Strategic innovation and technology adoption in an evolving industry

Darren Filson*, Richard T. Gretz

Department of Economics, Claremont Graduate University, 160 E. Tenth Street, Claremont, CA 91711, USA

Received 25 April 2003; received in revised form 29 July 2003; accepted 29 July 2003

Abstract

We introduce a racing model with multiple product generations, product innovation, spin-outs, and licensing. Industry conditions and innovation characteristics affect who wins the race and who markets the resulting product. Small firms market their innovations when they pioneer a new generation or improve quality in a young generation and license their innovations in mature generations. If old generation leaders ever market improvements in young generation goods, they do so early on. Leadership in mature generations persists. Tests on the rigid disk drive industry (1977–1997) provide empirical support. The results have implications for antitrust policies and policies governing employee non-compete agreements.

JEL classification: K31; L41; L63; O31; O38

Keywords: Industry dynamics; Employee mobility; Innovation market; Covenant not to compete; Spin-offs

1. Introduction

Effective policy analysis requires an effective positive model of behavior. In this paper, we consider some important facts about innovation that have been underemphasized in previous models. We introduce a model that considers these facts and...
test the model on the rigid disk drive industry during the period 1977–1997. Our results lead us to criticize certain features of antitrust policies and policies governing employee non-compete agreements.

Three facts have been under-emphasized in models of innovation races (Reinganum, 1989; Gans and Stern, 2000) and industry evolution (Jovanovic and MacDonald, 1994; Klepper, 1996; Filson, 2001, 2002; Franco and Filson, 2002): First, small firms rarely grow by attempting to compete head-to-head with industry leaders in a given product market. Instead, small firms grow by differentiating themselves from current industry leaders. For example, small software firms do not pursue competition with Microsoft in personal computer operating systems, word processors, or spreadsheets. Instead, they produce different products in an effort to become leaders in the market for their good.

Second, firms with the “best” technologies are not always the most profitable firms or the ones with the largest market shares. In many formal models of innovation and technology adoption, firm size and firm profits are monotonic functions of a summary technological “know-how” variable. In contrast, real-world large firms often appear to have relatively mediocre technology compared to their smaller competitors. Marketing and connections with important buyers often appear to trump technological know-how.

Third, many resources are mobile. This implies that small high-tech firms can be important even when they do not grow large. Models where employees leave existing firms to create new “spin-out” firms have been developed and tested by Franco and Filson (2002) and Klepper and Sleeper (2002). Acquisitions and licensing are also important in many industries (Salant, 1984; Gans and Stern, 2000). Large firms may acquire innovative small firms or license from them, as in the biotechnology industry.

We adapt the standard single-prize racing model (Reinganum, 1989) to allow for product differentiation, spin-out formation, and licensing. Doing so yields a richer description of initial industry conditions, outcomes, and policy impacts than previous models. Innovation involves new products instead of the cost reductions that are typically analyzed in racing models. This facilitates our empirical analysis, which focuses on product innovation. We consider two types of innovation: quality improvements and new product generations. Distinguishing between the two is useful in many industries. For example, in the computer industry, mainframes, minicomputers, desktops, laptops, notebooks, and hand-held devices are all product generations. New generations have different impacts on existing goods than quality improvements. The richness in our model involves some tradeoffs. For example, the model considers only one innovation at a time, and players do not look beyond the current race. Insights for industries with a sequence of innovations are obtained by considering how initial conditions at the beginning of each race change as the industry evolves.

\(^1\) Some attempts to formally explain different innovation strategies exist, such as Nelson (1988) and Eeckhout and Jovanovic (2002). However, the distinction between strategies is typically quite simple, such as that between “innovators” and “followers.”
Analytical results and numerical computations suggest several intuitive testable hypotheses, and tests on the rigid disk drive industry provide empirical support. The results clarify the role of small firms in evolving industries. New firms and industry laggards market their innovations only when they pioneer a new product generation or improve quality early on in a young generation. These firms license or sell their innovations to industry leaders when they improve quality in mature generations. This result is driven by changes in the business stealing effects of innovations as a product generation evolves. When a new generation is first introduced, demand is typically low and the generation poses little threat to existing firms. In mature generations, however, a competitor’s introduction of a good that improves even incrementally on a leader’s existing goods can have devastating effects on the leader. To prevent this, leaders buy the innovation from the innovator.

The results clarify how market leadership evolves in new product generations. In general, new generations pass through at most three stages. In the first stage, an entrant or an industry laggard pioneers the generation. In the second stage, the leader in the old generation either wins the race to improve quality in the new generation or licenses the quality improvement from the innovator. In the third stage, the leader in the new generation either wins the race or licenses the quality improvement. If the business stealing effect of the new generation on the old is high when it is first introduced, then the first stage is skipped. If the business stealing effect of the new generation is low early on but the new generation diffuses rapidly thereafter because of rapid exogenous growth in demand or rapid quality improvements not due to race-winning, then the second stage is skipped. New firms and laggards are most likely to grow large when they pioneer a new product generation and then experience the favorable shocks necessary to prevent old generation firms from taking over.

Our results have implications for antitrust policies. Antitrust policy in the U.S. favors competition between many small competitors. High product market concentration and price markups are causes for concern. The impacts of market structure and firm behavior on innovation have been considered in more cases since the introduction of the 1995 Department of Justice/Federal Trade Commission Antitrust Guidelines for the Licensing of Intellectual Property, but this consideration appears to have been one sided. Gilbert and Tom (2001) present evidence that the consideration of innovation effects has not affected the outcomes of most cases, but when it has affected outcomes it has led to challenges in more markets and broader remedies. In contrast, our results suggest that innovation effects should often be mitigating factors that cause the agencies to permit greater product market concentration and higher price markups. Persistent concentration and high markups in an environment with licensing can allow all innovators (including small start-ups) to appropriate a greater amount of the social benefits generated through innovation. This may lead to more rapid innovation, which benefits consumers.

---

2 Quality improvements not due to race-winning would include improvements through learning by doing and improvements in components produced by suppliers.
Further, our results suggest that analyses of entry barriers and potential competition in innovation markets should consider barriers to employee mobility and spin-outs. Thus, our paper adds to a recent literature on policies governing employee non-compete agreements in high-tech industries (Gilson, 1999; Cooper, 2001). Most state laws enforce non-compete agreements; California law does not. Our analysis suggests that non-compete agreements are socially harmful because they discourage the emergence of small start-ups that can compete in innovation races and market goods in new product generations.

1.1. Innovation and new product generations in the rigid disk drive industry

IBM introduced the first rigid disk drive in 1956. The first drives, 14'' in diameter, were either sold in mainframe computers or sold directly to computer users. When the minicomputer market emerged in the mid-1970s an original equipment market developed, and disk drive manufacturers began selling drives to computer manufacturers.

Our analysis covers the period 1977–1997. Innovation occurred rapidly during this period and took three main forms. First, several new product generations were introduced in the form of smaller diameter drives. When first introduced, the new drives served new customers; 8'', 5.25'', 3.5'', 2.5'', and 1.8'' drives were first used in minicomputers, desktops, laptops, notebooks, and hand-held devices, respectively. Second, several improvements in technical features improved storage capacities and access times. Third, several improvements in design and manufacturing techniques improved costs and reliability.

We focus on the first two forms of innovation (new diameters and improvements in storage capacities within a diameter) because our data is best-suited to address these two. Tables 1–5 examine the top ten storage capacity leaders in each diameter

<table>
<thead>
<tr>
<th>Number of:</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>79</td>
</tr>
<tr>
<td>Leaders</td>
<td>6</td>
</tr>
<tr>
<td>Leaders who are old generation large manufacturer/marketers</td>
<td>1</td>
</tr>
<tr>
<td>Leaders who are new generation large manufacturer/marketers</td>
<td>0</td>
</tr>
<tr>
<td>Leaders who are recent entrants</td>
<td>4</td>
</tr>
</tbody>
</table>
Each firm’s highest capacity drive in each diameter each year is used to determine the leaders (data sources are discussed in Section 3.1). Within the group of leaders, we focus on three subgroups: old generation manufacturer-marketers are firms whose sales exceeded $50 million 1983 dollars for at least 3 years at some point during their life and achieved $50 million 1983 dollars for at least 1 year before the firm introduced drives in the diameter; recent entrants are firms who entered in the past 3 years and have not yet achieved $50 million 1983 dollars in sales; new generation manufacturer-marketers are firms whose sales exceeded $50 million 1983 dollars for at least 3 years only after the firm introduced drives in the diameter.

In general, the new diameters are pioneered by recent entrants and small firms and then two patterns emerge. First, some old generation market share leaders become

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Storage capacity leaders in the 5.25&quot; diameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of:</td>
<td>Year</td>
</tr>
<tr>
<td></td>
<td>80</td>
</tr>
<tr>
<td>Leaders</td>
<td>3</td>
</tr>
<tr>
<td>Leaders who are old generation large manufacturer/marketers</td>
<td>0</td>
</tr>
<tr>
<td>Leaders who are new generation large manufacturer/marketers</td>
<td>0</td>
</tr>
<tr>
<td>Leaders who are recent entrants</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Storage capacity leaders in the 3.5&quot; diameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of:</td>
<td>Year</td>
</tr>
<tr>
<td></td>
<td>83</td>
</tr>
<tr>
<td>Leaders</td>
<td>3</td>
</tr>
<tr>
<td>Leaders who are old generation large manufacturer/marketers</td>
<td>1</td>
</tr>
<tr>
<td>Leaders who are new generation large manufacturer/marketers</td>
<td>0</td>
</tr>
<tr>
<td>Leaders who are recent entrants</td>
<td>2</td>
</tr>
</tbody>
</table>

(or fewer if less than ten make the diameter). Each firm’s highest capacity drive in each diameter each year is used to determine the leaders (data sources are discussed in Section 3.1). Within the group of leaders, we focus on three subgroups: old generation manufacturer-marketers are firms whose sales exceeded $50 million 1983 dollars for at least 3 years at some point during their life and achieved $50 million 1983 dollars for at least 1 year before the firm introduced drives in the diameter; recent entrants are firms who entered in the past 3 years and have not yet achieved $50 million 1983 dollars in sales; new generation manufacturer-marketers are firms whose sales exceeded $50 million 1983 dollars for at least 3 years only after the firm introduced drives in the diameter.

In general, the new diameters are pioneered by recent entrants and small firms and then two patterns emerge. First, some old generation market share leaders become
storage capacity leaders in the new diameter. Second, some of the early storage capacity leaders grow to become large manufacturer marketers. The relative importance of each of these two patterns varies by diameter, but in either case as the new product generation matures the list of storage capacity leaders and market share leaders becomes more similar. However, some recent entrants still make the list of storage capacity leaders even as the diameter matures.

Of course, generating product market revenue is not essential for generating value in high-tech industries. Many small firms profit by licensing their technology instead. For a sample of publicly traded U.S. specialized disk drive manufacturers, Table 6 provides simple OLS regressions of the natural log of market capitalization on market share and a normalized measure of storage capacity described in Section 3.1. The results suggest that market share leadership and technological leadership have independent effects on firm value.

Why do small firms pioneer new generations? Why do some small firms grow in some generations but not in others? What factors affect the rate of old generation innovation in the new generation? In the next two sections, we introduce and test a
Table 6
OLS regression: firm value on market share and drive quality, including only U.S. firms that specialize in the rigid disk drive industry

<table>
<thead>
<tr>
<th>Variable</th>
<th>Eq. (1)</th>
<th>Eq. (2)</th>
<th>Eq. (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>Coefficient</td>
<td>Coefficient</td>
</tr>
<tr>
<td>Intercept</td>
<td>12.48****</td>
<td>12.03****</td>
<td>11.99****</td>
</tr>
<tr>
<td></td>
<td>(0.40)</td>
<td>(0.25)</td>
<td>(0.25)</td>
</tr>
<tr>
<td>Market share</td>
<td>—</td>
<td>18.81***</td>
<td>16.86***</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>(2.31)</td>
<td>(2.25)</td>
</tr>
<tr>
<td>Quality</td>
<td>0.58***</td>
<td>—</td>
<td>0.35***</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td></td>
<td>(0.097)</td>
</tr>
<tr>
<td>YR1977</td>
<td>−1.79****</td>
<td>−1.56****</td>
<td>−1.57****</td>
</tr>
<tr>
<td></td>
<td>(0.51)</td>
<td>(0.42)</td>
<td>(0.39)</td>
</tr>
<tr>
<td>YR1978</td>
<td>−1.64***</td>
<td>−1.40*</td>
<td>−1.56**</td>
</tr>
<tr>
<td></td>
<td>(0.63)</td>
<td>(0.74)</td>
<td>(0.68)</td>
</tr>
<tr>
<td>YR1979</td>
<td>−1.83***</td>
<td>−1.10*</td>
<td>−1.38***</td>
</tr>
<tr>
<td></td>
<td>(0.59)</td>
<td>(0.65)</td>
<td>(0.51)</td>
</tr>
<tr>
<td>YR1980</td>
<td>−0.96*</td>
<td>−0.77*</td>
<td>−0.65</td>
</tr>
<tr>
<td></td>
<td>(0.55)</td>
<td>(0.42)</td>
<td>(0.42)</td>
</tr>
<tr>
<td>YR1981</td>
<td>−0.77</td>
<td>−0.35</td>
<td>−0.31</td>
</tr>
<tr>
<td></td>
<td>(0.55)</td>
<td>(0.40)</td>
<td>(0.39)</td>
</tr>
<tr>
<td>YR1982</td>
<td>−0.90</td>
<td>−0.44</td>
<td>−0.36</td>
</tr>
<tr>
<td></td>
<td>(0.58)</td>
<td>(0.42)</td>
<td>(0.46)</td>
</tr>
<tr>
<td>YR1984</td>
<td>−1.13**</td>
<td>−0.75**</td>
<td>−0.62*</td>
</tr>
<tr>
<td></td>
<td>(0.52)</td>
<td>(0.37)</td>
<td>(0.37)</td>
</tr>
<tr>
<td>YR1985</td>
<td>−1.40***</td>
<td>−1.00***</td>
<td>−0.97***</td>
</tr>
<tr>
<td></td>
<td>(0.50)</td>
<td>(0.36)</td>
<td>(0.37)</td>
</tr>
<tr>
<td>YR1986</td>
<td>−0.88*</td>
<td>−1.05**</td>
<td>−0.54*</td>
</tr>
<tr>
<td></td>
<td>(0.46)</td>
<td>(0.50)</td>
<td>(0.31)</td>
</tr>
<tr>
<td>YR1987</td>
<td>−0.67</td>
<td>0.073</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>(0.60)</td>
<td>(0.31)</td>
<td>(0.31)</td>
</tr>
<tr>
<td>YR1988</td>
<td>−0.68</td>
<td>−0.083</td>
<td>−0.02</td>
</tr>
<tr>
<td></td>
<td>(0.58)</td>
<td>(0.33)</td>
<td>(0.33)</td>
</tr>
<tr>
<td>YR1989</td>
<td>−0.54</td>
<td>−0.53</td>
<td>−0.48</td>
</tr>
<tr>
<td></td>
<td>(0.55)</td>
<td>(0.43)</td>
<td>(0.43)</td>
</tr>
<tr>
<td>YR1990</td>
<td>−0.63</td>
<td>−0.11</td>
<td>−0.41</td>
</tr>
<tr>
<td></td>
<td>(0.58)</td>
<td>(0.50)</td>
<td>(0.48)</td>
</tr>
<tr>
<td>YR1991</td>
<td>−0.51</td>
<td>−0.21</td>
<td>−0.30</td>
</tr>
<tr>
<td></td>
<td>(0.62)</td>
<td>(0.34)</td>
<td>(0.34)</td>
</tr>
<tr>
<td>YR1992</td>
<td>0.10</td>
<td>−0.21</td>
<td>−0.32</td>
</tr>
<tr>
<td></td>
<td>(0.54)</td>
<td>(0.32)</td>
<td>(0.34)</td>
</tr>
<tr>
<td>YR1993</td>
<td>−0.27</td>
<td>−0.50*</td>
<td>−0.50*</td>
</tr>
<tr>
<td></td>
<td>(0.60)</td>
<td>(0.30)</td>
<td>(0.30)</td>
</tr>
<tr>
<td>YR1994</td>
<td>0.79</td>
<td>−0.51</td>
<td>−0.35</td>
</tr>
<tr>
<td></td>
<td>(0.58)</td>
<td>(0.32)</td>
<td>(0.36)</td>
</tr>
<tr>
<td>YR1995</td>
<td>0.73</td>
<td>−0.49</td>
<td>−0.27</td>
</tr>
<tr>
<td></td>
<td>(0.55)</td>
<td>(0.34)</td>
<td>(0.37)</td>
</tr>
<tr>
<td>YR1996</td>
<td>0.68</td>
<td>−0.33</td>
<td>−0.20</td>
</tr>
<tr>
<td></td>
<td>(0.65)</td>
<td>(0.63)</td>
<td>(0.62)</td>
</tr>
<tr>
<td>Number of observations</td>
<td>181</td>
<td>142</td>
<td>137</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.33</td>
<td>0.55</td>
<td>0.60</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.24</td>
<td>0.48</td>
<td>0.52</td>
</tr>
</tbody>
</table>

*, **, and *** indicate significant at the 10%, 5% and 1% levels.
Year dummies are included with 1983 as the base year.
The dependent variable is the natural log of market capitalization (White standard errors in parentheses).
simple model of innovation with different product generations. The results explain why technology leadership is not always equivalent to market leadership and how the two are related over the evolution of a new product generation.

The model oversimplifies innovation in two main ways. First, it considers only one innovation at a time. All firms compete for the same innovation, and the winner can either license it or market the resulting product. In real markets multiple innovations occur in sequence and several may occur simultaneously. We implicitly assume that firms have a limited time horizon when making decisions or that they cannot forecast beyond the current innovation. Second, the model has a partial equilibrium setting with a few large incumbents. Adjusting the technology and market sizes of these large firms yields testable hypotheses that relate industry conditions to which types of firms innovate.

2. The model

The model is a partial equilibrium model of a single-prize innovation race. The product space has a horizontal dimension (product generation) and a vertical dimension (quality). Initially, there are two generations: $A$ is old and $B$ is young. There are two firms in each generation: firms 1$a$ and 2$a$ produce generation $A$ and firms 1$b$ and 2$b$ produce generation $B$. Each firm produces one good. Each good has a different level of quality: $y_{2a} > y_{1a}$ and $y_{2b} > y_{1b}$, where $y_i$ denotes firm $i$’s quality. We refer to firms 2$a$ and 2$b$ as leaders; 1$a$ and 1$b$ are laggards. Since we lack data on production costs, our model ignores cost differences. All firms have identical marginal costs of production, $c$.

Initially, there are two groups of consumers. Two groups is sufficient to allow for the possibility that some consumers tend to prefer goods in generation $A$ while others prefer goods in $B$. To simplify notation, we identify each group with the generation its members tend to prefer consuming. Each consumer purchases at most one good. Consumer $i$ in group $A$ buys good $j$ in generation $A$ if doing so maximizes $i$’s utility:

$$U_{ij} = \alpha_{aa} \theta_j - p_j + \epsilon_{ij},$$

(2.1)

where $\alpha_{aa}$ is a preference parameter, $\theta_j$ and $p_j$ are the quality and price of good $j$, and $\epsilon_{ij}$ is an individual-specific shock. When consumer $i$ in group $A$ buys a good in generation $B$, the good’s quality is weighted by $\alpha_{ab}$. Consumers in group $B$ have parameters $\alpha_{ab}$ and $\alpha_{ba}$. A consumer can purchase none of the industry’s goods, and this option has an expected utility of zero. Under the standard assumption in the discrete choice literature that the individual-specific shocks are independently and identically distributed according to the distribution $e^{-e^{-\eta}}$, the probability that a consumer in group $A$ purchases good 1$a$ is

$$\lambda_{a,1a} = \frac{e^{\alpha_{aa} \theta_{1a} - p_{1a}}}{1 + e^{\alpha_{aa} \theta_{1a} - p_{1a}} + e^{\alpha_{aa} \theta_{2a} - p_{2a}} + e^{\alpha_{aa} \theta_{1b} - p_{1b}} + e^{\alpha_{ab} \theta_{2b} - p_{2b}}}.$$  

(2.2)

The other probabilities, $\lambda_{a,2a}$, $\lambda_{a,1b}$, $\lambda_{a,2b}$, $\lambda_{b,1a}$, $\lambda_{b,2a}$, $\lambda_{b,1b}$, and $\lambda_{b,2b}$, are similar.
There are \( n_a \) consumers in group \( A \) and \( n_b \) in group \( B \). Firm \( j \) ’s profits are
\[
\pi_j^0 = \max_{p_j} (p_j - c)(n_a \lambda_{a,j} + n_b \lambda_{b,j}).
\] (2.3)
All firms choose their prices simultaneously and in equilibrium every firm best responds. Each incumbent \( j \) earns \( \pi_j^0 \) up to the point where some firm innovates.

2.1. Innovation

Beginning from the initial state, one of two possible opportunities for innovating occurs. First, there may be an opportunity to improve quality in the young generation: an opportunity to develop a good in \( B \) with quality \( \theta_{3b} > \theta_{2b} \). Second, there may be an opportunity to pioneer a new generation \( C \). Denote the quality of the good in \( C \) by \( \theta_{1c} \). With the introduction of \( C \), a group of \( n_c \) consumers who had no demand for goods in \( A \) or \( B \) enters the market. These group \( C \) consumers value quality in \( C \) using the parameter \( \alpha_{cc} \) and never purchase goods in \( A \) or \( B \) (\( \alpha_{ca} = \alpha_{cb} = -\infty \)). Consumers in groups \( A \) and \( B \) value quality in \( C \) using the parameters \( \alpha_{ac} \) and \( \alpha_{bc} \), respectively.

After the opportunity for innovating occurs, each incumbent can enter the race or not. If an incumbent enters the race, it pays a race entry cost \( f_e \). New firms may also enter the race. We assume the new firms are spin-outs because spin-outs are important in the disk drive industry, but this assumption is not critical for most of our analysis. Each of the four incumbents can generate one spin-out. This reflects the fact that few employees in any given firm acquire sufficient know-how to found their own firm. If a spin-out forms, it pays a spin-out formation cost \( f_s \) in addition to \( f_e \). After entry decisions are made, there are zero to eight firms in the race. If no firms enter the race the industry remains in the initial state.

The race has several features that are standard in the literature (Reinganum, 1989). First, the race takes place in continuous time. This allows us to ignore ties where two or more firms obtain the innovation at the same time. Second, once one firm innovates, the race is over. Third, the innovation production function is memoryless. Given that no firm has successfully innovated, firm \( i \)'s probability of succeeding in the next instant is \( h_i(x_i) \), where \( x_i \) is firm \( i \)'s investment in the next instant. Since the race is memoryless, the firm’s problem is a stationary one, and the optimal level of \( x_i \) does not change during the race. Each firm \( i \) chooses \( x_i \) to maximize its value.

If firm \( j \) wins the race it has two options. First, \( j \) can manufacture and market the new good. If \( j \) does so, all firms simultaneously choose their prices to maximize profits. In equilibrium, firm \( j \) earns \( \pi_j/r \) from its original good (if it is an incumbent,

\[\text{In the young generation, as quality rises relative to the old generation, more consumers switch from the old to the young goods. In reality, falling prices would contribute to this substitution. We ignore this effect because we lack data on prices, but the impact would be similar to that of a quality improvement.}\]

\[\text{These assumptions are sufficient for describing the main demand effects that occur when a new product generation is first introduced. First, some consumers switch from the old to the new generation. Second, some new consumers enter the market.}\]
zero otherwise) and $\pi^i_I/r$ from the new good, where $r$ is the continuous time discount rate and the superscript indicates which firm markets the innovation. In the case of incumbents, the new product does not replace the marketer’s old one but simply extends the marketer’s product line.

Alternatively, $j$ can license (sell) the innovation to another firm. When firm $j$ licenses the innovation, it gives up the profits from the new good in return for a one-time payment from the licensee. Its profits from its original good may also change when it licenses because product market prices depend on who markets the new good. We assume that the licensor selects the efficient licensee in the sense that there are no further gains from trade. Licensing involves a transaction cost $f_i$. Denote firm $j$’s per period profit from its original good when firm $k$ markets the new product by $\pi^k_I$. Gains from trade must be positive in order for $j$ to license, which implies that the following condition must be satisfied for at least one firm $k$:

$$\frac{\pi^k_k + \pi^k_I}{r} + \frac{\pi^i_j}{r} - f_i \geq \frac{\pi^i_I}{r} + \frac{\pi^j_j}{r}$$

(2.4)

We assume that the two parties Nash bargain over the gains from trade (net of $f_i$) and that the relative bargaining power of licensors and licensees may be unequal. The outside option is to allow $j$ to market the new product. When $j$ licenses to $k$, $j$ receives

$$\frac{\pi^i_j + \pi^i_I}{r} + \gamma \left[ \frac{\pi^k_k + \pi^k_I}{r} + \frac{\pi^i_j}{r} - f_i - \left( \frac{\pi^i_I}{r} + \frac{\pi^j_j}{r} \right) \right]$$

(2.5)

and $k$ receives

$$\frac{\pi^k_k}{r} + (1 - \gamma) \left[ \frac{\pi^k_k + \pi^k_I}{r} + \frac{\pi^i_j}{r} - f_i - \left( \frac{\pi^i_I}{r} + \frac{\pi^j_j}{r} \right) \right],$$

(2.6)

where $\gamma$ measures the relative bargaining power of the licensor.

We focus on subgame perfect Nash equilibria. To compute the subgame perfect Nash equilibria, we use generalized backward induction.

---

5It is useful to think of the innovation as a prototype of the new product. The transaction cost would include the cost of verifying that the new product works and that it is manufacturable at reasonable cost.

6We adopt this simple bargaining rule for two reasons. First, this specification facilitates the antitrust policy analysis we describe below, where we explain how the licensor’s bargaining power may affect entry into the innovation race. Second, some natural alternatives to this rule appear to simply yield special cases of this rule. For example, we tried allowing firms to bid competitively for the license. Perhaps surprisingly, in equilibrium no firm bids more than the innovator’s outside option. Thus, competitive bidding is identical to assuming $\gamma = 0$. Firms other than the equilibrium licensee do not have sufficient incentives to bid higher; as we discuss below, the licensee sets the highest price for the new good and minimizes business stealing effects for every firm. These results suggest that more complicated bargaining rules where, for example, the innovator’s outside option is to bargain with another firm, do not yield equilibrium outside option payoffs different from the competitive bidding payoffs: firms other than the equilibrium licensee prefer the innovator to bargain with the licensee.
2.2. Marketing the new product

In the marketing stage, one firm has the innovation, either by winning the race or by acquiring a license. All firms choose their prices simultaneously. All of the goods are substitutes (cross-price elasticities are non-negative) and all best response functions are upward sloping. The firms choose lower prices than a monopolist would because they ignore the negative impact of their price reductions on their competitors’ demands. Owing to these business stealing effects, firms who do not market the new product prefer whoever does to set a high price.

New firms choose the lowest price for the new product because they do not internalize any of the business stealing effects. If a new firm $i$ markets the new product it chooses its price $p_i$ by maximizing

$$\max_{p_i} n_i = (p_i - c)(n_a\lambda_{a,I} + n_b\lambda_{b,I} + n_c\lambda_{c,I}),$$

which yields the first-order condition

$$n_a\lambda_{a,I} + n_b\lambda_{b,I} + n_c\lambda_{c,I} - (p_i - c)[n_a\lambda_{a,I}(1 - \lambda_{a,I}) + n_b\lambda_{b,I}(1 - \lambda_{b,I})
+ n_c\lambda_{c,I}(1 - \lambda_{c,I})] = 0. \tag{2.7}$$

If an established firm $j$ markets the new product it considers the effect of $p_I$ on the demand for its original product. This adds the following expression to the first-order condition:

$$(p_j - c)[n_a\lambda_{a,j}\lambda_{a,I} + n_b\lambda_{b,j}\lambda_{b,I} + n_c\lambda_{c,j}\lambda_{c,I}]. \tag{2.8}$$

Expression (2.9) is positive, which implies that $p_I$ exceeds the $p_I$ an entrant would choose.

2.3. Licensing the innovation

In the licensing stage, firms look ahead and anticipate the outcome of the marketing stage. If the business-stealing effects are non-zero and the licensing transaction cost is sufficiently low, then new firms never market the new product. There are gains from trade from transferring the innovation to an incumbent because the incumbent marketer internalizes the business stealing effect between its original good and the new product. The incumbent marketer increases the prices of the two goods, which through the best-response functions causes the other incumbents to increase their prices. The profit of every good rises.

Intuition suggests that gains from trade are exhausted when the firm who would set the highest $p_I$ obtains the innovation, because the business stealing effects of the new product are minimized. We do not have a formal proof that this must always occur, but all of our numerical computations yield this outcome, which implies that if licensing occurs, the firm who obtains the license is the one who would set the highest $p_I$. In the discussion that follows, we assume that this always occurs. Note that this implies that licensing may or may not preserve leadership in a particular product generation. The firm who markets the innovation would be the one who
experiences the most negative externality from a low market price on the innovation. While this would ordinarily be one of the leaders, it may be the old or young generation leader. New firms never obtain the license.\footnote{In some industries, new firms may obtain licenses to help them enter a market. This is optimal only if the firm has some complementary assets that will give it a competitive advantage after acquiring the license. We do not consider complementary assets because they are not important in the disk drive industry: the most prominent entrants were spin-outs.}

Numerical results also show who licenses. If $f_i = 0$, licensing always occurs unless the firm who would set the highest $p_I$ wins the race. This is because there are gains from trade. If $f_i > 0$, the gains from trade must exceed $f_i$ in order for licensing to occur. If we list firms in order of how high they would set $p_I$, from highest to lowest, the firm at the top of the list is the licensee if licensing occurs. The gains from trade are highest in the transaction where the firm at the bottom of the list licenses. Thus, if we begin from a high no-licensing level of $f_i$ and reduce $f_i$, the lowest firm on the list (a new firm) licenses first, then as $f_i$ continues to fall eventually the second lowest firm licenses, and so on.

2.4. The innovation race

In the innovation race firms look ahead and realize what their payoffs will be after licensing and marketing occurs. For simplicity we assume that the licensing stage takes no time. Each firm $i$ in the race chooses its investment $x_i$ to maximize

$$
V_i^0 = \int_0^{\infty} e^{-rt} e^{-\sum_j h_j(x_j) t} \left[ \pi_i^0 - x_i + h_i(x_i) V_i^j + \sum_{j \neq i} h_j(x_j) V_j^j \right] dt,
$$

(2.10)

where $V_i^j$ is the value firm $i$ receives if $j$ wins. Simplifying,

$$
V_i^0 = \frac{\pi_i^0 - x_i + h_i(x_i) V_i^j + \sum_{j \neq i} h_j(x_j) V_j^j}{r + \sum_j h_j(x_j)}.
$$

(2.11)

Firm $i$’s first-order condition is

$$
-r - \sum_j h_j(x_j) + h_i'(x_i) \left[ r V_i^j - \pi_i^0 + \sum_{j \neq i} h_j(x_j) (V_i^j - V_j^j) + x_i \right] = 0.
$$

(2.12)

The model yields the familiar replacement and efficiency effects, $r V_i^j - \pi_i^0$ and $\sum_{j \neq i} h_j(x_j) (V_i^j - V_j^j)$ (see Reinganum, 1989).

If no licensing occurs, then for incumbents,

$$
V_i^j = \left( \pi_i^0 + \pi_i^j \right)/r,
$$

and

$$
V_i^j = \pi_i^j/r.
$$

(2.13)

For new firms, $\pi_i^0 = 0$, $V_i^j = \pi_i^j/r$, and $V_j^j = 0$.\footnote{In some industries, new firms may obtain licenses to help them enter a market. This is optimal only if the firm has some complementary assets that will give it a competitive advantage after acquiring the license. We do not consider complementary assets because they are not important in the disk drive industry: the most prominent entrants were spin-outs.}
If no licensing occurs, the replacement effect is highest for spin-outs and laggards 1a and 1b. The new product reduces the profits of all original products because of the business stealing effects. Thus, \( \pi_i^t < \pi_i^0 \) for incumbents. The replacement effect is highest for spin-outs because they have no original good. Industry laggards are hurt less than the leaders because their goods have smaller demands to begin with. In contrast, the value terms in the efficiency effect (the \( V_j^i - V_j^f \)) are highest for the leaders 2a and 2b. These are the firms that suffer the most when a competitor who sets a low \( p_I \) markets the new product.

Now consider the impact of licensing. First, consider the impact of licensing on the \( V_j^i \) terms, which appear in the efficiency effect. Suppose firm \( i \) is the licensee. Then the first-order condition implies that \( i \)'s incentive to win the race falls when it can license because \( V_j^i \) rises; firm \( i \) licenses from \( j \) only if the value \( i \) receives exceeds the value of letting firm \( j \) market the good. Note that all other incumbents are also better off when they lose the race because they prefer firm \( i \) to market the good rather than any of their other competitors (because \( i \) sets the highest \( p_I \)). Thus, other incumbents invest less. However, the incentives of a new firm cannot fall because \( V_j^i = 0 \). Thus, when licensing is possible the effects on the \( V_j^i \) terms favor new firms over incumbents and cause industry investment to fall.

Now consider the impact of licensing on the \( V_i^i \) terms. If firm \( i \) is the licensee then \( V_i^i \) does not change; if \( i \) wins the race it markets the new product. If \( j \) wins the race and \( i \) has all of the bargaining power, then firm \( j \) is indifferent between licensing and marketing. As a result, \( V_j^j \) does not change. Thus, if the licensee has all of the bargaining power then the \( V_i^i \) terms do not change. If the licensor has some bargaining power, then \( V_j^j \) rises for firms other than the licensee. This causes firms other than the licensee to have a greater incentive to invest. Thus, when licensing is possible the effects on the \( V_i^i \) terms lead to more investment as long as the licensors have bargaining power.

### 2.5. Spin-out formation and race entry

Incumbents enter the race if the value of expression (2.11) minus \( f_s \) exceeds the value of expression (2.11) when \( x_i = 0 \). Spin-outs form if expression (2.11) exceeds \( f_s + f_e \). If \( f_s \) and \( f_e \) are sufficiently low, spin-outs always form and all firms enter the race. As \( f_e \) rises, fewer firms enter the race. As \( f_s \) rises, fewer spin-outs form. When \( f_e \) and \( f_s \) are positive, there may be multiple equilibria in the race entry stage. For example, it may be optimal for three firms to enter the race but several configurations of three firms satisfy the equilibrium entry conditions. To break ties, we first select the equilibrium where the number of racers is maximized. Then within that group we select the equilibrium that maximizes total firm value. These assumptions are not critical for our results.

### 2.6. Numerical examples: parameterization

Obtaining analytical results for racing models with heterogeneous firms is difficult. Changing a parameter or the identity of the innovation marketer affects the entire
vectors of prices and investments. The systems of equations that determine equilibrium prices and investments (the first-order conditions) consist of several non-linear equations. These difficulties account for the tendency of the previous literature on racing to explore environments where either firms are identical or the differences are minimized (for example, one entrant and one incumbent). We present numerical results from our model. We attempt to choose parameters to fit the rigid disk drive industry.

In the rigid disk drive industry, customers are segmented according to the size of the computers they manufacture. Firms who make small computers cannot use large drives, but firms who make large computers use small drives if storage capacities and prices favor doing so. In the data, we measure storage capacities using areal densities, which are measured in megabytes per square inch. To convert areal density to megabytes we compute

$$a_i \pi (d_i/2)^2,$$

where \(a_i\) is the drive’s areal density and \(d_i\) is the diameter of the drive. To compare areal densities across diameters, large-computer manufacturers weight the areal density in a small drive by

$$d_i^2 / d_j^2 (p_1^2) \text{drop out). The average ratio (8/14, 5.25/8, 3.5/5.25, 2.5/3.5, 1.8/2.5) is roughly 0.66. Given this,

$$a_{aa} = a_{ab} = a_{ac} = \begin{bmatrix} 1 & (0.66)^2 & (0.66)^4 \\ -\infty & 1 & (0.66)^2 \\ -\infty & -\infty & 1 \end{bmatrix} \text{. (2.14)}$$

We set \(c = 1, r = 0.04, f_e = 0, f_s = 0, f_l = 50, \gamma = 0.5, \text{ and } \theta_{1a} = 1.\) The other qualities are functions of \(\theta_{1a}.\) For example, \(\theta_{1b} = \omega \theta_{1a},\) where \(\omega\) is a constant. The average areal density in a small diameter drive is 25% greater than that of the next larger drive, so as a base case \(\omega = 1.25.\) We consider the effects of changing \(\omega\) below. To compute \(\theta_{2a}\) and \(\theta_{2b},\) we compute the ratio of the highest areal density in a given diameter to the average areal density in that diameter on an annual basis. The ratio is an estimate of the gap between the storage capacity leader and a typical follower. On average (across diameters and years), the areal density of the leader is twice that of the areal density of the average firm. This implies that \(\theta_{2a} = 2 \theta_{1a}\) and \(\theta_{2b} = 2 \theta_{1b}.\) On average, the highest areal density in a diameter improves by 50% each year. Given this, \(\theta_{3b} = \theta_{2b} \times 1.5.\) The highest areal density of a new generation drive in the year it is introduced is roughly equal to the highest areal density of the adjacent larger diameter drive that year, so \(\theta_{1c} = \theta_{2b}.\)

We need to specify \(n_a, n_b,\) and \(n_c.\) These values represent the size of the main customer base for each generation. We set \(n_a = 100.\) As product generation \(B\) matures, \(n_b\) rises relative to \(n_a\) because of exogenous growth in the demand for smaller computers. We consider the effects of increasing \(n_b\) below. We discuss \(n_c\) below.

For now, assume that all firms have the same innovation production function:

$$h_i(x) = a x^{1/2} \text{ for all } i. \text{ We set } a = 0.01.\) Increasing a firm’s \(a\) causes it to invest more. Increasing the scale of all of the firms’ \(a\)’s makes all of the \(h_i\)’s rise and the efficiency effect becomes more important. This causes firms \(2a\) and \(2b\) to invest more.
2.7. Numerical examples: results

First, consider the effect of exogenous demand growth on innovation in generation B. Table 7 shows how different values of \( n_b \) affect which firms are most likely to win the race and whether licensing occurs. When \( n_b \) is low, spin-outs are most likely to innovate, the old generation leader is next most likely to innovate, and licensing never occurs. As \( n_b \) rises, the new generation leader \( 2b \) becomes more likely to innovate. When \( n_b \) is sufficiently high, firm \( 2b \) and spin-outs are most likely to innovate and everyone licenses to \( 2b \). In general, spin-outs tend to be among the most likely innovators because the replacement effect is highest for spin-outs. The efficiency effect is highest for the firm that experiences the highest business stealing effect, which is either \( 2a \) or \( 2b \).

Table 7
Investment and licensing under alternative values of \( n_b \)

<table>
<thead>
<tr>
<th>( n_b )</th>
<th>Ranking, from left to right, in order of who invests most</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_b = 1 ) Licensing</td>
<td>1 Spin-out 2a 1b 1a 2b</td>
</tr>
<tr>
<td>( n_b = 5 ) Licensing</td>
<td>2a 1b 2b 1a</td>
</tr>
<tr>
<td>( n_b = 10 ) Licensing</td>
<td>2b 2a 1b 1a</td>
</tr>
<tr>
<td>( n_b = 20 ) Licensing</td>
<td>2b Spin-out 1a 1b 2a</td>
</tr>
<tr>
<td>( n_b = 50 ) Licensing</td>
<td>2b Spin-out 1a 1b 2a</td>
</tr>
</tbody>
</table>

Second, consider the effect of relative quality. Table 8 shows that increasing the relative quality of the young generation has an effect similar to increasing \( n_b \). We increase \( \omega \), which increases \( \theta_{1b} \) and \( \theta_{2b} \). This affects which firms are most likely to win the innovation race and whether licensing occurs. When \( \omega = 1.5 \), the old generation leader \( 2a \) is most likely to innovate, followed by spin-outs. No licensing occurs. When \( \omega = 1.75 \), firm \( 2a \) and spin-outs are still the most likely firms to innovate, but now spin-outs license the innovation to \( 2a \). As \( \omega \) continues to rise, eventually firm \( 2b \) displaces \( 2a \) as the firm most likely to innovate, and all firms license to the new generation leader.

Third, consider which firms are most likely to pioneer new product generations. When a new product generation is first introduced, demand is typically very low relative to the demand for the older generations’
Given this, we set $n_b = 10$ and $n_c = 1$ to focus on cases where $n_c$ is small relative to $n_a$ and $n_b$. In this case, the new generation has a small impact on existing generations; the efficiency effect is low and the most likely innovators are determined by who has the strongest replacement effect. Thus, spin-outs are most likely to pioneer a new product generation, followed by the laggards in each generation, $1b$ and $1a$, and then the leaders, $2b$ and $2a$. No licensing occurs. For other values of $n_b$, the pattern is similar, as long as the new good has a small impact on existing goods. If the new good has a higher impact then the efficiency effect becomes more important, and the existing leaders become more likely to win the race or license from the winner. Which leader has the greater incentive depends on which one experiences the highest business stealing effect.

### 2.8. The know-how effect and spin-outs

In addition to efficiency and replacement effects, a “know-how effect” may also affect who wins the race by changing the innovation production function: a firm with higher know-how may have a higher probability of innovating for a given investment, and the marginal effect of an additional dollar spent may be higher at every point on the innovation production function. In our model, the high quality firms $2a$ and $2b$ would have higher $\alpha$ parameters. Their spin-outs may also have high $\alpha$’s if the spin-outs transfer firm-specific know-how from their parents (Franco and

---

8This is typical: new generations take time to diffuse before they pose a threat to old generations. In the disk drive industry, sales of new diameter drives were much lower than the sales of old diameter drives in the early years of life of the new diameter. For 5.25”, 3.5”, and 2.5” drives, revenue exceeded 2% of the revenue of larger drives for the first time in the fourth year after the diameter was first introduced. For 1.8” drives, revenue never exceeded this mark; for 8” drives, revenue exceeded this mark in the second year.
Filson (2002) and Klepper and Sleeper (2002) present evidence that this occurs). The know-how effect also explains why spin-outs may be more successful than other types of entrants: non-spin-outs lack the know-how of the best spin-outs. We consider the know-how effect in the empirical analysis below.

2.9. Robustness

The numerical results described above are based on particular demand functions and parameter values. As our discussant Matthew Mitchell emphasized, our specification yields implicit assumptions about business stealing effects. Investment and licensing in the model are driven entirely by business stealing effects; thus, the implicit assumptions are critical for hypothesis formation. In an attempt to make the link between the model and the hypotheses more transparent, we summarize the main implicit assumptions here and link them to the hypotheses presented in the next subsection. A wide range of demand specifications and parameter values yield the same implicit assumptions.

**Implicit Assumption 1:** Larger firms suffer more when a competitor introduces a superior product. Formally, the value terms in the efficiency effect are higher for leaders within each generation (see Section 2.4). This drives Hypotheses 1 and 10.

**Implicit Assumption 2:** Leaders in a generation become larger as the generation matures. In combination with Implicit Assumption 1, this implies that leaders in mature generations have greater incentives to internalize business stealing effects by winning races and obtaining licenses. This drives Hypotheses 6 and 8.

**Implicit Assumption 3:** Business stealing effects of new generations on older generations are low initially. footnote 8 establishes that this is reasonable. This drives Hypothesis 4.

**Implicit Assumption 4:** As the young generation evolves, its products improve and diffuse, and its leaders become larger, the business stealing effects within the young generation eventually become more important relative to those between the young and old generations. This drives Hypotheses 2, 3, and 9.

As discussed below, the remaining hypotheses (5 and 7) are based on an assumption about racing and the know-how effect, respectively.

2.10. Testable hypotheses

The first three hypotheses pertain to quality improvements in young generations. The numerical results suggest that the most likely incumbents to improve quality in the young generation are the leaders in the old and young generations. The know-how effect reinforces this conclusion, because it implies that high-quality firms are more likely to innovate. Tables 7 and 8 show that the stage of evolution of the young generation affects which leader is most likely to innovate. If leaders in the old generation ever innovate in the young generation, they do so early on and when the quality in the young generation is relatively low compared to the quality in the old generation.

**Hypothesis 1:** Old and young generation leaders are the most likely incumbents to improve quality in the young generation.
Hypothesis 2: Old generation leaders are more likely to innovate in the young generation early on.

Hypothesis 3: Old generation leaders are more likely to innovate in the young generation when the quality in the young generation is relatively low compared to the quality in the old generation.

The numerical results on races to pioneer new product generations yield:

Hypothesis 4: New firms and industry laggards are the most likely pioneers of new product generations.

Hypotheses 5–7 pertain to new firm success. Hypothesis 5 is based on an assumption of the model rather than a result, but we test it below. Hypothesis 6 is based on the results in Table 7, which show how new firms become less likely to win races as the generation matures. Hypothesis 7 is based on the know-how effect.

Hypothesis 5: Exit without licensing or marketing is associated with failure to innovate.

Hypothesis 6: New firms are more likely to fail when attempting to improve quality in mature generations (because leading incumbents have a greater incentive to win such races).

Hypothesis 7: Spin-outs are more likely to succeed than other new firms.

Hypothesis 8 considers whether new firms who successfully innovate grow large. The numerical results suggest that new firms have no chance of growing large if they innovate in mature generations because all innovators in mature generations license to the current leader.

Hypothesis 8: New firms are more likely to become large manufacturer-marketers if they pioneer a new product generation or improve quality early on in a young generation; they are more likely to license or be acquired if they improve quality in a mature generation.

The last two hypotheses pertain to licensing. The results in Tables 7 and 8 show that early in the life of a young product generation, old generation firms may improve quality in the young generation by licensing innovations from other firms, particularly entrants. Thus, licensing early on is associated with entry into the young generation. As the young generation matures, the licensees tend to be firms who already produce the young generation’s goods. In general, because licensees are generation leaders, they tend to be large:

Hypothesis 9: Licensing early on is associated with entry into the young generation.

Hypothesis 10: Licensees are likely to be large manufacturer-marketers.

3. Empirical results

3.1. Data

The main data source is the Disk/Trend Report on Rigid Disk Drives (1977–1997). On an annual basis, the Disk/Trend Report performs a census of all firms producing or attempting to develop rigid disk drives. Thus, many of the firms included are
start-ups or other entrants who end up never achieving success in their initial development efforts. The data includes model characteristics and introduction dates. The level of detail allows us to construct measures of product quality and keep track of which product generations each firm produces. The dataset contains 193 firms, 1,189 firm/year observations, and 11,644 model/year observations. Annual sales of disk drives are reported for several firms. Approximately half (49%) of the U.S. entrants in the period 1977–1997 did not achieve positive sales, and most who did achieve positive sales remained small.

We measure quality using areal densities. The areal density measures how much information can be stored on each square inch of disk (megabytes/in²). To compare across years, we normalize areal densities using z scores. First, we select each firm’s highest areal density drive in each diameter in each year. Then, for each diameter/year group, we compute the mean and standard deviation of this best-drive measure across firms. We use these means and standard deviations to compute z scores for each firm/diameter/year measure, and this is our main quality measure. In tests where the appropriate diameter is not clear, a firm’s diameter quality is averaged across the diameters the firm produces to obtain an average measure of the firm’s quality (this is used in Table 6, described in Section 1.1).

To test Hypotheses 1–3, we divide firms into old generation firms and young generation firms for each adjacent diameter pair: {14”, 8”}, {8”, 5.25”}, {5.25”, 3.5”}, {3.5”, 2.5”}, and {2.5”, 1.8”}. In each pair, the larger diameter is the old generation and the smaller one is the young generation. If a firm produces only the young drives in a given year it is coded as YPG (young product generation) for that pair/year; if it produces only the old drives it is coded as OPG (old product generation); if it produces both it is coded as BOTH.

To test Hypotheses 4–8, we distinguish between spin-outs and non-spin-outs. Spin-outs and their parents were identified using information in the Disk/Trend Report, press releases and articles provided by James Porter (editor of the Disk/Trend Report), the Directory of Corporate Affiliations, the International Directory of Company Histories, and Christensen (1993). There are 40 spin-outs formed in the period 1977–1997. To determine the parents we focus on the background of the spin-out founders and not on other employees, for which data is unavailable. The implicit assumption is that founders had considerable influence on the products and

---

9 While the data is quite detailed, we lack sufficient data to estimate cross-model or cross-generation elasticities. Our discussant Matthew Mitchell suggested that we might be able to use such estimates to test or parameterize the model. Unfortunately, we lack data on prices and on which consumers are new to the market and which switch generations. We could attribute any decline in old generation sales to substitution, but in the data, sales in all generations typically grow from year to year.

10 Sales of other products, including licenses and disk drive components, are not included in the measure of disk sales. Only sales of drives are counted.

11 Only drives that have been produced are used when making these calculations. Drives that have been announced but not yet put into production (and may never actually be produced) are not included.

12 Firms who produce neither the young or the old drives are left out of the analysis of the pair, but this is irrelevant; such firms rarely innovate in the young generation. For example, firms who produce only smaller diameter drives innovate in larger diameter drives in only four cases in the 20 year period we are examining.
strategies of the start-up; evidence from firm press releases and the Disk/Trend Report supports this assumption.

We focus on U.S. entrants. All but two of the spin-outs are U.S. firms. We use press releases obtained through Lexis-Nexis to identify licensing agreements and acquisitions, and coverage of U.S. firms exceeds coverage of non-U.S. firms. We exclude two entrants who make drives only for their own use. We do not count changes in ownership or name changes as new entry. We group the entrants into three categories. Those who eventually achieve $50 million 1983 dollars in at least 3 years are large manufacturer-marketers. Those who are not in this category but eventually license to or get purchased by a large manufacturer-marketer are licensors. Those who are not in either category and exit before 1997 are failures. Only one late entrant does not fall into one of these three categories; we exclude it. Although the first two categories do not describe all of the ways entrants might generate value, they do capture the two ways that our model emphasizes. They also capture the two main types of success in our data: failures typically exited without generating substantial revenue.

3.2. Quality improvements in young product generations

We use probit models to test Hypotheses 1–3. The dependent variable is a binary variable that takes the value 1 if, in a particular year, a firm introduces a young generation drive with an areal density higher than the industry-best areal density in the generation in the previous year. We consider only firms who were in the market in the previous year; this is necessary in order to measure the initial quality of the potential innovators.

We use the following independent variables. We use the firm's quality in the previous year to test Hypothesis 1; the model identifies leaders as high-quality firms. We use the age of the young generation to test Hypothesis 2. To test Hypothesis 3, we compute the megabytes associated with the highest areal density drive in each generation each year. Then we compute a ratio \( \frac{MB_{YPG}}{MB_{YPG} + MB_{OPG}} \) for each old/young pair (discussed in the previous subsection) in each year. As this ratio rises, the quality of the young generation drive rises relative to the quality of the old.

---

13 Other cutoffs and criteria could be used (market share instead of sales levels, for example). For this industry it does not matter much, as it is easy to distinguish firms who grew large from those who did not. Over the period we examine, 97% of industry sales are accounted for by large manufacturer-marketers as we have defined them. On an annual basis, the share of industry sales accounted for by large manufacturer-marketers is never below 91%.

14 Thus, we focus on only those firms who improve on last year's best quality. This is the type of innovation our model and most racing models emphasize. Lerner (1997) takes a different approach and measures whether each firm improves on its own drives each year. He finds that laggards are more likely to introduce improvements than leaders, but this mainly measures a tendency of laggards to catch up to leaders, not surpass them.

Note that this variable does not distinguish between whether the firm innovated or obtained a license to introduce an improved drive. This poses no problem for the hypothesis tests because the hypotheses also do not distinguish between the two possibilities.
generation drives. The lagged value of the megabyte ratio is used to measure the relative qualities at the start of the innovation race.

We interact the know-how, age, and megabyte ratio variables with dummy variables that indicate whether the firm produces drives in the young generation, old generation, or both. This allows us to assess whether an increase in age, for example, makes young generation firms more likely to innovate. As additional control variables, we include dummy variables for young and old generation firms and year effects. Table 9 reports summary statistics.

The results in Table 10 support the hypotheses. Higher initial quality makes the firm more likely to innovate (Hypothesis 1). This effect is stronger for firms that produced the young generation or both in the previous period, but all three effects are significantly positive. The marginal effects show that a one standard deviation increase in quality increases the probability that a firm innovates by 1.4, 3.7, and 3.9 percentage points for old, young, and both firms, respectively. An increase in the megabyte ratio or the diameter’s age makes young generation firms more likely to innovate (Hypotheses 2 and 3). For example, an increase in the megabyte ratio from 0.25 to 0.5 increases the likelihood that a young generation or a both firm innovates by 8.2 and 2.1 percentage points, respectively. In contrast, the likelihood that an old generation firm innovates falls by 1.8 percentage points. An increase in drive age of one year decreases the likelihood that a young generation firm innovates by 0.63 percentage points, but it decreases the likelihood that an old generation firm

Table 9
Summary statistics. 1934 observations

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The firm innovates in the young generation this year</td>
<td>0.082</td>
<td>0.27</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Independent variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The firm’s drive quality last year*</td>
<td>−0.054</td>
<td>0.97</td>
<td>−2.07</td>
<td>4.45</td>
</tr>
<tr>
<td>The megabyte ratio last year</td>
<td>0.33</td>
<td>0.077</td>
<td>0.22</td>
<td>0.61</td>
</tr>
<tr>
<td>The age of the young generation last year</td>
<td>5.34</td>
<td>3.78</td>
<td>0.00</td>
<td>16.00</td>
</tr>
<tr>
<td>Dummy variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The firm produced both the new and old generations last year</td>
<td>0.27</td>
<td>0.44</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>The firm produced only the new generation last year</td>
<td>0.33</td>
<td>0.47</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>The firm produced only the old generation last year</td>
<td>0.40</td>
<td>0.49</td>
<td>0.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*The firm’s drive quality is measured using its quality in the young generation if it produces the young generation; otherwise its quality in the old generation is used.
Table 10
Probit model: innovation in young product generations

<table>
<thead>
<tr>
<th>Variable</th>
<th>Probit estimation</th>
<th>Marginal effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
<td>Coefficient</td>
<td>Coefficient</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>$-1.93^{***}$</td>
<td>$-0.16^{***}$</td>
</tr>
<tr>
<td></td>
<td>$(0.42)$</td>
<td>$(0.046)$</td>
</tr>
<tr>
<td>YPG dummy</td>
<td>$-0.84^*$</td>
<td>$-0.070^*$</td>
</tr>
<tr>
<td></td>
<td>$(0.45)$</td>
<td>$(0.039)$</td>
</tr>
<tr>
<td>OPG dummy</td>
<td>$0.21$</td>
<td>$0.017$</td>
</tr>
<tr>
<td></td>
<td>$(0.64)$</td>
<td>$(0.054)$</td>
</tr>
<tr>
<td>YPG $\times$ lagged quality</td>
<td>$0.45^{***}$</td>
<td>$0.037^{***}$</td>
</tr>
<tr>
<td></td>
<td>$(0.074)$</td>
<td>$(0.00089)$</td>
</tr>
<tr>
<td>OPG $\times$ lagged quality</td>
<td>$0.17^*$</td>
<td>$0.014^*$</td>
</tr>
<tr>
<td></td>
<td>$(0.093)$</td>
<td>$(0.0080)$</td>
</tr>
<tr>
<td>BOTH $\times$ lagged quality</td>
<td>$0.46^{***}$</td>
<td>$0.039^{***}$</td>
</tr>
<tr>
<td></td>
<td>$(0.078)$</td>
<td>$(0.010)$</td>
</tr>
<tr>
<td>YPG $\times$ lagged megabyte ratio</td>
<td>$3.90^{***}$</td>
<td>$0.33^{***}$</td>
</tr>
<tr>
<td></td>
<td>$(1.16)$</td>
<td>$(0.11)$</td>
</tr>
<tr>
<td>OPG $\times$ lagged megabyte ratio</td>
<td>$-0.88$</td>
<td>$-0.074$</td>
</tr>
<tr>
<td></td>
<td>$(1.86)$</td>
<td>$(0.16)$</td>
</tr>
<tr>
<td>BOTH $\times$ lagged megabyte ratio</td>
<td>$1.00$</td>
<td>$0.083$</td>
</tr>
<tr>
<td></td>
<td>$(1.14)$</td>
<td>$(0.097)$</td>
</tr>
<tr>
<td>YPG $\times$ lagged drive age</td>
<td>$-0.075^{***}$</td>
<td>$-0.0063^{**}$</td>
</tr>
<tr>
<td></td>
<td>$(0.027)$</td>
<td>$(0.0025)$</td>
</tr>
<tr>
<td>OPG $\times$ lagged drive age</td>
<td>$-0.28^{**}$</td>
<td>$-0.023^{***}$</td>
</tr>
<tr>
<td></td>
<td>$(0.14)$</td>
<td>$(0.0078)$</td>
</tr>
<tr>
<td>BOTH $\times$ lagged drive age</td>
<td>$0.22^*$</td>
<td>$0.018^{**}$</td>
</tr>
<tr>
<td></td>
<td>$(0.13)$</td>
<td>$(0.0082)$</td>
</tr>
</tbody>
</table>

Number of observations: 1934
Log likelihood: $-440.49$
innovates by 2.3 percentage points. The likelihood that a both firm innovates rises by 1.8 percentage points.

3.3. New product generations

Tables 1–5 provide some support for Hypothesis 4; many of the technology leaders early on are entrants. To analyze this further, we focus on the firms Franco and Filson (2002) define as “early movers” in the new diameters: firms that introduced drives in the diameter within the first three quarters of the introduction of the diameter. Table 11 lists these firms by diameter. For incumbents, the firm’s quality in its closest larger diameter in the previous year is listed. For spin-outs, the parent’s quality in its closest diameter in the year before the spin-out is born is listed. These quality measures capture the technological position of the firm (or the parent, where the spin-out gets its know-how) before the introduction of the new diameter.

The first firms to introduce all of the new diameters were spin-outs: International Memories, Seagate, Rodime, Prairietek, and Integral Peripherals were the first firms to introduce 8″, 5.25″, 3.5″, 2.5″, and 1.8″ drives, respectively. It is extremely unlikely that this occurred by chance. Given that there are 193 firms in the data set, 40 of which are spin-outs, the probability of randomly drawing five spin-outs in five random draws is 0.00038. There are nine spin-outs in the group of 20 early movers. This is also unlikely to have occurred by chance; the probability of observing at least nine spin-outs in 20 random draws is 0.013.

Table 11 shows that incumbents who were early pioneers were primarily low quality laggards in either the old or young generation. Of the firms listed in Table 11,
Table 11
Early movers (Firms are in alphabetical order in each category)

<table>
<thead>
<tr>
<th>Diameter</th>
<th>Firm</th>
<th>Introduction date</th>
<th>Did the firm spin-out to introduce this diameter?</th>
<th>Parent if spin-out</th>
<th>Closest diameter previously produced</th>
<th>Closest diameter previously produced by parent</th>
<th>Quality in closest diameter previously produced</th>
<th>Parent quality in closest diameter previously produced</th>
</tr>
</thead>
<tbody>
<tr>
<td>8&quot;</td>
<td>BASF</td>
<td>Q4, 1979</td>
<td>No</td>
<td>—</td>
<td>14&quot;</td>
<td>—</td>
<td>−0.117</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>IBM</td>
<td>Q1, 1979</td>
<td>No</td>
<td>—</td>
<td>14&quot;</td>
<td>—</td>
<td>0.873</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>International memories</td>
<td>Q1, 1979</td>
<td>Yes</td>
<td>Memorex</td>
<td>—</td>
<td>14&quot;</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Micropolis</td>
<td>Q4, 1979</td>
<td>Yes</td>
<td>Pertec</td>
<td>—</td>
<td>14&quot;</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>New world computer</td>
<td>Q3, 1979</td>
<td>No</td>
<td>—</td>
<td>—</td>
<td>None</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Pertec</td>
<td>Q4, 1979</td>
<td>No</td>
<td>—</td>
<td>14&quot;</td>
<td>—</td>
<td>−1.378</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Shugart</td>
<td>Q4, 1979</td>
<td>No</td>
<td>—</td>
<td>14&quot;</td>
<td>—</td>
<td>−0.940</td>
<td>—</td>
</tr>
<tr>
<td>5.25&quot;</td>
<td>Computer memories</td>
<td>Q2, 1981</td>
<td>Yes</td>
<td>Pertec</td>
<td>—</td>
<td>8&quot;</td>
<td>—</td>
<td>0.265</td>
</tr>
<tr>
<td></td>
<td>International memories</td>
<td>Q1, 1981</td>
<td>No</td>
<td>—</td>
<td>8&quot;</td>
<td>—</td>
<td>−0.830</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>New world</td>
<td>Q3, 1980</td>
<td>No</td>
<td>—</td>
<td>8&quot;</td>
<td>—</td>
<td>−1.571</td>
<td>—</td>
</tr>
</tbody>
</table>
Table 11 (continued)

<table>
<thead>
<tr>
<th>Diameter</th>
<th>Firm</th>
<th>Introduction date</th>
<th>Did the firm spin-out to introduce this diameter?</th>
<th>Parent if spin-out</th>
<th>Closest diameter previously produced</th>
<th>Closest diameter previously produced by parent</th>
<th>Quality in closest diameter previously produced</th>
<th>Parent quality in closest diameter previously produced</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5&quot;</td>
<td>Control data</td>
<td>Q3, 1983</td>
<td>No</td>
<td>—</td>
<td>8&quot;</td>
<td>—</td>
<td>0.757</td>
<td>—</td>
</tr>
<tr>
<td>Micro-computer memories</td>
<td>Q1, 1984</td>
<td>Yes</td>
<td>Alpha Data</td>
<td>—</td>
<td>14&quot;</td>
<td>—</td>
<td>—</td>
<td>-1.019</td>
</tr>
<tr>
<td>Micro-science international</td>
<td>Q2, 1984</td>
<td>No</td>
<td>—</td>
<td>5.25&quot;</td>
<td>—</td>
<td>0.289</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Rodime</td>
<td>Q3, 1983</td>
<td>No</td>
<td>—</td>
<td>—</td>
<td>5.25&quot;</td>
<td>—</td>
<td>1.067</td>
<td>—</td>
</tr>
<tr>
<td>2.5&quot;</td>
<td>PrairieTek</td>
<td>Q4, 1988</td>
<td>Yes</td>
<td>Mini-scribe</td>
<td>3.5&quot;</td>
<td>—</td>
<td>0.594</td>
<td>—</td>
</tr>
<tr>
<td>1.8&quot;</td>
<td>Integral Peripherals</td>
<td>Q3, 1991</td>
<td>Yes</td>
<td>PrairieTek</td>
<td>2.5&quot;</td>
<td>—</td>
<td>0.707</td>
<td>—</td>
</tr>
</tbody>
</table>

An early mover is defined to be a firm that introduces a drive in the diameter within 3 quarters after the first introduction date. The introduction date is the date the product was first shipped. Announced products that were still in the development stage, and had not shipped, are not included.
only two, IBM and Control Data, were large manufacturer-marketers with high quality and high market shares at the time they introduced their drives. The average incumbent quality in the closest drive previously produced is $-0.21$. In the sample as a whole, by construction, quality has mean zero, a standard deviation of one, and approximately half the firms have qualities above zero. Thus, the average incumbent pioneer was clearly not a quality leader. On the whole, the results support Hypothesis 4: spin-outs and laggards are the main pioneers.

3.4. Entrant success vs. failure

We use probit models to test Hypotheses 5–7. The dependent variable is a binary variable that takes the value 1 if the entrant experiences either type of success described in Section 3.1 (licensing or growing large).\footnote{Our discussant Matthew Mitchell pointed out that in the model, entrants who lose their first race exit without selling anything, whereas in our tests here we use our “large manufacturer-marketer” cutoff to indicate success. Our approach can be justified in two ways. First, reality is not as stark as our model. While 49\% of entrants never achieved positive sales, our data includes many entrants who achieved very low sales and subsequently exited. We think these firms are properly categorized as race losers rather than winners. As Cockburn and Henderson (1994) note, even in patent races late finishers obtain some kind of marketable product, but losers do not do nearly as well as winners. A more complete model would allow for different degrees of success in the race. Second, our model suggests that if a new firm survives beyond its first race and grows, it must have won its first race. Thus, observing that a new firm grew to become large is an indication that the firm was a successful innovator in its first race.} We include the following independent variables. First, we use the quality of the entrant’s first drive to measure how successful it was in its first effort to innovate (to test Hypothesis 5). Second, we use the age of the product generation of the entrant’s first drive to test Hypothesis 6. Since $14'$ drives were introduced long before the beginning of our sample period, we consider the age of this generation separately, using 1976 as the initial year. Using different initial years for $14'$ drives has no effect on the results. Finally, we include a dummy variable that takes the value 1 if the entrant is a spin-out to test Hypothesis 7. Table 12 reports summary statistics.

The results in Table 13 support the hypotheses. Entrants with higher initial quality are more likely to experience success (Hypothesis 5). The estimated marginal effects in Eq. (1) show that a one standard deviation increase in quality increases the likelihood of experiencing success by 10 percentage points. Entrants who attempt to innovate in more mature product generations are less likely to experience success (Hypothesis 6). Eq. (1) shows that a 1-year increase in the age of the product generation reduces the probability of success by 11 percentage points. Eq. (2) shows the probability of success rises by 25 percentage points if the entrant is a spin-out (Hypothesis 7). Eq. (3) re-estimates Eq. (1) including only spin-outs to confirm that quality and the age of the product generation continue to have the predicted effects.

3.5. Entrant licensing vs. marketing

To test Hypothesis 8 we focus on only those entrants who experienced one of the two types of success. The dependent variable is a binary variable that takes the value
Table 13
Probit model: entrant success vs. failure

<table>
<thead>
<tr>
<th></th>
<th>Eq. (1)</th>
<th></th>
<th>Eq. (2)</th>
<th></th>
<th>Eq. (3) (Spin-outs only)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Probit estimation</td>
<td>Marginal effects</td>
<td>Probit estimation</td>
<td>Marginal effects</td>
<td>Probit estimation</td>
<td>Marginal effects</td>
</tr>
<tr>
<td>Variable</td>
<td>Coefficient</td>
<td>Coefficient</td>
<td>Coefficient</td>
<td>Coefficient</td>
<td>Coefficient</td>
<td>Coefficient</td>
</tr>
<tr>
<td>Constant</td>
<td>0.29</td>
<td>0.099</td>
<td>−0.13</td>
<td>−0.042</td>
<td>0.51</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>(0.30)</td>
<td>(0.10)</td>
<td>(0.38)</td>
<td>(0.12)</td>
<td>(0.38)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>Initial quality</td>
<td>0.31***</td>
<td>0.10***</td>
<td>0.25*</td>
<td>0.081*</td>
<td>0.27</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td>(0.047)</td>
<td>(0.15)</td>
<td>(0.048)</td>
<td>(0.17)</td>
<td>(0.067)</td>
</tr>
<tr>
<td>Generation age</td>
<td>−0.33***</td>
<td>−0.11***</td>
<td>−0.35***</td>
<td>−0.11***</td>
<td>−0.29**</td>
<td>−0.12***</td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td>(0.035)</td>
<td>(0.12)</td>
<td>(0.035)</td>
<td>(0.13)</td>
<td>(0.050)</td>
</tr>
<tr>
<td>14th generation age</td>
<td>−0.20</td>
<td>−0.069</td>
<td>−0.20</td>
<td>−0.064</td>
<td>−0.23</td>
<td>−0.088</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.048)</td>
<td>(0.14)</td>
<td>(0.046)</td>
<td>(0.16)</td>
<td>(0.064)</td>
</tr>
<tr>
<td>Spin-out</td>
<td>—</td>
<td>—</td>
<td>0.75*</td>
<td>0.25*</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.40)</td>
</tr>
<tr>
<td>Number of observations</td>
<td>63</td>
<td>63</td>
<td>37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>−33.45</td>
<td>−31.62</td>
<td>−21.29</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*, **, and *** indicate significance at the 10%, 5%, and 1% levels. The dependent variable takes the value 1 if the entrant becomes a large manufacturer/marketer or sells technology (standard errors in parentheses).
1 if the entrant became a large manufacturer-marketer, 0 if the entrant licensed. The independent variables are identical to those used in the previous subsection.

The results in Table 14 support the hypothesis: entrants who innovate in more mature product generations are less likely to grow large. Although the effects are statistically insignificant (the sample size is small), the point estimates are large. Eq. (1) shows that a 1 year increase in the age of the product generation reduces the probability that the entrant grows large (and increases the probability that the entrant licenses) by 14 percentage points. Entrants with higher initial quality are more likely to grow, but this effect is relatively weak. Eq. (2) shows that if the entrant is a spin-out its probability of growing large increases by 11 percentage points. Eq. (3) shows the considering only spin-outs does not change the conclusion about the effect of product generation age.

3.6. The role of licensing

To test Hypothesis 9, we sort licensees into those who did not produce the product generation in the year before obtaining the license and those who did. The average age of the product generation at the time of the license in the first group is 3.67 years; the average age in the second group is 5.89 years. The difference is significant at the 10% level. This supports the hypothesis: licensing early on is associated with entry into the product generation.

<table>
<thead>
<tr>
<th>Table 14</th>
<th>Probit model: type of entrant success</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eq. (1)</td>
</tr>
<tr>
<td></td>
<td>Probit estimation</td>
</tr>
<tr>
<td>Variable</td>
<td>Coefficient</td>
</tr>
<tr>
<td>Constant</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>(0.48)</td>
</tr>
<tr>
<td>Initial quality</td>
<td>0.039</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
</tr>
<tr>
<td>Generation age</td>
<td>−0.36</td>
</tr>
<tr>
<td></td>
<td>(0.25)</td>
</tr>
<tr>
<td>14th Generation age</td>
<td>−0.13</td>
</tr>
<tr>
<td>Spin-out</td>
<td>0.0058</td>
</tr>
<tr>
<td></td>
<td>(0.73)</td>
</tr>
<tr>
<td>Number of observations</td>
<td>22</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>−13.28</td>
</tr>
</tbody>
</table>

*indicates significance at the 10% level. The dependent variable takes the value 1 if the entrant becomes a large manufacturer/marketer, and 0 if it sells technology—failures are excluded (standard errors in parentheses).
To test Hypothesis 10, we note that 55% of the licensees in our sample are large manufacturer-marketers.¹⁶ In the sample of all firm-year observations, only 36% are large manufacturer-marketers. These proportions are significantly different at the 5% level. This supports the hypothesis: large manufacturer-marketers are over-represented in the group of licensees. However, large manufacturer-marketers account for 97% of industry sales over the period we examine. Thus, large firms are not over-represented relative to their sales.

4. Policy implications

In this section, we discuss the implications of our findings for antitrust policies governing licensing agreements and policies governing employee non-compete agreements. We use our model to examine the effects of prohibiting licensing and spin-outs. Our empirical results help clarify the role of licensing and the potential impacts of non-compete agreements and other barriers to employee departures.

Racing models typically yield ambiguous welfare results: there may be too little or too much investment relative to the social optimum. There may be too little because the innovation marketer does not appropriate all of the social benefit from the innovation as long as consumer surplus is positive. There may be too much because there is a single prize; the investment of the non-winners is wasted. Each firm invests without internalizing the negative externalities of its investment on its competitors (they are less likely to win).

One way to assess the relative importance of these two offsetting factors is to impose structure on consumer demand, as we have done. Anderson et al. (1992) describe how to compute consumer surplus in a discrete choice model with no consumer heterogeneity. It is straightforward to extend their approach to the case with three types of consumers. We assume that utility shocks are i.i.d. with distribution $e^{c−c_{ij}}$. The pre-innovation expected consumer surplus of consumers in group $A$ is

$$n_a \ln(1 + e^{\varphi_a \theta_{a,1} - p_{1a}} + e^{\varphi_a \theta_{a,2} - p_{2a}} + e^{\varphi_b \theta_{a,1b} - p_{1b}} + e^{\varphi_b \theta_{a,2b} - p_{2b}}).$$ (4.1)

The expected consumer surplus of consumers in group $B$ can be calculated in a similar fashion. Total consumer surplus is the sum of these two surpluses. Adding an extra good (the new product) adds an additional $e^{\varphi b - p}$ term to the expression in brackets, and if the good pioneers a new product generation then the surplus of group $C$ consumers is also added. The present value of consumer surplus can be computed in a manner similar to the present value of a firm (described above in Section 2.4).

¹⁶ We exclude broad cross-licensing agreements and acquisitions of large firms from the sample of licenses because these are not the type of transaction our model emphasizes. Including these would add more transactions between large firms. We also exclude a particular type of license that resulted from Rodime patenting the 3.5″ diameter design. All producers of 3.5″ drives had to obtain licenses from Rodime.
4.1. Antitrust policies governing licensing

Antitrust policy occasionally limits licensing from potential product market entrants to incumbents on the grounds that entry would improve product market competition. In our model, the welfare impacts of such policies depend critically on how they effect entry into the innovation race. If allowing licensing induces entry into the race, then industry investment rises. The social benefit of faster innovation may outweigh the social loss from market power in the product market. The results suggest that antitrust authorities should assess the impact of forbidding licensing on incentives to innovate.

Relative bargaining power is important. In Section 2.4 we argued that the incentive provided by the efficiency effect is reduced when licensing is allowed; the $V_j$ terms rise for every incumbent $i$. Further, if the licensee has all of the bargaining power, then the replacement effects do not change. This yields a general conclusion: if the licensee has all of the bargaining power, then licensing reduces investment. We also noted that as the licensor’s bargaining power rises, the replacement effect leads to increased innovation from firms other than the licensee. The game is one of strategic complements; thus, the licensee also increases its investment. This effect offsets the impact of the efficiency effect. The licensor with more bargaining power can appropriate more profits for itself, and this yields greater incentives to enter the race and invest.

Considering relative bargaining power yields a policy conclusion that is counter to the current approach of the U.S. antitrust agencies. The agencies view high payments from the licensee to the licensor as a sign that the licensor could have entered the product market. The agencies’ pro-competitive stance causes them to prefer entry to licensing, and as a result they frown on licensing agreements with high payments to the licensor. In contrast, our model suggests that licensing agreements where the licensor obtains a higher percentage of the gains from trade (a higher payment) may be associated with more competition in the race, greater investment in innovation, and higher welfare.

Payments to the licensor are higher if the licensor can profitably enter, but our analysis suggests that the ability of a firm to enter does not imply that entry generates higher welfare than the best licensing agreement. After the innovation is available, welfare is maximized by improving product market competition, and this is what the agencies focus on. In doing so, they ignore the impact of their policies on the incentives to innovate in the first place.

4.2. Policies governing employee non-compete agreements

A non-compete agreement, or covenant not to compete, is a contract entered into by an employee and employer whereby upon termination of employment the employee is restricted from competing in the same business as the employer for a

---

17 Gilbert and Tom (2001) provide an excellent summary of U.S. antitrust agencies’ policies and practices in recent cases involving innovation concerns.
particular period of time in a certain geographical location. Non-compete agreements are designed to protect employers from unfair competition caused by former employees working for a competitor or starting a similar business. This type of competition is deemed unfair because the former employee can use confidential knowledge of the former employer to gain an advantage in the market.

A small minority of states do not enforce non-compete agreements unless they are used in conjunction with a sale of a business, dissolution of a partnership, or other very specialized cases. This is important for our analysis because the vast majority of the spin-outs in the rigid disk drive industry were initially located in California, the most prominent non-enforcer. Thus, the rigid disk drive industry shows what can happen if employee mobility is permitted.

In typical racing models, an increase in the number of firms in the race causes a winner to emerge sooner (Reinganum, 1989). Our model is no exception to this general result, which implies that spin-outs lead to higher industry investment and higher welfare. Thus, in our model, preventing spin-outs through non-compete agreements or other means lowers total welfare, primarily because it reduces the present value of consumer surplus. However, preventing spin-outs may increase total firm value as the incumbents avoid having to compete with the spin-out entrants in the innovation race. Thus, firms may have an incentive to prevent spin-outs even when doing so reduces total welfare.

Our numerical results suggest that if the business stealing effects of the new product are small then total firm value may rise when spin-outs enter the innovation race. This occurs because the positive effect of generating additional customers outweighs the negative effect of enhanced competition in the product market. This is most likely to happen when the spin-out intends to pioneer a new product generation. This suggests that non-competes would have the primary objective of preventing competition in races to improve quality in existing product generations. When total firm value rises, a prospective spin-out might be able to buy out its non-compete agreement by offering its parent a share of its value. Of course, in reality it may not be possible for a firm to assess which type of innovation its departing employee intends to pursue. In this case, the firm may refuse to renegotiate. Thus, the presence of a non-compete agreement and the possibility of its enforcement may deter employees from founding firms even when they do not intend to compete head-to-head against their parents in the product market.

We find little evidence of spin-outs forming to compete head-to-head with their parents in the product market. Most of the spin-outs who grew to become large manufacturer-marketers were pioneers of new product generations or early followers. The most successful of this group was Seagate, the first firm to introduce a 5.25" drive and the market share leader at the end of our sample. Of the 11 spin-outs that grew large, only Quantum entered to compete head-to-head against its parent. Quantum entered with a low end 8" drive that imitated its parent Shugart Associates’ 8" drive.

On the whole, our theoretical and empirical results suggest that spin-out formation has beneficial effects. Our results add to a growing literature that questions whether non-compete agreements should ever be enforced. Gilson (1999)
argues that California’s policy of not enforcing non-compete agreements contributed to high employee mobility in Silicon Valley and that this mobility encouraged growth. Cooper (2001) argues that non-competes are a double-edged sword if all firms use them: each firm gets to keep its own employees but cannot get other firms’ employees. The resulting labor allocation is sub-optimal.

5. Conclusion

In this paper, we introduce a racing model with multiple product generations, product innovation, spin-outs, and licensing. Tests of the model’s predictions using data from the rigid disk drive industry (1977–1997) provide empirical support. In the model, new product generations pass through at most three stages. First, they are typically pioneered by spin-outs and lagging firms. Second, old product generation leaders innovate or license to become leaders in the new generation. Third, as the new generation matures, eventually the new generation leaders maintain their leadership, either through innovation or licensing. The first stage is skipped if the business stealing effect of the new generation goods on the old generation goods is high immediately. The second stage is skipped if the business stealing effect is low initially, but the market for the new generation goods grows rapidly thereafter, either through exogenous demand growth or rapid quality improvements unrelated to racing (such as learning by doing or innovations by component suppliers).

The results clarify the role of spin-outs and other entrants. If an entrant innovates in a mature product generation, it licences its innovation to a current market leader. Entrants market and grow large only if they enter new product generations early on and those generations experience the favorable shocks that allow the entrants to maintain their leadership in the face of potential entry from the old generation leaders.

Our results lead us to question certain aspects of U.S. antitrust policy on licensing and policies governing employee non-compete agreements. On antitrust policy, our model suggests that in innovative environments, high product market concentration encourages investment unless the leading firms’ market power prevents potential licensors from appropriating any of the gains from trade from licensing. When the licensor is able to appropriate gains from trade from licensing, high market power in the product market raises potential licensors’ payoffs from innovating. This provides incentives to enter the innovation race and invest and may increase welfare. On non-compete agreements, our model suggests that spin-outs have beneficial welfare impacts, and our empirical results support this view. Spin-outs in the rigid disk drive industry played the two roles described above: market pioneers in new generations and licensors in mature generations.

Current U.S. antitrust policy in innovative environments, as expressed in the 1995 Department of Justice/Federal Trade Commission Antitrust Guidelines for the Licensing of Intellectual Property, tends to implicitly assume that the number of firms in the product market today affects the number of firms in future innovation races. Thus, a horizontal merger between two innovative firms is questioned not just
because it reduces product market competition, but because it reduces competition for future innovations. Our analysis suggests that if spin-out formation is possible then this concern is unwarranted.

Extending the model to incorporate multiple innovations and diffusion would be a useful next step. Our model focuses on innovations that advance the technological frontier within a product generation or pioneer a new product generation. To simplify the analysis, we ignore introductions of less-than-frontier products and improvements that allow laggards to catch up with current industry leaders. We also ignore future innovations. Evidence shows that it is difficult to forecast the success of new product generations, but if it is possible to do so then future payoffs might dominate the short-term concerns we focus on. Incorporating these factors should yield insights into strategy and policy.

References


Gilbert, R.J., Tom, W., 2001. Is innovation king at the antitrust agencies?: the intellectual property guidelines five years later. Antitrust Law Journal 69, 43–86.


