Federal Reserve Bank of Minneapolis
Research Department Staff Report 272

July 2000

Knowledge Diffusion through Employee Mobility

April Mitchell Franco*

University of Iowa
and Federal Reserve Bank of Minneapolis

Darren Filson*

Claremont Graduate University

ABSTRACT

In high-tech industries, one important method of diffusion is through employee mobility: many of the entering firms are started by employees from incumbent firms using some of their former employers' technological know-how. This paper explores the effect of incorporating this mechanism in a general industry framework by allowing employees to imitate their employers' know-how. The equilibrium is Pareto optimal since the employees "pay" for the possibility of learning their employers' know-how. The model's implications are consistent with data from the rigid disk drive industry. These implications concern the effects of know-how on firm formation and survival.

*We would like to thank Kelly Bedard, Robert Brown, Hal Cole, Nancy Gallini, Eric Helland, Tom Holmes, Hugo Hopenhayn, Glenn MacDonald, Matt Mitchell and Edward C. Prescott for their help in developing this paper; Aida Budiman, Shubala Mani, and Bunchon Songsamphant for their research assistance; and James Porter, editor of the *Disk/Trend Report,* for the data set. All remaining errors are our own. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1. Introduction

In existing industry dynamic models (Gort and Klepper (1982) and Jovanovic and MacDonald (1994b)), the mechanism through which knowledge diffuses is abstract. However, there is evidence of a specific mechanism in high-tech industries, where technological knowledge is important and improvements occur continuously and rapidly. Many entrants are actually spin-outs, firms started by a former employee of an incumbent firm. In Christensen's (1993) study of the rigid disk drive industry, he documents approximately 40 spin-outs in a 20 year period. These firms account for approximately 25% of the entering firms. In fact, one firm, Shugart, had seven descendants, and of these, six were in operation in 1991 and included the U.S. original equipment market's four largest firms. This type of activity has also been documented in the semiconductor industry. In the period from 1955 to 1976, at least 29 entering firms had at least one founder who worked for Fairchild Semiconductor (Braun and MacDonald (1982)). Both of these are generic examples of high-tech industries. Here, we propose an industry model that is consistent with the feature that agents can imitate their employer's knowledge and examine its implications. By specifying the particular mechanism, we can study its effects in a perfectly competitive framework similar to Hopenhayn (1992). This would allow for better policy prescriptions as suggested in Irwin and Klenow (1994).

The model is related to Lucas' Span-of-Control (1978), since higher knowledge increases output. However, agents can improve their productivity either by working as a researcher and imitating their employer or by running a firm and hiring researchers to innovate. Essentially, there are two production processes at work: one which produces output and the other which produces knowledge.

The model has implications concerning the relationship between knowledge and firm formation and survival. First, more technologically advanced firms will produce spin-outs. Second, firms with higher technological know-how will survive in the following period with a higher probability than those with lower technological know-how. Finally, a spin-out's probability of survival into its second period depends on the know-how of its parent.

These implications are compared with data from the hard drive industry. The hard drive industry is an industry where technological know-how is important and rapid and where innovation is constant. First, we show that the model is relevant for this industry: spin-outs were the single most important type of entrant in the period 1977-1997 and used know-how learned from their parent firms. There is support for the model's implications on spin-out generation and firm survival. One interesting fact is that firm size is not significantly correlated with the probability of spin-out generation. This suggests that while larger firms may have more employees, only the better ones will generate spin-outs. Spin-out survival is found to be more closely related to some forms of know-how possessed by parent firms than others.

Several puzzling facts about the hard drive industry can be replicated and explained with a simulation of the model. For example, Lerner (1997) establishes that in the 1970's and 1980's, while the industry was expanding, profits were low and rose as the industry matured, even though the price was steadily declining. During this time, technological laggards who were close behind the leaders tended to catch up to leaders, and laggards who were further behind were more likely to exit. The simulation results suggest that the model can explain several of the broad trends that have occurred in the disk drive industry, and other high-tech industries with similar trends, over time.

In contrast to previous models of technological diffusion, despite the fact that technology *spills over* from firms to spin-outs, the competitive equilibrium is Pareto optimal. The employees "pay" for the possibility of imitating their employers' technological know-how, because imitation is locationally specific and the agents who benefit can be identified. In Jovanovic and MacDonald (1994a), imitation depends only on the distribution of know-how in the industry and not on the actions of the individuals. The lack of property rights can lead to an externality and a suboptimal equilibrium. The optimality result presented here is similar to that in Chari and Hopenhayn (1991). Unlike that model, the arrival of new technology is endogenous and depends on actions undertaken by the agents in the economy. This suggests that government policy is unwarranted to increase innovation and social welfare.

The paper is organized in the following manner. The second section presents the model and theoretical results. The third section describes the rigid disk industry and compares the data with the implications of the model. A simulation of the model is discussed in the fourth section. The final section concludes and explores possible avenues for future research.

2. The Model

The model describes the evolution of a single industry in a discrete time, infinite horizon environment. There is a continuum of *ex ante* homogeneous, infinitely lived agents in the industry. Each agent has a level of technological know-how, given by $\theta \in [\theta_L, \theta_H]$. The distribution of know-how at time t is given by $\nu_t(\Theta)$, a probability measure, where ν_0 is given and Θ is the set $[\theta_L, \theta_H]$.

At the beginning of each period, an agent observes his level of know-how and the distribution of know-how within the industry. Each period, each agent decides whether to work outside the industry, work in the industry as a researcher at an existing firm, or operate a firm in the industry. An agent who works outside the industry receives a wage, W^0. This outside wage is constant over time. If an agent works outside the industry, his know-how does not improve.

An agent who works as a researcher must decide for which firm to work. All researchers are assumed to be identical in the innovation production function, and firms differ only by the level of know-how of their founder. A researcher with know-how θ_r who works for a firm with know-how θ_f receives a wage $w(\theta_r, \theta_f, \nu)$. With probability λ, a researcher learns her employer's know-how. If the researcher worked for a firm with a lower level of know-how than her own, she will keep her original level of know-how. Otherwise, the researcher will imitate the firm's know-how and may use it in the following period.

The firm's choice variables are given by the vector (q, l), where q is the quantity produced and l is the innovative effort, given by the measure of researchers hired in each period. Firms produce a homogeneous product and face an inverse demand curve, $D(Q)$, where Q is the aggregate quantity produced in the industry. For simplicity, the demand curve is assumed to be constant over time.[1] The firm's net revenue is given by

$$p(\nu)q - c(q, \theta) - lw(\theta_f, \theta_r, \nu).$$

The price of the good produced by the industry, in equilibrium, is determined by the distribution of knowledge in the industry and the demand curve and is given by $p(\nu)$. The costs are decomposed into the cost associated with production of the good and that associated with innovation. The firm's cost function for production, $c(q, \theta)$, satisfies the following conditions: $c(0, \theta) = 0$, $\frac{dc(0,\theta)}{dq} = 0$, $\frac{dc(q,\theta)}{dq} > 0$, $\frac{dc(q,\theta)}{dqdq} > 0$, $\frac{dc(q,\theta)}{d\theta} < 0$, and $\lim_{q \to \infty} c'(q, \theta) = \infty$,

[1] This model can be incorporated into a general equilibrium model, as in Mitchell (1999), where the demand for the industry's good is unaffected by income and the wages paid outside the industry are constant.

$\forall\theta \in \Theta$. The cost of innovation is the product of the number of researchers a firm hires, l, and the wage rate paid by the firm, $w(\theta_f, \theta_r, \nu)$.

The transition function of the firm's θ is given by a cumulative distribution function $\Psi(\theta'|l,\theta)$ that measures the probability of obtaining future know-how θ' given current know-how, θ, and innovative effort, l. It is *not* dependent on the distribution of agents, ν. The properties of Ψ are

(i) Innovation is not guaranteed. ($\Psi(\theta|l,\theta) > 0$.)

(ii) Innovation is costly. ($\Psi(\theta|0,\theta) = 1$.)

(iii) There is no forgetting. ($\Psi(\theta'|l,\theta) = 0$ if $\theta' < \theta$.)

(iv) $\Psi(\theta'|l,\theta)$ is multiplicatively separable in l and θ. ($\Psi(\theta'|l,\theta) = F(\theta'|\theta)G(l)$, where $G(l)$ is the probability that the firm obtains a new level of know-how given its innovative effort and $F(\theta'|\theta)$ is the cumulative distribution function of the firm's next period know-how given this period's know-how, given that the firm obtains a new level of know-how.)

(v) Increasing effort and know-how improves prospects. (If $\widehat{\theta} \geq \theta$, then $F(\theta'|\widehat{\theta})$ first order stochastically dominates $F(\theta'|\theta)$. If $\widehat{l} \geq l$, then $G\left(\widehat{l}\right)$ first order stochastically dominates $G(l)$.)

(vi) $G(l)$ is concave. (For any two levels of effort, l_1 and l_2, and $\alpha \in [0,1]$, $G(\alpha l_1 + (1-\alpha) l_2)$ dominates $\alpha G(l_1) + (1-\alpha) G(l_2)$ in the first order stochastic sense.)

The first three and fifth assumptions are similar to those of Jovanovic and MacDonald (1994a), but the imitative possibilities are suppressed. This isolates the mechanism through which imitation occurs. The fourth assumption is used only to prove that the probability of survival is weakly increasing in know-how. The final condition on the innovation technology helps to guarantee that firms with the same technological know-how will choose to expend the same effort given the same distribution of know-how, instead of randomizing between different levels of effort.

Imitation between existing firms is not allowed in this model. Imitation occurs through researchers who work within the industry. Firms can only learn by hiring researchers. Researchers supply a homogeneous product to the firms. Any increase in θ is based on the firm's innovative effort, its previous θ and the stochastic innovative shock.

Before the complete agent's problem is presented, the law of motion for the distribution of knowledge is presented.

A. The Law of Motion

The law of motion depends on the actions of the agents in the economy. The knowledge of the agents who work outside the industry is unchanged. So, the future distribution will be unaffected by their actions. In the case of agents who work as researchers within the industry, the distribution will be unaffected by $1 - \lambda$ of these agents who fail to learn their employers' knowledge, while λ of these agents will learn their employers' knowledge and affect the next period's distribution as long as these researchers work for firms with knowledge that is greater than their own. Finally, the plant owners will affect the distribution given their choice of innovative effort.

The law of motion is written formally using the following three subsets. Which agents are members of which subsets is determined by their actions. The measure of agents who become firm owners is denoted by ν_P, the measure of agents who work as researchers within the industry by ν_R and the measure of agents who work outside the industry by ν_W.

Without loss of generality, each firm is assumed to hire only one type of researcher. So, all researchers at a particular firm will have the same level of knowledge. In order to keep account of how many agents are hired by which firms and both the firms' and the agents' type, the function z is used. The function $z(l, \theta_r, \theta_f)$ is the measure of firms with θ_f that hire l units of researchers with θ_r and has the following characteristics:

$$\int_{L\times\Theta} z(dl \times d\theta_r \times \theta_f) = \nu_P(\theta_f)$$

$$\int_{L\times\Theta} l\ z(dl \times \theta_r \times d\theta_f) = \nu_R(\theta_r)$$

$$\int_{L\times\Theta\times\Theta} z(dl \times d\theta_r \times d\theta_f) = \nu_P(\Theta)$$

$$\int_{L\times\Theta\times\Theta} l\ z(dl \times d\theta_r \times d\theta_f) = \nu_R(\Theta).$$

The fraction of agents with knowledge equal to θ is given by $\nu(\theta)$. So, $\nu_P(\theta_f)$ is the measure of plant owners with know-how θ_f, and $\nu_R(\theta_r)$ is the measure of researchers with knowledge θ_r. Recall that Θ is the set $[\theta_L, \theta_H]$.

For any set $A \subset \Theta$,

$$\begin{aligned}\Phi(\nu)(A) = \nu_W(A) + (1-\lambda)\int_{L\times A\times[\theta_r,\theta_H]} l\ z(dl \times d\theta_r \times d\theta_f) \\ + \int_{L\times A\times[\theta_L,\theta_r]} lz(dl \times d\theta_r \times d\theta_f) \\ +\lambda\int_{L\times[\theta_L,\theta_f]\times A} l\ z(dl \times d\theta_r \times d\theta_f) + \int \Psi(A\,|l,\theta)\,dz.\end{aligned}$$

The first branch represents the measure of agents who worked outside the industry and whose know-how was an element of the set A. The knowledge of these agents is unchanged. The second branch simply represents those agents who had knowledge which was an element of the set A, worked as researchers at firms with knowledge that was weakly greater than their own, and failed to learn their employer's knowledge. The third branch is the measure of agents with initial knowledge in the set A and who worked as researchers for firms with knowledge that was worse than their own. Since their knowledge was better than that of their employers, their knowledge in the following period will be unchanged. The fourth branch is the measure of agents who worked in the previous period for firms within the industry with knowledge that was weakly better than their own and imitated their employer's know-how given that their employer's know-how was an element in the set A. The final branch is the measure of those incorporated agents whose knowledge is within the set A, either as a result of innovative effort or because their knowledge in the previous period was an element in the set A and they failed to innovate.

B. The Agent's Complete Problem

The agent's value function is given by a solution to the functional equation

$$V(\theta,\nu) = \max \left\{ \begin{array}{l} W^0 + \beta V(\theta,\Phi(\nu)), \\ \max_{f\in[L,H]} \left\{ \begin{array}{c} [w(\theta,\theta_f) + \beta[\lambda V(\theta_f,\Phi(\nu)) \\ +(1-\lambda)V(\theta,\Phi(\nu))],\ \text{if } \theta_f \geq \theta, \\ [w(\theta,\theta_f) + \beta V(\theta,\Phi(\nu))],\ \text{otherwise} \end{array} \right\}, \\ \max_{(q,l)}\{p(\nu)q - c(q,\theta) - lw(\theta_r,\theta) \\ +\beta \int V(\theta',\Phi(\nu))d\Psi(\theta'\,|q,l,\theta)\} \end{array} \right. \tag{1}$$

where β is the discount factor and $V(\theta,\nu)$ is the value function. The first branch is the return from taking a job outside the industry. In this case, the agent's knowledge is unchanged in the following period. The second is the return from choosing to become a researcher in the industry. In this case, the agent's future knowledge becomes the same as that of his employer with probability λ if the agent's initial knowledge was less than his employer's knowledge. Otherwise, his knowledge remains unchanged. The last branch defines the return from becoming an incorporated agent. Here the agent's future knowledge, θ', is determined by the transition function Ψ.

C. Equilibrium

Equilibrium is determined by agents optimizing and, once they have chosen whether to work as workers or researchers or entrepreneurs, selecting optimal policies such that supply equals demand for the industry's product and for the researchers' product. More formally, equilibrium is a V and $\Phi(\nu)$ with optimal policies $(q, l, \nu_W, \nu_R, \nu_P, z)$ and prices $p(\nu)$ and wages, W^o and $w(\theta_r, \theta_f)$, such that the following are true:

(i) Agents' choices between running a firm, being a researcher and working outside the industry are optimal given their θ and the industry state, ν.

(ii) $z(l, \theta_r, \theta_f)$ is described by the maximizers defined by equation (1).

(iii) $p(\nu) = D[\int q(\theta, \nu) d\nu]$.

(iv) $\nu_{t+1} = \Phi(\nu_t)$.

(v) $\int l \; dz = \nu_R$.

(vi) $\int dz = \nu_P$.

In equilibrium, agents optimize. The price of the product produced by the industry is set equal to the inverse industry demand given the distribution of know-how. The next period's distribution of know-how in the continuum is determined by which firms in the industry innovated in the current period in addition to which researchers imitated their employer's knowledge in the current period, given that these agents are acting optimally. Supply of labor for a particular firm is set equal to labor demanded by the firm type specific wage, and the supply of firms is equal to the demand for firms.

This equilibrium is a special case of the one presented in Jovanovic and Rosenthal (1988). Since there is no aggregate uncertainty and the sufficient conditions for such equilibrium to exist are satisfied, this equilibrium exists.

D. Wage Structure

In this subsection, we consider the structure of wages within the industry and the evolution of both the distribution of knowledge and the price of the industry's product. All proofs are left to the Appendix.

Proposition 1: For any two firms i and j, with know-how θ_i and θ_j, respectively, which hire $l > 0$ of the same type of researchers with knowledge θ_r, such that $\theta_i \geq \theta_r$ and $\theta_j \geq \theta_r$,

$$w(\theta_r, \theta_i) + \beta\lambda V(\theta_i, \Phi(\nu)) = w(\theta_r, \theta_j) + \beta\lambda V(\theta_j, \Phi(\nu)).$$

This Proposition shows that agents who work as researchers would be willing to receive a lower wage for working at a firm with a higher technology compared to a firm with a lower technology. As long as agents can imitate their employer's technology with positive probability and value the future, the lower wages paid by firms with higher technology is compensated by the future return. Firms with the highest available technology offer the highest return to their employers in the future, by increasing their ability to leave with their employer's know-how and become entrepreneurs. If either the agent does not value the future or cannot imitate his employer's technology, then the wages paid by two firms with different technologies would be the same.

Proposition 2: If there are any two agents with θ_r, where one works outside the industry and the other works as a researcher at a firm with θ_f, then

$$w(\theta_r, \theta_f) + \beta\lambda V(\theta_f, \Phi(\nu)) = W^o + \beta\lambda V(\theta_r, \Phi(\nu))$$

where $\theta_f \geq \theta_r$.

Obviously, firms with the highest available technology will pay a wage significantly below the outside wage, as long as researchers have a positive probability of imitating and value the future. This difference is simply equal to the difference between the present expected value of having imitated the firm's know-how and the present expected value of the agent's original level of know-how. If workers do not value the future or cannot imitate their employers' technology, then the wage paid by any firm within the industry must be the same as the outside wage in order to attract workers. This suggests that employees are "paying" their employers for the possibility of learning their know-how in order to imitate it.

Because there is no forgetting for either a firm owner or any worker, the distribution of knowledge in the industry is weakly improving. Further, the distribution eventually approaches a steady state, which depends on the initial distribution. For some initial distributions, there may be no change. The price is weakly decreasing. This is due to the fact that the knowledge within the industry is improving and demand is constant.[2]

E. Optimality of the Competitive Equilibrium

In the language of dynamic programming, this is a stationary model, since the reward function, the discount factor and transition function are all constant and independent of time. The resulting equilibrium is unique in terms of prices and aggregate allocations. However,

[2] For a formal statement of these results, see the Appendix.

while the aggregate allocations are unique, the individual choices may not be. For instance, an individual agent may be indifferent between working for any one of the firms with positive demand for researchers and working at the outside alternative. Where such an agent chooses to work is not important; what is important is the aggregate allocation of agents to firms.

Optimality of the Competitive Equilibrium

The planner's problem is defined. The optimal program maximizes the present expected value of the sum of the consumer surplus and the producer surplus. Again, the single firm is assumed to be a price taker. The solution is a W and $\Phi(\nu)$ with the aggregate optimal policies $\{\nu_{Pt}, \nu_{Rt}, \nu_{Wt}, z_t\}_{t=0}^{\infty}$ such that

$$W(\nu) = \max \sum_{t=0}^{\infty} \beta^t \left(S\left(\int q(\theta, \nu_t)\, d\nu_{Pt} \right) - \int c(q, \theta)\, d\nu_{Pt} - W^o \nu_{Rt} - W^o \nu_{Pt} \right)$$

$$\text{s. t. } \{\nu_{Pt}, \nu_{Rt}, \nu_{Wt}, z_t\}_{t=0}^{\infty} \text{ is feasible given } \nu_0$$

where S is the consumer surplus given the aggregate output, $\int c(q, \theta)\, d\nu_{Pt}$ is the aggregate cost of producing the aggregate level of output, $W^o \nu_{Rt}$ is the opportunity cost of hiring researchers instead of allowing them to work at the outside option, and $W^o \nu_{Pt}$ is the opportunity cost of hiring managers for the plants.

Proposition 3: The competitive equilibrium is optimal and unique.

Since firms have property rights to their knowledge and there is no asymmetric information about firms' knowledge, the diffusion externality is internalized and the competitive equilibrium is optimal. In Jovanovic and MacDonald (1994a), the firm's learning technology is a function of the distribution of knowledge within the industry. As a result, the equilibrium is suboptimal. Firms would choose to engage in a lower than optimal level of learning because of the lack of property rights over knowledge.

Since there is no asymmetric information about firms' knowledge, the wage paid to researchers is a decreasing function of the firms' knowledge. This contrasts with Jovanovic and Nyarko (1995), where the wage paid by firms is the same. If this model allowed for asymmetric information over firms' knowledge, wages would not depend on the firm's knowledge, only the average level of knowledge across firms.

One further statement can be made about the equilibrium. Because λ represents the cost or barrier to imitation, we can say the following.

Corollary 1: Social welfare is increasing in λ.

This is a direct result of the fact that the optimality of the competitive equilibrium is independent of λ. If the social planner were allowed to pick the value of λ, she would set λ equal to one, since this would imply that there were no barrier to imitation. Diffusion of current knowledge would be immediate.

F. Firm Generation and Survival

Any agent who has not incorporated and has failed to improve his knowledge will not incorporate in the future. This is shown by using the planner's solution. Since an agent who either operated a firm or imitated her employer's knowledge can produce output more efficiently and improve the distribution of knowledge more than one who did neither, the planner would choose the former as a firm owner. Given this, there is a critical value, $\tilde{\theta}$, such that any agent with knowledge greater than this critical value will operate a firm. Because the distribution of knowledge is improving, the critical value is increasing over time. The critical value helps to characterize the value function. The value function is constant for any level of knowledge below the critical value and is increasing above it.[3]

The next three results are compared with the data from the rigid drive industry in the following section. The first result shows that firms with greater technological know-how are more likely to generate spin-outs.

Proposition 4: There exists a critical value, $\tilde{\theta}'$, such that for any firm i with $\theta_i \leq \tilde{\theta}'$, none of its researchers will become a firm owner in the following period and any firm j with $\theta_j > \tilde{\theta}'$, $\lambda l(\theta_j)$ of its researchers will become firm owners.

The next result shows that firms with higher know-how are more likely to survive.

Proposition 5: The probability that an agent who currently operates a firm will continue to operate a firm in the following period is weakly increasing in his know-how.

Even though the likelihood of survival is increasing in know-how, leapfrogging is still possible in the model because of the stochastic learning technology. Leapfrogging is discussed further below in Section 4.

The final result shows that the probability of a spin-out surviving beyond its first year is increasing in its parent's know-how.

Corollary 2: The probability that an agent who imitates his former employer's technology and starts up a firm will operate a firm in the following period is weakly increasing in his former employer's know-how.

[3] For a formal statement and proof, see the Appendix.

3. The Rigid Disk Drive Industry

A. The Data

The main data source is the *Disk/Trend Report* on Rigid Disk Drives (Porter (1977-1997)). The data set contains 192 firms, 1190 firm/year observations, and 11,644 model/year observations. The data include product characteristics and introduction dates. Annual sales of disk drives are reported for several firms.[4] Information on the backgrounds of founders of new firms is provided, and historical information and recent news are summarized for each firm. To determine spin-out-parent relationships, the histories from the *Disk/Trend Report* were supplemented with company press releases and articles provided by James Porter, the editor of the *Disk/Trend Report.* Other sources include the *Directory of Corporate Affiliations,* the *International Directory of Company Histories,* and Christensen (1993).

There are 40 cases of one or more employees leaving one or more rigid disk drive manufacturers to found a new firm in the period 1977-1997. Table 1 sorts the spin-outs by year of entry and lists the parent firms, the founding year of the spin-out, and the spin-out's life span and mode of exit.[5] To determine the parent firms, we focus on the background of the founders and not on other employees, for which data are unavailable. The implicit assumption is that founders had considerable influence on the products and strategies of the start-up; evidence from company press releases and the *Disk/Trend Report* supports this assumption. We categorize mode of exit three ways: firms are still active as of 1997, have been acquired, or have exited due to failure. The distinction between acquisition and failure

[4] 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.

[5] The analysis uses data starting in the late '70's, after the industry was well into its takeoff stage. All of the non-captive parent firms of the early start-ups in the data were also spin-outs (Christensen (1993)). Memorex, Pertec, and Storage Technology Corporation were IBM spin-outs, Shugart Associates was a Memorex spin-out, and Tandon was a Pertec spin-out.

is important because acquired firms are typically not failing. In our analysis below we treat exits due to acquisitions as censored observations.

In order to test the theoretical implications on spin-out generation and firm survival, we use the available data to construct two measures of know-how. *Technical know-how* measures the firm's technical expertise using areal densities. The areal density is the main measure of drive quality; it measures how much information can be stored on each square inch of disk. The areal density of the firm's best drive in each diameter in each year is divided by the highest areal density in that diameter in that year to generate a measure of the firm's know-how in each diameter relative to the best available know-how in that diameter.[6] Then this measure is averaged across diameters to obtain a single measure of the firm's technical know-how in each year. The firm-level measure is necessary because the theoretical results pertain to firm-level decision-making.[7] *Early mover know-how* is a dummy variable for firms that introduced a drive of a new diameter within the first year that drives of that diameter were shipped. Early mover know-how is a proxy for the product design, product reliability, and marketing know-how associated with designing, manufacturing, and marketing new drives. Only the major diameters introduced in 1977-1997 are considered: the 8", 5.25", 3.5", 2.5", and 1.8" drives.

Intellectual property rights are one issue that has not been specifically addressed in the model. Empirically, institutional barriers to imitation appear to be low in the disk drive

[6] Only drives that have been shipped are used when making these calculations. Drives that have been announced but not yet put into production are not included. We assume that improvements in technical know-how are rapidly embodied in new products. Lerner (1997) argues convincingly that this is the case in the disk drive industry.

[7] Lerner treats each diameter separately in most of his analysis but reports some results using this average measure.

industry. Lerner provides evidence that patents were not widely used to protect key aspects of drive technology, and examples in the *Disk/Trend Report* show that when patents were used, licensing was widespread. Further, covenants not to compete and trade secret laws were largely ineffective. Most of the firms in this study were located in California, where covenants not to compete were prohibited by law and not enforced by the courts. Trade secret laws did not create much of an employee mobility barrier because of large contract negotiation costs, difficulties with enforcing the laws, and the Silicon Valley culture.[8]

B. The Industry's History

We refer the interested reader to Christensen (1993, 1997), Lerner (1997), and the *Disk/Trend Report* for more complete descriptions of the industry's history. The industry began in 1956 when IBM introduced the first rigid disk drive. Followers began entering in the 1960's and were of two main types. Captive producers, such as Burroughs, Control Data, and Univac, were vertically integrated computer manufacturers that produced drives for in-house use. Plug-compatible market (PCM) firms were independent drive producers that made drives that were plug-compatible with IBM's computers. PCM firms sold drives directly to users of IBM computers. Christensen (1993) reports that many of the early PCM firms were IBM spin-outs, including Century Data, Memorex, Pertec and Storage Technology Corporation. When the minicomputer market began growing rapidly in the mid 1970's, an original equipment market (OEM) emerged. OEM firms served as either primary or secondary sources of drives for computer manufacturers.

Innovation and imitation in the disk drive industry occurred at an extremely rapid

[8] Gilson (1998) and Saxenian (1994) discuss covenants not to compete and trade secret law in the Silicon Valley environment.

rate from 1956 to 1997 and took several forms. First, several improvements in technical features improved capacities and access times. Second, several improvements in design and manufacturing techniques improved costs and reliability. Third, several architectural innovations occurred: drives with smaller diameters were introduced beginning with 8" and 5.25" drives in the late '70's and continuing with 3.5", 2.5", and 1.8" drives. When first introduced, the new drives served new buyers: 8", 5.25", 3.5", 2.5", and 1.8" drives were first used in minicomputers, personal computers, portable computers, notebook computers, and smaller portable devices, respectively. In response to the profit opportunities generated from rapid technological change and market growth, net entry occurred. Firm numbers continued to rise until the mid '80's and then leveled off a short time before falling in the early '90's (Lerner). The patterns for net entry and firm numbers are similar to those established in industries with new products in Gort and Klepper (1982).

Spin-outs: Importance and Imitation

In the model spin-outs are the only source of entry. Focusing on U.S. disk drive firms in the period 1976-1989, Christensen (1993) shows that while spin-outs were not the only source of entry, they were definitely the most important source. Only three out of 28 non-spin-out entrants survived until 1989, but 16 out of 40 spin-outs survived. Spin-outs accounted for all but four of the start-ups that were successful at generating revenue and accounted for 99.4 percent of the total cumulative revenues generated by the start-up group. By 1989 seven of the world OEM/PCM market's ten largest firms were spin-outs. Our data show that after 1989 only five spin-outs and two non-spin-outs entered. This implies that Christensen's detailed analysis describes the vast majority of the entrants.

Spin-outs were an important source of entry, but were they imitators? Did spin-outs imitate firm-specific know-how of their former employer, as in the model, or did they simply learn industry-specific know-how?[9] The *Disk/Trend Report* describes several examples of firm-specific technical know-how being imitated. For example, founders of Amcodyne and Areal Technology learned how to make high areal density drives from their parent firms, founders of Dastek used thin film head technology after learning from IBM, and founders of Tecstor modeled their drives after their parent's drives.

Early mover know-how was also imitated. Table 2 lists early movers by diameter. Almost all of the firms listed are either spin-outs, parents, or both, with the exception of BASF, New World Computer, and Control Data.[10] Many of the firms are related to each other. Table 3 lists the spin-outs from early mover parents along with whether the spin-out was an early mover and, if so, in which diameter. From Table 3, the probability that a randomly selected spin-out from an early mover parent is an early mover is $\frac{5}{15} = .33$. Of the 177 firms that were not spin-outs from early-mover parents, only 12 were early movers, resulting in a probability of .068. The two probabilities differ substantially, and the difference is significant at the 1% level: the t statistic is 3.5 and the critical value is 2.33.[11]

Other types of know-how related to product reliability, low cost, and marketing, which

[9] Some alternative models from the labor economics literature, though not specifically developed to explain spin-outs, may provide some possible alternative explanations of spin-out formation. The model presented here is similar to a stepping stone mobility model, in which an agent works at one firm and acquires skills that allow him to move up the career ladder, possibly at another firm. In contrast, a standard matching model suggests that employee departures occur as a result of a bad match. If such a model was used to describe spin-out formation, there would be no connection between the parent firm's know-how and the spin-out's know-how.

[10] Many of these firms were extremely successful. International Memories, the first mover in 8" drives, became one of the most prominent OEM manufacturers in the early '80's. Seagate, the first mover in 5.25" drives, rapidly became the most prominent OEM firm and continued to hold this position as of 1997.

[11] If we include Syquest, a spin-out of Seagate and a pioneer of the small disk cartridge drive market, as an early mover, the point estimates become even more compelling. Syquest is excluded because our test includes only early movers in the main diameters.

cannot be measured using our data, also appear to have been imitated. Conner Peripherals, Seagate, and Quantum are prominent examples of spin-outs with strengths in these areas whose parents had similar strengths.

In contrast to our model, in some cases spin-outs were both imitators and innovators. A particularly striking example of this is that the first firms to introduce the major new diameters were all spin-outs: International Memories, Seagate, Rodime, PrairieTek, and Integral Peripherals. Christensen points out that it may have been easier for spin-outs to find new customers, convince them to buy new drives, and maintain focus on a new, small market because spin-outs did not face the same opportunity costs as their larger parents who were focused on existing markets. This is an interesting departure from the model that could be explored in future work.

Spin-out Formation

According to the model spin-outs come from firms with relatively high know-how. This hypothesis is tested using several probit models. The dependent variable is a dummy variable that takes the value 1 if the firm generates a spin-out in the current year. In the model the probability of a firm generating a spin-out in period t depends only on the firm's know-how in period $t-1$ and the distribution of know-how. Therefore, the independent variables are lagged values of technical know-how and early mover know-how. We use year dummies to account for changes in the distribution of know-how from year to year.[12] Summary statistics are presented in Table 4.

The list of spin-out parents in Table 1 includes many of the most successful disk drive

[12] Note that year dummies capture all changes in industry-level variables from year to year. In the model, all industry-level changes depend on changes in the distribution of know-how. We use 1983 as a base year.

firms. If success was due to "know-how," however defined, then these firms clearly had high know-how. Thus, even without using explicit know-how measures, the hypothesis appears to have some support. The statistical analysis confirms this impression. The estimation results are reported in Table 5. In equation 5a both know-how coefficients are positive and significant.[13] This supports the hypothesis. The magnitude of the effects are quite large: at the mean values of the data, obtaining early mover know-how raises the probability of generating a spin-out from 0.13 to 0.19, and raising technical know-how from 0.25 to 0.75 raises the probability of generating a spin-out from 0.11 to 0.18.

Equation 5b is similar to equation 5a but includes two additional control variables: lagged sales growth and the lagged number of drives produced.[14] Including these variables allows us to test two alternative explanations for spin-out formation. Lagged sales growth allows us to test whether spin-outs formed as a result of employee exit from a failing firm.[15] The coefficient on lagged sales growth is positive, which suggests that spin-outs are more likely to come from firms that are doing well in the market rather than those that are retrenching

[13] This was checked for robustness by limiting the sample to U.S. firms. The motivation for considering only U.S. firms is that only U.S. firms generated spin-outs in the disk drive industry. It is likely that institutional differences between the U.S. and Japan, the country in which most foreign firms were based, made spin-out generation more likely in the U.S. The results are essentially unchanged.

[14] Sales growth is computed as follows:

$$g_t = \frac{s_t - s_{t-1}}{\frac{s_t + s_{t-1}}{2}},$$

where g_t denotes the growth rate and s_t denotes firm sales in period t. This formula ensures that g_t is finite if either s_t or s_{t-1} is 0. For new entrants that have 0 sales in two adjacent periods, g_t is set equal to zero. If a spin-out has two parents, the average of the parents' sales data is used.

[15] This test is partly motivated by a few cases in Table 1 in which spin-out formation occurred when a once prominent parent suddenly declined. In 1985-86, Computer Memories lost its largest customer when IBM decided to supply more of its needs in-house. As Computer Memories declined employees left and two spin-outs were formed: Peripheral Technology and Brand Technologies. In another case when Lapine failed after its brief success, employees abandoned it and founded new firms: Comport and Kalok. The learning described by the model may still have been present, but the departure was partly forced rather than entirely voluntary.

or declining. The lagged number of drives is a proxy for firm size. A simple hypothesis about spin-out generation is that spin-outs are more likely to come from larger firms simply because larger firms have more employees who can leave. This result does not support this simple hypothesis; the coefficient on lagged number of drives is negative and insignificant.[16]

Firm Survival

The second implication of the model is that the probability of a firm surviving until the following period is increasing in its current know-how. This hypothesis is tested using several probit models. If a firm exits in period t because it is acquired, the exit year is treated as censored; only failures count as exits.

The estimation results are reported in Table 6. In equation $6a$ both know-how coefficients are positive.[17] This supports the hypothesis. The magnitudes of both know-how effects are substantial. At the mean values of the data, obtaining early-mover know-how raises the probability of surviving from 0.78 to 0.82, and raising technical know-how from 0.25 to 0.75 raises the probability of surviving from 0.75 to 0.84.

In the theoretical model, once spin-outs have been formed they are assumed to evolve according to the same transition rules as other firms. To confirm that this is a reasonable assumption, Table 7 repeats the regression from Table 6 including only spin-outs. Equation $7a$ shows that the results are similar to Table 6. In equation $7b$, parent know-how measures are included as control variables. In the model, only the current know-how and the distribution of know-how affect firm survival. Clearly, this assumption is violated in the data because the coefficients on parent know-how are both significant. Parent know-how appears to have

[16] This was checked for robustness by limiting the sample to U.S. firms. The results were unaffected.
[17] The robustness is checked by limiting the sample to U.S. firms. The results are essentially unchanged.

persistent effects. Interestingly, the coefficient on parent technical know-how is negative. We discuss this result further below.

Spin-out Survival

The final implication states that a spin-out's likelihood of surviving beyond its first period is increasing in its parent's know-how. Unfortunately this does not lead to an interesting test, because all the spin-outs in the data except two survived beyond their first year. Instead, a more general hypothesis is tested: a spin-out's expected lifetime is increasing in its parent's know-how. This is done using duration models in which the spin-out's lifetime is a function of its parent's know-how. Summary statistics on spin-out lifetimes, parent know-how, and the other included variables are provided in Table 8.

Several duration models were estimated; the best fit was obtained from a Weibull survival function of the form

$$\exp(-(\phi t)^{\frac{1}{\sigma}})$$

where

$$\phi = \exp(-\beta' x_i).$$

The resulting hazard function, which gives the probability that a firm exits given that it has survived until time t, is given by

$$\frac{\phi}{\sigma}(\phi t)^{\frac{1}{\sigma}-1}.$$

The parameters β and σ are to be estimated, and x_i represents firm i's parent's know-how.[18]

[18] We treat spin-outs that were still alive in 1997 and those that were acquired before 1997 as censored observations. While checking the robustness of the results, we estimated a Markov chain model that allowed

Estimation results are reported in Table 9. In equation $9a$ only the two parent know-how measures are included. Although the signs are the same as in equation $7b$, both coefficients are insignificant. In equation $9b$ more precise estimates are obtained using a 3-year average of parent technical know-how. The coefficient is still negative, while the coefficient on early mover know-how is positive. The results confirm the conclusions from equation $7b$: spin-out survival is decreasing in parent technical know-how. However, as shown in Table 7, the probability of a spin-out surviving is increasing in its own technical know-how. The results suggest that technical know-how was more difficult to imitate than early mover know-how. Spin-outs that came from firms with high technical know-how were less likely to imitate successfully and therefore were less likely to survive, but if they were successful at learning this type of know-how, they were more likely to survive. Christensen's (1993) analysis supports this conclusion. Many of the advances that improved areal densities were extremely expensive and time-consuming to develop, and only the large established firms were successful with these development projects. New small firms that tried had an extremely high failure rate.

In equation $9c$ parent sales growth is added as an explanatory variable. As mentioned above, some spin-outs were formed when the parent was failing. The estimates of equation $9c$ show that spin-outs from failing firms were less likely to have long lives than those from growing firms. The coefficients on parent know-how do not change substantially.

Another type of spin-out formation involved entrepreneur mobility: in several cases

for the two types of censoring explicitly. The results did not change: it appears that general statements about how the probability of being acquired depends on know-how and our other controls cannot be made. This conclusion makes sense given the history of the industry. All types of firms have been acquired, including new small firms still in the development stage, large successful firms, and failing firms have occurred throughout the life cycle.

in Table 1, one or more of the spin-out founders were also founders of the parent firm.[19] This suggests that expertise in founding start-ups, as well as other types of know-how, was useful. This entrepreneurial know-how is less likely to diffuse in the manner featured in the model because it is more likely to be obtained from experience at founding start-ups than from working for other firms. In equation $9d$, we include an *entrepreneur dummy*. If one of the spin-out founders was also a founder of the parent firm, this dummy takes the value of 1. Interestingly, the coefficient on the entrepreneur dummy is negative. Although the estimate is imprecise, it suggests that past experience at founding start-ups may have a negative impact on the lifetime of a new start-up. This result is similar to that documented in the PC software industry by Prusa and Schmitz (1994). This may be the case in rapidly evolving industries: past experience at founding a start-up may not be as important as having the right know-how for the current environment.

4. Simulation

In this section, we simulate the model to show how it reconciles three main facts about the disk drive industry's evolution described by Christensen (1993, 1997) and Lerner (1997). First, entry and spin-out formation peaked in the early '80's. Second, industry profits were low in the late '70's and later rose in the '80's and '90's as the market matured, even though the price per megabyte was declining. Third, during this period laggards had a systematic tendency to innovate more than leaders.

[19] These spin-outs were Micropolis, Irwin International, Seagate, Applied Information Memories, Maxtor, Syquest, Epelo, Brand Technologies, Conner Peripherals, PrairieTek, Areal Technology, and Ecol.2. In two cases, Irwin International and Seagate, the spin-out founder sold the parent firm before founding the new firm. In the case of Brand Technologies, the spin-out founder left a failing firm to found a new one. In all of the other cases, the founder left a viable firm to found a spin-out.

Functional forms and parameter values used in the simulation were chosen to roughly match the broad trends in entry and profits in the data. Figures 1 through 8 graph various simulated series. There are three levels of know-how: low-tech, medium-tech, or high-tech, denoted by θ_l, θ_m, and θ_h, respectively. This isolates imitation: low-tech researchers can imitate medium-tech firms, but no other imitation occurs. No agent can improve his know-how by working for a low-tech firm, and the parameter values ensure that high-tech firms do not hire researchers.[20]

We assign the following values: $\theta_l = 1$, $\theta_m = 4$, $\theta_h = 5$, $W^0 = 0.15$, $\beta = 0.9$, and $\lambda = 0.1$. The production cost function is quadratic in output:

$$c(q;\theta) = \frac{q^2}{2\theta}.$$

The market demand function is linear:

$$Q = 2 - 2.5p.$$

The firm's transition function is specified as follows. Firms obtain a new θ with a probability that depends on their labor usage, determined according to the function

$$\begin{array}{ll} 0.4l^{0.9} & \text{if } 0.4l^{0.9} \leq 1, \\ 1 & \text{otherwise,} \end{array}$$

where l represents the firm's labor choice. Low-tech firms that obtain a new θ become medium-tech agents with probability .5, become high-tech agents with probability .1, and

[20]Because high-tech firms have the highest possible know-how, they have no incentive to hire researchers. However, under some parameter values equilibrium can involve all non-high-tech agents working for the high-tech firms for free in order to have the chance to imitate. In this case imitation occurs, but no innovation occurs.

remain low-tech agents otherwise. Medium-tech firms that obtain a new θ become high-tech agents with probability .5 and remain medium-tech agents otherwise.

Figure 2 shows that as know-how improves, net entry occurs and reaches its peak by the tenth period; this roughly matches the trend in net entry in the disk drive industry. In fact, the trend in net entry matches up well with that established in Gort and Klepper (1982), where they assembled data on the historical development of 46 new products. There, entering firms come with knowledge from outside the industry. Even with a significantly different mechanism for diffusion, the trends are still apparent. Figure 3 shows the percentage of agents that work as researchers. Low-tech researchers at medium-tech firms imitate at a rate of 10% and start up new medium-tech firms in the following period. Most spin-outs are formed in the fourth through tenth periods. These periods correspond to the early to mid 1980's, when most of the disk drive spin-outs were formed.

The simulation also matches the evolution in prices and profits in the hard drive industry. Figure 5 shows that as know-how improves, the price and average cost fall and the market quantity rises. This matches the pattern of falling cost per megabyte that occurred in the industry. We report two profit series that allow for different R&D accounting methods.[21] Figure 4 shows that R&D expenditures per firm, which depend on how many researchers are employed, are initially high and fall over time. Figure 6 graphs revenue minus production costs and R&D expenses, and figure 7 graphs revenue minus production costs. The basic pattern of rising average profits occurs in both series. Gort and Klepper (1982) claim that periods of high net entry, falling price, and rising quantity are typically associated with innovations

[21] Much R&D activity in firms in rapidly evolving industries is not considered separately from other costs; every employee may play a role in improving products and processes. Therefore, reported profits likely include some of what is R&D in the model as part of production costs.

generated outside the industry. The simulation demonstrates that this is not always the case; all of the entrants are spin-outs that obtain their know-how from their parent firms.

The results also match the pattern of laggards innovating more rapidly. Low-tech firms become high-tech at faster rates than medium-tech firms do in periods 2 and 3. This does not always occur; in periods 4 and 5, medium-tech firms innovate at faster rates. Leapfrogging can occur in the model because the firm's learning technology is stochastic and depends on investment levels: if low-tech firms invest much more than medium-tech firms, then they innovate at higher rates. Low-tech firms have higher costs of innovation because they must pay higher wages to researchers but may have higher expected benefits, because in comparison to medium-tech firms their values are much lower than those of high-tech firms. This can be seen in figure 8, which shows the value of agents by type. The difference in value between low-tech and medium-tech agents is always larger than that between medium-tech and high-tech agents.

5. Conclusion

Nearly 25% of the entering firms in the rigid disk drive industry were spin-outs (see Christensen (1993) and Porter (1977-1997)). This is a significant and important portion of entering firms. In order to understand the effects of this imitation mechanism on industry dynamics, a model was developed.

The model provides insight into the role that this particular mechanism plays in an industry's evolution and, along with the empirical and simulation results, challenges some long standing ideas. For example, the empirical evidence shows that, at least in the disk drive industry and possibly in other industries, existing firms provided a training ground for employees who later left to found new start-ups. This contrasts with Gort and Klepper's (1982) model, where new firms used technology from outside the industry. The theoretical implications that were compared with the data include new firm generation, firm survival and spin-out survival and their relation to know-how. The data show that firms with higher know-how are more likely to survive and to generate spin-outs, which is in agreement with the model. One surprising fact is that firm size is not a good predictor of spin-out generation. Finally, parental technical know-how is not a good predictor of spin-out survival, while early mover know-how is. This may capture the difficulty of imitating technical know-how.

In contrast to other models of diffusion, such as Jovanovic and MacDonald (1994a), the resulting competitive equilibrium is Pareto optimal. This result establishes that the equilibria of models of diffusion with active learning can be efficient when imitation is allowed. It also suggests that public policies that affect employee mobility could have important effects on firm entry and technological diffusion in industries where know-how is an important factor of production. Since social welfare is increasing with the probability of imitation, these policies may ultimately have a detrimental effect on social welfare. One possible avenue for future research would be to structurally estimate this model to see how different the probability of imitation is between Massachusetts and California, where the laws regarding employee mobility differ greatly. This could also provide us with some insight into how much this imitation possibility effects welfare.

While this paper only compared the model's implications with data from the rigid disk drive industry, there is reason to believe that the employee mobility model captures some of the salient features of other high-tech industries. The semiconductor and computer software industries are typified by high employee mobility (See Braun and MacDonald (1982), SEMI

(1986) and Wilson, Ashton and Egan (1980)). A stylized fact from both of these industries is that better firms tend to generate more spin-outs. This suggests that there are similarities between these industries and the hard drive industry. As in the rigid disk drive industry, the price of semiconductors has been decreasing over time, while the profits in the industry have increased. This can be explained using a simulation of the model. This suggests that this model is general enough to provide a framework to understand other high-tech industries.

Appendix

Proof of Proposition 1: All researchers are a homogeneous input in innovative effort. Consider a researcher m with knowledge θ_m working at an arbitrary firm i with know-how θ_i, where $\theta_i \geq \theta_m$. The worker's return for working at firm i is given by

$$w\left(\theta_m, \theta_i\right) + \beta[\lambda V\left(\theta_i, \Phi\left(\nu\right)\right) + \left(1-\lambda\right) V\left(\theta_m, \Phi\left(\nu\right)\right)].$$

In order for worker m to be weakly indifferent between working for firm i and any arbitrary firm j in the industry with know-how $\theta_j \geq \theta_m$, the following must hold:

$$w(\theta_m, \theta_i) + \beta[\lambda V(\theta_i, \Phi(\nu)) + (1-\lambda)V(\theta_m, \Phi(\nu))] \geq$$
$$w(\theta_m, \theta_j) + \beta[\lambda V(\theta_j, \Phi(\nu)) + (1-\lambda)V(\theta_m, \Phi(\nu))] \; \forall j \neq i.$$

This simplifies to

$$w(\theta_m, \theta_i) + \beta\lambda V(\theta_i, \Phi(\nu)) \geq w(\theta_m, \theta_j) + \beta\lambda V(\theta_j, \Phi(\nu)).$$

Next, consider the case of a researcher n working at an arbitrary firm j with θ_j, where $\theta_j \geq \theta_n$. Like researcher m, the following condition must be satisfied for him to be weakly indifferent between working at that firm j and an arbitrary firm i with $\theta_i \geq \theta_m$:

$$w(\theta_n, \theta_j) + \beta[\lambda V(\theta_j, \Phi(\nu)) + (1-\lambda)V(\theta_n, \Phi(\nu))] \geq$$
$$w(\theta_n, \theta_i) + \beta[\lambda V(\theta_i, \Phi(\nu)) + (1-\lambda)V(\theta_n, \Phi(\nu))] \; \forall i \neq j.$$

Again, this simplifies to

$$w(\theta_n, \theta_j) + \beta\lambda V(\theta_j, \Phi(\nu)) \geq w(\theta_n, \theta_i) + \beta\lambda V(\theta_i, \Phi(\nu)).$$

Recall that by assumption, both of these researchers have the same knowledge. By replacing θ_m and θ_n with θ_r, these two conditions imply Proposition 1, since both i and j are arbitrary.□

Lemma 1a: The value function is non-decreasing in θ.

Corollary 1a: The wage $w(\theta_r, \theta_f)$ is non-increasing in θ_f.

Proof of Proposition 2: First, consider the case where a firm with θ_f hires a researcher with θ_n, where $\theta_f \geq \theta_n$. In order for the researcher to strictly prefer working for the firm rather than working at the outside alternative, the following condition must hold:

$$w(\theta_n, \theta_f) + \beta \left[\lambda V(\theta_f, \Phi(\nu)) + (1-\lambda) V\left(\theta_n, \Phi(\nu)\right)\right] > W^o + \beta V(\theta_n, \Phi(\nu)).$$

This simplifies to

$$w(\theta_n, \theta_f) + \beta\lambda V(\theta_f, \Phi(\nu)) > W^o + \beta\lambda V(\theta_n, \Phi(\nu)).$$

Second, consider the case where an agent m with θ_m finds that the outside opportunity is more attractive than working for a firm with θ_f, where $\theta_f \geq \theta_m$. This implies that

$$w(\theta_m, \theta_f) + \beta \left[\lambda V(\theta_f, \Phi(\nu)) + (1-\lambda) V\left(\theta_m, \Phi(\nu)\right)\right] < W^o + \beta V(\theta_m, \Phi(\nu)).$$

This simplifies to

$$w(\theta_f) + \beta\lambda V(\theta_f, \Phi(\nu)) < W^o + \beta\lambda V(\theta_n, \Phi(\nu)).$$

Recall that by assumption, the agents' levels of know-how are equivalent. By replacing both θ_n and θ_m with θ_r, these two conditions imply the proposition.□

Proposition 1a:

(i) $\Phi(\nu)$ weakly dominates ν.

(ii) Given ν_0, the equilibrium sequence, $\{\nu_t\}$, converges to a distribution, ν^*.

Proof of Proposition 1a: (i) Recall that $\Phi(\nu)$ is given by

$$\begin{aligned}\Phi(\nu)(A) = \nu_W(A) &+ (1-\lambda)\textstyle\int_{L\times A\times[\theta_r,\theta_H]} l\ z\left(dl \times d\theta_r \times d\theta_f\right) \\ &+ \textstyle\int_{L\times A\times[\theta_L,\theta_r]} lz\left(dl \times d\theta_r \times d\theta_f\right) \\ &+\lambda \textstyle\int_{L\times[\theta_L,\theta_f]\times A} l\ z\left(dl \times d\theta_r \times d\theta_f\right) + \int \Psi\left(A\,|l,\theta\right) dz\end{aligned} \tag{2}$$

for some arbitrary set $A \in [\theta_L, \theta_H]$. Further, ν can be decomposed into

$$\nu(A) = \nu_P(A) + \nu_R(A) + \nu_W(A).$$

Next, it follows that

$$\nu_P(A) \geq \int_A \Psi[A\,|q(\theta,\nu), l, \theta] dz.$$

Then

$$\lambda\nu_R(A) \geq \lambda \int_{L\times[\theta_L,\theta_f]\times A} l\ z\,(dl \times d\theta_r \times d\theta_f).$$

Further, the first three branches of (2) are constant, thus implying that

$$\nu(A) \geq \Phi(\nu)(A)$$

in the first order stochastic sense.□

(ii) There exists a monotone sequence of distribution functions underlying $\{\nu_t\}$ called F_t with the following property for some θ_L and θ_H:

$$F_t(\theta_L) = 0 \text{ and } F_t(\theta_H) = 1, \text{ for } t = 1, 2, ...,$$

and by Corollary 2 to Theorem 12.9, (Helly's Theorem) in Stokey, Lucas, Prescott (1989), there exists a distribution function F with

$$F(\theta_L) = 0 \text{ and } F(\theta_H) = 1,$$

and $\{F_t\}$ converges weakly to F.□

Proposition 2a: The price for the industry's product must decrease, at least weakly, eventually, such that $p(\nu_0) \geq p(\nu^*)$.

Proof of Proposition 2a: Suppose that at time T, when the distribution is sufficiently close to ν^*, the price is above the price in the initial period, or

$$p(\nu_T) > p(\nu_0).$$

The distribution of knowledge is always improving, at least weakly, by part (i) of Proposition 1a. Then, since the cost of production is decreasing in knowledge, this implies that an individual firm's output at T cannot be lower than that at time 0. Further, if the price at time T is above the initial price, agents who were not operating firms in the initial period will find it profitable to enter and produce output. These two imply that aggregate output will be weakly greater at time T than in the initial period. Demand is constant and downward sloping, so an increase in aggregate output will decrease the price, which provides a contradiction.□

We show that the competitive equilibrium defined above is optimal. It is similar to the methodology presented in Stokey, Lucas and Prescott (1989). First, we define the single firm problem in which an *aggregate* firm maximizes producer surplus. Instead of individual agents making decisions as to whether or not to operate a firm, where to work, how much to produce and how much learning effort to engage in, a single firm which treats these agents as assets maximizes the total value of all these assets. This is done to eliminate the wages paid by firms within the industry; only the price of the industry's output and the outside wage are important. We show that the allocations from the single firm equilibrium are equivalent to the aggregate allocations from the competitive equilibrium. This is done in Lemma 2a.

Next, the planner's problem is defined explicitly. The planner maximizes both the consumer and the producer surplus. Then, the competitive equilibrium is shown to be optimal and unique in terms of price and aggregate allocations. In following the steps presented in

Stokey, Lucas and Prescott (1989), the conditions that satisfy both the single firm problem and the planner's problem are shown to be equivalent. Since the competitive equilibrium and the single firm equilibrium are equivalent, the competitive equilibrium is also optimal. The equilibrium is unique, because demand is downward sloping and single-valued and the cost function is decreasing in θ.

Let profit function for the 'aggregate' firm be given by

$$\Pi(p, \nu_P) = \max_q \int_\Theta (pq(\theta) - c(q(\theta), \theta))\, d\nu_P(\theta),$$

where the single firm is assumed to be a price taker with no monopoly power. The single firm problem maximizes its total expected value over time given the law of motion of the distribution of knowledge. It chooses a sequence of allocations $\{\nu_{Pt}, \nu_{Rt}, \nu_{Wt}, z_t\}_{t=0}^{\infty}$ along with the price sequence, $\{p_t\}_{t=0}^{\infty}$, which is derived from the inverse demand function, to maximize

$$\max \sum_{t=0}^{\infty} \beta^t (\Pi(p_t, \nu_{Pt}) + W^o \nu_{Wt})$$

$$s.t.$$

$$\nu_{t+1} = \Phi(\nu_t).$$

This sequence of allocations can be used to solve for the aggregate quantity of the industry's product, the total cost of production and the total amount of learning and imitation that occurs. The single firm problem internalizes the value of each of the researchers by making the future distribution a constraint of the problem instead of having the wage functions as in the competitive case. Essentially each of the agents in the continuum are treated as if they were assets owned by the aggregate firm. The equivalence between the competitive equilibrium and the single firm equilibrium is given by the following lemma.

Lemma 2a: The equilibrium aggregate allocations and the industry's product price from the single firm problem are equivalent to those in the competitive equilibrium.

Proof: We need to show that the aggregate allocations of producers, researchers and workers are the same in both cases. These are the same since the single firm faces the same incentives as the individual agents do in the competitive equilibrium. There are three cases to be considered. The first is the case where the single firm is choosing whether an individual agent should work as a researcher or operate a productive unit. In this case, the firm must weigh the return from production, the return from learning and imitation that occurs when the agent operates a plant and the return from having this individual imitate a plant operator's knowledge and increasing that plant operator's knowledge. In the competitive equilibrium, the trade-off is the same since an agent who works as a researcher is paid a wage that is both his marginal product from learning by his employer and the expected return from imitating his employer's knowledge, as shown in Proposition 2.

In the second case, the single firm weighs the return from having an individual agent operate a plant or work outside the industry. In this case, the return from operating a plant is the current profits and the increase in the single firm's knowledge from both learning and imitation, and the return from working outside is the wage, given by W^0, and leaves the single firm's knowledge unchanged. In the competitive case, the incentives are the same, since the return from imitation is paid to an agent who chooses to incorporate, as shown in Proposition 2.

In the final case, the single firm must weigh the return from having an individual agent work as a researcher or work outside the industry. Here the trade-off is between increasing the single firm's knowledge by both learning and imitation and the wage from having the agent work outside the industry. In the competitive case, the marginal benefit from increasing a firm's knowledge is paid to the researcher and the expected return from imitating the firm's knowledge is paid by the researcher, so that the incentives and trade-offs faced by an individual agent are the same.□

Proof of Proposition 3: The conditions that satisfy the single firm problem are

$$(I1) \qquad \{\nu_{Pt}, \nu_{Rt}, \nu_{Wt}, z_t\}_{t=0}^{\infty} \text{ is feasible,}$$

$$(I2) \qquad p_t = D\left(\int q(\theta, \nu_t)\, d\nu_{Pt}\right),$$

$$(I3) \qquad \begin{aligned} &E\left\{\textstyle\sum_{t=0}^{\infty} \beta^t (\Pi p_t, \nu_{Pt} + W^o \nu_{Wt})\right\} \\ &\geq E\left\{\textstyle\sum_{t=0}^{\infty} \beta^t (\Pi p_t, \nu'_{Pt} + W^o \nu'_{Wt})\right\}, \\ &\text{all feasible } \{\nu_{Pt}, \nu_{Rt}, \nu_{Wt}, z_t\}_{t=0}^{\infty}. \end{aligned}$$

In order for the planner's problem to be satisfied, the following condition must hold:

$$(I4) \qquad \begin{aligned} &E\left\{\textstyle\sum_{t=0}^{\infty} \beta^t \left(S\left(\int q(\theta, \nu_t)\, d\nu_{Pt}\right) - \int c(q, \theta)\, d\nu_{Pt} - W^o \nu_{Rt} - W^o \nu_{Pt}\right)\right\} \\ &\geq E\left\{\textstyle\sum_{t=0}^{\infty} \beta^t \left(S\left(\int q(\theta, \nu_t)\, d\nu'_{Pt}\right) - \int c(q, \theta)\, d\nu'_{Pt} - W^o \nu'_{Rt} - W^o \nu'_{Pt}\right)\right\} \end{aligned}$$

for any feasible $\{\nu'_{Pt}, \nu'_{Rt}, \nu'_{Wt}, z'_t\}_{t=0}^{\infty}$. First, we prove that if $\{\nu_{Pt}, \nu_{Rt}, \nu_{Wt}, z_t\}_{t=0}^{\infty}$ is a single firm equilibrium, then $\{\nu_{Pt}, \nu_{Rt}, \nu_{Wt}, z_t\}_{t=0}^{\infty}$ is a surplus-maximizing allocation. Suppose that given $(I1)$, $(I2)$, and $(I3)$, $(I4)$ does not hold. This implies that

$$\begin{aligned} &E\left\{\textstyle\sum_{t=0}^{\infty} \beta^t \left(S\left(\int q(\theta, \nu_t)\, d\nu_{Pt}\right) - \int c(q, \theta)\, d\nu_{Pt} - W^o \nu_{Rt} - W^o \nu_{Pt}\right)\right\} \\ &< E\left\{\textstyle\sum_{t=0}^{\infty} \beta^t \left(S\left(\int q(\theta, \nu_t)\, d\nu'_{Pt}\right) - \int c(q, \theta)\, d\nu'_{Pt} - W^o \nu'_{Rt} - W^o \nu'_{Pt}\right)\right\}. \end{aligned}$$

By rearranging the above condition, we have

$$(3) \qquad \begin{aligned} &E\left\{\textstyle\sum_{t=0}^{\infty} \beta^t \left(\begin{array}{c} \int c(q, \theta)\, d\nu'_{Pt} + W^o \nu'_{Rt} + W^o \nu'_{Pt} \\ - \int c(q, \theta)\, d\nu_{Pt} - W^o \nu_{Rt} - W^o \nu_{Pt} \end{array}\right)\right\} \\ &< E\left\{\textstyle\sum_{t=0}^{\infty} \beta^t \left(S\left(\int q(\theta, \nu_t)\, d\nu'_{Pt}\right) - S\left(\int q(\theta, \nu_t)\, d\nu_{Pt}\right)\right)\right\}. \end{aligned}$$

By the concavity of the consumer surplus function and $(I2)$, we have

$$S\left(\int q(\theta,\nu_t)\, d\nu'_{Pt}\right) - S\left(\int q(\theta,\nu_t)\, d\nu_{Pt}\right)$$
$$\leq p_t\left(\int q(\theta,\nu_t)\, d\nu'_{Pt} - \int q(\theta,\nu_t)\, d\nu_{Pt}\right).$$

By replacing this in the above inequality and rearranging, we have

$$(3a)\quad E\left\{\sum_{t=0}^{\infty}\beta^t\left(p_t\left(\int q(\theta,\nu_t)\, d\nu_{Pt}\right) - \int c(q,\theta)\, d\nu_{Pt} - W^o\nu_{Rt} - W^o\nu_{Pt}\right)\right\}$$
$$< E\left\{\sum_{t=0}^{\infty}\beta^t\left(p_t\left(\int q(\theta,\nu_t)\, d\nu'_{Pt}\right) - \int c(q,\theta)\, d\nu'_{Pt} - W^o\nu'_{Rt} - W^o\nu'_{Pt}\right)\right\}.$$

Recall that

$$\nu_{Wt} = 1 - \nu_{Rt} - \nu_{Pt}.$$

Hence, inequality $(3a)$ contradicts $(I3)$, which completes the proof.

Second, we prove that if $\{\nu_{Pt},\nu_{Rt},\nu_{Wt},z_t\}_{t=0}^{\infty}$ is a surplus-maximizing allocation and p is defined by condition $(I2)$, then $\{\nu_{Pt},\nu_{Rt},\nu_{Wt},z_t\}_{t=0}^{\infty}$ is an industry equilibrium. For any feasible $\{\nu'_{Pt},\nu'_{Rt},\nu'_{Wt},z'_t\}_{t=0}^{\infty}$, define $f:[0,1]\to\Re$ by

$$f(\gamma) = E\left\{\sum_{t=0}^{\infty}\beta^t\left(\begin{array}{c} S\left((1-\gamma)\int q(\theta,\nu_t)\, d\nu_{Pt} + \gamma\int q(\theta,\nu_t)\, d\nu'_{Pt}\right) \\ -\left[\begin{array}{c}(1-\gamma)\left(\int c(q,\theta)\, d\nu_{Pt} - W^o\nu_{Rt} - W^o\nu_{Pt}\right) \\ +\gamma\left(\int c(q,\theta)\, d\nu'_{Pt} - W^o\nu'_{Rt} - W^o\nu'_{Pt}\right)\end{array}\right]\end{array}\right)\right\}.$$

To find the maximizing allocation, we take the derivative of f and set it equal to zero, or

$$E\{\sum_{t=0}^{\infty}\beta^t(D\left(\begin{array}{c}(1-\gamma)\int q(\theta,\nu_t)\, d\nu_{Pt} \\ +\gamma\int q(\theta,\nu_t)\, d\nu'_{Pt}\end{array}\right)\left[-\int q(\theta,\nu_t)\, d\nu_{Pt} + \int q(\theta,\nu_t)\, d\nu'_{Pt}\right]$$

$$-\left[\begin{array}{c}-\left(\int c(q,\theta)\, d\nu_{Pt} - W^o\nu_{Rt} - W^o\nu_{Pt}\right) \\ +\left(\int c(q,\theta)\, d\nu'_{Pt} - W^o\nu'_{Rt} - W^o\nu'_{Pt}\right)\end{array}\right])\} = 0.$$

Since the surplus maximizing allocation occurs at $\{\nu_{Pt}, \nu_{Rt}, \nu_{Wt}, z_t\}_{t=0}^{\infty}$ by assumption, we can replace γ with zero. By the concavity of the consumer surplus function and $(I2)$, we have

$$E\{\sum_{t=0}^{\infty} \beta^t (p_t \left[-\int q(\theta, \nu_t)\, d\nu_{Pt} + \int q(\theta, \nu_t)\, d\nu'_{Pt}\right] - \left[\begin{array}{l} -\left(\int c(q,\theta)\, d\nu_{Pt} - W^o \nu_{Rt} - W^o \nu_{Pt}\right) \\ +\left(\int c(q,\theta)\, d\nu'_{Pt} - W^o \nu'_{Rt} - W^o \nu'_{Pt}\right) \end{array} \right])\} = 0.$$

By rearranging, we have

$$E\left\{\sum_{t=0}^{\infty} \beta^t \left(p_t \left[\int q(\theta, \nu_t)\, d\nu'_{Pt}\right] - \left(\int c(q,\theta)\, d\nu'_{Pt} - W^o \nu'_{Rt} - W^o \nu'_{Pt}\right)\right)\right\}$$
$$= E\left\{\sum_{t=0}^{\infty} \beta^t \left(p_t \left[\int q(\theta, \nu_t)\, d\nu_{Pt}\right] - \left(\int c(q,\theta)\, d\nu_{Pt} - W^o \nu_{Rt} - W^o \nu_{Pt}\right)\right)\right\},$$

which completes the proof. Since the competitive equilibrium and the single firm equilibrium are equivalent by Lemma 2a, the competitive equilibrium is also Pareto optimal. Uniqueness is established by the concavity of the cost function and the downward sloping, single valued demand curve.□

Proof of Corollary 1: Note that the optimality of the competitive equilibrium does not depend on what value λ takes on. The probability of not imitating is a social cost. Given the option, the social planner would choose to set λ equal to one, so that the social cost would be zero.□

Lemma 3a: Any agent who does not become an entrepreneur in a given period and does not learn her employer's knowledge will not become an entrepreneur in the following period.

Proof: Suppose the planner must assign tasks to two arbitrary agents with different levels of know-how. The first agent, denoted by m, was assigned to work as a plant manager in the previous period and has knowledge, given by θ_m, in the current period. Since the agent cannot forget by assumption, his knowledge is at least as good as in the previous period. The second agent, u, did not operate a plant and has knowledge given by θ_u. His knowledge is unchanged from the previous period. Further, let $\theta_m \geq \theta_u$. The planner must determine which of these agents will be assigned the position of managing a firm and which will not. Suppose the planner chooses to have agent u manage a firm and agent m to either work outside the industry or work as a researcher. In this case, the planner can improve the total surplus in the economy by assigning agent m to manage a firm instead of agent u. This improvement is a result of the increase in current profits and the increase in future knowledge. Since the cost of production is decreasing in knowledge, the profits from having agent m manage a plant will be higher than those from agent u, or

$$\begin{aligned} &\max_q \{p(\nu)\, q(\theta_m) - c(q(\theta_m), \theta_m)\} \\ &> \max_q \{p(\nu)\, q(\theta_u) - c(q(\theta_u), \theta_u)\}. \end{aligned}$$

The current profits are weakly higher if agent m operates a productive unit instead of agent u. Next, we turn to the effect on the future distribution of knowledge. The improvement in future knowledge is from the assumption that the learning technology by the firm is increasing in knowledge and that researchers may learn the plant manager's knowledge. Since the set of agents who have knowledge that is lower than θ_m is higher than those that have knowledge that is lower than θ_u and the plant manager's future knowledge is increasing in the plant manager's current knowledge, the future distribution will be higher if agent m is the plant manager than if agent u is, or

$$\begin{aligned} &\lambda \int_{L\times[\theta_L,\theta_m]\times\theta_m} l\ z\,(dl \times d\theta_r \times d\theta_f) + \int \Psi\,(\Theta\,|l, \theta_m)\, dz \\ &> \lambda \int_{L\times[\theta_L,\theta_u]\times\theta_u} l\ z\,(dl \times d\theta_r \times d\theta_f) + \int \Psi\,(\Theta\,|l, \theta_u)\, dz. \end{aligned}$$

As a result, the planner would choose to have agent m operate a plant instead of agent u. Since the competitive equilibrium is Pareto optimal, an agent who failed to improve his knowledge from the previous period and did not operate a plant will not do so in the current period. $\square$

Proposition 3a: There exists a critical value $\tilde{\theta}$ defined by

$$W^o + \beta V(\tilde{\theta}, \Phi(\nu)) = \max_{(q,l)} \left\{ \begin{array}{c} p(\nu)\, q - c(q, \tilde{\theta}) - lw(\tilde{\theta}) \\ +\beta \int V(\theta', \Phi(\nu))\, d\Psi(\theta' \left| l, \tilde{\theta} \right.) \end{array} \right\}$$

such that if an agent has θ such that $\theta > \tilde{\theta}$, then that agent will become an entrepreneur.

Proof: The proof is shown in two steps. First, the return from not incorporating for any arbitrary agent i with $\theta_i > \tilde{\theta}$ is compared to that of an agent with $\tilde{\theta}$. Next, the return from incorporating for these types of agents is compared. Since the returns in the first case are equivalent and the returns to any arbitrary agent i in the second case outweigh those to any agent with $\tilde{\theta}$, agent i will choose to operate a firm.

From Proposition 2, the return from working outside the industry must be equivalent to that from working within the industry for any firm f with $l(\theta_f) > 0$. It is sufficient to consider only the case where an agent works outside the industry. By Lemma 1a, the value function is non-decreasing in θ, so the present discounted value of working outside the industry for agent i with $\theta_i > \tilde{\theta}$ is at least equivalent to that of an agent with $\tilde{\theta}$, or

$$W^o + \beta V(\theta_i, \Phi(\nu)) \geq W^o + \beta V(\tilde{\theta}, \Phi(\nu)).$$

Next, since both $c(q, \theta)$ and $w(\theta)$ are decreasing in θ and $\Psi(\theta' \left| l, \theta \right.)$ is increasing in θ, the returns to incorporating for any agent i are greater than to any agent with $\tilde{\theta}$, or

$$\max_{(q,l)} \left\{ \begin{array}{c} p(\nu)\, q - c(q,\theta) - lw(\theta) \\ +\beta \int V(\theta', \Phi(\nu))\, d\Psi(\theta' | l, \theta) \end{array} \right\} >$$

$$\max_{(q,l)} \left\{ \begin{array}{c} p(\nu)\, q - c(q,\tilde{\theta}) - lw(\tilde{\theta}) \\ +\beta \int V(\theta', \Phi(\nu))\, d\Psi(\theta' \left| l, \tilde{\theta}\right.) \end{array} \right\}.$$

This establishes the result.□

Lemma 4a: $\tilde{\theta}$ is weakly increasing, or $\tilde{\theta}(\nu) \leq \tilde{\theta}(\Phi(\nu))$.

Proof: By Lemma 3a, any agent who did not run a firm and failed to learn in the previous period will not run a firm in the current period. This implies that the critical value, $\tilde{\theta}$, is non-decreasing.□

Lemma 5a: The value function is constant for $\theta \leq \tilde{\theta}$.

Proof: Any agent i with $\theta_i \leq \tilde{\theta}$, by Proposition 3a, will not run a plant. In fact, agent i will face the following problem:

$$V(\theta_i, \nu) = \max \left\{ \begin{array}{l} W^o + \beta V(\theta_i, \Phi(\nu)), \\ \max_{f \in [L,H]} \left\{ \begin{array}{l} [w(\theta_i, \theta_f) \\ +\beta \left(\begin{array}{c} \lambda V(\theta_f, \Phi(\nu)) \\ +(1-\lambda) V(\theta_i, \Phi(\nu)) \end{array} \right)], \text{ if } \theta_f \geq \theta_i \\ [w(\theta_i, \theta_f) + \beta V(\theta_i, \Phi(\nu))], \text{ otherwise} \end{array} \right\} \end{array} \right\}.$$

Lemma 3a states that any agent who works as a researcher and fails to learn her employer's know-how will not run a plant. This implies that the last term in each of the branches of the agent's original value function can be replaced. By expanding the following period's value function, $V(\theta_i, \Phi(\nu))$, it follows that

$$V(\theta_i, \Phi(\nu)) = \max \begin{bmatrix} W^o + \beta V(\theta_i, \Phi(\Phi(\nu))), \\ \max_{f \in [L,H]} \left\{ \begin{array}{l} [w(\theta_i, \theta_f) \\ +\beta \begin{pmatrix} \lambda V(\theta_f, \Phi(\Phi(\nu))) \\ +(1-\lambda) V(\theta_i, \Phi(\Phi(\nu))) \end{pmatrix}], \text{ if } \theta_f \geq \theta_i \\ [w(\theta_i, \theta_f) + \beta V(\theta_i, \Phi(\Phi(\nu)))], \text{ otherwise} \end{array} \right\} \end{bmatrix}.$$

By expanding the value function using this for the continuation value, it is obvious that, since $\tilde{\theta}$ is weakly increasing, the effect on θ_i can be made arbitrarily small.□

Lemma 6a: The value function is increasing for $\theta > \tilde{\theta}$.

Proof: By Proposition 3a, all agents with $\theta > \tilde{\theta}$ will run a production unit. So we need only to consider the last branch of (1), or

$$V(\theta, \nu) = \max_{(q,l)} \left\{ p(\nu) q - c(q, \theta) - lw(\theta) + \beta \int V(\theta', \Phi(\nu)) \, d\Psi(\theta' | l, \theta) \right\}.$$

By assumption, the cost of production is decreasing in θ, or $dc/d\theta < 0$. By Corollary 1a, the wage is non-increasing in the firm's θ. By the assumptions on the innovative technology, $\Psi(\theta' | l, \theta)$ is increasing in θ. Together, these facts imply that the value function is increasing.□

Proof of Proposition 4: Recall that $\tilde{\theta}$ is weakly increasing as shown in Lemma 4a. The next period's critical value is given by $\tilde{\theta}'$. Consider the case of an agent who works for a firm i with $\theta_i \leq \tilde{\theta}'$. In the case where the agent fails to learn her employer's θ, she will not run a production unit in the following period, by Lemma 3a. In the case where the agent does learn her employer's know-how and can imitate it, she will not run a production unit, since her new technology will fall below the critical value. Next, consider the case of an agent who works for a firm j with $\theta_j > \tilde{\theta}'$. In the case that she does learn her employer's knowledge, she will imitate her employer's knowledge and run a production unit, since this knowledge exceeds the critical value. Hence, λ of the firm's employees will run a plant and imitate its technological know-how.□

Proof of Proposition 5: Here, we must show that for any two agents, i and j, with know-how, θ_i and θ_j, respectively, such that $\theta_i \geq \theta_j \geq \tilde{\theta}$,

$$\Pr\{\theta_i' \geq \tilde{\theta}'\} \geq \Pr\{\theta_j' \geq \tilde{\theta}'\}. \tag{4}$$

There are two cases to consider. In the first case, agent i's current knowledge exceeds the next period's critical value for incorporating. In the second case, neither agent's current knowledge exceeds the next period's critical value, and we must show that the probability that agent i will innovate to a level that is higher than the next period's critical value weakly dominates the probability that agent j will.

We turn to the first case. If $\theta_i \geq \tilde{\theta}'$, then by assumption (iii) on the firm's learning function, $\theta_i' \geq \tilde{\theta}'$. So, $\Pr\{\theta_i' \geq \tilde{\theta}'\} = 1$, and the result is established in this case.

Next, we consider the second case. If $\tilde{\theta}' > \theta_i$, inequality (1) is equivalent to

$$(1 - F(\tilde{\theta}'|\theta_i))G(l(\theta_i)) \geq (1 - F(\tilde{\theta}'|\theta_j))G(l(\theta_j)) \tag{5}$$

where $l(\theta_i)$ and $l(\theta_j)$ denote the optimal labor choices. By assumption (iv), $F(\theta'|\theta_i)$ first order stochastically dominates $F(\theta'|\theta_j)$. A sufficient condition for inequality (1) to hold is $G(l(\theta_i)) \geq G(l(\theta_j))$. This holds if $l(\theta_i) \geq l(\theta_j)$.

A firm with know-how θ_i chooses $l(\theta_i)$ by maximizing

$$-lw(\theta_i) + \beta(1 - G(l))V(\theta_i, \Phi(\nu)) + \beta G(l) \int V(\theta_i', \Phi(\nu))dF(\theta_i'|\theta_i).$$

The first order condition is

$$-w(\theta_i) - \beta G'(l)V(\theta_i, \Phi(\nu)) + \beta G'(l) \int V(\theta_i', \Phi(\nu))dF(\theta_i'|\theta_i) = 0.$$

The firm with know-how θ_j solves a similar problem. Since $\tilde{\theta}'$ exceeds both θ_i and θ_j, if either agent is unsuccessful at innovating, that agent does not run a firm in the following period by Lemma 3a. By Lemma 5a, the value of an agent who does not run a firm is the same, regardless of his know-how, so $V(\theta_i, \Phi(\nu)) = V(\theta_j, \Phi(\nu))$. Then equilibrium in the labor market implies that $w(\theta_i) = w(\theta_j)$. By Proposition 2, this wage must be equal to that paid to agents who work outside, W^0.

Given that both the value of failing to learn and the wages paid by these agents are the same, the first-order condition implies that $l(\theta_i) \geq l(\theta_j)$ if

$$\int V(\theta_i', \Phi(\nu))dF(\theta_i'|\theta_i) \geq \int V(\theta_j', \Phi(\nu))dF(\theta_j'|\theta_j).$$

This inequality holds, since $F(\theta_i'|\theta_i)$ first order stochastically dominates $F(\theta_j'|\theta_j)$ and, by Lemma 1a, $V(\theta', \Phi(\nu))$ is weakly increasing in θ'.$\square$

References

Braun, Ernest and Stuart MacDonald, *Revolution in Miniature*, Cambridge, MA: Cambridge University Press, 1982.

Chari, V.V. and Hugo Hopenhayn, "Vintage Human Capital, Growth and the Diffusion of New Technology," *Journal of Political Economy,* 99, no. 6 (1991): 1142-1165.

Christensen, Clayton M. *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail.* Boston: Harvard Business School Press, 1997.

Christensen, Clayton M., "The Rigid Disk Drive Industry: A History of Commercial and Technological Turbulence," *Business History Review*, 67 (Winter 1993), 531-588.

Gilson, Ronald J., "The Legal Infrastructure of High Technology Industrial Districts: Silicon Valley, Route 128, and Covenants Not To Compete," SSRN Working Paper, July 1998.

Gort, Michael and Steven Klepper, "Time Paths in the Diffusion of Product Innovations," *The Economic Journal*, 92, no. 367 (September 1982): 630-653.

Hopenhayn, Hugo, "Entry, Exit and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60, no. 5 (1992), 1127-1150.

Irwin, Douglas A. and Peter J. Klenow, "Learning-by-Doing Spillovers in the Semiconductor Industry," *Journal of Political Economy*, 102, no. 6 (1994): 1200-1227.

Jovanovic, Boyan and Glenn M. MacDonald, "Competitive Diffusion," *Journal of Political Economy,* 102, no.11 (1994a): 24-56.

Jovanovic, Boyan and Glenn M. MacDonald, "The Life Cycle of a Competitive Industry," *Journal of Political Economy,* 102, no. 2 (1994b): 322-347.

Jovanovic, Boyan and Yaw Nyarko, "The Transfer of Human Capital," *Journal of Economic*

Dynamics and Control, 19 (July-September), no. 5-7, (1995) 1033-64.

Jovanovic, Boyan and R. W. Rosenthal, "Anonymous Sequential Games," *Journal of Mathematical Economics*, 17, no. 1 (1988), 77-87.

Lerner, Josh,"An Empirical Exploration of a Technology Race," *RAND Journal of Economics*, 28, no. 2, (Summer 1997): 228-247.

Lucas, Robert E. Jr., "On the Size Distribution of Business Firms," *The BELL Journal of Economics,* 9, no. 2, (August 1978): 508-523.

Mitchell, Matthew F., "Adoption and the Size Distribution of Establishments," mimeo, University of Minnesota, December 1999.

Porter, J. *Disk/Trend Report: Rigid Disk Drives.* Los Altos, CA: Disk/Trend, 1977-1997.

Prusa, Thomas J. and James A. Schmitz, Jr., "Can Companies Maintain Their Initial Innovative Thrust? A Study of the PC Software Industry," *The Review of Economics and Statistics*, 76 no. 3 (1994): 523-540.

Saxenian, Annalee, *Regional Advantage: Culture and Competition in Silicon Valley and Route 128,* Cambridge: Harvard University Press, 1994.

Semiconductor Equipment and Materials International, "Silicon Valley Genealogy," 1986.

Stokey, Nancy L., Robert E. Lucas, Jr. and Edward C. Prescott, *Recursive Methods in Economic Dynamics,* Cambridge: Harvard University Press, 1989.

Wilson, Robert W., Peter K. Ashton and Thomas P. Egan, *Innovation, Competition and Government Policy in the Semiconductor Industry,* Lexington, MA: Lexington Books, 1980.

Fig. 1. The Percentage of Agents of Each Type

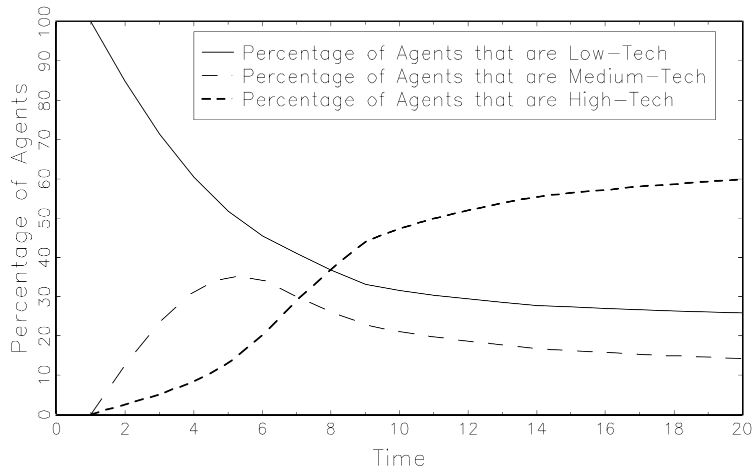

Fig. 2. The Percentage of Agents that are Incorporated and are of Each Type

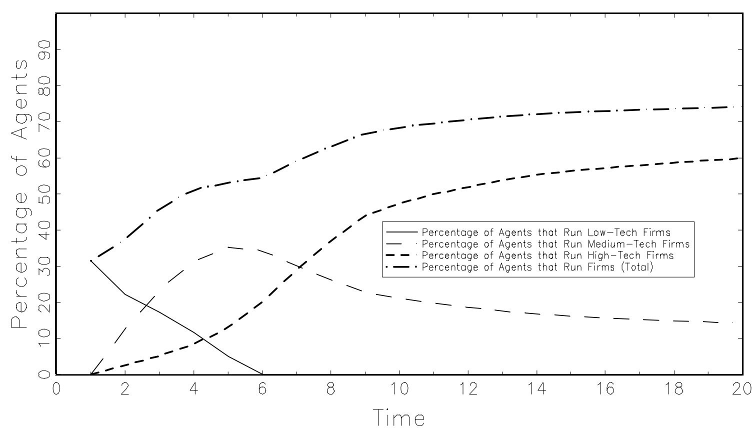

Fig. 3. The Percentage of Agents that are Researchers, By Type of Agent and Firm

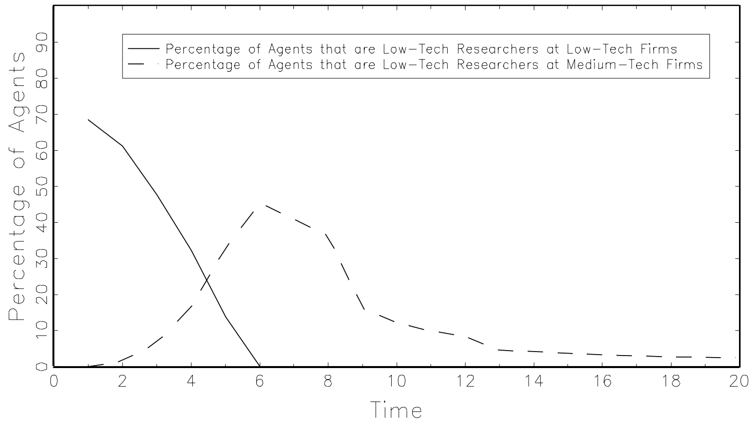

Fig. 4. R&D Expenditure per Firm

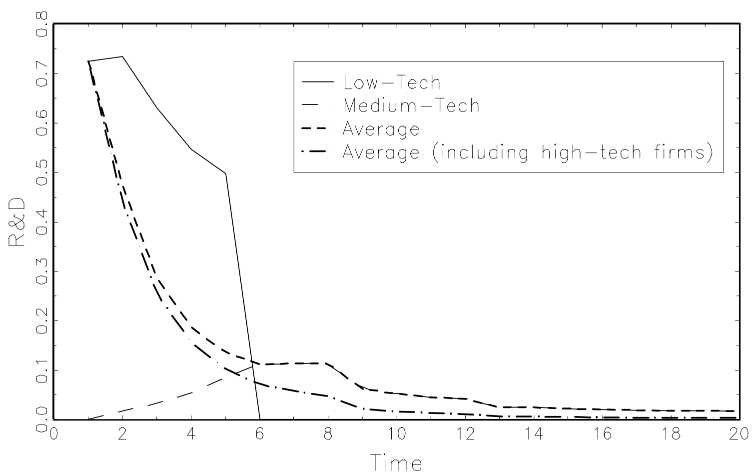

Fig. 5. Price, Market Quantity, and Average Production Cost

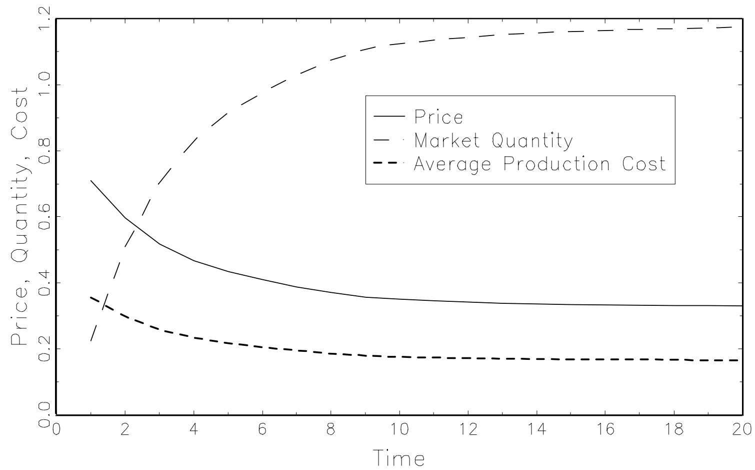

Fig. 6. Current Profits per Firm (includes R&D)

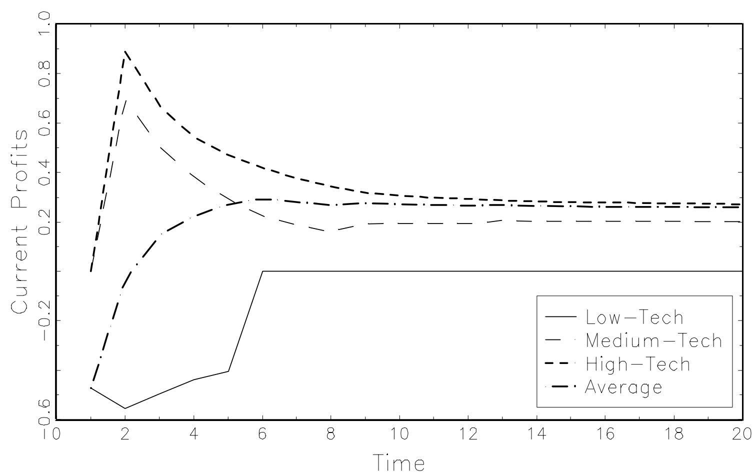

Fig. 7. Current Market Profits per Firm (excludes R&D)

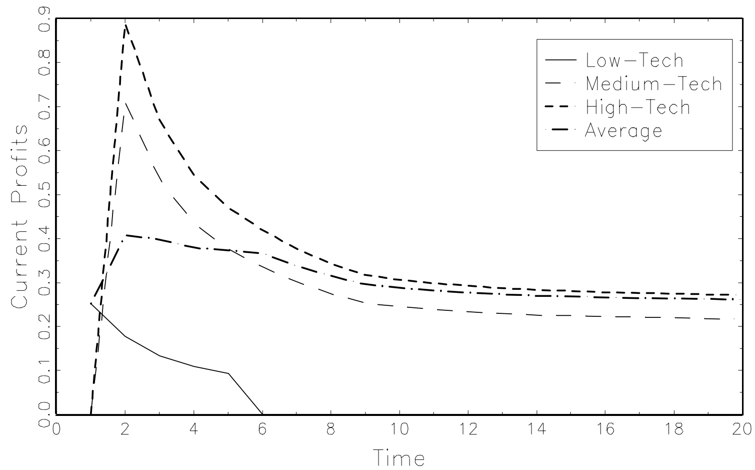

Fig. 8. The Value of Each Type of Agent

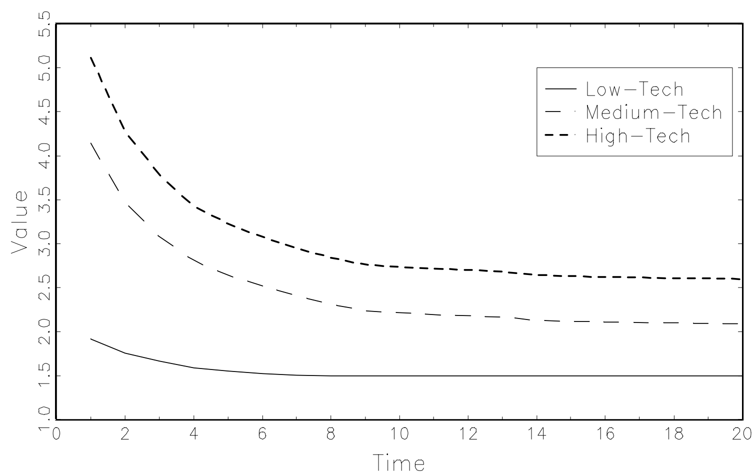

Table 1. Spin-outs, Parents, Founding Years, and Life Spans

Spin-Out	Parent(s)	Founding Yr.	Life Span
International Memories	Memorex	1977	8, Exited
Micropolis	Pertec	1977	19, Acquired
Dastek	IBM	1978	3, Acquired
Priam	Memorex	1978	12, Exited
Irwin International Industries, Inc.	Sycor	1979	3, Acquired
Seagate	Shugart Associates	1979	18, Still Active
Computer Memories	Pertec	1980	6, Exited
Ibis	Burroughs, Memorex	1980	10, Exited
Miniscribe	Storage Technology Corp.	1980	10, Acquired (by Maxtor)
Quantum	Shugart Associates	1980	17, Still Active
Rodime	Burroughs	1980	11, Exited
Rotating Memory Systems	Shugart Associates, Memorex	1980	2, Acquired
Amcodyne	Storage Technology Corp.	1981	5, Acquired
Atasi	International Memories	1981	6, Acquired
Evotek	Memorex, Data General	1981	2, Exited
Tecstor	Microdata	1981	6, Acquired
Applied Information Memories	Ibis	1982	3, Exited
Cogito	IBM	1982	6, Exited
Maxtor	Quantum	1982	14, Acquired
Microcomputer Memories	Alpha Data	1982	5, Exited
Microscience International	Datapoint	1982	10, Exited
Syquest	Seagate	1982	15, Still Active
Vertex Peripherals	Shugart Associates	1982	3, Acquired (by Priam)
Lapine	Irwin International	1983	4, Exited
Tulin	Ampex, Qume	1983	5, Exited
Epelo	Atasi	1984	1, Exited
Josephine County Technology	Tandon	1984	4, Exited
Micro Storage Corp.	Syquest	1984	2, Exited
Peripheral Technology	Computer Memories	1985	2, Acquired
Brand Technologies	Computer Memories	1986	6, Exited
Conner Peripherals	Seagate, Miniscribe	1986	10, Acquired (by Seagate)
PrairieTek	Miniscribe	1986	5, Exited
Comport	Lapine	1987	3, Exited
Kalok	Lapine	1987	7, Acquired
Areal Technology	Maxtor	1988	3, Acquired
Ecol.2	Areal Technology	1990	1, Exited
Integral Peripherals	PrairieTek	1990	7, Still Active
Orca Technology	Maxtor, Priam	1990	2, Exited
MiniStor	Maxtor	1991	4, Exited
Gigastorage International	Aura Associates	1993	4, Still Active

The exit date is the date the firm stops manufacturing and selling new drives. Spin-outs either exit through failure (denoted by exited in the life span column), are acquired (denoted by acquired), or are still active as of 1997 (denoted by still active). If the firm was acquired by another spin-out, we note the acquiring firm.

Table 2. The Early Movers, by Diameter
(firms are in alphabetical order in each category)

Diameter	Early Mover	Introduction Date
8”	BASF	Q4, 1979
	IBM	Q1, 1979
	International Memories	Q1, 1979
	Micropolis	Q4, 1979
	New World Computer	Q3, 1979
	Pertec	Q4, 1979
	Shugart Associates	Q4, 1979
5.25”	Computer Memories	Q2, 1981
	International Memories	Q1, 1981
	New World Computer	Q3, 1980
	Rodime	Q2, 1981
	Rotating Memory Systems	Q2, 1981
	Seagate	Q3, 1980
	Tandon	Q4, 1980
3.5”	Control Data	Q3, 1983
	Microcomputer Memories	Q1, 1984
	Microscience International	Q2, 1984
	Rodime	Q3, 1983
2.5”	PrairieTek	Q4, 1988
1.8”	Integral Peripherals	Q3, 1991

An early mover is defined to be a firm that introduces a drive in the diameter within 3 quarters after the first introduction. 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.

Table 3. Imitation of Early-mover Know-how in the Period 1977-1997

Early Mover Parent	**Spin-Out**	**Is the Spin-Out an Early Mover?**
Computer Memories, 5.25”	Brand Technologies	NO
	Peripheral Technology	NO
IBM, 8”	Cogito	NO
	Dastek	NO
International Memories, 8” and 5.25”	Atasi	NO
Pertec, 8”	Computer Memories	YES, 5.25”
	Micropolis	YES, 8”
PrairieTek, 2.5”	Integral Peripherals	YES, 1.8”
Seagate, 5.25”	Conner Peripherals	NO
	Syquest	NO, (but was the first mover in 4” removable cartridge drives)
Shugart Associates, 8”	Quantum	NO
	Rotating Memory Systems	YES, 5.25”
	Seagate	YES, 5.25”
	Vertex Peripherals	NO
Tandon, 5.25”	Josephine County Technology	NO

Table 4. Summary Statistics

	Mean	**Standard deviation**	**Minimum**	**Maximum**	**Cases**
Technical Know-How	0.44	0.25	0.0084	1.00	1039
Lagged Technical Know-How	0.45	0.24	0.0084	1.00	877
Early-Mover Know-How	0.15	0.36	0.00	1.00	1190
Lagged Number of Drives	11.31	15.06	1.00	119	886
Lagged Sales Growth	0.21	0.65	-2.00	2.00	846
U.S. Firm Dummy	0.62	0.49	0.00	1.00	1190
Spin-Out Generation Dummy	0.032	0.18	0.00	1.00	1190
Survival Dummy	0.91	0.29	0.00	1.00	1172

Definitions of technical know-how, early-mover know-how, and sales growth are provided in the text. Technical Know-How and Lagged Technical Know-How range from 0 to 1. Early-Mover Know-How is a dummy variable. Lagged Sales Growth ranges from -2 to 2.

Lagged Number of Drives measures the number of drives produced by the firm in the previous period. The U.S. Firm Dummy takes the value 1 if the firm is an American firm, and 0 otherwise. The Spin-Out Generation Dummy takes the value 1 if the firm generates a spin-out in the current period, and 0 otherwise. The Survival Dummy takes the value 0 if the firm exits through failure in the following period, and 1 otherwise.

Table 5. The Probability of Generating a Spin-Out as a Function of Know-How

Probit Model (Standard Errors in Parentheses)

	Equation 5a.	Equation 5b.
Variable	Coefficient	Coefficient
Constant	-2.67*** (0.44)	-2.58*** (0.45)
Lagged Technical Know-How	1.12*** (0.37)	1.08*** (0.40)
Early-Mover Know-How	0.49** (0.19)	0.47** (0.21)
Lagged Sales Growth	-	0.12 (0.15)
Lagged Number of Drives	-	-0.0096 (0.0080)
YR1978	0.44 (0.57)	0.57 (0.57)
YR1979	0.15 (0.61)	0.077 (0.62)
YR1980	0.82* (0.49)	0.80* (0.49)
YR1981	0.84* (0.48)	0.85* (0.48)
YR1982	0.83* (0.48)	0.81* (0.48)
YR1984	0.25 (0.51)	0.22 (0.51)
YR1985	-0.14 (0.59)	-0.13 (0.59)
YR1986	0.32 (0.50)	0.37 (0.50)
YR1987	-0.083 (0.58)	-0.037 (0.59)
YR1988	-0.059 (0.60)	0.053 (0.62)
YR1990	0.67 (0.49)	0.89* (0.51)
YR1991	-0.12 (0.59)	0.17 (0.62)
YR1993	0.0013 (0.60)	0.50 (0.66)
Number of Observations	673	556
Log Likelihood	-117.03	-112.19

The dependent variable is the spin-out generation dummy.

*Significant at the 10% level. **Significant at the 5% level. ***Significant at the 1% level.

When year dummies are included, 1983 is the base year.

Table 6. The Probability of Surviving to the Following Period as a Function of Know-How

Probit Model (Standard Errors in Parentheses)

	Equation 6a.
Variable	Coefficient
Constant	0.96*** (0.25)
Technical Know-How	0.99*** (0.25)
Early-Mover Know-How	0.27 (0.20)
YR1978	0.50 (0.50)
YR1979	0.11 (0.38)
YR1981	0.81* (0.48)
YR1982	0.30 (0.35)
YR1984	-0.21 (0.30)
YR1985	-0.17 (0.31)
YR1986	-0.37 (0.30)
YR1987	-0.010 (0.33)
YR1988	0.046 (0.33)
YR1989	-0.013 (0.33)
YR1990	-0.17 (0.32)
YR1991	-0.69** (0.29)
YR1992	-0.61** (0.31)
YR1993	-0.72** (0.31)
YR1994	-0.68** (0.33)
YR1996	-0.39 (0.37)
Number of Observations	918
Log Likelihood	-301.41

The dependent variable is the survival dummy. It is 0 if the firm exits through failure in the following period, and 1 otherwise.

*Significant at the 10% level.

**Significant at the 5% level.

***Significant at the 1% level.

When year dummies are included, 1983 is the base year.

Table 7 The Probability of Surviving to the Following Period as a Function of Know-How - Spin-Outs Only

Probit Model (Standard Errors in Parentheses)

	Equation 7a.	Equation 7b.
Variable	Coefficient	Coefficient
Constant	0.76 (0.59)	0.49 (0.73)
Technical Know-How	1.46** (0.61)	3.08*** (0.95)
Early-Mover Know-How	-0.029 (0.29)	-0.47 (0.42)
Parent Technical Know-How	-	-1.36** (0.62)
Parent Early-Mover Know-How	-	0.68* (0.37)
YR1984	-.10 (0.62)	-
YR1985	-0.11 (0.63)	0.64 (0.76)
YR1986	0.22 (0.71)	0.79 (0.82)
YR1987	-0.13 (0.66)	0.12 (0.73)
YR1988	0.32 (0.73)	0.65 (0.82)
YR1989	-0.21 (0.65)	0.11 (0.72)
YR1990	-0.60 (0.64)	-0.63 (0.70)
YR1991	-0.81 (0.62)	-0.85 (0.68)
YR1992	-0.70 (0.66)	-1.08 (0.71)
YR1994	-0.32 (0.74)	-0.35 (0.82)
YR1996	-0.28 (0.78)	-0.17 (0.97)
Number of Observations	184	150
Log Likelihood	-57.99	-40.13

The dependent variable is the survival dummy. It is 0 if the firm exits through failure in the following period, and 1 otherwise.
*Significant at the 10% level.
**Significant at the 5% level.
***Significant at the 1% level.
When year dummies are included, 1983 is the base year.

Table 8. Summary Statistics on Spin-outs

	Mean	**Standard deviation**	**Minimum**	**Maximum**	**Cases**
Spin-Out Life Span	6.60	4.80	1.00	19.00	40
Parent Technical Know-How	0.57	0.28	0.053	1.00	34
Average Parent Technical Know-How in the 3 years surrounding the spin-out's entry	0.48	0.23	0.019	1.00	38
Parent Early-Mover Know-How	0.39	0.50	0.00	1.00	40
Parent Sales Growth in the 3 years surrounding the spin-out's entry	0.085	0.67	-1.86	1.6	38
Entrepreneur Dummy	0.30	0.46	0.00	1.00	40
Number of Censored Observations	19				

Parent Technical Know-How and Average Parent Technical Know-How range from 0 to 1. Parent Early-Mover Know-How is a dummy variable. Parent Sales Growth ranges from –2 to 2. The Entrepreneur Dummy is a dummy variable.
The Censored Observations are spin-outs that have either been acquired or are still active at the end of the sample.

Table 9. Spin-out Life Span as a Function of Parent Know-how

Duration Model using Weibull Specification (Standard Errors in Parentheses)

	Equation 9a.	Equation 9b.	Equation 9c.	Equation 9d.
Variable	Coefficient	Coefficient	Coefficient	Coefficient
Constant	2.61*** (0.49)	2.73*** (0.45)	2.73*** (0.46)	2.84*** (0.46)
Parent Technical Know-How	-0.71 (0.64)	-		
Parent Early-Mover Know-How	0.68 (0.43)	0.78* (0.41)	0.66* (0.40)	0.79** (0.40)
Average Parent Technical Know-How in the 3 years surrounding the spin-out's entry	-	-1.18* (0.70)	-1.18* (0.72)	-1.30* (0.71)
Parent Sales Growth in the 3 years surrounding the spin-out's entry	-	-	0.55* (0.29)	0.59** (0.29)
Entrepreneur Dummy	-	-	-	-0.31 (0.39)
Sigma	0.73*** (0.22)	0.70*** (0.41)	0.65*** (0.19)	0.64*** (0.19)
Number of Observations	34	38	38	38
Log Likelihood	-33.03	-36.20	-33.55	-33.26

The dependent variable is the spin-out's life span (from Table 1).
The definitions of technical know-how, early mover know-how, other know-how, and the entrepreneur dummy are discussed in the text.
*Significant at the 10% level.
**Significant at the 5% level.
***Significant at the 1% level.