A Demand-Based View of Sustainable Competitive Advantage:

The evolution of substitution threats, resource rents

and competitive positions

Ron Adner and Peter Zemsky*

October 8, 2003

Abstract

We explore the sustainability of competitive advantage using an approach that is grounded in an explicit treatment of consumer choice. We address the evolution of substitution threats, resource rents, and competitive positions using a formal model of competition with differentiated products in which production technologies improve over time. We show how the interplay between improving technologies and consumers’ valuation of the resulting performance improvements affects which market segments new technologies are able to enter, how the rents from different types of resources change over time, and whether or not classic generic strategies remain viable. Our focus on consumer choice and value creation complements the traditional focus in the strategy literature on competition and value capture.

*Both authors at Insead. We thank seminar participants at Insead, New York University, Ohio State University, Strategy Research Forum, University of Michigan, University of Toronto and Wharton, as well as Pankaj Ghemawat, Javier Gimeno, Constance Helfat, Bruce Kogut, Bill Lazonic, Michael Leiblein, Dan Levinthal, Anita McGahan, Mary O’Sullivan, Subi Rangan and Louis Thomas for helpful comments.
1. Introduction

The drivers of sustainable competitive advantage are a focal point of debate in the strategy literature. By now, there is broad recognition that the concerns of the competitive strategy school (e.g., Porter, 1980; Ghemawat, 1991) regarding industry structure and a firm's choice of position within an industry, and the concerns of the resource-based school (e.g., Wernerfelt, 1984; Barney, 1991) regarding the value, uniqueness, inimitability and non-substitutability of resources, are most productively viewed as complementary approaches to understanding competitive outcomes (e.g., Henderson and Mitchell, 1997). Common to both of these approaches is a principal focus on firms’ supply-side interactions. The thrust of their concerns, whether couched in the language of mobility barriers (Porter, 1980), imperfect factor markets (Barney, 1986), or isolating mechanism (Rumelt, 1984) is with the safeguarding of competitive advantage through the exclusion of rivals from the pursuit of opportunities.

While attention has been focused on imitation and the intensity of competition among firms, the demand-side concerns of value and substitutability have received relatively less consideration. These demand-side factors, however, play an important and distinct role in determining competitive outcomes, and suggest a complementary logic for understanding, and responding to, the erosion of competitive advantage. In this paper we use a demand-based lens to examine the dynamics of sustainability in the face of improving technologies. We offer a unified approach to understanding three distinct questions: (i) How does the threat of substitute technologies change over time? (ii) How do resource rents change over time? (iii) How does optimal positioning within an industry change over time? At the center of our approach is an explicit treatment of consumer choice, which leads us to focus on a firm’s value creation – the difference between consumer’s willingness to pay for an offer and the firm’s production cost (Brandenburger and Stewart, 1996) – and how value creation changes over time.2

1 Clearly, the question of value has long been present in the literature, but it has tended to be treated as a background concern. As the Priem and Butler (2001)-Barney(2001) debates highlight, in the context of the resource-based view, value has largely been considered in terms of firms’ ability to capture value when resources are purchased in imperfect factor markets. In contrast, the assessment of how resources create value from a consumer perspective has received less attention. While the competitive strategy school has been more explicit in its treatment of valuation by consumers, which underlies the ideas of competitive positions and substitution, its research agenda too seems to have been primarily focused on firm-firm interactions.

2 Following Brandenberger and Stuart (1996) the term willingness to pay indicates the maximum price a consumer would be willing to pay for an offer. It is a distinct construct from market price, which reflects relative value capture between producers and consumers and is influenced by bargaining power and rivalry.
We are concerned with how consumers’ relative willingness to pay for competing product and service offers changes as these offers improve over time. That is, we are fundamentally concerned with the evolution of substitutability between different offers, where differences in offers can reflect firms having different production technologies, different resource endowments or different competitive positions. We consider factors that act to homogenize the value of firms’ offers even when firms are able to maintain unique industry positions and possess non-imitable resources. Specifically, we highlight the extent of consumers’ decreasing marginal utility (DMU) – the degree to which consumers have a decreasing willingness to pay for performance improvements – as a key driver of dynamics that links supply-side technological progress and demand-side value creation. We highlight consumer heterogeneity – the extent to which consumers in different segments differ in their valuation of offers – as a key driver of the viability of different competitive positions. Finally, we highlight the interaction of these factors with firms’ production technology (e.g., the extent of economies of scale) in governing the convergence of value creation. Our focus on consumer choice and value creation complements the traditional focus in the strategy literature on competition and value capture.

Using this approach, we are able to consider whether technology threats are temporary or permanent; how the rents from different categories of resources change over time and the extent to which demand-side factors lead to their erosion; and when an industry segments according to the classic generic strategies of Cost Leadership and Differentiation, with firms located at other positions “stuck in the middle” (Porter, 1980), and when such generic strategies are out-competed by a Generalist that dominates from the middle.

We study these dynamics using a formal analytic model. Our objective is to validate our approach with a simple tractable model, just as Lippman and Rumelt (1982) use a simple model to demonstrate the importance of barriers to imitation for explaining heterogeneous firm performance. The origin of dynamics in our model is that the performance of product and service offers improves over time. We interpret performance improvement broadly (e.g., the increasing speed of microprocessors; the increasing safety of automobiles; the increasing breadth and timeliness of financial information). To highlight the role of demand-side factors, we consider their effects in a ‘best-case’ supply-side world, one in which firms can possess valuable, unique and inimitable resources and where a constant rate of technological progress is assured. Our approach extends prior work that has elaborated a demand-based perspective on technology competition (Adner and Zemsky, 2001; Adner, 2002) and technol-
ogy evolution (Adner and Levinthal, 2001). We see ours as a tractable baseline model upon which future work can build and which contributes to understanding the formal foundations of strategy (e.g., Brandenburger and Stuart, 1996; Makadok, 2001; Makadok and Barney 2001; MacDonald and Ryall, forthcoming).

The paper proceeds as follows. Section 2 introduces the model and discusses the key assumptions. Section 3 formally defines competitive advantage and value creation in the context of the model and shows how these concepts govern competitive interactions at a point in time. Section 4 uses the model to examine the evolution of substitution threats from alternative technologies. Section 5 characterizes how the rents from different categories of resources change over time. The analysis in Sections 4 and 5 is for one market segment and a single resource. Section 6 shows how these analyses can be aggregated to look at the evolution of substitution threats and resource rents when there are multiple market segments and multiple resources. Section 7 turns to the evolution of competitive positions. Section 8 discusses the implications of our results. All proofs are in the appendices.

2. The Model

Our intent is to present a stylized model that is useful in characterizing the demand-side drivers of the sustainability of competitive advantage. We explore these dynamics in a best-case supply-side world as detailed below. Despite the simplicity of our model, we are still able to capture reasonably complex behaviors.

We borrow two important elements from traditional models of product differentiation (e.g., Hotelling 1929; Shaked and Sutton, 1987). The first is heterogeneity of both consumers and offers. Specifically, offers differ in their performance levels and consumers differ in their willingness to pay for a given level of performance. The second is a discrete choice setting where each consumer either buys a single unit of one offer or buys nothing at all. We deviate from standard differentiation models by grouping consumers into discrete market segments and by assuming that offers improve over time. The specifics of the model are as follows.
2.1. Supply-Side

There are two firms that each produce a single offer. These offers differ in their cost of production and in their level of performance due to differences in production technologies, resource endowments or competitive positions. We index the firms and their offers by \( i = 1, 2 \). The performance of offer \( i \) is denoted by \( x_i \) and the marginal production cost is denoted \( c_i \).

We model the performance of offer \( i \) as a function of its technology trajectory, \( b_i \), which governs the rate at which performance improves over time, \( t \), independently of a firm’s investments; and as a function of the firm’s specific investments in differentiation, which we denote by \( d_i \). Specifically, we have \( x_i(t) = b_i d_i t \). We assume that costs are increasing in \( d_i \) such that there is a trade-off between performance and costs. While we do not model cost trajectories explicitly, note that the cost of producing a given level of performance falls over time as \( t \) and hence \( x_i \) increases.

Note that the parameter \( d_i \) allows firms to choose their competitive position. When studying substitution threats and resource rents in Sections 4–6, we abstract from positioning considerations by setting \( d_i = 1 \). Conversely, when studying competitive positioning in Section 7, we abstract from differences in technology or resources by assuming that trajectories do not vary across firms (i.e., \( b_i = b \)). Other section-specific assumptions are as follows:

**Substitute Technologies:** In Section 4, in which we study the evolution of the threat posed by substitute technologies, we consider a situation where firm 1 is bringing a new offer to market. The key attribute of both offers improves exogenously over time as \( x_1(t) = b_1 t \) and \( x_2(t) = b_2(t + h) \), where \( h > 0 \) reflects the fact that firm 2 is using a more mature technology that has a head start in the market.

**Resource Rents:** In Section 5, in which we study the evolution of resource rents, we consider a situation where firms 1 and 2 are identical except that firm 1 has a unique and inimitable resource. If not for this unique resource, the firms would have the same costs (\( c_1 = c_2 \)) and their offers would follow the same trajectory (\( x_1(t) = x_2(t) = bt \)). We categorize resources according to the way they create value. Process resources give firm 1 lower production costs so that \( c_1 < c_2 \). Product resources increase firm 1’s performance level by some \( r > 0 \); that is \( x_1(t) = bt + r \) and \( x_2(t) = bt \). Timing resources give firm 1 a head-start of \( h \) in developing its offer; that is \( x_1(t) = b(t + h) \) and \( x_2(t) = bt \).
resources give firm 1 a better trajectory so that \( b_1 > b_2 \).

**Competitive Positioning:** In Section 7, in which we study the evolution of competitive positions, we consider a situation where firms choose cost-performance positions along a productivity frontier that is shifting outwards over time. In particular, an offer’s performance is given by \( x_i(t) = b d_i t \). We examine two technologies that differ in their cost structures. For technology \( M \) there are no fixed costs and marginal cost is given by \( c_i = c + d_i \), where \( c > 0 \) is the minimal cost to produce the offer. For technology \( F \) marginal costs are given by \( c_i = c + (1 - f)d_i \) and fixed costs are given by \( f K d_i \) where \( f \in [0, 1] \) reflects the scalability of the technology because it determines the sensitivity of costs to production volume.

Throughout the paper we assume a best-case supply-side world where technological progress is largely exogenous to the firm’s activity in the focal market segment. Such technological progress in the absence of activity in the focal segment can be observed when the technology is present in other market segments. An example of this is Christensen’s (1997) description of the hard disk drive market, in which 3.5 inch hard drives initially serve only notebook computer users and then, after continued development in this niche, eventually improve to the point at which they can attract consumers in the mainstream desktop segment. From the perspective of an observer of the mainstream desktop segment (which is the perspective we will take in this paper) the 3.5 inch technology improved over time despite the fact that it had no sales in the mainstream segment.

Exogenous technological progress can also be driven by spillovers and transfers from external sources such as basic science, military R&D and advances in other industries (e.g., engines, navigation systems, construction materials and techniques that have been transferred to civilian aviation from the military). Moreover, progress can be driven by the activities of suppliers of general-purpose inputs (e.g., improvements in computer aided manufacturing systems and worker education that allow downstream firms to offer better products and services). Finally, the assumption of exogenous technological progress may not be necessary for our results. Adner (2002) studies the threat of technology substitutes using a simulation model with endogenous technological progress and finds results that are broadly consistent with ours.
2.2. Demand-Side

Consumers are divided into discrete market segments based on their willingness to pay for offers. Through Section 6, we focus on a single segment. Denote by \( s \) the number of consumers in this segment. Denote by \( w_i \) the segment’s willingness to pay for offer \( i \). We decompose willingness to pay into two components. The first is the offer’s quality as perceived by the segment \( q(x_i) \). The second component is the segment’s taste for quality, which is parameterized by \( a \). Willingness to pay for offer \( i \) at time \( t \) is then \( w_i(t) = aq(x_i) \). We introduce decreasing willingness to pay for quality improvement by assuming that \( q(x_i) = x_i^\beta \) where \( \beta \in (0, 1) \).

The parameter \( a \) can be interpreted in several ways. If consumers are individuals, \( a \) can be interpreted as a taste for quality, or alternatively as an ability to pay that is (usually) increasing in income level. If consumers are organizations, \( a \) can be interpreted as the importance of the input (e.g., airplane engine performance might be more important for military than for civilian buyers) or as the frequency with which the offer will be used.

The parameter \( \beta \) determines the extent to which consumers have decreasing marginal utility (DMU) from performance improvements. For example, the utility from an additional megahertz of processing power was much higher when micro processor speeds were 100mhz than when they were 1000mhz. DMU is of interest because it varies across products and services. For example, while DMU is high for microprocessor speed (relative to the rate of technological advance), DMU seems less pronounced for the resolution of digital cameras, where the latest models still seem to command a significant price premium.

A key property of the model is that consumers’ willingness to pay increases over time but at a decreasing rate, (i.e. \( \frac{\partial w_i}{\partial t} > 0 \) and \( \frac{\partial^2 w_i}{\partial t^2} < 0 \)). We simplify our model by assuming

---

4 Although consumer heterogeneity is usually modeled as a uniform continuous distribution in industrial organization (e.g., most work on Hotelling’s linear city), we consider discrete segments. Discrete segmentation is a good representation of heterogeneity in many settings such as when the product is a component used in multiple end products (e.g., hard disk drives, which are used in notebook, desktop and mainframe computers). Other examples of discrete consumer heterogeneity are personal versus professional users, industry segments (in business-to-business markets) and national markets.

5 For expositional simplicity, we will sometimes drop the explicit reference to time dependence. Here, for example, we write \( q(x_i) \) for \( q(x_i(t)) \).

6 Note that, mathematically, for \( x_i > 1 \), DMU is decreasing in \( \beta \), while for \( x_i < 1 \) DMU is increasing in \( \beta \). We restrict attention to \( x_i > 1 \) so that \( \beta \) has an unambiguous effect on DMU.

7 The functional form \( q(x) = x^\beta \) that we use to model DMU is conservative in that consumer utility is unbounded and hence \( \lim_{x \to \infty} w_i = \infty \). Imposing an upper bound on willingness to pay would imply that DMU is even more pronounced than our functional form.
a constant rate of technological progress for each firm (i.e., \( x_i(t) = b_i t \)) and by assuming decreasing returns to performance improvements on the demand side (i.e., \( \beta < 1 \)). We could relax these assumptions by assuming that \( x_i = b_i t^\alpha \), where \( \alpha < 1 \) corresponds to a decelerating rate of technological progress (e.g., Foster 1986) and \( \alpha > 1 \) corresponds to accelerating rates (e.g., Moore’s Law). All our results hold in this more general model as long as \( \alpha \beta < 1 \) so that there are net decreasing returns. We note that although decreasing marginal utility is mathematically equivalent to a decreasing rate of technological advance, the managerial implications of the constraints are very different. We emphasize DMU because the challenges associated with supply-side decreasing returns to innovation effort are already well studied (e.g., Foster 1986; Utterback, 1994).

2.3. Pricing

Denote by \( p_i \) the price that firm \( i \) quotes to the segment. We assume that consumers buy one unit of the offer that gives them the greatest surplus (or delivered value), \( w_i - p_i \), unless this is negative for both offers, in which case they make no purchase. As in classic work on product differentiation, we assume Bertrand (price-based) competition and that prices form an equilibrium of the one-shot game played between the firms at a point in time (Tirole, 1997). Thus, we abstract away from repeated game effects. We assume that price discrimination across segments is possible. This results in a maximum level of rivalry where the price of a weaker offer is competed down to its marginal cost. Abstracting from repeated game effects and allowing price discrimination greatly simplifies the game theoretic interactions between firms. In Appendix I we consider the effects of modeling competition as Cournot rather than Bertrand and of suppressing firm’s ability to price discriminate. We find that our results are broadly robust and discuss the specifics in the appendix.

Besides the above simplifications, we abstract from other factors that are important for

\[ \frac{\partial^2 w_i}{\partial t^2} = - \frac{(1 - \alpha \beta) ab_i^\beta \alpha \beta}{t^{2-\alpha \beta}} \]

which is negative if and only if \( \alpha \beta < 1 \).

For simplicity, we assume that a consumer indifferent between offer 1 and her next best alternative (either buying offer 2 or not purchasing at all), buys offer 1.

Repeatedly playing the one-shot equilibrium of the stage game is always one possible equilibrium of the repeated game. It is the unique subgame perfect equilibrium if the firms only meet a finite number of times (Vives, 1999).
strategy dynamics such as evolving costs (e.g., due to economies of learning), the drivers of technology progress such as organizational capabilities and innovation incentives, and the possibility of multi-product firms. The tractability of our baseline model suggests it as a potential vehicle to explore these additional research avenues in future work.

3. Preliminaries

We begin the analysis by formally defining two strategy concepts—value creation and competitive advantage—in the context of our model. We then show how these concepts determine which firm serves the focal segment and the resulting profit at a given point in time. These concepts and the static analysis provide the necessary foundation for the dynamic analyses in the rest of the paper.

**Definition 3.1.** An offer’s **value creation** for a consumer in a given segment is the difference between the consumer’s willingness to pay and the marginal production cost. We denote this by $v_i(t) = w_i(t) - c_i$.

It is useful to separate competitive advantage into two parts, relative costs and relative differentiation. We focus our analysis on offer 1 and define the following:

**Definition 3.2.** The **cost (dis)advantage** of offer 1 is $A_c = c_2 - c_1$. The **differentiation (dis)advantage** of offer 1 is $A_d(t) = w_1(t) - w_2(t)$. The net **competitive advantage** is then $A_d(t) + A_c$.

Note that both an offer’s value creation and its differentiation advantage will vary across segments.

We now characterize the outcome of competition between firm 1 and firm 2 at a given point in time. Specifically, we take as given the value creation of each firm, $v_1$ and $v_2$. We are interested in whether firm 1 sells its offer to the segment and if it does, what level of profits it achieves. The answer depends on whether or not offer 2 has positive value creation for the segment. If it does not ($v_2 < 0$), then offer 2 is irrelevant because consumers are not willing to pay the marginal cost of production ($w_2 < c_2$). Firm 1 serves the market if its price is no greater than consumer’s willingness to pay, $p_1 \leq w_1$. The firm optimally sets the highest possible price $p_1^* = w_1$ and this is profitable if $w_1 > c_1$ or equivalently $v_1 > 0$. Profit
is then \( s(p^*_1 - c_1) = sv_1 \). Thus, firm 1’s ability to serve the segment and its profits depend on its absolute level of value creation.

Now suppose that firm 2 does have positive value creation in the focal segment so that \( v_2 = w_2 - c_2 > 0 \). Firm 2 will be willing to reduce its price to \( c_2 \) in order to capture the segment, which results in a delivered value of \( w_2 - c_2 \). For firm 1 to serve the segment it must then set a price so that \( w_1 - p_1 \geq w_2 - c_2 \). In this case, the optimal price is \( p^*_1 = w_1 - w_2 + c_2 \). Charging this price is only profitable if \( p^*_1 - c_1 > 0 \). Note that \( p^*_1 - c_1 = w_1 - w_2 + c_2 - c_2 = A_d + A_c = v_1 - v_2 \). If \( p^*_1 - c_1 > 0 \), the profit is \( s(p^*_1 - c_1) = s(A_d + A_c) = s(v_1 - v_2) \). Thus, in the competitive case, firm 1’s ability to serve the segment and its profits depend on its competitive advantage over firm 2, which is equivalent to the superiority of firm 1’s value creation.\(^\text{11}\)

Thus, when a rival begins to create value in the segment, the determinant of the focal firm’s profits shifts from its absolute level of value creation to its relative value creation. Formally:

**Proposition 3.3.** (i) Consider a segment where the alternative offer does not create value \((v_2 \leq 0)\). Consumers buy the focal offer if it has positive value creation \((v_1 \geq 0)\) and the profit is then proportional to the size of the segment and the value created: \( sv_1 \).

(ii) Now consider a segment where the alternative does create value \((v_2 > 0)\). Consumers buy the focal offer if it has superior value creation \((v_1 \geq v_2)\), which is equivalent to having a competitive advantage \((A_d + A_c \geq 0)\). The profit is proportional to the size of the segment and the competitive advantage: \( s(A_d + A_c) \).

This proposition is consistent with Brandenburger and Stuart (1996), who use cooperative game theory to argue that a firm’s added value places an upper bound on the rents that it can appropriate.

Given the Bertrand specification of competition in our model, in which firms set prices and then produce to meet demand, only the firm with greater value creation will serve the focal segment at any point in time. Under Cournot competition, in which firms set quantities and then price to clear the market, both firms can simultaneously serve the market (see Appendix I). In both cases, however, profits are directly related to a firm’s absolute and

\(^{11}\)Note that firm 1 is lowering its price in response to the threat of entry by firm 2. The threat is only realized when firm 2 comes to have superior value creation, \( v_2 > v_1 \).
relative value creation and the key qualitative results are the same. We present the model under the Bertrand assumption because it allows for a more straightforward exposition.

4. The Threat of Technology Substitutes

Industry analysis (Porter, 1980) is often criticized as taking too static a view of market boundaries (e.g., Grant, 1998), leaving undeveloped how the threat of substitutes changes over time. The importance of understanding shifts in market boundaries is highlighted, for example, by the phenomenon of disruptive technologies (Christensen, 1997). Disruptions occur when existing industry boundaries are redrawn by the entry of firms using new technologies that start in a niche segment and, as they improve, displace existing technologies from mainstream segments. The recent technology bubble, in which many promising new technologies turned out not to be disruptive, underscores the need to critically assess the substitution threat posed by new technologies.

We consider whether a new offer will displace an established offer from a given segment; if so, when this will occur; and finally, whether displacement will be permanent or transitory. We show how the answers depend on extent of DMU, consumer’s taste for quality, technology trajectories and cost positions. Note that one can interpret offers broadly, including competition between different strategies such as discount retailers and department stores.

We consider a situation where firm 1 is bringing a new type of offer to the market. We assume that firm 2 has a head-start in the market of $h > 0$. The performance of both offers improves over time as follows:

\[
\begin{align*}
    x_1(t) &= b_1 t, \\
    x_2(t) &= b_2 (t + h),
\end{align*}
\]

where $b_i$ is the technology trajectory along which offer $i$ is improving. Thus, the firm’s value creation increases over time as follows: $v_1(t) = ab_1^\beta t^\beta - c_1$ and $v_2(t) = ab_2^\beta (t + h)^\beta - c_2$.

At time $t = 0$, offer 1 must have a negative value creation since $v_1(0) = -c_1$. Let $t_0 > 0$ be such that $v_1(t_0) = 0$. We assume that at this time firm 2 is already creating value for the segment (i.e., $v_2(t_0) > 0$), which assures that offer 2 is already established in the segment. Finally, we assume that $t_0 b_1 > 1$ so that DMU is decreasing in $\beta$.

Since offer 2 already has positive value creation, from Proposition 3.3 we have that offer
1 can only enter the segment profitably if it has a competitive advantage, $A_c > 0$. The cost advantage $A_c$ is fixed while the differentiation advantage $A_d(t)$ changes over time as follows

$$A_d(t) = w_1(t) - w_2(t) = a(b_1 t)^\beta - a(b_2 (t + h))^{\beta}. $$

The timing of entry, which we denote by $t_E$, occurs when $A_d(t_E) + A_c = 0$. We begin by characterizing entry dynamics when the offers have the same trajectory.

**Proposition 4.1.** Suppose the technologies have the same trajectory ($b_1 = b_2$). Without a cost advantage ($A_c \leq 0$), the new firm never enters the segment. With a cost advantage ($A_c > 0$), the new firm displaces the established firm from some time $t_E$ onwards. The greater the extent of DMU and the smaller the segment’s taste for quality, the sooner the new firm enters the segment (i.e., $t_E$ is increasing in $\beta$ and $a$).

The intuition for Proposition 4.1 is as follows. With identical technology trajectories, firm 1 always has a differentiation disadvantage (i.e., $A_d(t) < 0$ for all $t$) due to firm 2’s head start, which implies that firm 2 has a differentiation advantage. However, firm 2’s differentiation advantage erodes over time due to DMU as firm 2’s performance advantage has less and less impact on consumers’ willingness to pay. In the limit, DMU completely erodes firm 2’s advantage (i.e., $\lim_{t \to \infty} A_d(t) = 0$). Thus, if firm 1 has a cost advantage, it will enter the segment eventually. What affects the time to entry? The lower the marginal utility from quality improvements, which depends on DMU and the taste for quality, the faster the differentiation advantage erodes and the sooner entry occurs. This proposition highlights the importance of cost advantage and consumers’ willingness to pay for performance improvements when identifying disruptive threats.

Now consider entry dynamics when technologies have different trajectories:

**Proposition 4.2.** Suppose the new technology has a better trajectory ($b_1 > b_2$). The new firm always displaces the established firm from some time $t_E$ onwards with $t_E$ decreasing in $A_c$. The effect of DMU and the taste for quality depends on the cost position. Specifically, with a cost advantage ($A_c > 0$) the time of entry is delayed as the taste for quality increases and as DMU decreases; conversely, with a cost disadvantage ($A_c < 0$) the time of entry is advanced as the taste for quality increases and as DMU decreases.

Because firm 2 has a head start, firm 1 starts with a differentiation disadvantage ($A_d(0) < 0$), but over time it develops a differentiation advantage due to its superior technology tra-
Proposition 4.3. Suppose the new technology has a worse trajectory \( b_1 < b_2 \). If the new firm does not have a sufficiently large cost advantage, it never enters the segment. If it does have a sufficiently large cost advantage, it enters the segment, but only for a limited interval of time. The duration of entry is decreasing in DMU and in the segment’s taste for quality.

When the new technology has a worse trajectory but a sufficient cost advantage, it enters the focal segment but only temporarily. The logic underlying these temporary threats is illustrated in Figure 4.1. Because of its superior technology trajectory and its head start, firm 2 always has a differentiation advantage over firm 1. At first, the differentiation advantage falls as DMU erodes the effect of firm 2’s head start. If firm 1 has a cost advantage that is large
enough to offset its differentiation disadvantage it enters the market at time $t_1$. However, firm 1’s competitive advantage is not sustainable. Over time, firm 2’s superior trajectory causes its differentiation advantage to grow. Eventually, firm 2’s differentiation advantage offsets its cost disadvantage and firm 2 re-enters the segment at time $t_2$. An example of a temporary threat is Bic’s entry into the mainstream razor market with its disposable offering, an incursion that was ultimately reversed by Gillette’s superior technology trajectory for cartridge razors. Distinguishing between permanent and temporary threats informs the choice of whether firms should respond to threats by continuing to invest along their existing trajectory or by embracing the new technology.

5. The Evolution of Resource Rents

A fundamental question in strategy is the sustainability of resource rents. The question of how the value of resources evolves, however, is largely unexplored. The resource based view argues that competitive advantage is conferred to those firms that possess valuable, rare, inimitable and non-substitutable resources. The received literature has primarily focused on firms’ abilities to acquire resources at favorable prices (e.g., Barney, 1986; Dierickx and Cool, 1989; Makadok, 2001) and on the threat of imitation (e.g., Lippman and Rumelt (1982), Barney (1991), Peteraf (1991)). Although the resource-based view is clearly sensitive to the issue of resource value and substitution, its treatment of these constructs is less well developed than its treatment of barriers to imitation (Priem and Butler, 2001). Further, within the resource based view, substitution is usually approached from a supply-side perspective, concerned with the emergence of equivalent, or possibly superior, resources for production (e.g., the substitution of charismatic leadership for formal planning systems, Barney 1991), rather than with buyers’ assessments of relative value in consumption.

In this section we identify different categories of resources based on how they affect a firm’s value creation and consider how rents from different types of resources vary over time. We highlight the role of DMU in both dampening the rate at which resource rents buildup and in hastening the rate at which rents decay. We start with the simplest case, in which only one firm has a resource and there is a single focal segment; the following section extends the analysis to a multi-resource, multi-segment setting.

Assume that firms 1 and 2, and their offers, are identical except that firm 1 has a unique and inimitable resource that improves its ability to create value for the focal segment. But
for this unique resource, the firms would have the same costs \(c_1 = c_2\) and their offers would follow the same technology trajectory \(x_1(t) = x_2(t) = bt\). Recalling the resource typology introduced in Section 2.1, consider the ways in which resources can impact firm 1’s competitive advantage:

- Resources can lower firm 1’s production costs, so that \(c_1 < c_2\). We term such resources **process resources**. Dell’s superior supply chain management in the PC industry could be characterized as a process resource.

- Resources can increase the performance of firm 1’s offer by \(r > 0\), so that \(x_1(t) = bt + r\). We term such resources **product resources**. McKinsey’s reputation in management consulting and Sony’s advantages in miniaturization could be characterized as product resources.

- Resources can give firm 1 a head start of \(h\) on developing its offer, so that \(x_1(t) = b(t + h)\) We term such resources **timing resources**. One can think of firms that are consistently fast to market such as 3M as possessing a timing resource.\(^{12}\)

- Resources can give firm 1 a better technology trajectory, so that \(b_1 > b_2\) when \(x_1(t) = b_1t\) and \(x_2(t) = b_2t\). We term such resources **innovation resources**. One could interpret Gillette’s superior product development process as rooted in an innovation resource.

The early rent dynamics for all resources are broadly the same: Because of its unique resource, firm 1 has a competitive advantage over firm 2 at all times, \(A_d(t) + A_c > 0\). Therefore, firm 1 will be the first to enter the segment and is never displaced. Let \(t_1\) be the time of entry for firm 1. Let \(t_2 > t_1\) be the time at which firm 2’s offer creates positive value in the segment. Although firm 1 is always dominant, its rents change over time:

**Proposition 5.1.** From the time of firm 1’s entry into the segment (i) the rents from cost-resources increase over time and then stabilize; (ii) the rents from both performance-resources and timing-resources first increase and then decrease over time; (iii) the rents from innovation resources increase over time, but at a decreasing rate.

\(^{12}\)Note that performance and timing resources are formally equivalent, differing only by a factor \(b\). In particular, the dynamics are identical for \(r = bh\) since \(x_1(t) = bt + r\) for an attribute resource and \(x_1(t) = b(t + h)\) for a timing resource.
Figure 5.1: The evolution of value creation and rents when firm 1 has a cost resource (for $c_1 = 2$, $c_2 = 3$, $\beta = .5$, and $b = 1$).

The intuition for the results is as follows. From time $t_0$ to $t_1$ neither firm is in the market because neither firm has positive value creation. From time $t_1$ to $t_2$, firm 1 is alone in the market because firm 2 does not create value and hence from Proposition 3.3 part (i), we have that firm 1’s rents are proportional to its absolute level of value creation and the size of the segment: $sv_1(t)$. Over time, as the quality of firm 1’s offer improves, its value creation and its rents increase (but at a decreasing rate due to DMU). Starting at time $t_2$, firm 2’s offer has improved sufficiently that it too has positive value creation in the segment, and hence Proposition 3.3 part (ii) applies. From this point on, firm 1’s rents are proportional to its relative value creation and the size of the segment: $s(v_1(t) - v_2(t)) = s(A_d(t) + A_c)$. The evolution of rents from time $t_2$ on depends on the type of resource.

For process resources, there is no differentiation advantage ($A_d(t) = 0$ for all $t$) and hence the competitive advantage is a constant ($v_1 - v_2 = A_c$), which implies that rents stabilize at $sA_c$. Figure 5.1 illustrates. Until time $t_1$, neither firm has positive value creation. From time $t_1$ to time $t_2$ only firm 1 has positive value creation and during this interval firm 1’s rent ($sv_1(t)$) increases with its value creation. After time $t_2$ both firms have positive value creation that is increasing at the same rate, and hence rents stabilize.

For both timing resources and product resources, firm 1 has a differentiation advantage
Figure 5.2: The evolution of value creation and rents when firm 1 has a performance resource or a timing resource (for $c_1 = c_2 = 2$, $\beta = .4$, $b = 1$, $r = h = 4$).

$(A_d(t) > 0)$, but this advantage erodes over time $(A'_d(t) < 0)$ due to DMU. Hence, firm 1’s rents decay starting at time $t_2$. Figure 5.2 illustrates. Thus, even a firm possessing an inimitable resource, such that it maintains performance superiority (i.e., $x_1(t) - x_2(t)$ is a constant over time), can see its resource rents decay over time as DMU erodes consumers’ willingness to pay for those differences.

With innovation resources, firm 1 again has a differentiation advantage, but in this case the advantage grows over time $(A'_d(t) > 0)$, but at a decreasing rate $(A''_d(t) < 0)$. Hence, firm 1’s rent continues to increase even after firm 2 has positive value creation, although the rate of profit growth decelerates. Figure 5.3 illustrates. This logic speaks to the difficulty firms have in maintaining profit growth rates—even in the best case of non-imitable innovation resources we observe decaying profit growth.

We close by considering the effect of consumers’ willingness to pay for quality on the net present value of resource rent streams.

**Proposition 5.2.** For all resource types, the net present value of the resource rent stream is decreasing in the extent of DMU and increasing in the taste for quality.

As the taste for quality increases, both $v_1(t)$ and $v_2(t)$ shift up proportionally, which speeds firm 1’s entry ($\partial t_1/\partial a < 0$) and increases rent prior to time $t_2$. For performance,
Figure 5.3: The evolution of value creation and rents when firm 1 has a innovation resource (for $b_1 = 2$, $b_2 = 1$, $c_1 = c_2 = 2$, $\beta = .5$).

timing and innovation resources, the taste for quality also increases subsequent rents by magnifying the differentiation advantage ($\partial A_d/\partial a > 0$). Conversely, the extent of DMU decreases both $v_1(t)$ and $v_2(t)$, which delays firm 1’s entry and (weakly) reduces rent at any point in time. This proposition suggests a need to control for DMU and taste for quality in empirical studies of profitability across segments and industries.

6. Competing with Resources across Segments

To this point we have restricted ourselves to examining sustainability in a single segment. These analyses provide building blocks that can be assembled to examine more complex settings. Most industries have multiple market segments and many resources. In this section, we consider an industry with two market segments and two competitors, each possessing a different resource. We are interested in how competition emerges across segments and its implication for the evolution of resource rents.

Consider the following setting: There is a high-end segment with a taste for quality $a_H = 2$ and a low-end segment with a taste for quality $a_L = 1.4$. In a loose analogy to Sony and Matsushita in consumer electronics, firm 1 has a product resource ($r = 3$) and firm 2
Figure 6.1: The evolution of firm rents when firm 1 has a performance resource and firm 2 has a cost resource in a two segment setting.

has a process resource ($A_c = 0.7$). Figure 6.1 shows how each firm’s rents change over time.

First note that both offers have positive value creation in the high-end segment before the low-end segment, because of the former’s higher taste for quality ($a_H > a_L$). In this example, firm 1 is first to create value in the high-end, starting at $t_1$. Until time $t_2$, only firm 1 creates value in the high-end segment. Firm 1’s rents, which are equal to its value creation for the segment, increase as its offer improves over time. At time $t_2$, firm 2’s offer begins to create value in the high-end segment as well. At this point firm 1’s rents, which are now its competitive advantage, begin to decay as DMU erodes its differentiation advantage. (To this point, the dynamics are the same as those illustrated in Figure 5.2.)

At time $t_3$, firm 1 returns to a period of profit growth as its offer begins to deliver sufficient value to serve the low-end segment. Hence, firm 1 starts earning rents from the low-end segment which are proportional to its absolute level of value creation and which increase as its offer improves over time. The increasing rents from the low-end segment more than offset the continued decay in rents from the high-end.

\[\beta = 0.5, b = 1, s = 100, c_1 = 3.5, c_2 = 2.8.\]
At time $t_4$, firm 2 begins to create value in the low-end, triggering a permanent decline in firm 1’s fortunes. Now, firm 1’s rents from both segments are decaying over time.

Note that until time $t_5$, firm 2 does not sell to either of these segments. It is the increasing threat of entry by firm 2 that limits firm 1’s rents. At time $t_5$, the threat is realized: DMU has eroded firm 1’s differentiation advantage sufficiently that firm 2’s cost advantage gives it superior value creation in the low-end segment. Consequently, at time $t_5$ firm 1 is displaced from the low-end segment. Then, at time $t_6$, the differentiation advantage has eroded so much that firm 2 displaces firm 1 from the high-end segment as well. After time $t_6$, firm 2’s rents continue to increase as firm 1’s differentiation advantage further decays, converging to a final profit level of $s_A = 70$ in each segment.

Thus, by assembling different elements of our simple model, one can characterize relatively complex dynamics that include shifts in market leadership across segments and shifts in firms’ absolute and relative profits over time. This section examined how firms’ competitive advantage varies across market segments and over time given exogenous resource endowments. Below we examine competitive advantage as an endogenous outcome when firms choose their competitive positions.

7. Competitive Positioning

Beyond the question of resources, firms’ value creation is substantially affected by their choice of competitive positions (e.g., Porter, 1980). Porter argues that firms face a choice between positioning as Cost Leaders or Differentiators, where the latter have higher quality offers and higher costs than the former. Those firms that do not choose one of these positions risk being “stuck in the middle” and being out competed.

One limitation of the generic strategy perspective is the observation that firms that pursue both cost and differentiation advantage simultaneously are sometimes very successful (e.g., Kim and Mauborne, 1997; Barney, 1997; Besanko et al. 2000). More generally, a weakness with the received literature on positioning is that “our understanding of the dynamic processes by which firms perceive and ultimately attain superior market positions is far less developed [than our understanding of advantage at a point in time]” (Porter, 1991, 14)

---

14 Although, per the discussion of the assumption of exogenous technological progress in Section 2.1, it might be selling its products to consumers in a (unanalyzed) niche segment. Also, recall that Firm 2’s complete absence from the segment despite positive (though inferior) value creation is an artifact of the assumption of Bertrand competition and is not the case under Cournot competition (see Appendix I).
Motivated by these critiques, we focus on three questions: When should firms choose to follow classic generic strategies? When should firms position in the middle? What causes new competitive positions to arise over time? Our results offer a logic with which to approach both cross-sectional comparisons as well as longitudinal patterns in positioning choices.\footnote{We do not address the generic strategy of Focus in this paper. To do so would require introducing niche segments based on horizontal differentiation, which is beyond the scope of this paper.}

Following Porter (1980), we assume a trade-off between product quality and production costs. Specifically, firms choose $d_i$, a level of differentiation, that determines both the quality of their offer and their costs.\footnote{Thus, $d_i$ can be interpreted as an investment in a performance resource.} As elaborated below, the precise effect of $d_i$ on costs depends on the nature of the technology that the firm uses. We consider the case of two market segments that vary in their taste for quality. Specifically, there is a low-end segment with a taste for quality of $a_L$ and a high-end segment with a taste for quality of $a_H > a_L$. We allow for as many as three firms to be active in the market.

Our analysis proceeds in three steps. Section 7.1 characterizes the static choice of positioning in a Porterian world where firms that do not follow Cost Leadership or Differentiation strategies are indeed stuck in the middle. Section 7.2 shows how classic generic strategies breakdown in the presence of a sufficiently scale intensive technology. Following these static analyses, Section 7.3 shows how new positions arise over time as technologies improve.

### 7.1. Segmentation and Generic Strategies

In this subsection we show formally how a market can be segmented by firms using classic generic strategies. We assume that firms only have access to production technology $M$ for which the quality of a firm’s offer is given by $x_i = bd_i$ and the marginal cost of production is $c_i = c + d_i$. Because cost and quality are both increasing in $d_i$, there exist production possibility frontiers along which production cost and willingness to pay are traded off (Porter 1996; Saloner, et. al., 2001), with a different frontier for each segment.\footnote{The familiar depiction of a single cost-willingness to pay frontier ignores consumer heterogeneity. We show two different frontiers because our segments differ in their taste for quality, and hence in their willingness to pay for a given performance level.} Figure 7.1 illustrates.

We define $d^*_H$ as the level of differentiation that maximizes value creation, and hence competitive advantage, for the high-end segment. Similarly, we define $d^*_L$ as the level of
differentiation that maximizes value creation for the low-end segment. Because the optimal level of differentiation is increasing in the segment’s taste for quality, we have $d^*_H > d^*_L$.

**Proposition 7.1.** Suppose an entrant faces a single Generalist incumbent serving both segments from a middle position $d_I$ (i.e., $d^*_L < d_I < d^*_H$). (i) The optimal position for the entrant is either as a Cost Leader serving only the low-end segment from the position $d^*_L$ or as a Differentiator serving only the high-end segment from the position $d^*_H$. (ii) The relative attractiveness of being a Cost Leader is increasing in the quality level of the incumbent ($d_I$) and the extent of DMU. The attractiveness of being a Differentiator is increasing in the technology trajectory ($b$) and the taste for quality in both the low end ($a_L$) and the high-end ($a_H$).

With the incumbent positioned in the middle, there is room to enter by exclusively targeting either segment. Recall from Proposition 3.3 that entry into a segment requires a competitive advantage for that segment. A Cost Leader has a competitive advantage in the low-end segment because its position $d^*_L$ gives it a cost advantage that more than offsets its differentiation disadvantage in the low-end. However, it cannot enter the high-end because

---

**Figure 7.1:** The cost-willingness to pay frontier for both segments when $\beta = .5$, $b = 1$, $a_L = 2$, and $a_H = 3$.

---

18That is $d^*_\theta = \text{arg max}_d (a_\theta (bd)^{\beta} - (c + d)) = [a_\theta \beta b^\beta]^{-\frac{1}{\beta}}$ for $\theta = H, L$. 

21
those consumers’ greater taste for quality magnifies its differentiation disadvantage in the high-end segment. In contrast, a Differentiator has a competitive advantage in the high-end because its position \( d_H^* \) gives it a differentiation advantage that more than offsets its cost disadvantage in the high-end segment, but it has a competitive disadvantage in the low-end segment.

Part (ii) of Proposition 7.1 addresses the relative attractiveness of the two generic strategies. On the demand-side, as performance becomes more important for consumers (either due to increases in the taste for quality or due to decreases in the extent of DMU), the more attractive it is to be a Differentiator. On the supply-side, one consideration in choosing a generic strategy is to move away from existing competition, a familiar result from industrial organization models. Thus, the higher the incumbent’s quality (\( d_I \)) the more attractive it is to be a Cost Leader. In addition, a better technology trajectory makes it easier to increase performance and hence more attractive to be a Differentiator.

**Corollary 7.2.** Suppose two entrants face a single Generalist incumbent positioned at some \( d_L^* < d_I < d_H^* \). The Generalist is “stuck in the middle” in that firms will enter as both Cost Leaders and Differentiators, which leaves the Generalist with a competitive disadvantage in both segments. Further, with firms positioned at both \( d_L^* \) and \( d_H^* \) there is no position at which a new firm can profitably enter using technology \( M \).

Thus, we offer a formal characterization of generic strategies and the condition of being stuck in the middle.

### 7.2. De-segmentation and Positioning in the Middle

We now identify conditions under which a firm can profitably enter an industry by simultaneously serving both segments from a single position. From Corollary 7.2, we know that this cannot occur if firms are limited to technology \( M \). Extend the firms technology possibilities to include a “fixed-cost” technology \( F \). As with technology \( M \), the quality of a firm using technology \( F \) is \( x_i = bd_i \). In contrast to technology \( M \), with technology \( F \) a fraction \( f \) of the costs associated with differentiation are fixed. The parameter \( f \in (0, 1) \), which we refer to as the “scalability” of the technology, splits the effect of differentiation into a fraction \( fd \) that increases fixed costs and a fraction \( (1 - f)d \) that increases marginal costs.\(^{19}\) The

\(^{19}\)Recall that for technology \( M \) all of the costs associated with differentiation increase marginal costs, which take the form \( c + d_i \). Note that technology \( M \) is a special case of technology \( F \) with \( f = 0 \).
marginal cost is then $c + d(1 - f)$ and the fixed cost is $dfK$.\textsuperscript{20} Note that the more scalable the technology, the lower the variable costs. An example of differentiation with low scalability is adding leather to a car’s interior, which increases the production cost for each car. An example of differentiation with high scalability is adding a fuel cell engine to a car where the required R&D investment is a fixed cost that does not increase with production volumes. The fixed cost parameter $K$ affects the level of fixed costs required for differentiation. We restrict attention to $s < K$.\textsuperscript{21}

**Proposition 7.3.** Consider a potential entrant facing incumbents at $d^*_L$ and $d^*_H$ using technology $M$. There exists a critical value $\bar{K} > s$ such that for $K \geq \bar{K}$, entry is never profitable. For $K < \bar{K}$ we have: (i) for low levels of scalability ($0 < f < f_1$), profitable entry is not possible and the market remains segmented by firms pursuing Cost Leadership and Differentiation; (ii) for intermediate levels of scalability ($f_1 < f < f_2$), entry as a Generalist (using technology $F$) is profitable and the entrant’s optimal position $d^*_E$ allows it to dominate both segments from the middle (i.e., $d^*_L < d^*_E < d^*_H$); (iii) for high levels of scalability ($f_2 < f$), entry as a Generalist (using technology $F$) is profitable and the entrant dominates both segments with a quality level higher than theDifferentiator’s (i.e., $d^*_E > d^*_H$).

The balance of two countervailing forces determines the possibility of “dominating from the middle.” On the one hand, a firm’s ability to exploit economies of scale acts to increase the attractiveness of serving both segments as a Generalist using technology $F$. On the other hand, heterogeneity across market segments acts to increase the attractiveness of the specialist strategies of Cost Leadership and Differentiation using technology $M$ because a firm can optimally trade-off marginal cost and performance (as in Proposition 7.1). With low scalability ($0 < f < f_1$), the economies of scale are insufficient to offset the advantages of fine tuning the offer to a single segment and the Generalist is unable to profitably enter the market. With access to a sufficiently scalable technology ($f_1 < f$), the Generalist is able to out compete the specialists in both segments.

Note that a successful Generalist might or might not be located in the middle, depending on the level of scalability of its technology. For intermediate levels of scalability ($f_1 < f_2$), the Generalist is able to dominate both segments.

\textsuperscript{20}This specification of technology $F$ draws on Sutton (1991)’s work on endogenous fixed costs.

\textsuperscript{21}The restriction means that technology $F$ is unattractive to a firm that serves a single segment but potentially attractive to a firm serving both segments. Specifically, the difference between the cost of serving a single segment with technology $F$ and the cost with technology $M$ is $s(c + d_i(1 - f)) + dfK - s(c + d_i) = d_i f(K - s) > 0$. 

23
the Generalist locates in the middle and offering a compromise product. For a sufficiently high level of scalability \((f_2 < f)\), its cost of increasing performance is so low that the performance of the Generalist’s offer exceeds that of a Differentiator. An example of a Generalist leveraging high fixed costs to target the mass of the market is Barnes and Noble book superstores which offered higher performance (e.g., wider selection, knowledgeable staff and in-store cafe) than the differentiated independent booksellers that had dominated the high-end of the market prior to its entry.

### 7.3. The Evolution of Competitive Positions

Thus far, we have identified three strategies that can be viable at any point in time: Cost Leader, Differentiator and Generalist. We now consider the emergence of firms using these strategies in a dynamic setting where technologies improve over time. Since \(x_1(t) = bd_t\), the cost-willingness to pay frontiers are shifting outward over time. When \(t = 0\), consumers have a zero willingness to pay for offers, while the marginal cost of production is always at least \(c\). Hence, no firm can profitably enter the market.

We are interested in which strategy is used by the firm that pioneers the market and in whether that strategy is sustainable as technology matures. We identify three regimes.

**Proposition 7.4.** Suppose there is a pool of potential entrants with access to technologies \(M\) and \(F\) and \(K < \bar{K}\). (i) For low levels of scalability \((0 < f < f_1)\), a Differentiator pioneers the market and is later joined by a Cost Leader. (ii) For intermediate levels of scalability \((f_1 < f < f_3)\), a Differentiator pioneers the market and is later displaced by a Generalist; for \(f_1 < f < f_2\), the Generalist has lower quality than the Differentiator and for \(f_2 < f < f_3\) the Generalist has higher quality. (iii) For high levels of scalability \((f_3 < f)\), a Generalist is the first and only firm to enter the market.

Consider first the case where technology \(F\) is not very scalable \((f < f_1)\). From Proposition 7.3, we know that a Generalist strategy is never used. Because it is easier to create value in the high-end due to its greater taste for quality, a Differentiator is able to enter the market when a Cost Leader’s offer still has negative value creation. As technology improves further, the Cost Leader strategy becomes viable as well and the two strategies coexist in the market.

For intermediate levels of scalability \((f_1 < f < f_3)\), the Generalist strategy dominates in the long-run. Initially, however, the willingness to pay of the low-end segment is too low to
Figure 7.2: The evolution of competitive positions for different levels of scalability and consumer heterogeneity for $\beta = .5$ (solid lines) and $\beta = .4$ (dashed lines) and $K = 1.25s$. 

justify the broad market deployment that is the hallmark of a Generalist. The Differentiator, unencumbered by fixed costs and focused only on the high-end, is then the first to create value and therefore pioneers the market. Over time, with further technology improvements, the willingness to pay of the low-end segment increases sufficiently that the Generalist strategy becomes viable and it displaces the Differentiator. 

For high levels of scalability ($f_3 < f$), the Generalist’s marginal costs are so low that serving both segments is profitable early on, leaving no room for other strategies. 

What factors determine which regime characterizes a given market?

**Corollary 7.5.** The critical thresholds $f_1 < f_2 < f_3$ from Propositions 7.3 and 7.4 are increasing in consumer heterogeneity ($\frac{a_H-a_L}{a_H}$) and the extent of fixed costs ($K$) and they are decreasing in the size of the segments ($s$). The extent of DMU decreases $f_1$, increases $f_3$ and does not affect $f_2$. 

Consumer heterogeneity ($\frac{a_H-a_L}{a_H}$) reflects the extent to which the segments differ in their taste for quality. As heterogeneity increases the returns to targeting individual segments increase and so the Generalist strategy becomes less attractive. Thus, as Figure 7.2 illustrates, the thresholds $f_1$ and $f_3$ both increase in consumer heterogeneity. Now consider the
effects of DMU. On the one hand, DMU acts to mask heterogeneity between segments by reducing the difference in optimal quality levels \( (d^*_H - d^*_L) \), which shrinks region I, in which the Differentiator is joined by a Cost Leader (i.e., \( f_1 \) falls in DMU). On the other hand, DMU acts to lower overall willingness to pay, which makes it less attractive to serve both segments early on, which shrinks region III, in which the market is pioneered by a Generalist (i.e., \( f_3 \) increasing in DMU). The dashed lines in Figure 7.2 show the effects of an increase in DMU (moving from \( \beta = .5 \) to \( \beta = .4 \)).

Finally, consider the effects of market size \( (s) \) and fixed costs \( (K) \). The larger the size of the market, the more attractive is the Generalist strategy due to its scale economies. Hence, growing market size can trigger a shift to a Generalist strategy. Conversely, the larger the fixed costs, the less attractive is the Generalist strategy. The spectacular failure of many internet companies highlights the importance of these factors. For example, the large fixed costs in advertising and warehouses required for firms such as Etoys and WebVan were too large relative to the size of their markets.

8. Discussion

Our approach to sustainability starts with product market competition and improving technologies. The novelty is that we introduce an explicit treatment of how technology improvements affect consumer choice among competing offers. This leads us to focus on demand-side drivers: consumer taste for quality, the degree to which utility from quality improvements decreases as offers improve, and the extent of consumer heterogeneity. We combine these elements in a simple model that we use to address a wide range of issues related to the sustainability of competitive advantage. Our goal is to complement, not displace, the traditional focus in strategy on competitors and imitation with a focus on consumers and value creation.

The strategy field has traditionally approached the question of firm heterogeneity from a supply-side perspective, concerned with the ways in which firms can or cannot replicate each others’ resources, capabilities and positions. Of particular concern has been the role of imitation in undermining firms’ unique advantages and the role of isolating mechanisms in limiting this erosion. In contrast, this paper considers firm heterogeneity from the perspective of consumers. It shows how differentiation advantage can erode, not because imitation leads to the convergence of resources and absolute levels of performance, but because consumer
valuation of firm differences declines due to decreasing marginal utility.

A limitation of our model is that firms are treated as relatively static actors who do not make additional investments to affect the dynamics to which they are subject (e.g., through innovation strategies or additional investments in resources). Although such responses are outside the scope of this paper, the analysis does provide a useful platform for considering how firm actions might affect these dynamics. The demand-side variables that we highlight can inform the choice of strategic responses in different settings: the characterization of temporary versus permanent substitution threats (Proposition 4.3) offers guidance as to whether firms should respond to the threat of substitute technologies by embracing the new technology or by improving their existing technology; the characterization of positioning speaks to when incumbents may want to consider moving from specialist to generalist positions (Propositions 7.3 and 7.4).

While DMU is characteristic of many (though not all) settings, it has largely been ignored in strategy studies. We note that the assumption of DMU is critical to our results and that without it, many of our observed dynamics would disappear. This, of course is precisely the point: to highlight the role of demand-side factors in determining sustainability, and to assure that these demand-side threats are not overlooked. We thus hope first, to encourage the field to be sensitive to their existence; and second, in those settings where DMU is important, to develop intuitions regarding their effects on sustainability.

While we have presented a tractable model that could be extended in future theoretical work, we see empirical testing and industry cases as critical to pushing forward this research agenda. Our model and propositions suggest new ways to approach longitudinal studies of business strategy. We specify the effects of, and interactions among, several understudied independent variables: extent of decreasing marginal utility ($\beta$); segment taste for quality ($a$); and the extent of consumer heterogeneity ($a_H-a_L$), technology trajectories ($b_i$); scalability and extent of fixed costs ($f$ and $K$). Our analysis argues for strategy studies to control for these variables.

Further, many of our propositions are potentially testable. Consider, for example, a substitute technology with a superior trajectory displacing an incumbent technology— which segment does it enter first? An implication of Proposition 4.2 is that if the substitute technology has a cost advantage it will start in the low-end segment and move up into segments with greater and greater quality; conversely, if it has a cost disadvantage it will start in the high-end and then move down market. The analysis of the evolution of resource rents sug-
gests hypotheses about how cross-firm profit comparisons depend on the type of resources that firms possess (Proposition 5.1) and about how cross-segment profit comparisons depend on the extent of DMU and the taste for quality in each segment (Proposition 5.2). The analysis of competitive positioning predicts how technology scalability and consumer heterogeneity jointly affect the emergence of different competitive positions over time and can be used to generate hypotheses for both cross sectional and longitudinal industry studies (Proposition 7.4). Collectively, these results raise the broad empirical question of whether and when imitation or substitution poses the greater threat to profits.

Considering the interaction of firm strategy and the demand context suggests new dimensions of firm strategy. In particular, important features of demand are subject to influence by firms. Firms can affect DMU by spurring the developments of complementary products that benefit from higher levels of performance of their own offer. Examples include Hewlett Packard’s decision to sell low priced digital cameras to spur the sales of more advanced printers and inks, and Intel’s strategic venture fund, whose investment mandate is to look beyond direct financial returns and support firms whose offers increase demand for advanced microprocessors. Competitive advantage is always relative to market segmentation. Hence, the issue of how firms segment and potentially de-segment markets seems crucial for understanding the sustainability of competitive advantage. More broadly, firms face a choice about the allocation of resources between supply-side investments that improve performance and investments to improve the demand context.

Although competitive interactions, internal resources and the demand environment have long been viewed as central to business strategy, the impact of demand-side variables has remained under-explored in the received literature. They warrant further exploration.
References


9. Appendix I: Robustness

In the main body of the paper, we considered the simplest possible model of competition: Bertrand competition with price discrimination across segments. Here we relax these assumptions and redo much of the analysis.

9.1. Cournot Competition

In this section we consider a Cournot duopoly with linear demand. Specifically, demand in the focal segment for each firm’s output is given by the inverse demand curve \( p_i = w_i - Q/s \) where \( Q \) is total output of both firms, \( s \) is a parameter of market size and now \( w_i \) is the maximum willingness to pay.\(^ {22} \) As in the base model, let \( c_i \) be the marginal cost of each firm and maintain the definitions \( v_i = w_i - c_i \), \( A_c = c_2 - c_1 \) and \( A_d = w_1 - w_2 \). We restrict attention to \( v_1 > 0 \) so that firm 1 has positive value creation.

In the Bertrand model, firm 1 enters the segment if and only if it has superior value creation, \( v_1 > v_2 \), and if it enters its profits are either \( s v_1 \) or \( s(v_1 - v_2) = s(A_c + A_d) \) depending on whether or not firm 2 creates value (see Proposition 3.3). The analogous results for the Cournot model are as follows.

**Proposition 9.1.** In the Cournot model, firm 1’s profits are

\[
\pi_1(v_1, v_2) = \begin{cases} 
0 & v_1 \leq \frac{1}{2} v_2 \\
\frac{1}{s}(v_1 + A_d + A_c)^2 & \frac{1}{2} v_2 < v_1 < 2v_2 \\
\frac{1}{4} sv_1^2 & 2v_2 \leq v_1
\end{cases}
\]

and its sales are \( q_1 = \sqrt{s \pi_1} \).

**Proof** Firm \( i \) seeks to maximize the profit function \( \pi_i = q_i(p_i - c_i) = q_i(v_i - (q_i + q_j)/s) \). The first order conditions are \( v_i - 2q_i - q_j = 0 \) and there is a unique solution with \( q_1 > 0 \) and \( q_2 > 0 \) given by \( q_i^* = s(2v_i - v_j)/3 \) as long as \( \min v_i > \frac{1}{2} \max v_i \). For \( v_i > 2v_j \) we get \( q_i^* = sv_i/2 \) and \( q_j^* = 0 \). Profits are then given by \( \pi_1(v_1, v_2) = (v_i - q_i^* - q_j^*)q_i^* \). \( \blacksquare \)

\(^{22}\)These demand curves arise, for example, when there are a continuum of consumers in the segment who have a willingness to pay for product \( i \) of \( w_i - r \) where \( w_i \) is a constant across consumers and \( r \) is uniformly distributed between 0 and \( w_i \) with density \( s \) (see Katz and Shapiro, 1985). Thus, in the Cournot model \( w_i \) is the maximum willingness to pay in the segment, whereas in our base model \( w_i \) gives the willingness to pay of all consumers.
There are some key commonalities between the Bertrand and the Cournot case. In particular, a firm’s ability to enter a segment depends critically on relative value creation \((v_1 \text{ versus } v_2)\) and profits vary with both relative value creation \((v_1 - v_2 = A_d + A_e)\) and with absolute value creation \((v_1)\). The main difference is that with Bertrand, each segment is winner-take-all, while with Cournot the firms co-exist in the segment if their value creation is sufficiently similar (i.e., for \(\frac{1}{2}v_2 < v_1 < 2v_2\)). The other main difference is that firm output is increasing in value creation, rather than fixed at \(s\), and hence profits are quadratic, rather than linear, in value creation.

We now turn to the threat of substitution. Recall that with Bertrand competition the threat of substitution depends critically on relative trajectories. With a superior trajectory \((b_1 > b_2)\), the firm using the new technology eventually enters the segment and displaces the incumbent firm with the entry time falling in the cost advantage \(c_2 - c_1\). In addition, whether or not the new firm has a cost advantage determines the direction of the effect of \(a\) and \(\beta\) on the entry timing. With an inferior trajectory \((b_1 < b_2)\), any threat is only temporary.

With Cournot one can distinguish between the time \(t_E\) at which the new firm first enters the segment and starts taking market share from the incumbent, and the time \(t_D\), the time at which the incumbent is displaced from the segment. Beyond this distinction, the results under Cournot are very similar to those under Bertrand:

**Proposition 9.2.** (i) Suppose \(b_1 > \frac{1}{2}b_2\). The new firm always enters the segment from some time \(t_E\) onwards with \(t_E\) increasing in \(c_1\) and decreasing in \(c_2\). The effect on \(t_E\) of DMU and the taste for quality depends on the cost position. Specifically, \(\frac{\partial t_E}{\partial a} > 0\) iff \(c_2 - 2c_1 > 0\) and \(\frac{\partial t_E}{\partial \beta} > 0\) iff \(c_2 - 2c_1 > R_E\) for some critical value \(R_E\). (ii) Suppose \(b_1 < \frac{1}{2}b_2\). The new firm enters the segment only if \(c_2 - 2c_1\) is sufficiently large, with any entry being only temporary.

**Proof** (i) Suppose \(b_1 > \frac{1}{2}b_2\). From Proposition 9.1, the new firm enters the segment iff \(v_1(t) > \frac{1}{2}v_2(t)\). Because \(v_1'(t) > \frac{1}{2}v_2'(t)\) for \(b_1 \geq \frac{1}{2}b_2\), there exists an unique value \(t_E\) such that \(v_1(t_E) = \frac{1}{2}v_2(t_E)\) and the new firm enters the segment for all \(t > t_E\). Using the implicit function theorem on the defining equality \(2a(b_1t)^{\beta} - a(b_2(t + h))^{\beta} - 2c_1 + c_2 = 0\), we have

\[
\frac{\partial t_E}{\partial c_1} = \frac{2}{2v_1'(t_E) - v_2'(t_E)} > 0, \quad (9.1)
\]

\[
\frac{\partial t_E}{\partial c_2} = -\frac{1}{2v_1'(t_E) - v_2'(t_E)} < 0, \quad (9.2)
\]
\[
\frac{\partial t_E}{\partial a} = -\frac{2(b_1 t_E)^\beta - (b_2 (t_E + h))^\beta}{2v'_1(t) - v'_2(t)} = \frac{(c_2 - 2c_1)/a}{2v'_1(t) - v'_2(t)}, \quad (9.3)
\]
\[
\frac{\partial t_E}{\partial \beta} = -\frac{2a(b_1 t_E)^\beta \ln(b_1 t_E) - (2a(b_1 t_E)^\beta + c_2 - 2c_1) \ln(b_2 (t_E + h))}{2v'_1(t) - v'_2(t)}. \quad (9.4)
\]

We have that \(\partial t_E/\partial a\) has the same sign as \(c_2 - 2c_1\). It is straightforward to show that \(\partial t_E/\partial \beta\) is increasing in \(c_2 - 2c_1\) and positive iff \(c_2 - 2c_1\) is greater than some cutoff \(R_E\). (ii) The proof follows closely the logic used in the proof of Proposition 4.3 and is omitted. ■

**Proposition 9.3.** (i) Suppose \(b_1 > 2b_2\). The new firm always displaces the established firm from the segment from some time \(t_D > t_E\) onwards with \(t_D\) increasing in \(c_1\) and decreasing in \(c_2\). The effect on \(t_D\) of DMU and the taste for quality depends on the cost position. Specifically, \(\partial t_E/\partial a > 0\) iff \(2c_2 - c_1 > 0\) and \(\partial t_E/\partial \beta > 0\) iff \(2c_2 - c_1 > R_D\) for some critical value \(R_D\). (ii) Suppose \(b_1 < 2b_2\). The established firm is displaced from the segment only if \(2c_2 - c_1\) is sufficiently large, with any displacement being only temporary.

**Proof** (i) Suppose \(b_1 > 2b_2\). From Proposition 9.1, the established firm exists the segment iff \(v_1(t) > 2v_2(t)\). Because \(v'_1(t) > 2v'_2(t)\) for \(b_1 \geq 2b_2\), there exists an unique value \(t_D\) such that \(v_1(t_D) = 2v_2(t_D)\) and the established firm exits the segment for all \(t > t_D\). Using \(a(b_1 t_E)^\beta - 2a(b_2 (t_E + h))^\beta - c_1 + 2c_2 = 0\) and the implicit function theorem we have

\[
\frac{\partial t_E}{\partial c_1} = \frac{1}{v'_1(t_E) - 2v'_2(t_E)} > 0, \quad (9.5)
\]
\[
\frac{\partial t_E}{\partial c_2} = -\frac{2}{v'_1(t_E) - 2v'_2(t_E)} < 0, \quad (9.6)
\]
\[
\frac{\partial t_E}{\partial a} = -\frac{(b_1 t_E)^\beta - 2(b_2 (t_E + h))^\beta}{v'_1(t_E) - 2v'_2(t_E)} = \frac{(2c_2 - c_1)/a}{v'_1(t_E) - 2v'_2(t_E)}, \quad (9.7)
\]
\[
\frac{\partial t_E}{\partial \beta} = -\frac{a(b_1 t_E)^\beta \ln(b_1 t_E) - (a(b_1 t_E)^\beta + c_2 - 2c_1) \ln(b_2 (t_E + h))}{v'_1(t_E) - 2v'_2(t_E)}. \quad (9.8)
\]

We have that \(\partial t_E/\partial a\) has the same sign as \(2c_2 - c_1\). It is straightforward to show that \(\partial t_E/\partial \beta\) is increasing in \(2c_2 - c_1\) and positive iff \(2c_2 - c_1\) greater than some cutoff \(R_D\). (ii) The proof follows closely the logic used in the proof of Proposition 4.3 and is omitted. ■

We now turn to the evolution of resource rents. Under Bertrand competition firm 2, which is assumed to have no resources, earns no rents and firm 1 with a resource earns \(sv_1\) for \(v_2 < 0\) and \(s(v_1 - v_2) = s(A_c + A_d)\) when \(v_2 > 0\). Under Cournot competition, even without any resources, firms earn some rents in a duopoly and hence resource rents need
to be taken net of these duopoly profits. Moreover, because the Cournot model allows for market growth while the Bertrand model does not, comparison is best done in terms of rents divided by output.

**Lemma 9.4.** (i) The rents earned by firm 1 normalized by its output are

\[ R_s = \frac{\pi_1(v_1, v_2) - \pi_1(v_2, v_2)}{q_1} \]

\[ = \begin{cases} 
\frac{1}{2} v_1 & \text{if } v_2 < 0 < v_1, \\
\frac{1}{2} v_1 - \frac{2 v_2^2}{9 v_1} & \text{if } 0 < 2 v_2 < v_1, \\
\frac{4}{3} \frac{A_c + A_d}{2 - v_2/v_1} & \text{if } 0 < v_1 < 2 v_2. 
\end{cases} \]

(ii) For process resources, \( R_s \) increases over time and converges to \( A_c \). (iii) For performance and timing resources, \( R_s \) at first increases over time, but ultimately converges to 0. (iv) For innovation resources, \( R_s \) is increasing over time.

**Proof** (i) For all four resource types, \( v_1(t) > v_2(t) \) for all \( t > 0 \) and hence firm 1 enters the segment first and is never displaced. Without its resource, firm 1’s value creation is the same as firm 2’s and hence its profits are \( \pi_1(v_2, v_2) = \frac{1}{3} s(v_2)^2 \) for \( v_2 > 0 \) and 0 otherwise; see Proposition 9.1. The rents from firm 1’s resource are then

\[ R = \pi_1(v_1, v_2) - \pi_1(v_2, v_2) \]

\[ = \begin{cases} 
0 & \text{if } v_1 < 0, \\
\frac{1}{3} s v_1^2 & \text{if } v_2 < 0 < v_1, \\
s \left( \frac{1}{4} v_1^2 - \frac{1}{2} v_2^2 \right) & \text{if } 0 < 2 v_2 < v_1, \\
\frac{4}{3} s(A_c + A_d)v_1 & \text{if } 0 < v_1 < 2 v_2. 
\end{cases} \]

For \( 2v_2 < v_1 \), we have that \( q_1 = sv_1/2 \) and for \( 2v_2 > v_1 \) we have \( q_1 = s(2v_1 - v_2)/3 \). Dividing each region of \( R \) by the appropriate value of \( q_1 \) gives the result.

(ii) With a process resource we have \( \frac{\partial}{\partial t} \left( \frac{1}{2} v_1 \right) = \frac{1}{2} b \beta / t^{1-\beta} > 0, \) \( \frac{\partial}{\partial t} \left( \frac{1}{2} v_1 - \frac{2 v_2^2}{9 v_1} \right) = \frac{1}{4} \left( \frac{v_2 - v_2^2}{2 - v_2 / v_1} \right)^2 \) and hence, \( \frac{\partial R_s}{\partial t} > 0 \) for \( v_1 > 0 \). Finally, since \( \lim_{t \to \infty} \frac{v_2}{v_1} = 1 \), we have \( \lim_{t \to \infty} R_s = \frac{4}{3} A_c. \)

(iii) With a timing or product resource we have \( \frac{\partial}{\partial t} \left( \frac{1}{2} v_1 \right) = \frac{1}{2} b \beta / (t + h)^{1-\beta} > 0 \) and hence \( R_s \) is initially increasing. Since \( \lim_{t \to \infty} \frac{v_2}{v_1} = 1 \) and \( \lim_{t \to \infty} (A_c + A_d) = 0 \), we have that \( R_s \)
converges to 0.

(iv) With an innovation resource we have 
\[ \frac{\partial}{\partial t} \left( \frac{1}{2}v_1 \right) = \frac{1}{2} b_1 \beta / t^{1-\beta} > 0 \] and 
\[ \frac{\partial}{\partial t} \left( \frac{1}{2}v_1 - \frac{v_2}{v_1} \right) = \frac{\beta}{t^{1-\beta}} \left( \frac{b_1}{2} + \frac{v_2}{v_1} (b_1 - \frac{4}{9} b_2) \right) > 0. \]
Since 
\[ \frac{\partial}{\partial t} (v_1 - v_2) > 0 \] and 
\[ \frac{\partial}{\partial t} (v_1 \cdot v_2) = (b_1 - b_2) c \beta t^{1-\beta} > 0, \]
\[ \frac{\partial}{\partial t} \left( \frac{4}{3} \left( \frac{v_1 - v_2}{2v_2/v_1} \right) \right) > 0. \]
Thus, \[ \frac{\partial R_1}{\partial t} > 0 \] for \( v_1 > 0. \)

The expressions characterizing per output rents given in part (i) of the proposition are sufficiently similar to the expressions under Bertrand that the evolution of per output rents follows the same path: process resources increase up to a constant determined by \( A_c; \) timing and performance rents first increase and then decay to zero; innovation rents increase.

Turning to total rents we have:

**Proposition 9.5.** (i) For all resource types, the net present value of the resource rent stream is decreasing in the extent of DMU and increasing in the taste for quality. (ii) With no upperbound on total sales, all rent streams increase over time without bound except for timing and product resources when \( \beta \leq 1/2 \) in which case rents converge to a constant and the constant is 0 for \( \beta < 1/2. \)

**Proof** (i) Follows easily from the expression for \( R. \) (ii) For cost and innovation resources it is straightforward to show that \( q_1 \) is increasing over time and grows without bound. Since rents per unit sale are increasing for these resources, total rents increase without bound. For timing and product resources, we have \( v_1 < 2v_2 \) for \( t \) sufficiently large and thus we are interested in the limit \( \lim_{t \to \infty} ((t + h)^\beta - t^\beta) / ((t + h)^\beta - c), \) which is 0 for \( \beta < 1/2, \) \( h/2 \) for \( \beta = 1/2 \) and infinity for \( \beta > 1/2. \)

Part (i) is the same as in Bertrand. In part (ii) we see that the assumption that sales grow over time with value creation introduces a strong force for increasing rents over time. For example, as the market grows a firm with a process resource can leverage that advantage over a larger and larger number of consumers. Introducing limits to market size would reverse this result.

Analysis of positioning is more complicated with Cournot than Bertrand because of the possibility that multiple firms can co-exist in a given segment and is beyond the scope of this paper. Nonetheless, we expect that the viability of the generalist strategy will show a similar dependence on fixed costs and consumer heterogeneity and hence that similar dynamics will occur.
9.2. Price Discrimination

We now consider sustainability in settings where firms cannot price discriminate across segments. We find that rivalry is reduced and that competitive outcomes in one segment depend in part on value creation in other segments. As before, we suppose that there is Bertrand competition, a focal market segment of size $s$, and that firm $i$'s value creation is given by $v_i = w_i - c_i > 0$. We now add a second (niche) segment of size $s_N$ and we denote the value creation of firm $i$ in this segment by $v_{Ni} = w_{Ni} - c_i$. Firm $i$ charges a single price $p_i$ to both segments. We simplify the analysis by supposing that $v_{N2} \geq v_2$ and $v_{N1} = 0$, which assures that firm 2 is present in the secondary segment and faces the decision as to whether to expand its scope to include the focal segment as well.

With price discrimination, the segments can be analyzed separately and firm 1 has profits from the focal segment if and only if $v_1 > v_2$ (see Proposition 3.3). In contrast, without price discrimination we have:

Lemma 9.6. (i) For $v_2 \leq \frac{s_N}{s+s_N} v_{N2}$, the unique pure strategy equilibrium is for firm 2 to stay out of the focal segment. (ii) For $v_2 > v_1 + \frac{s_N}{s+s_N} v_{N2}$, the unique pure strategy equilibrium is for firm 2 to enter the focal segment. (iii) Firm 1 is weakly better off without price discrimination in that its profits are either higher or the same as when there is price discrimination.

Proof There are two possible pure strategy equilibria: firm 2 serves only the secondary segment and firm 2 serves both segments. Consider first the equilibrium where firm two serves only the secondary segment, in which case its price is $p_2 = w_{N2}$ and its profits are $\pi_2 = s_N v_{N2}$ while firm 1 charges a price $p_1 = w_1$. Firm 2’s optimal deviation is to charge a price just below $w_2$ and its profits are arbitrarily close to $(s + s_N)v_2$. The deviation is unprofitable if and only if $v_2 \leq \frac{s_N}{s+s_N} v_{N2}$, which is a necessary and sufficient condition for the existence of the equilibrium.

Now consider the equilibrium where firm 2 serves both segments. Then $p_1 = c_1$ and $p_2 = w_2 - w_1 + c_1$. Firm 2’s profits are $\pi_2 = (s + s_N)(v_2 - v_1)$. Firm 2’s optimal deviation is to price at $p_2 = w_{N2}$ which results in profits $s_N v_{N1}$ and the deviation is unprofitable if and only if $v_2 > v_1 + \frac{s_N}{s+s_N} v_{N2}$, which is a necessary and sufficient condition for the existence of the equilibrium. Since the two conditions are disjoint, the pure strategy equilibria are unique, when they exist. This completes the proof of parts (i) and (ii).
Part (iii): If \( v_2 \geq v_1 + \frac{s_N}{s+N}v_{N2} \) then firm 1 has zero profits with or without price discrimination. If \( v_2 \leq \frac{s_N}{s+N}v_{N2} \), then firm 1’s profits without price discrimination are \( sv_1 \) which is greater than \( s(v_1-v_2) \) its profits with price discrimination. If \( v_1 < v_2 < v_1 + \frac{s_N}{s+N}v_{N2} \), then profits with price discrimination are zero, while profits are at least zero without price discrimination. Finally, there may be that \( \frac{s_N}{s+N}v_{N2} < v_2 < v_1 \), in which case there is a mixed strategy equilibrium without price discrimination. Since firm 2 never prices below \( c_2 \) in any equilibrium, a price of \( p_1 = w_1 - w_2 + c_2 \) assures firm 1 a profit of at least \( s(v_1-v_2) \), which is its profit with price discrimination. We conclude that firm 2’s profits are weakly greater without price discrimination.

Without the ability to price discriminate, firm 2 faces a trade-off. It can stay in the secondary segment and earn high margins or it can lower its prices to expand its volume by entering the focal segment. The larger is the size \( s_N \) of the secondary segment and the greater its value creation \( v_{N2} \) there, the more likely that firm 2 stays out of the focal segment.\(^{23}\) Note that it is now possible for a firm to earn profits even if it does not have a competitive advantage in any segment, but only if its competitor is focused on exploiting other opportunities and it has positive value creation in the segment. Consequently, in part (iii) we see that firm 1 is potentially better off, and certainly no worse off, without price discrimination.

We now consider the implications of Lemma 9.6 for our analysis of sustainable competitive advantage in settings with multiple segments and no price discrimination. The threat posed by a substitute technology is reduced by an inability to price discriminate because firms using it are less eager to enter new segments because doing so involves sacrifice of margins in existing segments. Similarly, we expect that removing the possibility of price discrimination raises resource rents because firms have less incentive to attack each others’ segments. Although the effect on positioning is more complex, we expect that the threat posed by Generalists is reduced because they benefit from the opportunity to price discriminate across their multiple segments while specialist firms do not.

Adner and Zemsky (2003) study substitution threats in a model that incorporates Cournot competition and no price discrimination simultaneously. No new issues arise from making both assumptions simultaneously.

\(^{23}\)For \( \frac{s_N}{s+N}v_{N1} < v_2 < v_1 + \frac{s_N}{s+N}v_{N1} \), a pure strategy equilibrium does not exist and instead there is a mixed strategy equilibrium where firm 1 serves the focal segment with some probability between 0 and 1.
10. Appendix II: Proofs

Proof of Proposition 4.1 Suppose \( b_1 = b_2 \). We have \( A_d(t) = ab_1^\beta(t^\beta - (t + h)^\beta) \) and hence \( A_d(t) < 0 \) and \( A_d'(t) > 0 \) for all \( t > 0 \) and \( \lim_{t \to \infty} A_d(t) = 0 \). If \( A_c \leq 0 \), then \( A_d(t) + A_c < 0 \) for all \( t > 0 \) and entry does not occur. Conversely, if \( A_c > 0 \), then there exists a \( t_E \) such that \( A_d(t) + A_c > 0 \) iff \( t > t_E \) where \( t_E \) satisfies \( a(b_1 t_E)^\beta - a(b_2 (t_E + h))^\beta + A_c = 0 \). Using this equality and the implicit function theorem,

\[
\frac{\partial t_E}{\partial a} = -\frac{(b_1 t_E)^\beta - (b_2 (t_E + h))^\beta}{A_d'(t)} = \frac{A_c/a}{A_d'(t)},
\]

which is positive since \( A_c > 0 \). Similarly, we have

\[
\frac{\partial t_E}{\partial \beta} = -\frac{a(b_1 t_E)^\beta \ln(b_1 t_E) - (a(b_1 t_E)^\beta + A_c \ln(b_2 (t_E + h)))}{A_d'(t)},
\]

which is positive since \( A_c > 0 \) and \( \ln(b_2 (t_E + h)) > \ln(b_1 t_E) > 0 \) as \( b_1 t_E > b_1 t_0 > 1 \).

Proof of Proposition 4.2 Suppose \( b_1 > b_2 \). Since \( A_d(0) + A_c < 0 \), \( A_d'(t) > 0 \) and \( \lim_{t \to \infty} A_d(t) = \infty \), there exists a \( t_E > t_0 \) such that \( A_d(t) + A_c > 0 \) iff \( t > t_E \) where \( t_E \) satisfies \( a(b_1 t_E)^\beta - a(b_2 (t_E + h))^\beta + A_c = 0 \). Thus, \( \frac{\partial t_E}{\partial a} \) is given by (10.1), which is positive iff \( A_c > 0 \). Finally, \( \frac{\partial t_E}{\partial \beta} \) is given by (10.2), which is positive if \( A_c > 0 \) since in this case \( b_1 t_E < b_2 (t_E + h) \), while conversely \( \frac{\partial t_E}{\partial \beta} < 0 \) if \( A_c < 0 \) since in this case \( b_1 t_E > b_2 (t_E + h) \).

Proof of Proposition 4.3 Suppose \( b_1 < b_2 \). Then \( A_d(t) < 0 \) for all \( t \geq 0 \) and there exists a \( t^* \) such that \( A_d'(t) > 0 \) iff \( t < t^* \). If \( A_c < -A_d(t^*) \), firm 1 never enters the segment. If \( A_c \geq -A_d(t^*) \) then firm 1 enters the segment at some \( t_1 \in (0, t^*) \). As \( \lim_{t \to \infty} A_d(t) = -\infty \), entry is reversed at some time \( t_2 > t^* \). We have that \( \frac{\partial t_1}{\partial a} \) and \( \frac{\partial t_1}{\partial \beta} \) are given by (10.1) and (10.2) with \( t_E \) replaced by \( t_j \) for \( j = 1, 2 \). The numerator of both equalities are positive since \( A_c > 0 \). Since \( A_d'(t_1) > 0 \), \( \frac{\partial t_1}{\partial a} > 0 \) and \( \frac{\partial t_1}{\partial \beta} > 0 \). Since \( A_d'(t_2) < 0 \), \( \frac{\partial t_1}{\partial a} < 0 \) and \( \frac{\partial t_1}{\partial \beta} < 0 \). Hence, the interval of entry \([t_1, t_2]\) is falling in \( a \) and \( \beta \).

Proof of Proposition 5.1 For all four resource types, \( A_d(t) + A_c > 0 \) for all \( t > 0 \) and hence firm 1 enters the segment first and is never displaced. We have that \( t_i \) is defined by \( v_i(t_i) = 0 \). From Proposition 3.3, we have that firm 1’s rent (profit) from its resource is

\[
\pi_1(t) = \begin{cases} 
0 & \text{if } t \leq t_1, \\
s v_1(t) & \text{if } t \in (t_1, t_2], \\
s(v_1(t) - v_2(t)) & \text{otherwise.}
\end{cases}
\]
For all resource types, \( v'_1(t) > 0 \) as performance increases in \( t \) and hence rents are always increasing for \( t \in (t_1, t_2) \). For a process resource, \( v_1(t) - v_2(t) = A_c \) such that rents are constant for \( t > t_2 \). For all other resources, \( v_1(t) - v_2(t) = \lambda_d(t) \), which varies with time. For performance and timing resources, \( \lambda'_d(t) > 0 \) and rents decline for \( t > t_2 \). For innovation resources, \( \lambda''_d(t) > 0 \) and rents increase over time but at a decreasing rate since \( \lambda''_d(t) < 0 \).

**Proof of Proposition 5.2** We have that \( v_1(t) \) is increasing and \( v_1(t) - v_2(t) \) is nondecreasing in \( a \). Hence, we have that \( \pi_1(t) \) is nondecreasing in \( a \) for all \( t \) and increasing for some \( t \), which implies that the net present value of resource rents is increasing in \( a \). An analogous argument holds for \( \beta \).

**Proof of Proposition 7.1** (i) Denote by \( d_E \) the position of the entrant. For all \( d_E > d_l \) we have that \( v_L(d_E) < v_L(d_l) \) and (by Proposition 3.3) the entrant does not serve the low-end segment. Hence, the optimal positioning for \( d_E > d_l \) is the one which maximizes value creation and rents from the high-end segment, which is \( d_E = d_E^* \). Similarly, if \( d_E < d_l \) the entrant does not have superior value creation in the high-end segment and hence the optimal positioning is \( d_E = d_L^* \). (ii) The profit from being a Differentiator is \( \pi_H = s(v_H(d_H^*) - v_H(d_l)) \) and the profit from being a Cost Leader is \( \pi_L = s(v_L(d_L^*) - v_L(d_l)) \). The relative attractiveness of being a Cost Leader is then \( \pi_L - \pi_H \). The results then follow from evaluating the partial derivatives (e.g., \( \partial \pi_L / \partial d_l > 0 \) or \( \partial \pi_H / \partial d_l \)).

**Proof of Proposition 7.3** Given Corollary 7.2, the entrant must use technology \( F \). It cannot be that the entrant serves only one of the segments since its costs would be greater than those of the incumbent for any level of \( d \): \( s(c + df) + fdK > s(c + d) \) for \( K > s \).

Moreover, given the scale economies in technology \( F \), either the potential entrant serves all customers (in both segments), or it stays out of the market altogether. We proceed by assuming that the entrant is serving both segments using technology \( F \) and check whether or not this is profitable.

Incumbents reduce their prices to marginal cost in an effort to fight off entry. Following the logic used to derive Proposition 3.3, the profits of the entrant for any given \( d \) are

\[
\pi_E(d) = s(v_H^F(d) + v_L^F(d) - \max\{v_H^*, 0\} - \max\{v_L^*, 0\}),
\]

where \( v_H^* = v_H(d_H^*) \) is the value creation of the Differentiator in the high-end and \( v_L^* \) is the value created by the Cost Leader in the low-end and where \( v_H^F(d) = a_0(bd)^\theta - [c + df + dfK/(2s)] \) for \( \theta = H, L \) is the value created by the entrant when it serves a customer in
segment θ. The level of differentiation which maximizes the entrant’s profits is then

\[ d^*_E(f) = \left[ \frac{a_H + a_L}{2 - f(2 - K/s)} \right]^{\frac{1}{1-\beta}}. \]

Let \( v^*_E(f) = v^*_H(d^*_E(f)) + v^*_L(d^*_E(f)) \) be the entrants maximum possible value creation for one customer from each segment. We have

\[ v^*_E(f) = \gamma \left( \frac{a_H + a_L}{(2 - 2f + fk)^{\beta}} \right)^{\frac{1}{\beta}} - 2c, \]
\[ v^*_H = \gamma (a_H)^{\frac{1}{1-\beta}} - c, \]
\[ v^*_L = \gamma (a_L)^{\frac{1}{1-\beta}} - c, \]

where \( \gamma = (1 - \beta)(b\beta)^{\frac{\beta}{1-\beta}}. \) Note that \( d^*_E/f > 0 \) and \( \partial v^*_E/f > 0. \)

Given that there are incumbents at \( d^*_L \) and \( d^*_H \) we assume that \( v^*_L, v^*_H \geq 0. \) Then \( \pi_E(d^*_E) = s(v^*_E(f) - v^*_L - v^*_H). \) Let \( f_1 \) be such that \( v^*_E(f_1) = v^*_L + v^*_H, \) which yields

\[ f_1 = \left( 1 - \frac{1}{2} \left( \frac{a_L + a_H}{(a_L)^{1/(1-\beta)} + (a_H)^{1/(1-\beta)}(2 - K/s)^{1-\beta}} \right)^{\frac{1}{1-\beta}} \right) / \left( 1 - \frac{K}{2s} \right). \]

Let \( f_2 \) be such that \( d^*_E(f_2) = d^*_H, \) which yields

\[ f_2 = \frac{1}{2} \left( \frac{a_H - a_L}{a_H} \right) / \left( 1 - \frac{K}{2s} \right). \]

It follows that \( f_1 < f_2 \) and that there exists a \( \tilde{K} \in (s, 2s) \) such that \( f_1 < 1 \) iff \( K < \tilde{K}. \) For \( K > \tilde{K}, f_1 > 1 \) and a Generalist is never viable. ■

**Proof of Proposition 7.4** This proof builds closely on the arguments and definitions in the proof of Proposition 7.3. The strategies that exist in the market at any point in time are those that have positive and superior value creation. Recall that \( v^*_E = \gamma (a_{\theta})^{\frac{1}{1-\beta}} - c \) for \( \theta = H, L \) where \( \gamma = (1 - \beta)(b\beta)^{\frac{\beta}{1-\beta}}. \) Hence, \( v^*_H > v^*_L \) and both are increasing over time with \( b(t) \) from an initial value of \( v^*_L = v^*_H = -c. \) Let \( t_H \) be the critical time at which \( v^*_H = 0 \) and a Differentiator becomes willing to enter the market. At this time, \( \gamma = c/(a_{\theta})^{\frac{1}{1-\beta}} \) and hence
\(v_E^*(f) > 0\) is equivalent to

\[
f > f_3 = \left( 1 - \left( \frac{a_L + a_H}{2a_H} \right)^{1/\beta} \right) \left( 1 - \frac{K}{2s} \right)
\]

where \(f_3 > f_2\). Thus, for \(f > f_3\) the market is pioneered by a Generalist, otherwise by a Differentiator.

We have that \(v_E^*(f) > v_L^* + v_H^*\) is equivalent to \(f > f_1\) where \(f_1\) is independent of \(t\). For \(f < f_1\), the Generalist never enters and the Differentiator is joined by a Cost Leader. For \(f > f_1\), the Differentiator is displaced by the Generalist before the Cost Leader would have entered. ■

**Proof of Proposition 7.5** The comparative statics follow from the expressions for \(f_1\), \(f_2\) and \(f_3\) in Propositions 7.3 and 7.4. ■