Optimal Enforcement of Non-compete Covenants

Suman Ghosh* Kameshwari Shankar†

February 13, 2015

Abstract

Non-compete covenants or Covenant Not to Compete (CNC) are clauses in employment contracts in which the employee agrees not to gain employment with a competitor firm. In this paper, we study the efficiency aspects of such contracts by incorporating the effect of labor mobility restrictions on knowledge transfer across firms, investment decisions by firms and investment by workers. Following research that shows state-wise variations in the degree of CNC enforcement, we allow the extent of CNC enforcement in courts to vary as a matter of regulatory policy and derive the optimal level of enforcement. We also look at how regulations around CNCs should be optimally designed when employers can use collusive agreements, such as “no poaching” agreements, as an alternative to non-compete clauses. Given recent allegations of employer collusion among large Silicon Valley firms, we argue for a cautious approach in designing policies on CNC enforcement.

Keywords: labor mobility, human capital, noncompete clauses, collusion
JEL Classification Codes: J24, J41, J63, K31

1 Introduction

As knowledge-intensive industries have grown in importance in the last few decades, several recent studies have focused on the role of worker mobility as an important conduit through which knowledge is transferred across firms (Moen, 2005; Franco and Filson, 2006.) In a recent paper, Tambe and Hitt (2014) show that over the last two decades, productivity spillovers through worker mobility across IT firms have contributed 20 to 30 percent as much to productivity growth as the firms’ own

*Corresponding Author- Address: Florida Atlantic University; Department of Economics; 777 Glades Road; Boca Raton; FL 33131. Email: sghosh@fau.edu.
†Department of Economics and Business, Colin Powell School of Civic and Global Leadership, NAC 5/139C, City College of New York, New York NY 10031. Email: kshankar@ccny.cuny.edu.
IT investments. Shankar and Ghosh (2013) explain how worker turnover plays a unique role in optimally allocating worker knowledge in high-technology industries where firm-specific technological shocks create a continuous cross-section of expanding and contracting firms.

Given the apparent benefits of worker turnover among knowledge-intensive firms, many researchers have argued in favor of the free flow of ideas facilitated through employee mobility in these industries (Saxenian, 2004; Gilson, 1999; Hyde, 2003). Employers, on the other hand, have generally attempted to restrict such transfer of knowledge via exiting workers. One such contractual restriction that firms and workers often enter into is a non-compete agreement, or Covenant Not to Compete (CNC) clause. These agreements prevent employees from accepting employment in a competing firm for a specified period of time after they leave the current employer. Hence by restricting worker mobility, they limit the flow of knowledge to competitors.

There are widespread differences from state to state within the US in the extent to which CNCs can be legally enforced. Some states such as California have made such contracts completely unenforceable. At the other extreme, Texas and Massachusetts are known for being very permissive in their enforcement of such contractual employment restrictions. State-wise variation in the enforcement of CNCs has provided fertile ground for researchers to analyze the effects of labor mobility restrictions. Not surprisingly, several papers, such as Fallick et. al. (2006) and Marx et. al. (2009), show that more permissive legal enforcement of CNCs reduces worker mobility among firms. To the extent that labor turnover plays a positive role in the efficient allocation of worker knowledge and ideas across firms, these studies would suggest that CNCs have a negative effect on growth and productivity. Gilson (1999) and Hyde (2003), for example, attribute the success of Silicon Valley firms in California and correspondingly the failure of technology firms in the Route 128 cluster of Massachusetts to differences in CNC enforcement across the two states. Samila and Sorenson (2011) find that permissive enforcement of CNCs has an adverse effect on innovation. The research we have described thus far highlights the negative effect of CNCs on the incentive for workers to invest in developing new ideas and on the optimal transfer of such ideas through worker mobility.

On the other hand, it is arguable that CNCs can also have positive output effects by limiting turnover and allowing firms to protect their human capital investments and intellectual property thereby enhancing such investments by firms. Kim and Marschke (2005) show that turnover of scientists reduces R&D expenditures by the firm. The hold-up of firms’ investments in training and general human capital accumulation caused by labor market competition more generally has also
been studied extensively in the literature.\textsuperscript{1} However, the role of CNCs in mitigating this investment hold-up has received very little attention. Thus our first objective in this paper is to provide an integrated theoretical framework to explore all efficiency aspects of CNCs, namely, the efficient transfer of knowledge through labor mobility and optimal investment decisions by firms as well as workers.

Secondly, our analysis of labor mobility trade-offs provides novel insights regarding how strongly CNCs should be enforced by law. A key drawback of the current debate over state enforcement of CNCs is that regulation has been treated as a binary choice. In contrast, the actual extent of CNC enforcement by courts varies greatly across states. Yet, much of the research on CNCs has focused only on states that fall at either end of the enforcement spectrum - states such as California with no enforcement and states like Massachusetts with extremely permissive enforcement - leaving out the vast majority of states in the middle.\textsuperscript{2} Bishara (2011) gives a detailed account of the variation in enforcement policy across the 50 states. He rightly points out that researchers have been using a simplistic analysis of CNCs by cherry-picking certain states for analysis while ignoring the fact that strength of enforcement varies across states. In line with Bishara’s critique, we argue that posing the problem in this binary fashion in terms of existence or absence of CNC enforcement severely limits the analysis and ignores the role that policies surrounding CNC enforcement can play in balancing firm and worker investments and the flow of knowledge through worker turnover.\textsuperscript{3}

Further, restricting the policy analysis to two extremes - to enforce or not to enforce CNCs - is even more problematic when we recognize that firms often tailor CNCs in their employment contract to balance counteracting effects of labor mobility restrictions. Thus, we adopt a unique approach to this issue by characterizing the firm’s choice of CNCs in the employment contract as a continuous variable that balances this trade-off rather than as a discrete decision process. By doing so, we account for differences in the strength of CNCs in the firm’s employment contract. Similarly we also allow the strength of regulatory enforcement to vary continuously to derive an optimal enforcement level that maximizes surplus. This allows us to provide more nuanced policy prescriptions about how CNC enforcement should be optimally designed. Hence we argue, contrary to some researchers,

\begin{itemize}
  \item \textsuperscript{1}Acemoglu and Pischke (1998, 1999a and 1999b) are seminal papers on this issue.
  \item \textsuperscript{2}Marx et. al. (2009) look at Michigan which reversed its ban on non-compete agreements and hence established a highly permissive environment for CNCs.
  \item \textsuperscript{3}Bishara et. al. (2012) is a recent attempt in this direction to study CEO compensation by incorporating “degree of enforcement” of CNCs.
\end{itemize}
that zero legal enforcement of non-competes is sometimes sub-optimal.\footnote{Hyde (2010), for example, argues that these clauses should be banned across the United States, as they are in California.}

Specifically, we argue the following. Output is jointly determined by investment from firms and investment from workers as well as the match quality between firms and workers. When human capital investments are general, firms cannot appropriate all gains from their investment. Hence they under-invest. On the other hand, outside wages for workers grow more slowly than output due to uncertain match quality with a new firm. This leads workers also to under-invest in human capital. CNCs restrict the worker from joining a competing firm and hence effectively convert general human capital acquired by workers into firm-specific human capital. As a result the market wage falls and this further dampens incentives for workers to invest. Firm investments, however, improve as CNCs allow firms to appropriate more of the returns from their investment. This investment trade-off is present in both profits to the firm and overall surplus although the firm investment effect is relatively stronger on profits than on surplus. As a result the optimal employment contract entails a weakly less restrictive CNC than that chosen by the firm. At the same time we find that zero legal enforcement of CNCs is efficient if and only if the marginal product of worker investment is sufficiently higher than the marginal product of firm investment. In all other cases, the optimal employment contract includes a CNC with positive strength and further if firm investment is productive enough relative to worker investment, the firm always chooses the optimal employment contract eliminating any need for regulation.

To our knowledge, the only other paper that explores the firm-worker investment trade-off in this manner is Garmaise (2011). He also shows that CNC enforcement by courts encourages firm investment in its managers, but lowers managers’ own investment in human capital. However, unlike our paper, he treats the existence of CNCs as an exogenous variable, whereas we allow the decision to use CNCs in employment contracts to be determined in equilibrium. Further, as with other papers on CNC enforcement, Garmaise also treats CNC enforcement as a binary variable and hence he does not address the question of optimal CNC enforcement as we do here.

Finally we also examine some unintended consequences of regulating CNCs, namely the use of “no poaching” agreements between firms. Recently, a number of high-tech companies in Silicon Valley including Apple, Google, Intel, Adobe and eBay were accused of colluding to not hire each others’ employees.\footnote{“Justice Department Requires eBay to End Anticompetitive “No Poach” Hiring Agreements,” \textit{US Department of Justice Release}, May 1, 2014.} While the conspiracy was clearly illegal, as we argue in this paper, restrictions...
on labor turnover can improve efficiency. When an alleged anti-competitive act also has potentially positive effects on surplus, the legal process should examine whether there are other less egregious ways of limiting turnover. One such alternative is a non-compete agreement between the firm and its employees. Since the strength of CNCs in the employment contract can be tailored in a continuous manner to achieve more optimal labor mobility restrictions, it is clearly preferable to no poaching agreements that severely restrict turnover possibilities for the employee. We show that regulations around CNC enforcement can, in some cases, facilitate collusion and further that the relative productivity of human capital investments by firms and workers can affect the firm’s choice between no poaching agreements and CNCs when regulation limits the enforceability of CNCs. To the extent that strict regulations around CNCs in states such as California make it difficult for firms to utilize such employment contracts to protect their investments, we argue that firms may have had a greater incentive to use no poaching agreements. As a result, regulatory restrictions on CNCs may have had the opposite effect of reducing labor mobility rather than enhancing it.

The paper is structured as follows. Section 2 describes the basic model for our follow-up analysis. In Section 3 we solve the equilibrium employment contract and compare it with the optimal contract. Section 4 analyzes the impact of CNC regulations when the firm can enter into no poaching agreements with its competitors. Section 5 concludes. All proofs are in the appendix.

2 Model

We consider a two period model of production. At the beginning of Period 1 a single firm hires a worker of unknown match quality. The Period 1 employment contract includes a hiring wage and a CNC which places restrictions on the worker’s future employment in a new firm. The strength of these restrictions is captured by the parameter $z \in [0, 1]$. There is significant variation in the terms of the CNC that firms can adopt. For example a non-compete agreement typically specifies the duration and geographic scope for which the contract is valid. The duration can range from a few months to several years. Similarly, the geographic scope may also vary across employment contracts. In some cases, the worker’s mobility may be restricted within the county where the firm operates, or, as in the case of many CEO contracts, the scope may be as broad as any global region where the firm has headquarters. In our model, a lower $z$ represents fewer restrictions in the employment contract. At the two extremes, $z = 0$ is equivalent to no CNCs in the contract, while $z = 1$ effectively prevents the worker from using any of her human capital investments in
a new firm and hence eliminates the possibility of turnover. Production occurs after the worker is hired in Period 1. At the end of Period 1, the worker and the firm simultaneously invest in human capital.\textsuperscript{6} We assume that both worker and firm make investments independently. The cost of making the investment is \( c(I) \) for both worker and firm but the marginal return from the two types of investment differs. In order to obtain closed form solutions, we assume a quadratic cost function, \( c(I) = \frac{1}{2} I^2 \).

In Period 2, the match quality between the current firm \( i \) and the worker \( k \) is revealed. Let us denote this match quality between worker \( k \) and firm \( i \) by \( r_{ki} \) which is uniformly distributed over the unit interval, i.e. \( r_{ki} \sim U [0, 1] \). The worker’s match quality with a new firm is still unknown. Period 2 output is determined by the investment levels of both workers and firms, as well as the firm-worker match. There are two raiding firms that make competitive wage-offers to the worker based on unknown match quality and human capital investment conditional on the CNC agreement between the worker and her current employer.

In Period 1 the worker’s output is represented by \( y \). The output of worker \( k \) in the incumbent firm \( i \) in Period 2 is:

\[
y_{2ki} = y + (v_w I_w + v_f I_f) r_{ki},
\]

where \( I_f > 0 \) and \( I_w > 0 \) are the investment levels chosen by the worker and the firm respectively and \( v_w \) and \( v_f \) are the marginal returns to worker and firm investments respectively.

If worker \( k \) moves to a competitor firm, \( j \), her expected output there is given by

\[
y_{2kj} = y + (1 - z) (v_w I_w + v_f I_f).
\]

The timing of the game is as follows. In Period 1, firm \( i \) offers a take-it-or-leave-it employment contract to worker \( k \) which specifies a hiring wage and a CNC of strength \( z \). If the worker accepts the offer, then both the firm and the worker simultaneously invest in human capital. In Period 2, the match quality between firm \( i \) and worker \( k \) is publicly revealed. Firm \( i \) makes a wage offer to the worker. Two identical raiding firms whose match quality with the worker is unknown simultaneously make wage offers to poach the worker. The worker decides which wage offer to accept and the game ends. The equilibrium concept we use is Subgame Perfect Nash Equilibrium.

\textsuperscript{6}Firm investments in employee human capital may take the form of training while worker investments may take the form of learning a new skill or a new computing language. See Garmaise (2011) for a distinction between such investments. Related to this, see Almedia and Carneiro (2009) on firm investments in human capital.
We make a few preliminary observations based on the set-up. First, we normalize the worker’s reservation wage in Period 1 to zero. This means that the Period 1 wage in the employment contract is trivially zero. Thus the Period 1 profits to firm $i$ is simply $\pi_{1i} = y$. However as we show in our analysis below, the CNC strength has Period 2 effects on the firm’s profits, and hence this is a non-trivial decision for the firm. Second, if the raiding firms want to poach the worker, then they must match firm $i$’s offer. Suppose, given CNC strength $z$ and match quality $r$, firm $i$ makes a wage offer $w_{2ki}$ and $R_i$ is the measure of workers retained by firm $i$, then $\pi_{2i}(z) = \int_{R_i} (y_{2ki} - w_{2ki}) dr$. The total profits across both periods is then

$$\pi_i(z) = y + \int_{R} (y_{2ki} - w_{2ki}) dr.$$  \hspace{1cm} (1)

### 3 Investment, Turnover and CNC Agreements

In this section we describe the equilibrium investment levels chosen by the firm and the worker given the turnover restrictions imposed by the CNC agreed to at the time of hire. Since firms invest to maximize profits and workers invest to maximize their wages, neither worker nor firm investments are optimal for maximizing surplus. However, as the CNC agreement in the employment contract becomes more restrictive, the firm can appropriate more of its investment, and this improves firm investment levels. At the same time greater restrictions on worker mobility distort turnover outcomes and depress worker incentives to invest. This trade-off suggests that the strength of restriction in CNCs can play a role in improving overall surplus by balancing investment incentives for firms and workers. In what follows, we first describe the equilibrium investment choice by the firm and workers and the wage and turnover outcomes for a CNC of given strength, $z$. We then derive the optimal CNC strength, $z^*$, that would maximize surplus conditional on the equilibrium investment and turnover outcomes following the CNC. Finally we look at how the profit maximizing choice of CNC strength, $z^*$, deviates from this optimal. This allows us to address the role that regulation can play in improving surplus.

#### 3.1 Period 2 Investment and Turnover

Let us start by taking the CNC enforcement strength, $z$, as given and describe the incentives facing firms and workers to invest in Period 2. We assume that all investments are general except for the firm-specificity introduced by the CNC contract.
Since the raiding firms are identical, they will compete for a worker so that given \( z \), \( I_f \) and \( I_w \), the worker’s outside wage is her expected output in a raiding firm, i.e. \( w_2 = y + \frac{1}{2} (1 - z) (v_w I_w + v_f I_f) \). Output in the current firm given match quality \( r \) is \( y + (v_w I_w + v_f I_f) r \). So turnover occurs if and only if \( r < \frac{1}{2} (1 - z) \), i.e. \( R = \{ r_1 \} \leq r \leq \frac{1}{2} \}. \) If there is turnover, then firm profits are zero. If \( r \geq \frac{1}{2} (1 - z) \) and there is no turnover, then profits are \( (v_w I_w + v_f I_f) \left( r - \frac{1}{2} (1 - z) \right) \). So expected Period 2 profits in firm \( i \), wages and output are, respectively, the following.

\[
\pi_{2i} = (v_w I_w + v_f I_f) \left( r - \frac{1}{2} (1 - z) \right) dr - c(I_f), \quad (2)
\]

\[
w_2 = y + \frac{1}{2} (1 - z) (v_w I_w + v_f I_f),
\]

\[
Y_2 = y + (v_w I_w + v_f I_f) \left( \frac{1}{2} (1 - z) dr + \int \frac{1}{2} r dr \right).
\]

Equilibrium firm investment, \( I_f^* \), equates the marginal effect of investment on profit to the marginal cost of investment. So \( I_f^* \) solves,

\[
c' (I_f^*) = \frac{v_f}{8} (1 + z)^2. \quad (3)
\]

The worker’s equilibrium investment level, \( I_w^* \), equates the marginal increase in wages from worker investment to the marginal cost of investment.

\[
c' (I_w^*) = \frac{1}{2} (1 - z) v_w. \quad (4)
\]

On the other hand, surplus maximizing investment levels for the firm and worker respectively are \( I_f^0 \) and \( I_w^0 \) which solve

\[
c' (I_f^0) = \frac{v_f}{8} \left( 4 + (1 - z)^2 \right), \quad (5)
\]

\[
c' (I_w^0) = \frac{v_w}{8} \left( 4 + (1 - z)^2 \right). \quad (6)
\]

Looking at equations (5) and (6), the surplus maximizing investment level for the worker is greater than the firm’s if and only if \( v_w > v_f \). However, comparing equilibrium investment levels in equations (3) and (4) we see that this is not the case. Since \( \frac{1}{2} (1 - z) < (2 - (1 - z))^2 \), even if \( v_w > v_f \), worker investment levels can be lower than the firm’s in equilibrium. The proposition below describes the equilibrium investments and turnover in Period 2 and compares it to the surplus-maximizing outcome.
Proposition 1 If firms use CNCs to restrict the transfer of human capital and CNC strength is z, then the following is true in the Period 2 equilibrium.

a) Turnover occurs if and only if \( r < \frac{(1-z)}{2} \).
b) \( I_w^* < I_w^o \) and the under-investment by workers increases as z increases.
c) \( I_f^* < I_f^o \) and the under-investment by firms decreases as z increases.

Proposition 1 highlights inefficiencies in both firm and worker investment levels in the presence of CNCs. Both the firm and the worker under-invest. The worker under-invests since retention wages rise slower than output with investment. Further as z increases, the retention wage becomes less sensitive to investment so that \( I_w^* \) decreases as z increases. Looking at optimal investment in (6), we see that \( I_w^o \) is also decreasing in z. However the rate at which \( I_w^o \) falls is lower than the fall in \( I_w^* \). This is because a higher z makes profits more sensitive to worker investment and this partially offsets the negative effect on retention wages in total surplus.

Part c) shows that the firm also invests less than the surplus maximizing level since the retention wage, \( \frac{(1-z)}{2} (v_w I_w + v_f I_f) \), increases with investment and hence firms cannot appropriate all gains in output from their investment. But a higher z alleviates this investment hold-up. From (3) and (5) we see that whereas \( I_f^* \) increases with z, the optimal investment falls as z increases thus bringing the two investment levels closer. This result conforms to previous research on the effects of turnover on firms' incentives to provide general training to its workers. For example, Moen and Rosen (2004) and Acemoglu (2007) show that firms under-invest in training when they cannot write long-term wage contracts with their employees. Similarly, Stevens (1996, 2001) argues that the possibility of workers being poached by competing firms leads to lower than efficient levels of worker training. Our result extends this well-established finding by allowing the employment contract between the worker and the firm to influence the extent of the under-investment.

3.2 Equilibrium and Optimal Employment Contracts

As the analysis in the previous section shows, a higher level of z implying greater restrictions on worker mobility enhances firm investment, but depresses worker investment. This affects both firm profits as well as surplus. Further z also affects worker turnover as the cut-off match quality for retaining a worker \( \frac{1}{2} (1 - z) \) is decreasing in z. Since the firm only earns positive profits on the workers it retains, the turnover effect of CNCs on profits is always favorable. By contrast, CNCs distort the allocation of firms with workers based on match quality. Comparing actual output
across the two firms in Period 2, we see that a worker is on average more productive in the other firm if and only if \( y + r (v_w I_w + v_f I_f) < y + \frac{1}{2} (v_w I_w + v_f I_f) \), i.e. \( r < \frac{1}{2} \). For \( z > 0 \), \( \frac{1}{2} (1 - z) > \frac{1}{2} \) and hence CNCs produce less than optimal turnover and this distortion worsens as \( z \) increases.

Since the trade-off in profits faced by the firm is different from the trade-offs in total surplus, the firm’s choice of \( z \) in the Period 1 employment contract deviates from the second-best efficient level that maximizes total surplus. Below, we first describe the equilibrium, \( z^* \). We then show the conditions under which regulatory policies on the enforcement of CNCs can impact total surplus.

First let us look at how the firm chooses \( z \) to maximize its own profits. The derivative of \( \pi_{2i}(z) \) with respect to \( z \), given that \( I_f^* \) is subsequently chosen to maximize profits is

\[
\frac{d}{dz} \pi_{2i}(z) = v_w \frac{dI_w^*}{dz} \int_{\frac{1 - z}{2}}^1 \left( r - \frac{(1 - z)}{2} \right) dr + (v_w I_w^* + v_f I_f^*) \left( \frac{1 + z}{4} \right).
\]

Since \( \frac{dI_w^*}{dz} < 0 \), the firm loses profits through lower worker investment. However, as the second term in the above derivative shows, a higher \( z \) lowers worker turnover and hence increases the range of match quality where the firm can earn positive profits on the worker. Thus the firm chooses \( z^* \) to balance out the worker investment effect and the turnover effect. The following proposition describes the equilibrium employment contract chosen by the firm in Period 1. We define \( \lambda = \frac{v_w}{v_f} \) which represents the marginal rate of technical substitution between worker and firm investment.

**Proposition 2** The firm always chooses a CNC with strength, \( z^* > 0 \) in its employment contract. There exists \( \lambda \) such that \( z^* = 1 \) if and only if \( \lambda \leq \lambda \). For every \( \lambda > \lambda \), \( z^* \in (0, 1) \) and it solves

\[
\frac{d\pi_{2i}(z)}{dz} \bigg|_{z^*} = 0.
\]

The above proposition describes the effect of \( \lambda \) on how restrictive the firm’s choice of CNC is. If worker investment is relatively less important, the firm chooses the strongest possible restrictions on worker mobility through its CNC.

We now turn to the optimal CNC strength that maximizes total surplus. The derivative of total surplus, \( S_2(z) = Y_2(z) - c (I_f^* - c (I_w^*)) \) with respect to \( z \) is:

\[\text{10}\]
\[
\frac{dS_2(z)}{dz} = \frac{dI^*_w}{dz} \left\{ v_w \left( \int_0^{\frac{1-z}{2}} \frac{1}{2} (1-z) \, dr + \int_{\frac{1-z}{2}}^{1} r \, dr \right) - c' \left( I^*_w \right) \right\} - \frac{1}{4} \left\{ (v_w I^*_w + v f I^*_f) (1-z) \right\} \\
+ \frac{dI^*_f}{dz} \left\{ v_f \left( \int_0^{\frac{1-z}{2}} \frac{1}{2} (1-z) \, dr + \int_{\frac{1-z}{2}}^{1} r \, dr \right) - c' \left( I^*_f \right) \right\}.
\]

As the above expression shows, there are three effects of a higher \( z \) on total surplus. First as with profits, lower worker investment decreases output and hence lowers surplus. Second, turnover efficiency is worsened as workers are misallocated across firms with respect to match quality. Finally, surplus is improved as firms invest more in the presence of stronger CNCs.

**Proposition 3** Let \( z^o \) represent the optimal CNC strength. There exist \( \tilde{\lambda}_1^o \) and \( \tilde{\lambda}_2^o \) where \( \tilde{\lambda}_1^o < \tilde{\lambda}_2^o < \tilde{\lambda} \) such that the following is true in the optimal employment contract in Period 1:

1. If \( \lambda \leq \tilde{\lambda}_1^o \), then \( z^o = z^* = 1 \).
2. If \( \tilde{\lambda}_1^o < \lambda < \tilde{\lambda}_2^o \), then \( 0 < z^o < 1 \), where \( z^o \) solves \( \frac{d}{dz} S_2(z^o) = 0 \) and \( z^o < z^* \).
3. If \( \lambda \geq \tilde{\lambda}_2^o \), then \( z^o = 0 \), i.e. CNCs are not optimal.

Comparing propositions 2 and 3 we see that, when \( \lambda \geq \tilde{\lambda}_2^o \) the firm chooses a CNC contract even though employment restrictions are never optimal. Further, even if the existence of CNCs may be optimal as in part b) of Proposition 3, the firm places inefficiently high employment restrictions. This provides a justification for regulatory restrictions on the enforcement of CNCs observed in various states across the US. At the same time, the results also demonstrate the importance of a nuanced approach to regulation rather than an outright prohibition on enforcing such agreements. Since labor mobility restrictions can serve an efficient role by enhancing firm investments, states should allow firms to negotiate CNCs with their employees. Further, as part (a) of the above Proposition shows, if firm investment is productive enough relative to worker investment, regulatory restrictions may be altogether unnecessary.

The result from Proposition 3 also addresses the widely accepted argument that strict regulations on CNC that make them unenforceable under California law had a role to play in the success of the high-technology industry in Silicon Valley. There is a general belief that much of the innovation in the high-technology industry is employee-driven. To the extent that worker investment is more productive implying a high \( \lambda \) in our model, our result in part c) would imply that CNCs...
should not be enforced. At the same time, it is important to recognize that when firms have an
incentive to restrict worker mobility even when it is not efficient, they may find other retention
strategies in the face of regulatory restrictions on CNCs. One such strategy is a no poaching agree-
ment with potential raiders in the labor market. In the following section, we analyze how strict
regulations around CNCs, such as those in California, may harm rather than help labor mobility
by incentivizing collusion among firms in the labor market.

4 No Poaching Agreements

Over the last few years the Antitrust Division of the US Department of Justice has investigated
claims of employer collusion in the recruiting practices of several high-technology firms. The in-
vestigation led to a series of civil lawsuits and most recently a class action lawsuit against large
Silicon Valley firms including Apple, Google, Intel, eBay and Adobe. Although explicitly illegal
these firms allegedly used no poaching agreements to reduce labor market competition and hence
drive down worker wages. In this section, we analyze the firm’s decision to collude with raiding
firms in the presence of regulatory restrictions on CNCs.

Although a no poaching agreement, as with any firm-level collusion, is illegal and can never be
enforced in court, in our view it is much more difficult to enforce laws against such practices relative
to CNCs. There are several reasons for this. First, many collusive agreements are likely to be tacit
and even if there is explicit communication between the parties, it is often difficult to establish
a paper trail to prove collusion.\(^7\) Second, lawsuits based on non-compete agreements are usually
initiated by firms against workers, whereas litigation on no poaching agreements typically involve
class action lawsuits brought on by employees. The latter is procedurally much more difficult since
common harm needs to be proved first. Thus with this presumption, we show that regulatory
restrictions on CNCs can have unintended consequences of incentivizing no poaching agreements
which may hurt labor mobility and surplus.

Suppose the firm can enter into a no poaching agreement with the raiding firms in Period 2.
This is effectively the same as the firm choosing \(z = 1\) in Period 1 since it completely eliminates
any possibility of turnover and drives worker wages to \(w_2 = y\). Since worker wages under collusion
do not depend on investment, workers do not invest at all, i.e. \(I^*_w (1) = 0\). In this section, we

\(^7\) The class action lawsuit against Apple, Google, Intel, eBay and Adobe sought billions of dollars in damages under
antitrust law. Yet the companies settled for an estimated $300 million and denied violating any laws. (“Tech Giants
consider two questions concerning the existence of such collusive agreements between firms. First, we look at how regulations on CNCs can incentivize no poaching agreements. Second, we look at how optimal CNC enforcement changes when no poaching agreements are taken into account.

We begin by taking the regulatory constraint as given in order to analyze how it affects the incidence of no poaching agreements. Let $z^R$ denote the maximum CNC strength enforceable under state law. Then firms are constrained to choose $z < z^R$. Under this notation, the two extreme enforcement regimes are (1) “no enforcement” or $z^R = 0$ which effectively bans non-compete agreements, and (2) “perfect enforcement” or $z^R = 1$, in which case there are no regulatory restrictions on CNCs. Note that under perfect CNC enforcement the firm always weakly prefers CNCs over no poaching agreements since CNCs offer a better tool to balance out the positive and negative effects of employment restrictions on profits.\footnote{The assumption here is that as long as CNCs are legal, firms can negotiate the CNC contract in a transparent way with their employees. On the other hand, the illegality of no poaching agreements inherently makes it difficult to negotiate worker turnover in a continuous manner without increasing the visibility of the collusive act.} Similarly if regulations are weak so that the regulatory constraint does not bind, i.e. if $z^* \leq z^R$, the firm never chooses to collude with its the competitors. Let us define $z^{R*} = \min \{z^*, z^R\}$ which is the equilibrium CNC strength chosen by the firm under the regulatory constraint $z^R$. We use the term weak enforcement to characterize low values of $z^R$ and permissive enforcement for high values of $z^R$.

In order to analyze the collusive agreement, we frame the problem as a repeated game between the three firms - the investing firm and the two raiding firms - as well as the worker. For simplicity, let us call the investing firm, $i$ and denote the raiding firms by $j$. Let $\delta \in (0, 1)$ be the discount rate for future profits. The stage game is then the same as the model laid out in Section 2. Thus, firm $i$’s cumulative profits across infinite stages is $\Pi_i = \sum_{t=0}^{\infty} \delta^t \pi_{it}(z)$, where $\pi_{it}(z)$ is the profit defined in (1). Similarly the cumulative profits to the raiding firms is $\Pi_j = \sum_{t=0}^{\infty} \delta^t \pi_{jt}(z)$, where $\pi_{jt}(z) = \int_{R_j} (y_{2kj} - w_{2kj}) dr$; $R_j$ represents the measure of workers poached by raiding firm $j$ at wage $w_{2kj}$.

Under the Nash equilibrium (NE) of the stage game described in the previous sections, and given the regulatory enforcement of CNCs $z^R$, firm $i$ invests $I_f(z^{R*})$ and workers invest $I_f(z^{R*})$. Firm $i$’s stage game profits are thus $\pi_i(z^{R*}) = y + \pi_{2i}(z^{R*})$, where $\pi_{2i}(z^{R*})$ is as described in (2). On the other hand, the two raiding firms do not make any profits due to labor market competition, i.e. in the stage game NE, $\pi_j = 0$. 

$\delta$
In the collusive agreement, the raiding firms agree to not poach any worker from firm \( i \). In return, firm \( i \) agrees to share a portion of its collusive gains with the raiding firms. Let \( \alpha \in (0, 1) \) represent the proportion of collusion gains kept by firm \( i \). Then in a no poaching agreement, the raiding firms receive an equal portion of the remaining \((1 - \alpha)\) share. We solve for a Subgame Perfect Nash Equilibrium in \( \alpha \) and the raiding firm’s decision to poach a worker. Under trigger strategies a deviation by any firm will result in the stage game NE being played for all subsequent periods. Thus for a collusive equilibrium to exist, \( \alpha \) must be such that no firm has an incentive to deviate. Below we develop the conditions that enable such an \( \alpha \) to exist.

First note that a necessary condition for collusion to occur is that the Period 2 profits to firm \( i \) from colluding must be greater than its profits from not colluding and instead using a non-compete agreement under the regulatory restriction, i.e. \( \pi_{2i} (z^R) \leq \pi_{2i}(1) \). From Proposition 2, we know that for \( \lambda \leq \hat{\lambda} \), \( z^* = 1 \), which means that \( \pi_{2i}(1) \geq \pi_{2i}(z^R) \) for all \( z^R \). Now let us see what happens to \( \pi_{2i}(z^R) - \pi_{2i}(1) \) as \( \lambda \) increases above \( \hat{\lambda} \). Here the firm’s profits are concave in \( z^R \). Note that after applying the envelope theorem, a higher \( z^R \) affects \( \pi_{2i}(z^R) \) only through an increase in the value of worker investment. Since the firm chooses its investment to maximize profits, any affect of \( \lambda \) on firm investment is internalized through profit maximization. Thus \( \frac{d\pi_{2i}(z^R)}{d\lambda} > 0 \) for all \( z^R < 1 \). As worker investment is zero under collusion, \( \pi_{2i}(1) \) is independent of \( \lambda \). As a result, for every \( z^R < 1 \), as \( \lambda \) increases above \( \hat{\lambda} \), the difference \( \pi_{2i}(z^R) - \pi_{2i}(1) \) increases. When \( \lambda \) is sufficiently high, the difference is positive for all \( z^R \) making a no poaching agreement unprofitable despite CNC regulations. We define this cut-off level of \( \lambda \) as \( \tilde{\lambda} \). Similarly, when \( \hat{\lambda} < \lambda < \tilde{\lambda} \), collusion is not profitable if enforcement is very permissive, i.e. if \( z^R \) is high. Let us define this cut-off as \( \hat{z}^R \). The lemma below states this result.

**Lemma 1** There exists \( \tilde{\lambda} > \hat{\lambda} \) such that the following is true:

a) If \( \lambda \leq \hat{\lambda} \), then \( \pi_{2i}(1) \geq \pi_{2i}(z^R) \) for all \( z^R > 0 \).

b) If \( \hat{\lambda} < \lambda \leq \tilde{\lambda} \), there exists \( z^R < z^* \) such that \( \pi_{2i}(1) \geq \pi_{2i}(z^R) \) if and only if \( z^R \leq \hat{z}^R \), where \( \hat{z}^R \) solves \( \pi_{2i}(z^R) = \pi_{2i}(1) \).

c) If \( \lambda > \tilde{\lambda} \), \( \pi_{2i}(1) < \pi_{2i}(z^R) \) for all \( z^R \geq 0 \) and hence the firm never chooses a no poaching agreement.

Lemma 1 shows that if worker investment is very important to output then regulation does not affect the incentives to collude. This is because very high restrictions on employee mobility depress worker incentives for investment which has a significant negative impact on output and
profits. When this is the case, the firm always finds it more profitable to allow some degree of worker mobility. Moreover, from Lemma 1, we see that firm $i$ makes higher profits with collusion if $\lambda \leq \lambda$ or if $\lambda < \lambda \leq \lambda^C$ and $z^R \leq z^R$. This is however not sufficient to guarantee the existence of a collusive equilibrium as firms may have an incentive to deviate from the agreement. Since collusion never exists for $\lambda > \lambda^C$, below we describe the existence of a collusive equilibrium when $\lambda \leq \lambda^C$.

Given that firm $i$ keeps a share $\alpha \in (0, 1)$ of the gains from the no poaching agreement for itself and splits the remaining $(1 - \alpha)$ share equally among the two raiding firms, the payoff to firm $i$ in the collusive equilibrium is $\Pi_i^c = \frac{1}{1 - \delta} \alpha \pi_i (1)$. Similarly the collusion payoff to each raiding firm is $\Pi_j^c = \frac{1}{\pi(1 - \delta)} (1 - \alpha) \pi_i (1)$.

Let us first consider a deviation by firm $i$. If firm $i$ kept the entire share of profits to itself then its profits for that period are $\pi_i (1)$. However under the trigger strategy that dissolves the collusion if a deviation occurs, the firm earns the continuation payoff from the NE in the stage games. Thus its profits from a deviation are $\pi_i (1) + \frac{\delta}{1 - \delta} \pi_i (z^R)$. So a deviation is unprofitable if and only if $\Pi_i^c > \pi_i (1) + \frac{\delta}{1 - \delta} \pi_i (z^R)$, i.e.

$$\alpha > \frac{(1 - \delta) \pi_i (1) + \delta \pi_i (z^R)}{\pi_i (1)}.$$  (7)

Now let us consider a deviation by one of the raiding firms. Given that firm $i$ follows its collusive strategy and invests $I_f^c (1)$, the worker’s output in a raiding firm is $y + (1 - z^R) v_f I_f^c (1)$. So if a raiding firm deviates and poaches the worker at wage $y$, given that the other raiding firm and firm $i$ follow the collusive strategy, it earns a one-time profit of $(1 - z^R) v_f I_f^c (1)$. Hence, a deviation is unprofitable if and only if $\Pi_j^c > (1 - z^R) v_f I_f^c (1)$, or

$$\alpha < 1 - \frac{2 (1 - \delta) (1 - z^R) v_f I_f^c (1)}{\pi_i (1)}.$$  (8)

Thus in order for collusion to be possible in equilibrium, both (7) and (8) must simultaneously hold, i.e. $\frac{(1 - \delta) \pi_i (1) + \delta \pi_i (z^R)}{\pi_i (1)} < 1 - \frac{2 (1 - \delta) (1 - z^R) v_f I_f^c (1)}{\pi_i (1)}$, or

$$\left[ \pi_i (1) - \pi_i (z^R) \right] - 2 \frac{(1 - \delta)}{\delta} (1 - z^R) v_f I_f^c (1) > 0.$$

The expression on the left hand side then denotes the range of values that $\alpha$ can take in a successful collusive agreement. Let us represent this function by $\Omega (z^R)$ which then represents how easy employer collusion is, i.e.

$$\Omega (z^R) = \left[ \pi_i (1) - \pi_i (z^R) \right] - 2 \frac{(1 - \delta)}{\delta} (1 - z^R) v_f I_f^c (1).$$  (9)
Looking at $\Omega(z^R)$, the first term in square brackets, represents the total profit gained through collusion, while the second term denotes the deviation profits to the raiding firm. It is clear that the likelihood of collusion depends on CNC regulations, $z^R$, the discount rate, $\delta$, and the relative worker investment productivity, $\lambda$. Moreover the marginal effect of $z^R$, $\frac{d}{dz^R} \Omega(z^R)$, also depends on $\lambda$.

Let us begin by looking at how $z^R$ affects $\Omega(z^R)$. There are two effects of weakening non-compete enforcement. First a decrease in $z^R$ lowers the baseline profit of Firm $i$ without collusion, (i.e. $d_i(0; z^R) > 0$). This increases the profitability of collusion. However, there is also a counteracting effect as deviations by the raiding firm become more profitable with weaker CNC enforcement. Since baseline profits are sensitive to worker investment, while the raiding firm’s deviation profits are not, the strength of the two effects depends on $\lambda$. In general we find that when $\lambda$ is high, i.e. worker investment is relatively more important, weak enforcement of CNCs (i.e. low $z^R$) makes collusion more likely to occur. In contrast, when $\lambda$ is low so that firm investment is relatively important, both very permissive CNC enforcement (or high $z^R$) and very weak enforcement may favor collusion.

Proposition 4 below explains how $\lambda$ and $\delta$ determine the impact of $z^R$ on collusion. In order to restrict the number of cases to consider, we only look at the cases where $\Omega(0) > 0$, i.e. where collusion exists under no enforcement. For every $\lambda \leq \bar{\lambda}^c$, there is a high enough $\delta$ where this will be true. So we define $\delta^L(\lambda)$ as the cut-off where $\Omega(0; \delta^L(\lambda)) = 0$; then we assume $\delta > \delta^L(\lambda)$. As mentioned before collusion is not possible for any $\delta$ when $\lambda > \bar{\lambda}^c$. We leave out that case in the proposition below.

**Proposition 4** There exist cut-offs $\bar{\lambda}_1^L < \bar{\lambda}_2 < \bar{\lambda}$ such that the following is true about the existence of a collusive equilibrium.

a) If $\lambda \leq \bar{\lambda}_1$ and $\delta > \delta^L(\lambda)$, collusion always exists for every $z^R \in [0, 1)$.

b) If $\bar{\lambda}_1 < \lambda < \bar{\lambda}_2$, and $\delta^L(\lambda) < \delta \leq \delta^H(\lambda)$ then collusion exists if and only if either $z^R < \bar{z}^R_1$ or $z^R > \bar{z}^R_2$, where $\delta^H(\lambda) \in (\delta^L(\lambda), 1)$, $\bar{z}^R_1 \in (0, 1)$ and $\bar{z}^R_1 \in [\bar{z}^R_2, 1)$. If $\delta > \delta^H(\lambda)$ then collusion always exists.

c) If $\bar{\lambda}_2 \leq \lambda < \bar{\lambda}$, and $\delta^L(\lambda) < \delta \leq \delta^H(\lambda)$ then collusion exists if and only if $z^R < \bar{z}^R_1$. If $\delta > \delta^H(\lambda)$ then collusion always exists.

d) If $\bar{\lambda} \leq \lambda \leq \bar{\lambda}^c$, and $\delta > \delta^L(\lambda)$ then collusion exists if and only if $z^R < \bar{z}^R_1$, where $\bar{z}^R_1 < \bar{z}^R$. The cut-offs for $\lambda$, $\delta$ and $z^R$ are defined in the appendix.

First from Proposition 4 it is straightforward to see that a low $\lambda$ and high $\delta$ facilitate collusion.
This is because a low $\lambda$ makes firm investment very productive relative to worker investment and hence provides large gains from collusion. At the same time a high $\delta$ makes deviations less attractive as the value of future profits is high relative to current gains. This is shown in Figure 1 a). Similarly as $\lambda$ increases, i.e. worker investment becomes more productive relative to firm investment, the range of parametrizations that can support collusion becomes smaller. Finally, as parts b), c) and d) demonstrate, when $\lambda$ is high but $\delta$ is relatively low, weak CNC enforcement facilitates collusion by making the gains from collusion very high.

Parts b) and c) also highlight the role of $\lambda$ in a differentiated impact of weak and permissive CNC enforcement on the likelihood of collusion. Specifically it states that when $\lambda$ is relatively low, collusion can exist with both very low and very high $z_R$, whereas at higher levels of $\lambda$, permissive CNC enforcement can prevent collusion. To see why, let us examine the effect of CNC enforcement starting from a point of perfect enforcement ($z_R = 1$). From (9), $\Omega (1) = 0$ implying that firm $i$ is indifferent between colluding and not colluding. Now consider the effect a small decrease in CNC enforcement.

First let us look at the lower range of $\lambda$ presented in part b). Since $z_R$ is high, worker investment is insignificant even in the absence of collusion, while firm $i$’s investment is substantial. A low $\lambda$ implies that firm investment is highly productive relative to worker investment. This, combined with relatively large firm investments, makes firm $i$’s baseline profits very sensitive to $z_R$. As a result, the increase in collusive profits from a drop in $z_R$ is larger than the increase in deviation gains for raiders, and the overall effect makes collusion possible with very permissive CNC enforcement. Figure 1 b) illustrates the trade-off between collusion gains and deviation incentives at weak and permissive enforcement levels when $\lambda$ is in this range.

Conversely when $\lambda$ is in the range presented in part c), worker investment is significantly more productive than firm investment. Here, very permissive CNC enforcement in the presence of a high $\lambda$, leads to smaller collusive gains and larger deviation profits. As a result very permissive CNC enforcement makes collusion difficult to sustain. Figure 1 c) graphically represents this outcome.

Proposition 4 also provides a number of interesting policy implications. First the result that weak enforcement of CNCs increases the likelihood of collusion speaks to the alleged no poaching agreements among Silicon Valley firms. The fact that California law makes non-competes unenforceable may have made collusion more attractive to employers. At the same time, part b) of the proposition suggests that very permissive enforcement of CNCs may also sometimes favor collusion if worker investment is relatively less productive. The large number of start-ups and spin-offs from
employees in Silicon Valley may be indicative of a high \( \lambda \). To the extent that this is true, our results suggest that collusion in Silicon Valley is less likely to have occurred if the California courts had enforced CNCs. At the same time, we should apply caution in designing CNC regulations in other states where industry growth depends primarily on firm-level worker training. In those cases, both very permissive and very weak enforcement can facilitate employer collusion.

Finally we derive optimal CNC regulation when we account for the possibility of employer collusion.

**Proposition 5** When the firm can use a no poaching agreement to restrict turnover the following is true about optimal CNC regulation (denoted by \( z^o \)). If \( \lambda < \tilde{\lambda}_1 \) or \( \delta > \delta^H (\lambda) \) then CNC regulation has no affect on surplus. In every other case, \( z^o = \tilde{z}_1 \geq z^o \).

Proposition 5 suggests that the optimal CNC contract entails weakly higher restrictions on labor mobility when firms can use no poaching agreements. Thus based on the proposition, we advocate a cautionary approach to regulating non-compete agreements. As the alleged collusion between large technology firms in California illustrates, we find that focusing regulation on employment contracts may incentivize other types of market conduct that have much worse impacts on social surplus than non-compete agreements.

5 Conclusion

Much of the previous research on CNCs has focused on their negative impact on labor mobility and the free flow of knowledge across firms. Under this reasoning, there has been a general consensus to restrict their enforcement under state laws. However, the literature on firm training has recognized the problem of investment hold-up created by worker turnover. Yet the role of CNCs in alleviating this hold-up of firm investment has not received much attention among researchers. In this paper, we provide an integrated theoretical framework for analyzing the trade-offs presented by CNCs as they lower worker investment but improve firm investment.

By posing this problem in a wider framework than the current literature does, we first solve the equilibrium employment contract that firms sign with their workers. Unlike earlier papers that viewed the firms’ CNC decision as binary choice of whether or not use such contracts, we derive the strength of labor mobility restrictions chosen in the CNC in equilibrium. Similarly, given that regulatory restrictions surrounding CNC enforcement also vary considerably across jurisdictions, we
also derive the socially optimal level of labor mobility restrictions. After comparing the equilibrium and optimal outcomes, we find that in some cases the firm restricts labor mobility more than is optimal suggesting the need for regulation to serve as a corrective mechanism in those situations.

Finally, given recent events surrounding the use of no poaching agreements by important Silicon Valley firms such as Google and Apple, we further extend our inquiry to account for employer collusion in designing optimal CNC regulations. We find surprising lessons from a policy perspective given this possibility. We show that zero enforcement of CNCs, such as in California, may worsen labor mobility by incentivizing employer collusion through no poaching agreements. Thus we argue for a cautious approach towards the enforcement of CNCs by taking all significant market ramifications of such regulations into account.

6 Appendix

Proof of Proposition 1. a) Turnover occurs if and only if the expected output for a worker in a new firm is greater than her output in the current firm, i.e. if and only if \( y + \frac{(1-z)}{2} (v_w I_w + v_f I_f) \geq \frac{(1-z)}{2} \) or \( r (v_w I_w + v_f I_f) \), or \( r \leq \frac{(1-z)}{2} \).

Without CNCs, the actual output of the worker in a new firm is \( y + \frac{1}{2} (v_w I_w + v_f I_f) \) and hence turnover is efficient if and only if \( r \leq \frac{1}{2} \). Since \( \frac{(1-z)}{2} \leq \frac{1}{2} \), turnover is suppressed even though it is efficient for \( \frac{(1-z)}{2} < r \leq \frac{1}{2} \).

b) The worker’s equilibrium investment level is \( c'(I_w^*) = \frac{1}{2} (1-z) v_w \). The surplus is \( S_2 = y + \frac{1}{8} (v_w I_w + v_f I_f) (4 + (1-z)^2) - c(I_f) - c(I_w) \). So the surplus maximizing investment level for workers is \( c'(I_w^*) = \frac{v_w}{8} (4 + (1-z)^2) \). \( c'(I_w^*) - c'(I_w^*) = \frac{v_w}{8} (1+z)^2 > 0 \). Hence \( I_w^* > I_w^* \).

\[ \frac{d}{dz} [c'(I_w^*) - c'(I_w^*)] = \frac{v_w}{4} (1+z) > 0 \]. Hence \( I_w^* - I_w^* \) is increasing in \( z \).

c) The expected Period 2 profit of the firm are \( \pi_{2i} = \frac{1}{8} (v_w I_w + v_f I_f) (1+z)^2 - c(I_f) \). Hence the equilibrium level of firm investment is \( c'(I_f^*) = \frac{v_f}{4} (1+z)^2 \) and the surplus maximizing level firm investment is \( c'(I_f^*) = \frac{v_f}{8} (4 + (1-z)^2) \). \( c'(I_f^*) - c'(I_f^*) = \frac{v_f}{2} (1-z) > 0 \), so \( I_f^* > I_f^* \).

\[ \frac{d}{dz} [c'(I_f^*) - c'(I_f^*)] = -\frac{v_f}{2} < 0 \], so \( I_f^* - I_f^* \) is decreasing in \( z \).

Proof of Proposition 2. The expected Period 2 profits of the firm given equilibrium investment and its higher order derivatives with respect to \( z \) are as follows.

\[
\pi_{2i}^* = \frac{(1+z)^2}{128} v_f^2 \left[ 8 (1-z) x^2 + (1+z)^2 \right],
\]

\[
\frac{d\pi_{2i}^*}{dz} = \frac{(1+z) v_f^2}{32} \left[ 2 (1-3z) x^2 + (1+z)^2 \right],
\]

19
\[ \frac{d^2 \pi_{2i}^*}{dz^2} = \frac{v_f^2}{32} \left\{ -4 (1 + 3z) \lambda^2 + 3 (1 + z)^2 \right\}, \]

\[ \frac{d^3 \pi_{2i}^*}{dz^3} = \frac{v_f^2}{32} \left\{ -12 \lambda^2 + 6 (1 + z) \right\}. \]

First note that for \( z \leq \frac{1}{3} \), profits are always increasing. So \( z^* > \frac{1}{3} \). Let us look at \( z > \frac{1}{3} \).

At \( z = \frac{1}{3} \), \( \frac{d^2 \pi_{2i}^*}{dz^2} = \frac{v_f^2}{32} \left\{ -12 \lambda^2 + 8 \right\} > 0 \) if and only if \( \lambda^2 < \frac{2}{3} \). At \( z = 1 \), \( \frac{d^3 \pi_{2i}^*}{dz^3} = \frac{v_f^2}{32} \left\{ -12 \lambda^2 + 12 \right\} > 0 \) if and only if \( \lambda^2 < 1 \).

1) \( \lambda^2 \leq \frac{2}{3} \), then \( \frac{d^2 \pi_{2i}^*}{dz^2} \geq 0 \) so that \( \frac{d^2 \pi_{2i}^*}{dz^2} \) is increasing in \( z \). At \( z = \frac{1}{3} \), \( \frac{d^2 \pi_{2i}^*}{dz^2} = \frac{v_f^2}{32} \left\{ -8 \lambda^2 + 16 \frac{3}{4} \right\} \geq 0 \) if and only if \( \lambda^2 \leq \frac{2}{3} \) which is true. This means that \( \frac{d^2 \pi_{2i}^*}{dz^2} \geq 0 \) for all \( z \) and hence \( \frac{d^2 \pi_{2i}^*}{dz^2} \geq 0 \) for all \( z \). So \( z^* = 1 \).

2) \( \frac{2}{3} < \lambda^2 \leq 1 \), then \( \frac{d^2 \pi_{2i}^*}{dz^2} \) reaches a minimum at \( z = 2 \lambda^2 - 1 \). \( \frac{d^2 \pi_{2i}^*}{dz^2} \) at \( z = \frac{1}{3} \) \( < 0 \) for all \( \lambda^2 > \frac{2}{3} \). \( \frac{d^2 \pi_{2i}^*}{dz^2} \) at \( z = 1 \) \( < 0 \) if and only if \( \lambda^2 > \frac{3}{4} \).

2.1) \( \frac{2}{3} < \lambda^2 \leq \frac{3}{4} \), then \( \frac{d^2 \pi_{2i}^*}{dz^2} \) \( > 0 \) and \( \frac{d^2 \pi_{2i}^*}{dz^2} \) \( > 0 \) since \( \frac{d^2 \pi_{2i}^*}{dz^2} \) \( > 0 \) and \( \frac{d^2 \pi_{2i}^*}{dz^2} \) \( > 0 \) \( < 0 \) so that \( z^* = 1 \). \( \lambda^2 \leq \frac{3}{4} \) \( < 0 \). \( \frac{d^2 \pi_{2i}^*}{dz^2} \) \( < 0 \) if and only if \( \lambda^2 \leq \frac{3}{4} \).

2.2) \( \frac{3}{4} < \lambda^2 \leq 1 \), then \( \frac{d^2 \pi_{2i}^*}{dz^2} < 0 \) and \( \pi_{2i}^* \) is concave but \( \frac{d^2 \pi_{2i}^*}{dz^2} \) \( > 0 \). So again \( z^* = 1 \).

3) \( \lambda^2 > 1 \), then \( \pi \) is concave in \( z \) and \( \frac{d^2 \pi_{2i}^*}{dz^2} \) \( > 0 \) and \( \frac{d^2 \pi_{2i}^*}{dz^2} \) \( > 0 \) hence \( z^* \) solves \( \frac{d^2 \pi_{2i}^*}{dz^2} = 0 \).

Hence let us define, \( \tilde{\lambda} = 1 \). \( \blacksquare \)

**Proof of Proposition 3.** The expected Period 2 surplus at \( I_w^* \) and \( I_f^* \) and its higher order derivatives with respect to \( z \) are as follows.

\[ S_2(z) = y + \frac{v_f^2}{128} \left[ (8 \lambda^2 + 1) (3 + z^2 - 3 z - 3z) + 5z \right], \]

\[ \frac{d}{dz} S_2(z) = \frac{v_f^2}{128} \left[ - (8 \lambda^2 + 1) (3z^2 - 2z) - 24 \lambda^2 + 5 \right], \]

\[ \frac{d^2}{dz^2} S_2 = \frac{v_f^2}{64} (1 - 3z) [8 \lambda^2 + 1]. \]

Let us first look at the shape of \( \frac{d}{dz} S_2 \). This is concave in \( z \) with a max at \( z = \frac{1}{5} \). At \( z = 0 \), \( \frac{d}{dz} S_2 = \frac{v_f^2}{128} \left[ -24 \lambda^2 + 5 \right] > 0 \) if and only if \( \lambda^2 < \frac{5}{24} \). At \( z = 1 \), \( \frac{d}{dz} S_2 = \frac{v_f^2}{128} \left[ -32 \lambda^2 + 4 \right] > 0 \) if and only if \( \lambda^2 < \frac{1}{8} \). At \( z = \frac{1}{3} \), \( \frac{d}{dz} S_2 = \frac{v_f^2}{24} \left[ -4 \lambda^2 + 1 \right] > 0 \) if and only if \( \lambda^2 < \frac{1}{4} \).
1) $\lambda^2 < \frac{1}{8}$, then $\frac{d}{dz} S_2 > 0$ for all $z$. Hence $z^o = 1$. Define $\hat{\lambda}_1^e = \frac{1}{8}$.

2) $\frac{1}{8} < \lambda^2 < \frac{5}{24}$, then $\frac{d}{dz} S_2 < 0$ at $z = 1$. So $S_2$ is increases first at an increasing rate, then at a decreasing rate, reaches a maximum and then falls. So there is a $z^o$ where it reaches a maximum which solves $\left[ \frac{d S_2}{dz} \right]_{z^o} = 0$.

3) $\frac{5}{24} < \lambda^2 < \frac{1}{4}$, then $S_2$ has a global maxima and a minima, i.e. two stationary points. So we check the maximum value against the value of the surplus at $z = 0$. At $z = 0$, $S_2(0) = y + \frac{16}{128} \left[ 8 \lambda^2 + 1 \right]$. To find the maximum value, we solve the quadratic equation, $2z_{\text{max}} \left( 8 \lambda^2 + 1 \right) - 3z_{\text{max}}^2 \left( 8 \lambda^2 + 1 \right) - 24 \lambda^2 + 5 = 0$. $\frac{d}{dz} \left[ S_{\text{max}} - S_2(0) \right] = \frac{v_f}{128} \left[ 8 \left( 1 - z \right) \left( 3 + z^2 \right) \right] - \frac{v_f}{128} 24 < 0$. At $\lambda^2 = \frac{5}{24}$, $S_{\text{max}} > S_2(0)$. At $\lambda^2 = \frac{1}{4}$, $S_{\text{max}} < S_2(0)$. Hence there exists $\hat{\lambda}_2^o \in \left( \frac{5}{24}, \frac{1}{4} \right)$ such that $S_{\text{max}} - S_2(0) > 0$ if and only if $\lambda < \hat{\lambda}_2^o$. Then for $\lambda < \hat{\lambda}_2^o$, $z^o$ solves $\left[ \frac{d S_2}{dz} \right]_{z^o} = 0$ and for $\lambda > \hat{\lambda}_2^o$, $z^o = 0$.

4) $\lambda^2 > \frac{1}{4}$, then $\frac{d}{dz} S_2 < 0$ for all $z$, so $z^o = 0$. ■

**Proof of Lemma 1.** From Proposition 2, if $\lambda \leq 1$, then $z^* = 1$, so that the profits from collusion are always higher than baseline profits in the presence of any CNC regulation. Hence $z^* = 1$. If $\lambda > 1$, $\pi_{2i}(z)$ is concave in $z$. We check if $\pi_{2i}(1) \geq \pi_{2i}(0)$. This is true if and only if $\lambda^2 \leq \frac{15}{8}$. Hence for $1 < \lambda^2 \leq \frac{15}{8}$, there exists $z^R$ such that if $z^R < z^R$, $\pi_{2i}(1) \geq \pi_{2i}(0)$. However if $\lambda^2 > \frac{15}{8}$ or $z^R \geq z^R$, then $\pi_{2i}(z^R) > \pi_{2i}(1)$. $z^R$ solves $\pi_{2i}(z^R) = \pi_{2i}(1)$, i.e. $(1 + z^R)^2 \left[ 8 \left( 1 - z^R \right) \lambda^2 + (1 + z^R)^2 \right] - 16 = 0$. ■

**Proof of Proposition 4.** Looking at (9),

$$\Omega(z^R) = \left[ \pi_{2i}(1) - \pi_{2i}(z^R) \right] - \frac{2(1-\delta)}{\delta} (1 - z^R) v_f I_f^t(1)$$

$$= v_f \left\{ 1 - \frac{(1 + z^R)^2}{128} \left[ 8 \left( 1 - z^R \right) \lambda^2 + (1 + z^R)^2 \right] - \frac{(1-\delta)}{\delta} \right\}$$

$$\Omega(0) \geq 0 \text{ if and only if } \delta \geq \frac{128}{143 - 8\lambda^2}. \text{ So define } \delta^L(\lambda) = \frac{128}{143 - 8\lambda^2}.$$ 

First we restrict $\lambda^2 < \hat{\lambda} = 1$. We look at the slopes of $\Omega(z^R)$ at the two extremes, i.e. $\frac{d\Omega(z^R)}{dz^R} \bigg|_{z^R=0}$ and $\frac{d\Omega(z^R)}{dz^R} \bigg|_{z^R=0} < 0$ if and only if $\delta > \frac{32}{33 + 2\lambda^2}$ and $\frac{d\Omega(z^R)}{dz^R} \bigg|_{z^R=0} < 0$ if and only if $\delta > \frac{4}{5 - \lambda^2}$. Following the shape of $\pi_{2i}(z)$ derived in the Proof of Proposition 2, and comparing the various $\delta$ cut-offs we get the following cases to consider.

a) $\lambda^2 \leq \frac{2}{3}$. Then $\frac{d\Omega(z^R)}{dz^R} \leq 0$. Since $\Omega(0) \geq 0$, this means that $\Omega(z^R) \geq 0$ for all $z^R$ and hence collusion is always possible. So we define $\lambda_1^e = \frac{2}{3}$.

b) $\frac{2}{3} < \lambda^2 < \frac{3}{4}$. Then $\frac{d\Omega(z^R)}{dz^R}$ first decreases, reaches a minimum, then increases reaches a maximum and then decreases again.
b.1) If $\lambda^2 < \frac{11}{16}$, then $\frac{128}{143 - 8\lambda^2} < \frac{32}{33 + 2\lambda^2}$ and $\frac{4}{5 - \lambda^2} < \frac{128}{143 - 8\lambda^2}$. So for all $\delta > \frac{128}{143 - 8\lambda^2}$, $\left.\frac{d\Omega(z^R)}{dz^R}\right|_{z^R = 1} < 0$.

First consider $\frac{128}{143 - 8\lambda^2} < \delta < \frac{32}{33 + 2\lambda^2}$. Here $\frac{d\Omega(z^R)}{dz^R}$ is positive at $z^R = 0$, so we first check the sign of $\min \left[ \frac{d\Omega(z^R)}{dz^R} \right]$ which is decreasing in $\delta$. $\min \left[ \frac{d\Omega(z^R)}{dz^R} \right]$ is decreasing in $\lambda^2$. At $\lambda^2 = \frac{11}{16}$, it is negative and at $\lambda^2 = \frac{2}{3}$ it is positive. So define $\min \left[ \frac{d\Omega(z^R)}{dz^R} \right]_{\delta = \frac{128}{143 - 8\lambda^2}, \lambda'}$.

If $\lambda > \lambda'$, then $\min \left[ \frac{d\Omega(z^R)}{dz^R} \right]_{\delta = \frac{128}{143 - 8\lambda^2}} < 0$. This means that $\min \left[ \frac{d\Omega(z^R)}{dz^R} \right] < 0$. Now let us check the sign of $\left[ \min \Omega(z^R) \right]$ which is increasing in $\delta$. $\left[ \min \Omega(z^R) \right] = \frac{128}{143 - 8\lambda^2}$ is decreasing in $\lambda$. At $\lambda'$, $\left[ \min \Omega(z^R) \right] = \frac{128}{143 - 8\lambda^2} > 0$ and at $\lambda = \frac{11}{16}$ it is negative. So define $\lambda_1$ as $\left[ \min \Omega(z^R) \right] = \frac{128}{143 - 8\lambda^2}, \lambda_1 = 0$. Then for $\lambda' < \lambda < \lambda_1$, collusion is always possible. For $\lambda_1 < \lambda < \frac{11}{16}$, $\left[ \min \Omega(z^R) \right] = \frac{128}{143 - 8\lambda^2}, \lambda_1 = 0$. Then for $\lambda' < \lambda < \lambda_1$, collusion is always possible. For $\lambda_1 < \lambda < \frac{11}{16}$, $\left[ \min \Omega(z^R) \right] = \frac{128}{143 - 8\lambda^2}, \lambda_1 = 0$. Then for $\lambda' < \lambda < \lambda_1$, collusion is always possible. For $\lambda_1 < \lambda < \frac{11}{16}$, $\left[ \min \Omega(z^R) \right] = \frac{128}{143 - 8\lambda^2}, \lambda_1 = 0$. Then for $\lambda' < \lambda < \lambda_1$, collusion is always possible. For $\lambda_1 < \lambda < \frac{11}{16}$, $\left[ \min \Omega(z^R) \right] = \frac{128}{143 - 8\lambda^2}, \lambda_1 = 0$. Then for $\lambda' < \lambda < \lambda_1$, collusion is always possible. For $\lambda_1 < \lambda < \frac{11}{16}$, $\left[ \min \Omega(z^R) \right] = \frac{128}{143 - 8\lambda^2}, \lambda_1 = 0$. Then for $\lambda' < \lambda < \lambda_1$, collusion is always possible. For $\lambda_1 < \lambda < \frac{11}{16}$, $\left[ \min \Omega(z^R) \right] = \frac{128}{143 - 8\lambda^2}, \lambda_1 = 0$. Then for $\lambda' < \lambda < \lambda_1$, collusion is always possible. For $\lambda_1 < \lambda < \frac{11}{16}$, $\left[ \min \Omega(z^R) \right] = \frac{128}{143 - 8\lambda^2}, \lambda_1 = 0$. Then for $\lambda' < \lambda < \lambda_1$, collusion is always possible. For $\lambda_1 < \lambda < \frac{11}{16}$.

b.2) $\frac{11}{16} < \lambda^2 < \frac{17}{21}$, then $\frac{128}{143 - 8\lambda^2} > \frac{32}{33 + 2\lambda^2}$ and $\frac{4}{5 - \lambda^2} < \frac{128}{143 - 8\lambda^2}$.

For $\delta > \frac{128}{143 - 8\lambda^2}$, $\frac{d\Omega(z^R)}{dz^R} < 0$ at both extremes. So we check $\max \left[ \frac{d\Omega(z^R)}{dz^R} \right]$ which is decreasing in $\delta$.

$\max \left[ \frac{d\Omega(z^R)}{dz^R} \right]_{\delta = \frac{128}{143 - 8\lambda^2}} > 0$. At $\delta = 1$, $\max \left[ \frac{d\Omega(z^R)}{dz^R} \right] < 0$. So define $\delta'$ as $\max \left[ \frac{d\Omega(z^R)}{dz^R} \right]_{\delta'} = 0$.

If $\delta < \delta'$, then we have to check $\min \left[ \Omega(z^R) \right]$. $\min \left[ \Omega(z^R) \right]$ is increasing in $\delta$. At $\delta = \frac{128}{143 - 8\lambda^2}$, then $\min \left[ \Omega(z^R) \right] = 0$. If $\delta = \delta'$, then $\min \left[ \Omega(z^R) \right] > 0$. So define $\delta^H(\lambda)$ as $\min \left[ \Omega(z^R) \right]_{\delta^H(\lambda)} = 0$. Then if $\delta \leq \delta^H(\lambda)$, then collusion is possible if and only if $\lambda \notin (\tilde{z}_1^R, \tilde{z}_2^R)$. If $\delta^H(\lambda) < \delta < \delta'$, then collusion is always possible.

If $\delta > \delta'$, then $\frac{d\Omega(z^R)}{dz^R} < 0$, so that collusion is always possible.
b.3. $\frac{17}{24} < \lambda^2 < \frac{3}{4}$, then $\frac{128}{143-8\lambda^2} > \frac{32}{33+2\lambda^2}$ and $\frac{4}{5-\lambda^2} > \frac{128}{143-8\lambda^2}$.

If $\frac{128}{143-8\lambda^2} < \delta < \frac{4}{5-\lambda^2}$, then $\Omega(z^R)$ first decreases and then increases. This means that there is $z_1^R$ such that collusion is possible if and only if $z^R < z_1^R$.

If $\delta > \frac{4}{5-\lambda^2}$, consider $\max\left[ \frac{d\Omega(z^R)}{dz^R} \right]$. At $\max\left[ \frac{d\Omega(z^R)}{dz^R} \right] = 0$. Hence $\max\left[ \frac{d\Omega(z^R)}{dz^R} \right] < 0$.

This means that $\frac{d\Omega(z^R)}{dz^R} < 0$ for all $z^R$ and hence collusion is always possible.

Define $\delta^H(\lambda) = \frac{4}{5-\lambda^2}$.

c) $\frac{3}{4} \leq \lambda^2 < 1$. Then $\frac{d^2\Omega(z^R)}{dz^R} \geq 0$ and $\frac{128}{143-8\lambda^2} < \frac{4}{5-\lambda^2}$.

(i) If $\frac{128}{143-8\lambda^2} \leq \delta \leq \frac{4}{5-\lambda^2}$, then $\frac{d\Omega(z^R)}{dz^R} \bigg|_{z^R=1} \geq 0$. This means that there exists $z_1^R \in (0, 1)$ which solves $\Omega(z_1^R) = 0$ such that $\Omega(z^R) > 0$ if and only if $z^R < z_1^R$.

(ii) If $\delta > \frac{4}{5-\lambda^2}$, then $\frac{d\Omega(z^R)}{dz^R} \bigg|_{z^R=1} < 0$, then $\Omega(z^R) > 0$ for all $z^R$ and hence collusion always exists.

Define $\delta^H(\lambda) = \frac{4}{5-\lambda^2}$.

d) $1 \leq \lambda^2 < \frac{15}{8}$, then from Lemma 1, a necessary condition for collusion is $z^R < z_1^R$. In this range, $\pi_2i \left( z^R \right)$ is increasing and concave in $z^R$. So $\Omega(z^R)$ is convex in $z^R$. At $z_1^R$, $\Omega(z^R) < 0$. So for $\delta > \frac{128}{143-8\lambda^2}$, there exists $z_1^R$ such that $\Omega(z^R) > 0$ if and only if $z^R < z_1^R$.

Thus we define $\lambda_1^2 = \frac{2}{3}$, $\lambda_2^2 = \frac{17}{24}$ and from Proposition 2, $\tilde{\lambda} = 1$. ■

**Proof of Proposition 5.** From Proposition 4, for $\lambda \leq \lambda_1^2$, collusion always exists and hence regulation has no effect on surplus which is always $S(1)$. $\tilde{\lambda}_1^2 > \frac{1}{4}$, so $S(z^R)$ is decreasing in $z^R$.

Hence for $\lambda_1^2 < \lambda < \tilde{\lambda}^2$, $\tilde{S}^O = z_1^R > S^O = 0$. ■

**References**


