th>
<th>Product 1</th>
<th>Product $2'$</th>
<th>Product $2''$</th>
<th>Product 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skill 1</td>
<td></td>
<td></td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>Skill 2</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Skill 3</td>
<td>×</td>
<td>×</td>
<td></td>
<td>×</td>
</tr>
</tbody>
</table>

Production technologies are as follows:

- For product 1, skills 2 and 3 are substitutes: $R_1(1,1) = 0.8$, $R_1(1,0) = R_1(0,1) = 0.7$, and $R_1(0,0) = 0$. Note that this implies that Condition 1 holds, which, in turn, implies that $2''$’s performance leads to counter-intuitive belief updating about skill 3 (see Lemma 3).
- For $i \in \{2', 2''\}$, skills are complements: $R_i(1,1) = 0.8$, $R_i(1,0) = R_i(0,1) = 0.1$, $R_i(0,0) = 0$;
- For product 3, skill 3 is much more important than skill 2: $R_3(1,1) = 0.9$, $R_3(1,0) = 0.8$, $R_3(0,1) = 0.1$, $R_3(0,0) = 0$.

The prior beliefs about skills 2 and 3 are $\alpha_2 = \alpha_3 = 0.2$. Under these assumptions, $\Pr(q_2' = S|q_1) = \Pr(q_{2''} = S|q_1)$, as required by weak comparability, if $\alpha_1 = 0.3$. Intuitively, since products 1 and $2'$ use the same skills, success of product 1 is relatively good news for product $2'$. For $2''$ to have the same
probability of success, conditional on $q_1$, the prior probability of skill 1 must be larger than those of skills 2 and 3.

The probability of the success of product 3, conditional on different histories, become

- $\Pr(q_3 = S | q_1 = S) \simeq 0.61$;
- $\Pr(q_3 = S | q_1 = q_2 = S) \simeq 0.69$;
- $\Pr(q_3 = S | q_1 = q_2'' = S) \simeq 0.31$.

Success of $2'$ reveals good news about product 3, while success of $2''$ reveals bad news about product 3. More importantly, the performance of product $2''$ has larger absolute impact than that of $2'$. By Lemma 5, the information value of $2''$ therefore exceeds that of $2'$. Hence, the second-period product that has less overlap with product 3 has a greater information value in this example.

Intuitively, the success of product 1 implies favorable beliefs about skills 2 and 3. Success of product $2'$, which uses the same skills, merely provides some additional information pointing in the same direction. It increases $\Pr(q_3 = S | q_1 = S)$, but only by a small amount. In contrast, success of $2''$ is very bad news for product 3. Because skills 1 and 2 are strongly complementary for product $2''$, the firm is likely to have both skills if $2''$ succeeds. But because skills 2 and 3 were substitutes for product 1, success of $2''$ leads to a strong downward revision of the belief about skill 3. That, in turn, is bad news about product 3, whose success probability depends strongly on having skill 3. Conversely, failure of product $2''$ leads to a strong upward revision of the belief about skill 3.

Let us now turn to cases in which one of the second-products overlaps partially with product 3, while the other does not overlap with 3 at all.

**Proposition 4.** Let $n$ and $n'$ denote the skills used by product 3. Consider a product $2''$ that uses skills $n$ and $n'' \neq n'$ (partial skill overlap with product 3), and a product $\hat{2}$ that uses neither skill $n$ nor skill $n'$ (no skill overlap with product 3).
(a) If product 1 overlaps neither with product $\hat{2}$ nor with product 3, then the product that has partial skill overlap with product 3 has greater information value than product that has no skill overlap with product 3: $\Delta_{2''} > \Delta_2 = 0$.

(b) If product 1 overlaps with both product $\hat{2}$ and product 3, then the product that has partial skill overlap with product 3 (product $2''$) can have smaller or greater information value than the product that has no skill overlap with product 3 (product $\hat{2}$). This is true even if $2''$ and $\hat{2}$ are strongly comparable.

The result in part (a) follows from Proposition 1. If the product that is available in the second period has no overlap with product 3 and does not affect the beliefs about the skills used by product 3 indirectly (through overlap with product 1) either, then its information value is zero. The information value of a product that has partial overlap with product 3 is positive, on the other hand (see Proposition 1). It follows that the product with partial overlap has a greater information value than the product without any overlap.

Part (b) extends the result from Proposition 1 that, even in the absence of any direct overlap with product 3, product 2 has a positive information value if it overlaps with 1, and 1 overlaps with 3, because 2’s performance then leads the firm to update its beliefs about one of the skills used by product 3. In this case, we obtain a striking result. Even if the two second-period products under consideration are strongly comparable, the product that has zero skill overlap with product 3 can have a higher information value than the product that has partial skill overlap.

Intuitively, this surprising result can happen because the performance of a product with partial skill overlap may cause intuitive belief updating about one the skills used by product 3 but counter-intuitive belief updating about its other skill. The two belief updating effects thus partially offset each other, and product 2’s performance may have little impact on 3’s predicted success probability overall. The performance of the product that has no skill overlap with 3, on the other hand, can cause strong belief updating about one of the skills used by 3 (provided there is a pattern of partial
overlaps with product 1). The following example provides an illustration.

**Example B** Suppose that product 1 was introduced successfully \( q_1 = S \) and that the products have the following skill requirements:

<table>
<thead>
<tr>
<th></th>
<th>Product 1</th>
<th>Product 2</th>
<th>Product 2''</th>
<th>Product 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skill 1</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>Skill 2</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skill 3</td>
<td></td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>Skill 4</td>
<td></td>
<td></td>
<td></td>
<td>×</td>
</tr>
</tbody>
</table>

Production technologies are as follows:

- For product 1, skills 2 and 3 are substitutes: \( R_1(1,1) = 0.6, R_1(1,0) = R_1(0,1) = 0.5 \) and \( R_1(0,0) = 0 \);

- For \( i \in \{2, 2'', 3\} \), \( R_i(1,1) = 0.8; R_i(1,0) = R_i(0,1) = 0.4; R_i(0,0) = 0 \).

Assume further that \( \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0.25 \). It is then easy to check that products \( \hat{2} \) and \( 2'' \) are strongly (and weakly) comparable. Moreover, it can be shown that \( \Delta_{\hat{2}} - \Delta_{2''} = \frac{69}{40000} \pi_3 \). The information value of the product that does not overlap with product 3 on any of its skills is thus greater than that of the product that has partial skill overlap with product 3.

Intuitively, because skills 1 and 2 are substitutes for product 1, and neither \( \hat{2} \) nor \( 2'' \) uses skill 2, success of either \( \hat{2} \) or \( 2'' \) leads to negative belief updating about skill 2. But for \( 2'' \), success also leads to positive updating about skill 3, counterbalancing the negative effect on skill 2. By contrast, for \( \hat{2} \), negative updating about skill 2 is the only learning effect relevant for product 3. As a result, launching product \( \hat{2} \) is more informative than launching product \( 2'' \).
5.3 Discussion

Our analysis in this section has focused on the comparative statics of information value. The results on the effects of skill overlap (in Propositions 3 and 4) were obtained under the assumption that all second-period products under consideration are estimated to be equally likely to succeed given product 1’s performance. In this case, holding the potential profit $\pi_2$ fixed, the firm always wants to launch the product with the higher information value or no product at all (if the expected direct profit from product launch is negative and dominates the information value).

More generally, a firm may of course face the choice between products that differ in terms of both their expected direct profits and their information values. Our analysis implies that, in such situations, direct profits should be weighted against the value of the information that a product’s observed performance would generate. Formally, suppose the firm faces the choice between two different products, product $2'$ and product $2''$, with associated information values $\Delta_2'$ and $\Delta_2''$, in the second period. Introducing product $2'$ is then optimal, given the second-period cost realizations, if

$$\Pr (q_{2'} = S | q_1) \pi_{2'} - c_{2'} + \delta \Delta_2' > \max \{0, \Pr (q_{2''} = S | q_1) \pi_{2''} - c_{2''} + \delta \Delta_2'' \}.$$  

It is easy to see from this condition that it may well be optimal to introduce the product with the lower direct profit potential if doing so generates more useful information about the firm’s skills, as captured by the information values. The information value consideration is more likely to dominate as future decisions become more important relative to the current decision, that is, for larger $\delta$ and larger $\pi_3$ (see also Proposition 2).

As we have shown, the information gained from a new product introduction depends on the skill requirements of present, past, and potential future products in complex and sometimes surprising ways. In particular, the amount of information that a new product introduction provides about the success probability of a potential future product can be non-monotonic in the degree of skill overlap between the
products. Maybe most strikingly, launching a product that has no skill overlap whatsoever with potential future products can generate useful information, sometimes more so than a product with skill overlap.

Our insights are driven by the multi-dimensional nature of firm skills. When skills are multi-dimensional, it makes sense to expect that the firm’s beliefs about its different skills are not distributed independently of each other (i.e., that they are correlated). In our model, although prior beliefs are assumed to be independently distributed, beliefs are correlated after the first product introduction because each product relies on a combination of multiple skills. The production functions by which the skills translate into success probabilities shape the joint distribution of beliefs. In other situations, the beliefs about different skills may be correlated simply due to the nature of these skills. For example, it may be reasonable to expect that firms with chemical skills often also have strong engineering skills, but may lack design skills.

Once the beliefs about different skills are correlated, be it due to past product performance observations or for other reasons, each product’s performance can lead to belief updating for a wide range of skills beyond those used by the product itself. It becomes crucial to assess what the firm would learn from either the success or the failure of a new product about all of its skills, not just those used by the product.

In our model, additional information about skills is valuable because it allows the firm to make more informed and thus better decisions about future product introductions. For simplicity, it was assumed that the firm’s decision in the next period is simply whether or not to introduce a product whose skill requirements are known ex ante. However, the reasoning extends to situations in which the firm is uncertain about which product opportunities will be available in the future, as long as it is able to assign probabilities to different products. In this case, the optimal strategy depends on an expected information value, where the expectation is taken over different possible future product introduction opportunities.

Similarly, the reasoning in our model extends to situations in which the firm can choose between multiple product introductions in future periods. In this case,
the information value again reflects how the observed performance of the current product allows the firm’s to make more profitable decisions in the future. However, the information gained through a product introduction may now affect the choice between different potential future products.

Suppose, for instance, that the firm faces the choices between two different products, product 3′ and product 3′′, in the third period. Given the history Ω_3 and the cost realizations observed at the beginning of the third period, launching product 3′ is optimal if

\[ \Pr(q_{3'}) = S | \Omega_3 \pi_{3'} - c_{3'} > \max \{ 0, \Pr(q_{3''} = S | \Omega_3 \pi_{3''} - c_{3''} \} , \]

and similarly for product 3′′. Using this optimal decision rule, we can compute, for every history Ω_3, the probability with which each possible period 3 decision (introduction of product 3′, introduction of product 3′′, or no product launch) is optimal and the expected profit of each decision conditional on it being optimal. As in our baseline model, launching product 2 adds another performance observation to the history, which allows the firm to learn about (some) of its skills. This learning allows to firm to improve the quality of its decision in the third period if either product 3′ or product 3′′ use a skill, or several skills, for which product 2’s performance leads to belief updating. Importantly, product 2 has a positive information value even if its performance leads the firm to revise its beliefs about the success probability of only one of products available in the third period.

6 Conclusion

Firms often face considerable uncertainty about their resources and capabilities, which makes it difficult to develop a successful product strategy. Our paper develops a model of product strategy in which a firm learns about its own capabilities from observed successes and failures in different markets. We show that, due to the multi-
dimensional nature of capabilities, such learning can exhibit complex and sometimes surprising patterns.

Our paper has several managerial implications. First, when assessing their firms’ capabilities, managers should try to also assess the level of uncertainty about those capabilities. Which are the skills about which the firm is highly confident, for instance, because these skills have been proven repeatedly by multiple product successes? On the other hand, which are the skills that the firm might have, but cannot be sure of based on current evidence? When deciding whether to launch a new product, managers should then consider both the expected direct profit from that product, based on current skill beliefs, and the expected value of the information that the new product will generate about the firm’s skills. This information value depends, among other things, on the product’s overlap with past and potential future products in terms of skill requirements.

Future research could extend the multi-skill, multi-product model used in this paper to study other interesting phenomena. For example, one could model customer beliefs about the firm’s private skill information in order to study the firm’s brand extension decisions (e.g., Wernerfelt 1988; Miklós-Thal 2012; Moorthy 2012). One could also incorporate competition and investment in skill formation into the model (e.g., Fudenberg and Tirole 1984; Ericson and Pakes 1995; Selove 2014a, 2014b). Finally, one could allow certain skills to become more important over time or allow skills to become obsolete, as new technologies arrive and customer preferences change (e.g., Christensen 1997; Chandy and Tellis 1998; Adner and Zemsky 2005).
References


http://www.economist.com/node/21542796


Appendix A: A general result on belief-updating

We show that the direction of belief-updating for skill $n$ following the success or failure of product $t$ depends on whether the following condition holds:

**Condition A.**

$$\frac{\Pr(q_t = S, \Omega_t | \theta_n = 1)}{\Pr(q_t = S, \Omega_t | \theta_n = 0)} > \frac{\Pr(\Omega_t | \theta_n = 1)}{\Pr(\Omega_t | \theta_n = 0)}$$

where $\Pr(\Omega_t)$ is the probability of the observed set of successes and failures for a given set of product launches prior to period $t$.\(^{19}\) This condition states that the ratio of the probability of the observed sequence of successes and failures if the firm has skill $n$ relative to the same probability if the firm does not have skill $n$ becomes greater following a successful launch of product $t$. Intuitively, if this condition holds, then the success of product $t$ implies we now have “more favorable” information about skill $n$ than we previously did, given the firm’s history of product successes and failures prior to time $t$.

Appendix B proves the following lemma, by applying Bayes’ rule and rearranging terms.

**Lemma A.** If the firm launches product $t$, beliefs about skill $n$ are updated as follows:

1. If Condition A holds, intuitive belief-updating occurs.
2. If Condition A is reversed, counter-intuitive belief-updating occurs.
3. If the two sides of the inequality in Condition A are equal, no belief-updating occurs.

The results in subsection 4.1 make use of Lemma A to derive conditions under which intuitive or counter-intuitive belief-updating occurs.

\(^{19}\)If no products have been launched, we define $\Pr(\emptyset) = 1$.\)
Appendix B: Proofs

Proof of Lemma A  Bayes’ rule states that the posterior probability of having skill $n$ given history $\Omega_t$ is:\footnote{Note that if $\Omega_t = \emptyset$, we can define $\Pr(\emptyset | \theta_n = 1) = \Pr(\emptyset | \theta_n = 0) = 1$, and the right side of (5) then equals the prior value of having skill $n$, which is $\alpha_n$.}

$$E[\theta_n | \Omega_t] = \frac{\alpha_n \Pr(\Omega_t | \theta_n = 1)}{\alpha_n \Pr(\Omega_t | \theta_n = 1) + (1 - \alpha_n) \Pr(\Omega_t | \theta_n = 0)}$$

(5)

Similarly, the posterior probability of having skill $n$ given history $\Omega_t$ and given that product $t$ succeeds is:

$$E[\theta_n | q_t = S, \Omega_t] = \frac{\alpha_n \Pr(q_t = S, \Omega_t | \theta_n = 1)}{\alpha_n \Pr(q_t = S, \Omega_t | \theta_n = 1) + (1 - \alpha_n) \Pr(q_t = S, \Omega_t | \theta_n = 0)}$$

(6)

Thus, following history $\Omega_t$, product $t$’s success increases the posterior probability of having skill $n$ if (6) is greater than (5). Cross-multiplying and canceling like terms shows that this inequality is equivalent to:

$$\Pr(q_t = S, \Omega_t | \theta_n = 1) \Pr(\Omega_t | \theta_n = 0) > \Pr(\Omega_t | \theta_n = 1) \Pr(q_t = S, \Omega_t | \theta_n = 0)$$

(7)

Note this condition is equivalent to Condition A. Therefore, following a product success, intuitive belief-updating, counter-intuitive belief updating, and no belief-updating are, respectively, equivalent to Condition A holding, Condition A being reversed, and the two sides of Condition A being equal.

The proof for the case of belief-updating after product failures is similar.

QED

Proof of Lemma 1  Suppose that product $t$ uses only skills that were not used by any previously introduced products. This implies that $q_t = S$ and $\Omega_t$ are independent events. Therefore, for any skill $n$,

$$\Pr(q_t = S, \Omega_t | \theta_n) = \Pr(q_t = S | \theta_n) \Pr(\Omega_t | \theta_n).$$

(8)

Condition A is thus equivalent to:

$$\frac{\Pr(q_t = S | \theta_n = 1) \Pr(\Omega_t | \theta_n = 1)}{\Pr(q_t = S | \theta_n = 0) \Pr(\Omega_t | \theta_n = 0)} > \frac{\Pr(\Omega_t | \theta_n = 1)}{\Pr(\Omega_t | \theta_n = 0)}.$$

(9)
If \( n \in S(t) \), which implies that \( n \) has not been used by any product \( \tau < t \), (9) simplifies to

\[
\frac{\Pr(q_t = S|\theta_n = 1)}{\Pr(q_t = S|\theta_n = 0)} > 1.
\]

(10)

Assumption 1 implies that this condition holds for any \( n \in S(t) \). Lemma A then implies the results.

QED

Proof of Lemma 2  Recall that we denote \( I(1) = \{n, n'\} \) and \( I(2) = \{n', n''\} \). For the skill used only by product 2 (skill \( n'' \)) in period 2, Condition A is:

\[
\frac{\Pr(q_2 = S, \Omega_2|\theta_{n''} = 1)}{\Pr(q_2 = S, \Omega_2|\theta_{n''} = 0)} > \frac{\Pr(\Omega_2|\theta_{n''} = 1)}{\Pr(\Omega_2|\theta_{n''} = 0)} \tag{11}
\]

We now show that this condition holds. Because skill \( n'' \) does not affect product 1’s success probability, we have:

\[\Pr(\Omega_2|\theta_{n''} = 1) = \Pr(\Omega_2|\theta_{n''} = 0).\]

(12)

Thus, the right-hand-side of (11) equals one. We now need to show that the left-hand-side of (11) is greater than one, so that condition (11) holds.

Note first that, if we condition on all three relevant skills, the following holds:

\[\Pr(q_2 = S, \Omega_2|\theta_n, \theta_{n'}, \theta_{n''}) = \Pr(q_2 = S|\theta_{n'}, \theta_{n''}) \Pr(\Omega_2|\theta_n, \theta_{n'})\]

(13)

Assumption 1 implies that, for any given values of \( \theta_n \) and \( \theta_{n'} \), the right side of (13) is greater for \( \theta_{n''} = 1 \) than for \( \theta_{n''} = 0 \), and therefore:

\[\Pr(q_2 = S, \Omega_2|\theta_{n''} = 1, \theta_n, \theta_{n'}) > \Pr(q_2 = S, \Omega_2|\theta_{n''} = 0, \theta_n, \theta_{n'}).\]

(14)

Because this inequality holds for any particular values of \( \theta_n \) and \( \theta_{n'} \), it also holds if we integrate over all possible values of these skills, and therefore:

\[\Pr(q_2 = S, \Omega_2|\theta_{n''} = 1) > \Pr(q_2 = S, \Omega_2|\theta_{n''} = 0).\]

(15)

Hence, the left-hand-side of (11) is greater than one. We have shown that Condition A holds for skill \( n'' \) in period two, and so Lemma A guarantees that intuitive updating
occurs for this skill.

We now need to show that intuitive updating also occurs in period two for the skill used by both products (skill $n'$). In period two, the right side of Condition A for this skill is:

$$\frac{\Pr(\Omega_2|\theta_{n'} = 1)}{\Pr(\Omega_2|\theta_{n'} = 0)} = \frac{\alpha_n \Pr(\Omega_2|\theta_{n'} = 1, \theta_n = 1)}{\alpha_n \Pr(\Omega_2|\theta_{n'} = 0, \theta_n = 1)} + (1 - \alpha_n) \frac{\Pr(\Omega_2|\theta_{n'} = 1, \theta_n = 0)}{\Pr(\Omega_2|\theta_{n'} = 0, \theta_n = 0)}$$

(16)

while the left side of Condition A for this skill is:

$$\frac{\Pr(q_2 = S, \Omega_2|\theta_{n'} = 1)}{\Pr(q_2 = S, \Omega_2|\theta_{n'} = 0)} = \frac{\alpha_n \Pr(q_2 = S|\theta_{n'} = 1, \theta_n = 1) + (1 - \alpha_n) \Pr(q_2 = S|\theta_{n'} = 1, \theta_n = 0)}{\alpha_n \Pr(q_2 = S|\theta_{n'} = 0, \theta_n = 1) + (1 - \alpha_n) \Pr(q_2 = S|\theta_{n'} = 0, \theta_n = 0)}$$

(17)

where we define:

$$\Gamma_1 \equiv \alpha_n \Pr(q_2 = S|\theta_{n'} = 1, \theta_n = 1) + (1 - \alpha_n) \Pr(q_2 = S|\theta_{n'} = 1, \theta_n = 0)$$

(18)

and

$$\Gamma_0 \equiv \alpha_n \Pr(q_2 = S|\theta_{n'} = 0, \theta_n = 1) + (1 - \alpha_n) \Pr(q_2 = S|\theta_{n'} = 0, \theta_n = 0)$$

(19)

Assumption 1 implies that $\Gamma_1 > \Gamma_0$. Therefore, for skill $n'$ in period two, the left side of Condition A is greater than the right side. We have shown that Condition A holds for skill $n'$ in period two, and so Lemma A guarantees that intuitive updating occurs for this skill.

QED

Proof of Lemma 3  Recall that we denote $I(1) = \{n, n'\}$ and $I(2) = \{n', n''\}$. In period two, for the skill used only by product 1 (skill $n$), the right side of Condition A is:

$$\frac{\Pr(\Omega_2|\theta_n = 1)}{\Pr(\Omega_2|\theta_n = 0)} = \frac{\alpha_{n'} \Pr(\Omega_2|\theta_{n'} = 1, \theta_n = 1) + (1 - \alpha_{n'}) \Pr(\Omega_2|\theta_{n'} = 0, \theta_n = 1)}{\alpha_{n'} \Pr(\Omega_2|\theta_{n'} = 1, \theta_n = 0) + (1 - \alpha_{n'}) \Pr(\Omega_2|\theta_{n'} = 0, \theta_n = 0)}$$

(20)

while the left side of Condition A is:

$$\frac{\Pr(q_2 = S, \Omega_2|\theta_n = 1)}{\Pr(q_2 = S, \Omega_2|\theta_n = 0)} = \frac{\alpha_n \Gamma_1 \Pr(q_2 = S|\theta_{n'} = 1, \theta_n = 1) + (1 - \alpha_n') \Gamma_0 \Pr(q_2 = S|\theta_{n'} = 0, \theta_n = 1)}{\alpha_n \Gamma_1 \Pr(q_2 = S|\theta_{n'} = 1, \theta_n = 0) + (1 - \alpha_n') \Gamma_0 \Pr(q_2 = S|\theta_{n'} = 0, \theta_n = 0)}$$

(21)

where $\Gamma_1$ and $\Gamma_0$ are defined as in (18) and (19).

For Condition A to hold, we need that (21) is greater than (20), which is equivalent to saying the numerator of (21) times the denominator of (20) is greater than the
numerator of (20) times the denominator of (21). Performing these multiplication operations and canceling like terms, we are left with the following condition:

\[
\left( \Gamma_1 - \Gamma_0 \right) \Pr(\Omega_2|\theta_{n'}) = 1, \theta_n = 1) \Pr(\Omega_2|\theta_{n'} = 0, \theta_n = 0) > \left( \Gamma_1 - \Gamma_0 \right) \Pr(\Omega_2|\theta_{n'} = 1, \theta_n = 0) \Pr(\Omega_2|\theta_{n'} = 0, \theta_n = 1)
\]

Because \( \Gamma_1 > \Gamma_0 > 0 \), this condition is equivalent to:

\[
\Pr(\Omega_2|\theta_{n'} = 1, \theta_n = 1) \Pr(\Omega_2|\theta_{n'} = 0, \theta_n = 0) > \Pr(\Omega_2|\theta_{n'} = 1, \theta_n = 0) \Pr(\Omega_2|\theta_{n'} = 0, \theta_n = 1)
\]

which, in turn, is equivalent to Condition 1.

We have shown that, in period 2, for the skill that is used by product 1 but not by product 2 (skill \( n \)), Condition A is equivalent to Condition 1. Therefore, Lemma A implies that intuitive updating occurs for this skill if Condition 1 holds, and counter-intuitive updating occurs for this skill if Condition 1 is reversed.

QED

**Proof of Lemma 4** To simplify notation, we focus on the case in which product 1 succeeds, and products 1 and 2 have the same low importance skill and same high importance skill:

\[
\Omega_2 = \{ q_1 = S \}
\]
\[
n = L(1) = L(2)
\]
\[
\hat{n} = H(1) = H(2)
\]

The other cases (product 1 fails, and/or product 1’s low importance skill is product 2’s high importance skill) can be proven with the same approach.

**Part (a):** In period 2 for skill \( \hat{n} \), the right side of Condition A is:

\[
\frac{\Pr(q_1 = S|\theta_{\hat{n}} = 1)}{\Pr(q_1 = S|\theta_{\hat{n}} = 0)} = \frac{\alpha_n R_1(1, 1) + (1 - \alpha_n) R_1(1, 0)}{\alpha_n R_1(0, 1) + (1 - \alpha_n) R_1(0, 0)}
\]

The left side of Condition A is:

\[
\frac{\Pr(q_2 = S, q_1 = S|\theta_{\hat{n}} = 1)}{\Pr(q_2 = S, q_1 = S|\theta_{\hat{n}} = 0)} = \frac{\alpha_n R_1(1, 1) R_2(1, 1) + (1 - \alpha_n) R_1(1, 0) R_2(1, 0)}{\alpha_n R_1(0, 1) R_2(0, 1) + (1 - \alpha_n) R_1(0, 0) R_2(0, 0)}
\]
For Condition A to hold, (25) must be greater than (24), which is equivalent to saying the numerator of (25) times the denominator of (24) must be greater than the numerator of (24) times the denominator of (25).

Multiplying the numerator of (25) by the denominator of (24) gives:

\[
\alpha^2 R_1(1,1) R_1(0,1) R_2(1,1) + \alpha_n(1 - \alpha_n) R_1(1,1) R_1(0,0) R_2(1,1) + \\
\alpha_n(1 - \alpha_n) R_1(1,0) R_1(0,1) R_2(1,0) + (1 - \alpha_n)^2 R_1(1,0) R_1(0,0) R_2(1,0)
\]  

(26)

Multiplying the numerator of (24) by the denominator of (25) gives:

\[
\alpha^2 R_1(1,1) R_1(0,1) R_2(0,1) + \alpha_n(1 - \alpha_n) R_1(1,1) R_1(0,0) R_2(0,0) + \\
\alpha_n(1 - \alpha_n) R_1(1,0) R_1(0,1) R_2(0,1) + (1 - \alpha_n)^2 R_1(1,0) R_1(0,0) R_2(0,0)
\]  

(27)

For Condition A to hold, (26) must be greater than (27). Because Assumption 1 implies \( R_2(1,1) > R_2(0,1), R_2(1,1) > R_2(0,0), \) and \( R_2(1,0) > R_2(0,0) \), the first, second, and fourth terms of (26) are strictly greater than the respective corresponding terms of (27). Because Assumption 1 also implies \( R_2(1,0) \geq R_2(0,1) \), the third term of (26) is weakly greater than the third term of (27). Therefore, (26) is strictly greater than (27).

We have shown that, in period 2, for product 2’s high importance skill, Condition A holds. Therefore, Lemma A implies that intuitive updating occurs for this skill.

**Part (b):** In period 2 for skill \( n \), the right side of Condition A is:

\[
\frac{\Pr(q_1 = S| \theta_n = 1)}{\Pr(q_1 = S| \theta_n = 0)} = \frac{\alpha_n R_1(1,1) + (1 - \alpha_n) R_1(0,1)}{\alpha_n R_1(1,0) + (1 - \alpha_n) R_1(0,0)}
\]  

(28)

The left side of Condition A is:

\[
\frac{\Pr(q_2 = S, q_1 = S| \theta_n = 1)}{\Pr(q_2 = S, q_1 = S| \theta_n = 0)} = \frac{\alpha_n R_1(1,1) R_2(1,1) + (1 - \alpha_n) R_1(0,1) R_2(0,1)}{\alpha_n R_1(1,0) R_2(1,0) + (1 - \alpha_n) R_1(0,0) R_2(0,0)}
\]  

(29)

For Condition A to hold, we need that (29) is greater than (28). Multiplying the numerator of (29) by the denominator of (28) gives:

\[
\alpha^2 R_1(1,1) R_1(1,0) R_2(1,1) + \alpha_n(1 - \alpha_n) R_1(1,1) R_1(0,0) R_2(1,1) + \\
\alpha_n(1 - \alpha_n) R_1(1,0) R_1(0,1) R_2(0,1) + (1 - \alpha_n)^2 R_1(0,1) R_1(0,0) R_2(0,1)
\]  

(30)
Multiplying the numerator of (28) by the denominator of (29) gives:

\[ \alpha_{\hat{n}}^2 R_1(1,1)R_1(1,0)R_2(1,0) + \alpha_{\hat{n}}(1 - \alpha_{\hat{n}})R_1(1,1)R_1(0,0)R_2(0,0) + \alpha_{\hat{n}}(1 - \alpha_{\hat{n}})R_1(0,1)R_1(1,0)R_2(1,0) + (1 - \alpha_{\hat{n}})^2 R_1(0,1)R_1(0,0)R_2(0,0) \]  

(31)

For Condition A to hold, (30) must be greater than (31). Because Assumption 1 implies \( R_2(1,1) > R_2(1,0) \) and \( R_2(0,1) > R_2(0,0) \), the first and fourth terms of (30) are greater than the respective corresponding terms of (31). Therefore, a sufficient condition for (30) to be greater than (31) is that the second term plus the third term of (30) exceeds the second term plus the third term of (31), which is equivalent to the following:

\[ R_1(1,1)R_1(0,0)\left[ R_2(1,1) - R_2(0,0) \right] > R_1(0,1)R_1(1,0)\left[ R_2(1,0) - R_2(0,1) \right] \]  

(32)

Because Assumption 1 implies \( [R_2(1,1) - R_2(0,0)] > [R_2(1,0) - R_2(0,1)] > 0 \), a sufficient condition for (32) to hold is:

\[ R_1(1,1)R_1(0,0) > R_1(0,1)R_1(1,0) \]  

(33)

which is equivalent to Condition 1. We have shown that, if Condition 1 holds, then Condition A must also hold (and Lemma A implies intuitive updating must occur) for product 2’s low importance skill in period 2.

On the other hand, suppose Condition 1 is reversed. If we hold all other parameters constant and let \( R_2(0,1) \to R_2(0,0) \) and \( R_2(1,1) \to R_2(1,0) \), then in the limit, (30) will be less than (31) if the following holds:

\[ R_1(1,1)R_1(0,0)\left[ R_2(1,0) - R_2(0,0) \right] < R_1(0,1)R_1(1,0)\left[ R_2(1,0) - R_2(0,0) \right] \]  

(34)

which is equivalent to Condition 1 being reversed. We have shown that, if Condition 1 is reversed and we let \( R_2(0,1) \to R_2(0,0) \) and \( R_2(1,1) \to R_2(1,0) \), then Condition A must also be reversed (and Lemma A implies counter-intuitive updating must occur) for product 2’s low importance skill in period 2.

QED
Proof of Proposition 1  By the law of iterated expectations,

$$\Pr(q_3 = S|q_1) = \Pr(q_2 = S|q_1) Pr(q_3 = S|q_1) + (1 - \Pr(q_2 = S|q_1)) \Pr(q_3 = S|q_1, F).$$  

(35)

Therefore, using (3) and (35),

$$\Delta_2 = \frac{2}{\pi_3} \Pr(q_2 = 1|q_1) E[\Pi_3|q_1, S] + (1 - \Pr(q_2 = 1|q_1)) E[\Pi_3|q_1, F] - E[\Pi_3|q_1]$$

$$= \Pr(q_2 = S|q_1) Pr(q_3 = S|q_1, S)^2 + (1 - \Pr(q_2 = S|q_1)) Pr(q_3 = S|q_1, F)^2 - Pr(q_3 = S|q_1)^2$$

$$= \Pr(q_2 = S|q_1) Pr(q_3 = S|q_1, S)^2 + (1 - \Pr(q_2 = S|q_1)) Pr(q_3 = S|q_1, F)^2$$

$$- [\Pr(q_2 = S|q_1) Pr(q_3 = S|q_1, S) + (1 - \Pr(q_2 = S|q_1)) Pr(q_3 = S|q_1, F)]^2,$$

which by Jensen’s Inequality is nonnegative because the function $f(x) = x^2$ is strictly convex in $x$, and is strictly positive if and only if $\Pr(q_3 = S|q_1, S) \neq \Pr(q_3 = S|q_1, F)$. By Lemmas 2 to 4, generically $E[\theta_n | q_1, q_2] \neq E[\theta_n | q_1]$ for at least one of the skills used by product 3 if one or both of the conditions in Proposition 1 holds, otherwise $E[\theta_n | q_1, q_2] = E[\theta_n | q_1]$ for both of the skills that product 3 uses. Hence, generically $\Pr(q_3 = S|q_1, S) \neq \Pr(q_3 = S|q_1, F)$ if one or both conditions hold, which, as shown above, implies that $\Delta_2 > 0$.

QED

Proof of Proposition 2  From (3), $E[\Pi_3|\Omega]$ is proportional to $\pi_3/2$ for any $\Omega$. $\Delta_2$ can therefore be written as

$$\Delta_2 = \left[\Pr(q_2 = S|q_1) Pr(q_3 = S|q_1, S)^2 + (1 - \Pr(q_2 = S|q_1)) Pr(q_3 = S|q_1, F)^2 - Pr(q_3 = S|q_1)^2\right] \frac{\pi_3}{2}.$$

Since $\pi_3 > 0$, $\Delta_2 > 0$ if and only if the sum between brackets is positive. It follows that $\Delta_2$ is a strictly increasing function of $\pi_3$ if $\Delta_2 > 0$.

QED

---

21If there is partial skill overlap between all three product pairs with each pair overlapping on a different skill, or if all three products use the same two skills, then beliefs about the two skills that product 3 uses may be updated in opposite directions. In principle, the positive updating on one skill could exactly offset the negative updating on the other skill in the overall success probability of product 3 in such cases; however, this outcome is “degenerate” in the sense that for almost all parameter values, product 2’s performance will affect the success probability of product 3 when it leads to updating about both of the skills that product 3 uses.
Proof of Lemma 5  The information value $\Delta_2$ can be expressed as

$$\Delta_2 = (\Pr(q_2 = S|q_1) \Pr(q_3 = S|q_1, S)^2 + (1 - \Pr(q_2 = S|q_1)) \Pr(q_3 = S|q_1, F)^2 - \Pr(q_3 = S|q_1)^2) \frac{\pi_3}{2}. \quad (36)$$

Rearrange (35),

$$\Pr(q_3 = S|q_1, F) = \frac{\Pr(q_3 = S|q_1)}{1 - \Pr(q_2 = S|q_1)} - \frac{\Pr(q_2 = S|q_1)}{1 - \Pr(q_2 = S|q_1)} \Pr(q_3 = S|q_1, S) \quad (37)$$

and substitute for $\Pr(q_3 = S|q_1, F)$ in (36), which leads to

$$\Delta_2 = (\Pr(q_2 = S|q_1) \Pr(q_3 = S|q_1, S)^2 + \frac{1}{1 - \Pr(q_2 = S|q_1)} [\Pr(q_3 = S|q_1) - \Pr(q_2 = S|q_1) \Pr(q_3 = S|q_1, F)]) \Pr(q_3 = S|q_1, S) \quad (37)$$

Taking the partial derivative with respect to $\Pr(q_3 = S|q_1, S)$ yields

$$\frac{\partial \Delta_2}{\partial \Pr(q_3 = S|q_1, S)} > 0 \text{ if } \Pr(q_3 = S|q_1, S) > \Pr(q_3 = S|q_1),$$

$$\frac{\partial \Delta_2}{\partial \Pr(q_3 = S|q_1, S)} < 0 \text{ if } \Pr(q_3 = S|q_1, S) < \Pr(q_3 = S|q_1).$$

Hence, if $q_2 = S$ leads to positive updating of 3’s success probability, i.e., if $\Pr(q_3 = S|q_1, S) > \Pr(q_3 = S|q_1) > \Pr(q_3 = S|q_1, F)$, then the information value $\Delta_2$ is increasing in $\Pr(q_3 = S|q_1, S)$. Conversely, if $q_2 = S$ leads to negative updating of 3’s success probability, i.e., if $\Pr(q_3 = S|q_1, S) < \Pr(q_3 = S|q_1) < \Pr(q_3 = S|q_1, F)$, then the information value $\Delta_2$ is decreasing in $\Pr(q_3 = S|q_1, S)$. The statement in the lemma follows directly from these observations. QED

Proof of Proposition 3  Assume that product 3 uses skills $n$ and $n'$, product 2' uses the same skills $n$ and $n'$, and product 2'' uses skills $n$ and $n''$.

(a) We first show that $\Pr(q_3 = S|q_1, q_{2'} = S) > \Pr(q_3 = S|q_1, q_{2''} = S)$, which using Bayes' rule can be rewritten as

$$\frac{\Pr(q_3 = S, q_{2'} = S|q_1)}{\Pr(q_{2'} = S|q_1)} > \frac{\Pr(q_3 = S, q_{2''} = S|q_1)}{\Pr(q_{2''} = S|q_1)} \quad \iff \quad \Pr(q_3 = S, q_{2'} = S|q_1) > \Pr(q_3 = S, q_{2''} = S|q_1).$$

where the last step follows from strong comparability of 2' and 2''.
Without loss of generality, assume that \( n \) is the high-importance skill of products \( 2' \) and \( 2'' \). Let \( \sigma(x, y, z) \equiv P[\theta_n = x, \theta_{n'} = y, \theta_{n''} = z|q_1] \) for all eight possible permutations of \( x, y, \) and \( z \). Using this notation, the difference \( \Pr(q_3 = S, q_{2'} = S|q) - \Pr(q_3 = S, q_{2''} = S|q_1) \) can be decomposed into eight terms, which after canceling zero-terms leads to

\[
\sigma(1, 1, 0)[R_2(1, 1) - R_2(1, 0)]R_3(1, 1) + \sigma(1, 0, 1)[R_2(1, 0) - R_2(1, 1)]R_3(\theta_n = 1, \theta_{n'} = 0)
\]

\[
+ \sigma(0, 1, 0)[R_2(0, 1) - R_2(0, 0)]R_3(\theta_n = 0, \theta_{n'} = 1) + \sigma(0, 0, 1)[R_2(0, 0) - R_2(0, 1)]R_3(0, 0),
\]

where \( R_3(\theta_n = 1, \theta_{n'} = 0) \) is equal to \( R_3(1, 0) \) if \( n = H(3) \) and equal to \( R_3(0, 1) \) otherwise.

Requirement (iii') in the definition of strong comparability implies that

\[
\sigma(x, 0, 1) = \sigma(x, 1, 0) \text{ for any } x \in \{0, 1\}.
\]

(38) can thus be rewritten as

\[
\sigma(1, 1, 0)[R_2(1, 1) - R_2(1, 0)] [R_3(1, 1) - R_3(\theta_n = 1, \theta_{n'} = 0)]
\]

\[
+ \sigma(0, 1, 0)[R_2(0, 1) - R_2(0, 0)] [R_3(\theta_n = 0, \theta_{n'} = 1) - R_3(0, 0)].
\]

Because \( R_2 \) and \( R_3 \) are strictly monotonic in both arguments, both terms in (40) are strictly positive. Thus, \( \Pr(q_3 = S|q_1, q_{2'} = S) > \Pr(q_3 = S|q_1, q_{2''} = S) \). By Lemma 5, this implies that \( \Delta_{2'} > \Delta_{2''} \) if \( \Pr(q_3 = S|q_1, q_{2''} = S) \geq \Pr(q_3 = S|q_1) \).

For the remainder of the proof, we therefore consider cases in which \( \Pr(q_3 = S|q_1, q_{2''} = S) < \Pr(q_3 = S|q_1) \). Consider the expression

\[
\Pr(q_3 = S|q_1, q_{2'} = S) + \Pr(q_3 = S|q_1, q_{2''} = S) - 2\Pr(q_3 = S|q_1).
\]

Decomposing each term in the same manner as above into eight terms, using the
function $\sigma(x, y, z)$, and simplifying using (39), (41) can be rewritten as

$$2\sigma(1, 0, 0)[R_2(1, 1) - R_2(1, 0)] [R_3(1, 1) - R_3(\theta_n = 1, \theta_{n'} = 0)]$$
$$+ \sigma(1, 1, 0)[R_2(1, 1) - R_2(1, 0)] [R_3(1, 1) - R_3(\theta_n = 1, \theta_{n'} = 0)]$$
$$+ \sigma(0, 1, 0)[2R_2(1, 1) - R_2(0, 1) - R_2(0, 0)] [2R_3(1, 1) - R_3(\theta_n = 0, \theta_{n'} = 1) - R_3(0, 0)]$$
$$+ 2\sigma(0, 1, 1)[R_2(1, 1) - R_2(0, 1)] [R_3(1, 1) - R_3(\theta_n = 0, \theta_{n'} = 1)]$$
$$+ 2\sigma(0, 0, 0)[R_2(1, 1) - R_2(0, 0)] [R_3(1, 1) - R_3(0, 0)]$$

which is strictly positive because all individual terms are strictly positive. Thus, (41) is positive, which means that

$$\Pr (q_3 = S|q_1, q_2' = S) - \Pr (q_3 = S|q_1) > \Pr (q_3 = S|q_1, q_2'' = S) - \Pr (q_3 = S|q_1) > 0.$$  

By Lemma 5, it follows that $\Delta_2' > \Delta_2''$.

(b) The example in the main text proves that $\Delta_2'$ can be smaller than $\Delta_2''$ when products 2' and 2'' are weakly but not strongly comparable. The following example shows that the opposite outcome is possible as well. Suppose that products 1, 2', 2'' and 3 use skills as depicted in the following table.

<table>
<thead>
<tr>
<th>Skill</th>
<th>Product 1</th>
<th>Product 2'</th>
<th>Product 2''</th>
<th>Product 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>3</td>
<td>×</td>
<td>×</td>
<td></td>
<td>×</td>
</tr>
</tbody>
</table>

All four products use the following additive and symmetric technology: $R_i(1, 1) = 0.8; R_i(1, 0) = R_i(0, 1) = 0.4; R_i(0, 0) = 0$. Moreover, $\alpha_1 = \alpha_2 = 1/2$. In this case, $E(\theta_1|q_1 = S) = E(\theta_2|q_1 = S) = 3/4$, which implies that products 2' and 2'' are weakly comparable if and only if $\alpha_3 = 3/4$. However, strong comparability is not satisfied. For instance, the probability that the firm has both of product 2''s skills is $\Pr [\theta_1 = \theta_2 = 1|q_1 = S] = 1/2$, whereas the probability that the firm has both of product 2''s skills is $\Pr [\theta_2 = \theta_3 = 1|q_1 = S] = 9/16$. Consistent with intuition, following a success of product 1, product 2' has a higher information value than product 2'': $\Delta_2' = \frac{9}{1200} \pi_3 > \frac{1}{1200} \pi_3 = \Delta_2''$.  

QED
Proof of Proposition 4  Part (a) follows from Proposition 1. For (b), $\Delta_0 > 0$ again follows from Proposition 1. The fact that $\Delta_0$ can exceed $\Delta_2$ is straightforward. The statement that $\Delta_2$ can exceed $\Delta_0$ is shown by example in the text.

QED