Covenants not to Compete, Labor Mobility, and Industry Dynamics

April M. Franco and Matthew F. Mitchell*

September 22, 2005

Abstract

Conventional wisdom among legal scholars is that contractual restrictions on employee mobility affect turnover and led to the overtaking of Massachusetts' Route 128 by Silicon Valley. We study a model of employee mobility in the spirit of Pakes and Nitzan (1984) to see when this can be the case. We show that, in fact, with certain frictions taken into account, a model of employee mobility can not only replicate the overtaking by Silicon Valley, but it can also help to explain Route 128's early dominance. Further, the model explains the relative success of firms that start as, or form, spin-outs.

*University of Iowa. Contact email: matthew-mitchell@uiowa.edu. We thank participants at the 2004 Olin IO Conference and the Iowa Macro Lunch for suggestions.

1 Introduction

One of the most common statutory forms of intellectual property protection is that which protects firms from competition from former employees. States provide varying degrees of protection. In some states, it is possible to add a limited provision of non-competition to an employment contract (a non-compete clause or "covenant not to compete," or CNC), whereby the employee is forbidden from working for any other firm in the same industry (including a start-up of her own) for a fixed length of time, while in other states, most notably California, these contracts are unenforceable.

Starting with Gilson (1998), and Hyde (2003), there has been a longstanding conjecture in the legal and sociological scholarship that suggests that outcomes will differ depending on the enforcement of CNC's. In particular, Gilson and Hyde suggest that the main reason for the success of the high technology industrial district in Silicon Valley and the failure of the one in Massachusetts' Route 128 was the differential enforcement of CNC's. They argue that the different legal environments led to more turnover, and ultimately more firms in California.

We formalize such a story can in an economic model of employee mobility. There are several stylized facts, documented by authors such as Saxenian (1994), that the model has the potential to explain. Broadly speaking, Massachusetts, the region with CNC's, was relatively successful early on, but was later leapfrogged by California, the region without CNC's. In 1965, Route 128 had approximately three times more total technology employment compared to Silicon Valley. But by 1975, Silicon Valley's employment had quintupled, while Route 128 had only tripled. This meant that Silicon Valley's total employment was 15% larger. In the years that followed, from 1975 to 1990, Silicon Valley had tripled Route 128's new job creation. If attention were focused only on semiconductor and electronic component jobs, the same pattern appears. Note that the fields that contributed to Silicon Valley's strongest growth were software and multimedia, but excluding even these areas, Silicon Valley exported double the quantity of electronics when compared to Route 128. The overtaking by Silicon Valley was remarkable.

A primary reason for the growth of Silicon Valley was the emergence of spin-outs. A "spin-out" is defined as a new firm that is founded by employees of an incumbent firm in the same industry. Unlike a "spin-off," the choice to form the firm in a spin-out is made by the employee, not the employer. A subset of firms in Silicon Valley were responsible for the emergence of many new firms. 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.[1] The descendants of Memorex and Shugart were the most productive, accounting for 64% of the cumulative revenues of the start-ups. 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)). Fairchild itself was a spin-out of Shockley.[2]

The role of spin-outs can be seen clearly by the relative success of firms involved in the spin-out process. Several papers have documented the fact that firms that are spin-outs are more successful than other entrants, in terms of survival and performance, and the firms that generate spin-outs also dominate other incumbents (See Franco and Filson (2000), Agarwal, Echambadi, Franco and Sarkar (2004) for evidence from the hard drive industry, and Klepper (2002) for evidence from the automobile industry).

Our goal is to examine the idea that legal restrictions might explain the observed differences. In order to do so, we start with a situation where a worker has private information over whether or not he has learned the production process of an employer. We study an optimal contracting problem, determining if the contract calls for the employee to start a new firm, and what compensation should be. We find important differences between environments where CNC's are allowed (so that the employer can always get at least as much surplus as if the employer doesn't start a firm) and ones where CNC's are not allowed (so that the employee can always get as much as if he starts a firm). In particular, the contract always maximizes joint surplus in the former case, but in the latter there is sometimes spin-out formation when it is not in the interest of the joint surplus maximization. Moreover, turnover is uniformly higher without CNC's, as suggested by Gilson (1998) and Hyde (2003). These difference arise due to the asymmetric information we assume.

Our point of departure in Section 2 is an environment similar in spirit to

[1] Shugart was a spin-out of IBM.

[2] Robert Noyce, one of the "treacherous eight" who founded Fairchild, is credited with being the co-inventor of the integrated circuit, for which Jack Kilby received the Nobel prize in Physics in 2000, ten years after Noyce's death.

Pakes and Nitzan (1984). In Pakes and Nitzan, firms cannot use CNC's; the only way to protect themselves from competition from former employees is to compensate those employees sufficiently that they do not want to start a new firm. This is, however, sufficient to make moot the force stressed in papers like Gilson (1998) and Hyde (2003). There are several key differences in our approach: first, we have an asymmetry of information; the employer does not know if the employee has learned. Second, we study optimal contracts, and do not endow the employee and employer with enough commitment power for workers to take below-market wages in the first period as a sort of bond posting, to be recouped later, as a way to enforce employee mobility restrictions without formal CNC's. As a result, whereas in Pakes and Nitzan, the legal structures we consider do not affect turnover, in our environment they can. We can therefore study the optimal contract under different legal structures, which affect the outside options of the parties.

In Section 3, we embed the contracting problem into an industry equilibrium model to determine the life-cycle effects of having two regions with differential enforcement of CNC's. We consider the life-cycle of firms, as well as their size distribution in cross section, when there are two regions, one where CNC's are enforced and one where they are not. We compare our results to the history of Silicon Valley and Route 128. While authors such as Gilson (1998) and Hyde (2004) stress that legal differences led to Silicon Valley's growth, the model can also use the different legal environments to explain the early advantage of Route 128, where CNC's were available. We find that the region with CNC's initially dominates the non-CNC region, because the value that firms can appropriate in the CNC region is higher than that in the non-CNC region. The reason is a standard one in the economics literature: greater protection of intellectual property creates a greater reward. In the early period, when there are few firms to create spin-outs and new firms must be non-spin-outs, the greater intellectual property protection in Massachusetts invites more entry. However, over time, the non-CNC region can overtake the CNC region. We model competition between spin-outs and the firms that form them as relatively "tough" competition relative to competition with other firms. As a result, firms in the CNC region who form spin-outs charge a lower price. If demand is sufficiently elastic, this leads to enough extra quantity to make up for the initially lower number of firms in the non-CNC region. We also show that as long as demand is sufficiently inelastic, the size of firms that either create spin-outs or are spin-outs is larger than those incumbents which do not generate spin-outs. This is in accord with cases like the hard-drive industry described above.

Pakes and Nitzan (1984) is part of a larger literature studying the degree to which intellectual property can have value without statutory protection. Other papers include Anton and Yao (1994), Baccara and Razin (2002), and Boldrin and Levin (2003). We incorporate information frictions to a similar environment, to see how effective future employment offers can be in protecting the value of firms. We show that future employment may not be very useful, relative to CNC's, in maintaining the value of innovation under the set of frictions we assume.

2 A Contracting Problem

2.1 The Players and the Objective

An employer hires an employee. During the first period of employment, the employee may or may not learn something valuable. An employee that learns something can potentially start a firm of their own. For now, we focus on the contracting problem at the start of the second period; that is, after the learning may or may not have happened. This contrasts with the Pakes and Nitzan (1984) approach, where a two period contract may allow for low wages in the initial period (backloading) as a way to protect the employer. We assume that there is no power to commit to second period payments or actions in the first period, and therefore the long term contract cannot be written. The intuition for our result will hold with any (less severe) restriction on backloading, such as the ones suggested by Pakes and Nitzan.

In the second period, the employee can either leave ($a = 1$) or stay ($a = 0$). The employee may have learned ($\theta = 1$) or not ($\theta = 0$); learning occurs with probability λ (which is common knowledge after the employee begins to work at the firm), and the outcome θ is private information of the employee. The profits from operating firms in the second period are $\pi_0(a, \theta)$ for the employer and $\pi_1(a, \theta)$ for the employee, gross of any payments they receive from one another. We assume that

Assumption 1. $\pi_1(a, 0) = \pi_1(0, \theta) = 0$

The employee can only make profits if he both learns and starts a firm. The employee also may get a payment $w(\theta)$ in the second period. This payment can be either positive or negative, so that it is in fact a transfer to the original firm.

For the employer, we assume

Assumption 2. $\pi_0(a, 0) = \pi_0(0, \theta) = \bar{\pi}_0$

Assumption 2 implies that the employee's decision only matters if he has learned, and his learning only matters if he forms a firm. As such, it is without loss of generality to normalize $a(0) = 0$. Finally, for there to be any tension in the problem, we let $\pi_0(1, 1) < \pi_0(0, 0)$; in other words, an employee leaving the firm lowers the profits of the parent firm.[3]

Rather than take a stand on surplus division, we study optimal contracts. The optimal contract orders action a to maximize expected combined profits:

$$\lambda\left(\pi_0(a(1), 1) + \pi_1(a(1), 1)\right) + (1 - \lambda)\left(\pi_0(a(0), 0) + \pi_1(a(0), 0)\right) \quad (1)$$

Under our assumptions, this is equal to

$$\lambda\left(\pi_0(a(1), 1) + \pi_1(a(1), 1)\right) + (1 - \lambda)\bar{\pi}_0$$

Since the last term is a constant, maximizing this is equivalent to maximizing

$$S(a(1)) \equiv \pi_0(a(1), 1) + \pi_1(a(1), 1)$$

Define the solution to unconstrained maximization of S by $a^*(1)$; $a^*(1) = 1$ if and only if joint profits are greater when the employee leaves (i.e. $\pi_0(1, 1) + \pi_1(1, 1) > \pi_0(0, 1)$).

Since θ is private information of the employee, there are constraints in the maximization of S by the optimal contract. First, there is an incentive compatibility constraint for the employee:

$$\theta = \arg\max_{\hat{\theta} \in \{0,1\}} \left(\pi_1(a(\hat{\theta}), \theta) + w(\hat{\theta})\right) \quad (2)$$

There are also participation constraints which differ depending on the legal environment assumed; we discuss those in the next sections. The optimal contracting problem is

$$\max_{a(1), w(\theta)} S(a(1))$$

subject to (2) and these participation constraints.

[3] Without this assumption, it is clear that employees should always leave.

2.2 Legal Environments

2.2.1 Covenants not to Compete allowed

When covenants not to compete are allowed, the participation constraint for the employee is

$$\pi_1(a(\theta), \theta) + w(\theta) \geq 0 \ \forall \theta \tag{3}$$

Since the employee can be excluded, the contract simply needs to guarantee that he can make at least his outside option from quitting but not starting up a firm, which is normalized to zero. Since the employee can never be compelled to work, this constraint must be met for each θ.

The participation constraint for the employer is

$$\lambda \left(\pi_0(a(1), 1) - w(1)\right) + (1 - \lambda) \left(\pi_0(a(0), 0) - w(0)\right) \geq \pi_0(0, 0) \tag{4}$$

Since this constraint is in expectation, we are allowing the firm to commit to payments conditional on reports of θ when the second period begins. The employer can always employ the worker in the first period (for the outside option), and then exclude the employee in the second period when CNC's are allowed. In that case, he makes $\pi_0(0, 0)$ in the second period. Exclusion is not the only option; that is, we allow $a(1) = 1$, and the employee forms a firm. However, to the extent that this action lowers expected profits for the employer (the left hand side of (4)), the employer must be compensated. It is as if the employer had written in a CNC, but then the employer pays a buyout $w(1) < 0$ to eliminate enforcement of the clause.

When CNC's are allowed, the optimal contract can in fact implement $a^*(1)$:

Proposition 1 *Suppose the participation constraints are (3) and (4). Then an optimal contract* $(a(1), w(\theta))$ *satisfies* $a(1) = a^*(1)$ *and*
(a) If $\pi_0(1, 1) + \pi_1(1, 1) < \pi_0(0, 1)$, $w(\theta) = 0$.
(b) If $\pi_0(1, 1) + \pi_1(1, 1) > \pi_0(0, 1)$,
(i) $w(0) \in [0, \pi_1(1, 1) + w(1)]$
(ii) $w(1) \geq -\pi_1(1, 1)$
(iii) $\lambda(\pi_0(1, 1) - \pi_0(0, 0)) \geq \lambda w(1) + (1 - \lambda) w(0)$

Proof. If we can show that the unconstrained maximum $a^*(1)$ is feasible, then clearly it is optimal when there are constraints. We need to show is that (2), (3), and (4) hold, and that $w()$ is as specified.

For (a), $\pi_1(a(\hat{\theta}), \theta) + w(\hat{\theta}) = 0$ for all reports $\hat{\theta}$, so clearly it is incentive compatible and satisfies the participation constraint for the worker. Note that, given $a(1) = 0$, participation of the worker requires $w(\theta) \geq 0$. For the firm, the constraint (4) reduces to

$$\pi_0(0,0) - \lambda w(1) - (1-\lambda)w(0) \geq \pi_0(0,0)$$

This implies $w(\theta) = 0$.

For (b), the participation constraint for the employee requires $w(0) \geq 0$ and $w(1) \geq -\pi_1(1,1)$. Truthful reporting of $\theta = 0$ implies

$$\pi_1(a(0), 0) + w(0) = w(0) \geq \pi_1(a(1), 0) + w(1) = w(1) \tag{5}$$

Participation of employer requires:

$$\lambda(\pi_0(1,1) - w(1)) + (1-\lambda)(\pi_0(0,0) - w(0)) \geq \pi_0(0,0)$$

so

$$\lambda(\pi_0(1,1) - \pi_0(0,0)) \geq \lambda w(1) + (1-\lambda)w(0)$$

i.e. expected compensation must be no bigger than expected losses would be if the worker were allowed to leave. Since $w(0) \geq w(1)$ from (5),

$$\lambda(\pi_0(1,1) - \pi_0(0,0)) \geq w(1)$$

The left hand side is negative, so $w(1) < 0$.

Truthful reporting of $\theta = 1$ implies

$$\pi_1(a(1), 1) + w(1) = \pi_1(1,1) + w(1) \geq \pi_1(a(0), 1) + w(0) = w(0)$$

so $\pi_1(1,1) + w(1) \geq w(0)$. ■

The employee cannot be compensated beyond the outside option when $\theta = 0$ and CNC's are allowed. In the optimal contract, either (in (a)) the employee can never start a firm and is compensated the outside option, or (in (b)), the firm allows its employee to leave if it learns, but the employee pays compensation, since learning creates value. That compensation cannot be too large, or else the employee would rather forgo the value created by leaving, and cannot be too small, since it must compensate for (expected) lost revenue of the firm. The easiest case to see is where $w(0) = 0$, so that the restrictions are simply $w(1) \geq -\pi_1(1,1)$ (compensation cannot exceed employee's profits) and $\pi_0(1,1) - \pi_0(0,0) \geq w(1)$ (compensation must be sufficient to cover lost profits of the employer). This represents, intuitively, a buyout of the CNC clause by the employee, in the amount of $-w(1)$.

2.2.2 Covenants not to Compete not allowed

If CNC's are not allowed, the employee is free to leave in the second period. In other words, his participation constraint is

$$\pi_1(a(\theta), \theta) + w(\theta) \geq \pi_1(1, \theta) \ \forall \theta \tag{6}$$

In effect, his outside option is that he can leave and earn $\pi_1(1, \theta)$. This differs from (3) only when $\theta = 1$, so that the worker has learned something of value.

On the other hand, the employer is not under any obligation to employ the worker, or enter into any particular contract for the second period; he could have just employed the worker in period one, paying the outside option, and then let the worker leave:

$$\lambda (\pi_0(a(1), 1) - w(1)) + (1-\lambda) (\pi_0(a(0), 0) - w(0)) \geq \lambda \pi_0(1, 1) + (1-\lambda) \pi_0(1, 0) \tag{7}$$

In this case, the employer takes the chance that the worker has learned, and that his profits will be lowered to $\pi_0(1, 1)$ if he has.

Unlike the last case, the optimal contract cannot always implement $a(1) = a^*(1)$. In particular, when $\pi_0(1, 1) + \pi_1(1, 1) < \pi_0(0, 1)$, so that joint surplus would be maximized by $a(1) = 0$, the contract sometimes allows the worker to leave.

Proposition 2 *Suppose the participation constraints are (6) and (7). Then the optimal contract* $(a(1), w(\theta))$ *satisfies*
(a) If $\pi_0(1, 1) + \pi_1(1, 1) < \pi_0(0, 1)$,
(i) if $\pi_1(1, 1) \leq \lambda(\pi_0(0, 0) - \pi_0(1, 1))$, $a(1) = 0$, $w(\theta) \in [\pi_1(1, 1), \lambda(\pi_0(0, 0) - \pi_0(1, 1))]$
(ii) if $\pi_1(1, 1) > \lambda(\pi_0(0, 0) - \pi_0(1, 1))$, $a(1) = 1$, $w(\theta) = 0$
(b) If $\pi_0(1, 1) + \pi_1(1, 1) > \pi_0(0, 1)$, $a(1) = 1$ *and* $w(\theta) = 0$.

Proof. We start with (b), where the full information outcome is implemented. Clearly the policy is incentive compatible and satisfies the participation constraint if and only if $w(0) = w(1) \geq 0$. Participation for the employer requires

$$\lambda (\pi_0(1, 1) - w(1)) + (1 - \lambda) (\pi_0(0, 0) - w(0)) \geq \lambda \pi_0(1, 1) + (1 - \lambda) \pi_0(1, 0)$$

When $w(0) = w(1)$, this is satisfied only if $w(1) = w(0) \leq 0$, so it must be that $w(\theta) = 0$.

For (a), first consider the situation where $\pi_1(1,1) \leq \lambda(\pi_0(0,0) - \pi_0(1,1))$. Since the employee is never allowed to earn profits ($a(1) = 0$), it must be the case that $w(0) = w(1)$ for incentive compatibility, and $w(1) \geq \pi_1(1,1)$ for employee participation. This leaves employer participation as

$$\begin{aligned}
\lambda\left(\pi_0(0,0) - w(1)\right) + (1-\lambda)\left(\pi_0(0,0) - w(0)\right) &\geq \lambda\pi_0(1,1) + (1-\lambda)\pi_0(1,0) \\
\pi_0(0,0) - w(1) &\geq \lambda\pi_0(1,1) + (1-\lambda)\pi_0(1,0) \\
\lambda(\pi_0(0,0) - \pi_0(1,1)) &\geq w(1)
\end{aligned}$$

so, combining, $\pi_1(1,1) \leq w(\theta) \leq \lambda(\pi_0(0,0) - \pi_0(1,1))$

If $\pi_1(1,1) > \lambda(\pi_0(0,0) - \pi_0(1,1))$, then the same steps imply that it is impossible for all the constraints to be met with $a(1) = 0$. The only alternative is to set $a(1) = 1$; in that case, the same steps as in the proof of (b) imply that $w(\theta) = 0$. ■

Notice that, when $\pi_0(1,1) + \pi_1(1,1) < \pi_0(0,1)$, spin-outs can arise from firms where $\pi_1(1,1) > \lambda(\pi_0(0,0) - \pi_0(1,1)))$. In other words, in that case, spin-outs come from firms with relatively low λ. Firms that are likely to have employees who learn pay to keep the employees from competing. If the firm were to follow a policy of always retaining workers, incentive compatibility requires that the worker be paid regardless of θ, and the payment must be enough to retain the worker when $\theta = 1$. For low λ, the worker is not likely to have learned, so the firm would rather "take its chances" with the facing a spin-out if $\theta = 1$, even though that case results in a reduction of profits and no compensation.

When $\lambda = 0$ or $\lambda = 1$, there is no private information, and hence no difference between the choice $a(1)$ under the two legal environments. This is just as in Pakes and Nitzan.

2.3 Implications

In this environment, CNC's lead to implementing the full-information surplus maximizing outcome. Without CNC's, there is a region where that does not occur. We explore this implication in more detail below, in trying to explain the early advantage of Massachusetts in the computer industry; our story is that higher returns to entry, stemming from the ability to use CNC's, made entry more attractive in Massachusetts.

Moreover, as stressed in the legal literature, turnover is greater without CNC's. When $\pi_0(1,1) + \pi_1(1,1) < \pi_0(0,1)$ and $\pi_1(1,1) > \lambda(\pi_0(0,0) -$

$\pi_0(1,1)$), the optimal contract with CNC's (and the full information joint profit maximizing outcome) would keep the employee from starting a competing firm, but that outcome is impossible without CNC's.

Note that, in contrast with the conventional wisdom as in Gilson, CNC's (weakly) raise the profit for the employer regardless of whether we are in case (a) or (b). Moreover, the benefit of CNC's is strict when we are in case (a), regardless of whether $\pi_1(1,1)$ is greater or less than $\lambda(\pi_0(0,0) - \pi_0(1,1))$. In other words, regardless of whether CNC's increase turnover, they still benefit the employer. This contrasts Gilson, who has in mind that turnover generates a spillover that makes California attractive to all firms. We will show in the next section that the rise of California can be explained without such a spillover.

Finally, it should be said that we did not allow the contract to order the original firm to *exit* for some reports of θ. If this were possible, then there is essentially no friction from the asymmetric information; the planner can make either firm the residual claimant on all joint profits. We think our restriction is natural in this setting for two reasons. First, such a contract would be hard to describe if the employer had many lines of business. Second, and perhaps more importantly, while retaining employees is allowed even though it may restrict competition, it would likely be the case that regulatory authorities would take a dim view of a contract that pre-specified exit – especially since it would be made in exchange for a payment from the worker who wanted to start a firm.

2.4 Backloading Wages

Suppose that, in an initial period of employment where the learning took place, the employee could pay for the right to learn through below market wages. Such an allowance wouldn't change anything in terms of the world with CNC's, since our optimal contract already does as well as the full information optimum. However, this sort of bond-posting by the employee (which can be interpreted as backloaded wages, since the employee is paid less initially but more if he is retained) would make mobility more costly after learning took place, and as a result could be useful in avoiding $a(1) = 1$ when $a^*(1) = 0$. In fact, by backloading the wages enough, the participation constraint in this case is loosened to the point where it is the same as the case where CNC's are enforced.

In the model of Pakes and Nitzan, where learning is sure, the ability to write two period contracts affects surplus division but not whether or not employees leave; it corresponds to our model with $\lambda = 1$. Pakes and Nitzan can therefore never get differences in turnover based on restrictions like CNC's, even without two period contracts.

In the next section we develop an industry equilibrium model that incorporates the contracting problem described by the propositions of this section. In doing so, we are assuming that backloading is impossible. Wages before learning must therefore be equal to the outside option. This leaves the second period of employment, after learning, to fall within the framework we have studied in this section.

3 An Equilibrium Model of spin-outs

In this section we embed the contracting problem of the previous section into an industry equilibrium model. The goal is to see how spin-out formation can be affected by legal restrictions, and how industry evolution can in turn differ across locations with different restrictions.

3.1 Preferences

A representative consumer has income I in each of two periods to spend on an industry of differentiated products. When a mass N of products are available, the representative consumer has utility function

$$\left(\int_0^N c_j^\rho dj \right)^{1/\rho}$$

Where $\rho \in (0, 1)$, i.e. demand is elastic.[4]

3.2 Technology

Products are produced in two locations. Location 1 allows covenants to not compete. Location 2 does not allow such covenants.[5] There are two

[4]We could allow the consumer to substitute intertemporally; it would only strengthen our results, as the consumer would consume more in period 2 (when goods are relatively cheap), further aiding in the growth of the region that does not allow CNC's.

[5]There is no difficulty in adding more locations, but since only two legal environments are to be considered, there are only two meaningful regions in the model.

periods. In period one, each product is produced by one firm, which are called entrants. Let E_i be the number of such firms in location i. The number of products is then $N = E_1 + E_2$. Firms hire a single worker (regardless of the amount produced), and then produces output at constant marginal cost m, which can be thought of as a capital cost. As described in the previous section, workers are paid their outside option (normalized to zero) in the first period.

In the second period, the employer in the first period can either retain the worker and pay a wage w, or allow the worker to leave and potentially form a spin-out producing an identical product to its employer. The two compete by simultaneously choosing quantities, in Cournot competition. Our assumption implies that spin-outs and the firms that create them are tougher competitors for one another than they are for the average firm in the industry. We think this is a reasonable depiction of such firms, since the spin-outs are often making a particularly similar product to its parent firm. The elasticity parameter ρ dictates this difference: for ρ near 1, demand is nearly perfectly elastic across products, so there is little difference between competition the competition between spin-outs and their parents, and the competition across firms that are unrelated. When ρ is near zero, competition between unrelated firms becomes nearly Cobb-Douglas, while spin-outs and their parents remain perfect substitutes. This makes competition between spin-outs and their parents relatively tough.

Once again, λ denotes the probability of a worker learning; once again this outcome is private information of the employee. When the firm initially enters, it draws λ from some distribution described by c.d.f. $F(\lambda)$. Denote by V_i be the value of starting a new firm in period one at location i. Rather than focus on any specific model of firm entry, we take entry as exogenously determined by $E_i = G(V_i)$, an increasing function. We view this as a minimal structure: higher value to entry in a given region induces more entrants. Since spin-outs do not form new products, we always have N products available.

Our model will not imply that V_i is equated across regions. We have in mind that each region might have a fixed factor (for instance land) that becomes more expensive as E_i increases, but is not modelled. The value V_i is *gross* of payments to this factor, so that even though V_i differs across regions, the net reward to entry, after payments to the fixed factor, might be identical.

3.3 Equilibrium

3.3.1 Pricing and Profits of Firms

In period one, there are only the entrants, so each product is monopolized. The monopolistic competition outcome has prices equal to the usual markup rule:

$$p(1) = m/\rho \tag{8}$$

The one in parenthesis denotes the fact that the industry is monopolized.

In period two, either the employee is retained, in which case the firm remains a monopolist producer of the product and the price is $p(1)$, or the employee leaves. If the employee leaves, he can start a new firm and compete Cournot with the original firm. Denote by q and q_{-1} the quantities produced by the two firms in the duopoly, so $c_j = q + q_{-1}$. The first order condition for q in the duopoly situation is

$$A\left((\rho - 1)(q + q_{-1})^{\rho-2} q + (q + q_{-1})^{\rho-1}\right) = m \tag{9}$$

where A is a constant. Solving (9) and the analogous equation for q_{-1}, and then substituting into the demand function $p_j = Ac_j^{\rho-1}$ gives

$$p(2) = 2\frac{m}{\rho + 1} \tag{10}$$

for duopoly provided products.

Denote by $c(1, N, D)$ and $c(2, N, D)$ the quantity consumed for products provided by monopoly and duopoly, respectively, when a total of N products are available, D of which are duopolies, the rest of which are monopolized. The first order condition for the consumer gives

$$\frac{c(1, N, D)}{c(2, N, D)} = \left(\frac{p(2)}{p(1)}\right)^{\frac{1}{1-\rho}} = \left(\frac{2\rho}{\rho + 1}\right)^{\frac{1}{1-\rho}} \tag{11}$$

Let $\Pi(1, N, D)$ be monopoly profits for a given product when a mass N of products is provided, D of which are duopolies and let $\Pi(2, N, D)$ be the duopoly profits, gross of any payments to the employee. Then

$$\begin{aligned} \Pi(1, N, D) &= (p(1) - m)c(1, N, D) = m\frac{1-\rho}{\rho}c(1, N, D) \\ \Pi(2, N, D) &= (p(2) - m)\frac{c(2, N, D)}{2} = m\frac{1-\rho}{1+\rho}\frac{c(2, N, D)}{2} \end{aligned}$$

An important and useful fact is that the ratio of monopoly to duopoly profits is independent of N and D. It results from the fact that, according to (11), the ratio of consumption in the industries with different market structure depends only on ρ:

$$\Pi(1,N,D)/\Pi(2,N,D) = 2\frac{\rho+1}{\rho}\left(\frac{2\rho}{\rho+1}\right)^{\frac{1}{1-\rho}} = 2^{\frac{2-\rho}{1-\rho}}\left(\frac{\rho}{\rho+1}\right)^{\frac{\rho}{1-\rho}} > 2 \quad (12)$$

Spin-out formation depends solely on λ and this ratio, so we can now study spin-outs independently from N and D.

3.3.2 Spin-out Formation

Now we can discuss the issue of spin-out formation for each location and λ. In the language of section 2, we have $\pi_0(0,\theta) = \pi_0(a,0) = \Pi(1,N,D)$. If the employee leaves, $\pi_0(1,1) = \pi_1(1,1) = \Pi(2,N,D)$. Since, in every case, industry profits go down as a result of the spin-out ($\Pi(1,N,D) > 2\Pi(2,N,D)$), we are always in the case (a) of Propositions 1 and 2, where $\pi_0(1,1) + \pi_1(1,1) < \pi_0(0,1)$. According to Proposition 1, where CNC's are allowed, spin-outs are never formed ($S_1 = 0$). According to Proposition 2, in region 2, spin-outs are formed if

$$\pi_1(1,1) > \lambda(\pi_0(0,0) - \pi_0(1,1)).$$

Substituting with the values for $\Pi(1,N,D)$ and $\Pi(2,N,D)$ we have

$$\Pi(2,N,D) > \lambda(\Pi(1,N,D) - \Pi(2,N,D))$$

or

$$\Pi(1,N,D)/\Pi(2,N,D) < \frac{1+\lambda}{\lambda}$$

Replacing from the formula in (12) and solving for λ, a spin-out is formed by an employee who learns at a firm where

$$\lambda < 1/(2^{\frac{2-\rho}{1-\rho}}\left(\frac{\rho}{\rho+1}\right)^{\frac{\rho}{1-\rho}} - 1) \equiv \bar{\lambda}$$

The critical value $\bar{\lambda}$ is an increasing function of ρ :

In the extreme cases, we see that

$$\lim_{\rho \to 0} 1/(2^{\frac{2-\rho}{1-\rho}} \left(\frac{\rho}{\rho+1}\right)^{\frac{\rho}{1-\rho}} - 1) = \frac{1}{3} \tag{13}$$

The implication of (13) is that, for all ρ between zero and one, there are λ that lead to spin-outs in region 2; namely, firms with $\lambda < 1/3$ always let workers go, regardless of ρ. At the other extreme

$$\lim_{\rho \to 1} 1/(2^{\frac{2-\rho}{1-\rho}} \left(\frac{\rho}{\rho+1}\right)^{\frac{\rho}{1-\rho}} - 1) = \frac{1}{4e^{-\frac{1}{2}} - 1} < 1$$

Here the implication is that firms which draw sufficiently high λ will not allow spin-outs, even in region 2.

The number of duopoly industries is also the number of spin-outs in location 2. That number is

$$D = S_2 = E_2 \int_0^{\bar{\lambda}} \lambda dF(\lambda).$$

The number of monopolized industries is everything else; i.e. $E_1 + E_2 \left(1 - \int_0^{\bar{\lambda}} \lambda dF(\lambda)\right)$.

Firms with $\lambda > \bar{\lambda}$ pay to keep their workers from leaving; we know from Proposition 2 that this payment w is greater than or equal to $\Pi(2, N, D)$ and less than or equal to $\lambda(\Pi(1, N, D) - \Pi(2, N, D))$.We don't take a stand on surplus division, and treat the wage as a parameter falling in this interval. All of our analysis is valid for any such wage.

3.3.3 Equilibrium Conditions

For any N and D, consumers simply solve

$$\max_{c(1,N,D),c(2,N,D)} \left((N-D)c(1,N,D)^{\rho} + Dc(2,N,D)^{\rho}\right) \tag{14}$$

$$\text{s.t.}$$

$$(N-D)p(1)c(1,N,D) + Dp(2)c(2,N,D) = I$$

For any set of parameters, an equilibrium is a list of $\{E_i, V_i, S_2, c(1,N,D), c(2,N,D), \Pi(1,N,D), \Pi(2,N,D)\}$ where $c(1,N,D)$ and $c(2,N,D)$ solve the consumers problem in (14), given I and the prices in (8) and (10), and

$$\begin{aligned}
V_1 &= \Pi(1,N,0) + \beta\Pi(1,N,S_2) \\
V_2 &= \Pi(1,N,0) + \beta\begin{pmatrix} \int_0^{\bar{\lambda}} (\lambda\Pi(2,N,S_2) + (1-\lambda)\Pi(1,N,S_2))dF(\lambda)+ \\ (1-F(\bar{\lambda}))(\Pi(1,N,S_2) - w) \end{pmatrix} \\
E_i &= G(V_i) \\
N &= E_1 + E_2 \\
D &= S_2 = E_2\int_0^{\bar{\lambda}} \lambda dF(\lambda) \\
\Pi(1,N,D) &= m\frac{1-\rho}{\rho}c(1,N,D) \\
\Pi(2,N,D) &= m\frac{1-\rho}{1+\rho}\frac{c(2,N,D)}{2}
\end{aligned}$$

Where $\beta \in (0,1)$ is the discount factor for firms.

3.4 Analysis

3.4.1 Value of Firms

A classic question in the literature on innovation is the connection between the reward to innovation (here, entry in the first period) under various legal protections for intellectual property. Note that, for region 2, either $\lambda > \bar{\lambda}$, and the employee retains the worker with $w > 0$ (as implied by Proposition 2), or the firm allows the worker to leave, and the firm makes lower expected profits in period two(as implied by (12)). This implies that $V_1 > V_2$. But how much?

The worst case scenario is where $\lambda = 1$ and $w = \Pi(1, N, D) - \Pi(2, N, D)$,so that the employee takes all of the surplus in the second period, leaving the firm with only the duopoly profits. In that case, the firm would make monopoly profits with CNC's, but makes only 25% of that as a duopoly firm. Further, the 25% that remains is only a function of the form of competition; with Bertrand competition, the firm could still make monopoly profits in the second period with CNC's, and make no profits without CNC's. In other words, given the frictions we have assumed, CNC's can have a large effect on the incentive to enter.

3.4.2 Life-Cycle of Output by Region

Since entry is increasing in value, the fact that $V_1 > V_2$ implies

Life-Cycle Result 1 (initial dominance by CNC region): $E_1 > E_2$

Initially, there are more firms in location 1, where CNC's are allowed. All products are priced identically in period one, so differences in firm numbers are the only source of differences in output across the two regions. As a result, region 1 produces more output in period 1.

In the second period, output in region 1 is

$$Y_1 \equiv E_1 c(1, N, D)$$

and output in region 2 is

$$Y_2 \equiv E_2 \left(\int_0^{\bar{\lambda}} \left(\lambda c(2, N, D) + (1 - \lambda) c(1, N, D) \right) dF(\lambda) + (1 - F(\bar{\lambda})) c(1, N, D) \right)$$

There are three terms. The terms in the integrand reflects entrants that allow spin-outs ($\lambda < \bar{\lambda}$). With probability λ, these become duopolized and sell $c(2, N, D)$, and with probability λ they still produce $c(1, N, D)$. The final term is entrants that have high enough λ that they do not allow spin-outs, and hence continue to produce $c(1, N, D)$.

Output in region 2 divided by output in region 1 is

$$\frac{Y_2}{Y_1} = \frac{E_2}{E_1} \left(\int_0^{\bar{\lambda}} \left(\lambda \left(\frac{2\rho}{\rho + 1} \right)^{\frac{-1}{1-\rho}} + (1 - \lambda) \right) dF(\lambda) + (1 - F(\bar{\lambda})) \right)$$

which simplifies to

$$\frac{Y_2}{Y_1} = \frac{E_2}{E_1}\left(\left(\left(\frac{2\rho}{\rho+1}\right)^{\frac{-1}{1-\rho}} - 1\right)\left(\int_0^{\bar{\lambda}} \lambda dF(\lambda)\right) + 1\right) \tag{15}$$

In period two, region 2 adds spin-outs S_2. Moreover, in any industry that has a spin-out and therefore becomes a duopoly, the price is lower. This leads to greater output from those products in region 2 where spin-outs occur. We know from (11) that output in duopolized industries is always greater than in monopolized industries. It is easy to compute that

$$\lim_{\rho\to 0}\left(\frac{2\rho}{\rho+1}\right)^{\frac{-1}{1-\rho}} = \infty$$

so that, when demand is nearly unit elastic, output from products that are duopolized is *arbitrarily* bigger than output in monopolized industries, and the middle term $\left(\frac{2\rho}{\rho+1}\right)^{\frac{-1}{1-\rho}} - 1$ in (15) approaches infinity.

Moreover, since, for all ρ, $\bar{\lambda} \geq 1/3$, we know that the term $\left(\int_0^{\bar{\lambda}} \lambda dF(\lambda)\right)$ is strictly positive if $F(1/3) > F(0)$. Then if E_1/E_2 is uniformly bounded for all ρ, it must be the case that Y_2/Y_1 approaches infinity as ρ approaches zero. We have

Life Cycle Result 2 (eventual overtaking by no-CNC region) Suppose $F(1/3) > F(0)$ and $\frac{E_1}{E_2}$ is bounded. Then Y_2/Y_1 is arbitrarily large as ρ gets close to zero.

One natural concern is whether or not, as ρ becomes extreme, we should expect any entry in location 2. However, note $\Pi(1, N, D)/\Pi(2, N, D)$ in (12) is bounded above (by 4) for all ρ between zero and one; the ratio of profits in the two regions is not diverging, even though output is. As such, it is natural to assume that E_1/E_2 is bounded for such ρ; a sufficient condition is that $G(V)$ is a *strictly* increasing function. In that case, if there are any spin-outs ($F(1/3) > F(0)$), then output in region 2 is arbitrarily bigger than that of region 1 as ρ becomes close to zero.

The model predicts that, for sufficiently tough competition between spin-outs and their parents, that regions with CNC's initially dominate, but are eventually overtaken by regions which don't allow CNC's.

3.4.3 Size Distribution of Firms

The relative size of firms that do not create or enter as spin-outs, relative to those that do, is $2c(1, N, D)/c(2, N, D)$. This an increasing function of ρ :

Whenever $\rho < .633$ firms that either are spin-outs or create spin-outs are larger than their counterparts that do not. This is the stylized fact of the hard drive industry: firms that create or enter as spin-outs are relatively prominent. Once again, the requirement is that demand be elastic enough across products, or, in other words, that competition between spin-outs and their parents be relatively tough. On the other hand, since

$$\lim_{\rho \to 1} \left(\frac{2\rho}{\rho + 1} \right)^{\frac{1}{1-\rho}} = e^{-\frac{1}{2}},$$

for ρ near one, every entrant in region 2 leads to $\sqrt{e}$ (approximately1.6) times the output of an entrant in region 1, with each of the two firms producing $\sqrt{e}/2$ of the monopoly firm's output.

4 Conclusion

The conventional wisdom of legal scholars on the importance of CNC's for spin-out formation can be rationalized in a standard model of employee mobility, but not without some frictions added. We have incorporated such frictions and shown that, in fact, the model can explain higher turnover in places where CNC's are not allowed, and in fact can have the spin-outs be concentrated in a few firms. In addition, we have used the standard economic intuition that higher returns generate greater innovation to explain why, in the early stage, the region which enforces CNC's can have greater firm numbers.

Note that this outcome requires several ingredients. Wages could not be backloaded, so that the employee gets paid below his outside option while learning, in exchange for higher wages if he stays for the second period. Whether employees had learned or not was private information. And the contract could not pre-specify exit by the employer (in exchange for a payment) if the employee learned, so that the employee could obtain higher profits.[6]

The model can deliver both the life cycle facts (initial dominance by the CNC region and eventual overtaking by the non-CNC region), as well as the size distribution fact (large firms are one that either start as spin-outs, or ones that create spin-outs). The essential ingredient is that competition between spin-outs and their parent firm be relatively tough, compared to competition between unrelated firms.

In California, the debate focusing on the enforcement of CNC's has recently resurfaced, prompted by the California District Court's Electro Optical decision, in which the court determined than if a former employee's position at a "direct competitor" will require "inevitable disclosure" of the former employer's trade secrets, the former employee can be prevented from working for said competitor (Sandberg, 2000.) This has lead to several companies in California to file suits against their competitors. However, when the State Supreme Court depublished the ruling, the decision could not be used as precedent in other cases.

There are several points raised by the model that suggest future possibilities. Since the benefits of CNC protection are decreasing in the degree of competition, firms in concentrated industries would be most interested in locating in CNC enforcing regions, while firms in more competitive industries would be less willing to go out of their way to get the protection. In industries where spin-outs are especially important, this force might be important

[6] We focus on a particular type of asymmetric information, where the employee knows the outcome of the learning uncertainty. We could do the reverse: consider the case where only the *employer* knows. This is clearly important in many industries where employees seek letters of recommendation to attain future employment; the job market for economists is one example.

in determining the location of industries across locations.

References

[1] Agarwal, Rajshree, Raj Echambadi, April Franco, and MB Sarkar (2004), "Knowledge Transfer Through Inheritance: Spin-out Generation, Development and Survival," *Academy of Management Journal* 47 (4), pp. 501-523.

[2] Anton, James, and Dennis Yao (1994), "Expropriation and Inventions: Appropriable Rents in the Absence of Property Rights," *American Economic Review,* Vol. 84 (1), pp. 190-209.

[3] Baccara, Mariagiovanna, and Ronny Razin (2002), "From Thought to Practice: Appropriation and Endogenous Market Structure with Imperfect Intellectual Property Rights," Mimeo, New York University.

[4] Boldrin, Michele, and David Levine (2003), "Perfectly Competitive Innovation," Mimeo, University of California, Los Angeles.

[5] Braun, Ernest and Stuart MacDonald (1982), *Revolution in Miniature*, Cambridge, MA: Cambridge University Press.

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

[7] Franco, April and Darren Filson (2000), "Spin-outs: Knowledge Diffusion through Employee Mobility," Federal Reserve Bank of Minneapolis Staff Report 272.

[8] Gilson, Ronald J. (1999), "The Legal Infrastructure of High Technology Industrial Districts: Silicon Valley, Route 128, and Covenants Not To Compete," *New York University Law Review* 74 (3), 575-629.

[9] Hyde, Alan (2003), *Working in Silicon Valley: Economic and Legal Analysis of a High-Velocity Labor Market*, M.E. Sharpe.

[10] Klepper, Steven (2002), "The capabilities of new firms and the evolution of the US automobile industry," *Industrial and Corporate Change*, 11 (4), 645-666.

[11] Moen, Jarle (2000), "Is Mobility of Technical Personnel a Source of R&D Spillovers?" NBER Working Paper No. W7834.

[12] Pakes, Ariel and Saul Nitzan (1984), "Optimal Contracts for Research Personnel, Research Employment and Establishment of 'Rival' Enterprises," *Journal of Labor Economics,* 1 (4), 345-65.

[13] Sandberg, Brenda (2000), "After Uproar, California Supremes Depublish Trade Secrets Ruling," *The Recorder*, April 19.

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