Restrictions on Executive Mobility and Reallocation: The Aggregate Effect of Non-Competition Contracts *

Liyan Shi†

January 15, 2018

Click Here for the Most Recent Version

Abstract

This paper assesses the aggregate effect of non-competition employment contracts, agreements that exclude employees from joining competing firms for a duration of time, in the managerial labor market. These contracts encourage firm investment but restrict manager mobility. To explore this tradeoff, I develop a dynamic contracting model in which firms use non-competition to enforce buyout payment when their managers are poached, ultimately extracting rent from outside firms. Such rent extraction encourages initial employing firms to undertake more investment, as they partially capture the external payoff, but distorts manager allocation. I show that the privately-optimal contract over-extracts rent by setting an excessively long non-competition duration. Therefore, restrictions on non-competition can improve efficiency. To quantitatively evaluate the theory, I assemble a new dataset on non-competition contracts for executives in U.S. public firms. Using the contract data, I find that executives under non-competition are associated with a lower separation rate and higher firm investment. I also provide new empirical evidence consistent with non-competition reducing wage-backloading in the model. The calibrated model suggests that the optimal restriction on non-competition duration is close to banning non-competition.

*I am indebted to Hugo Hopenhayn for his invaluable guidance and support. I am grateful to Andrew Atkeson, Pablo Fajgelbaum, François Geerolf, and Pierre-Olivier Weill for their direction and advice. I also thank Saki Bigio, Ariel Burstein, Ioannis Kospentaris, Dennis Kuo, Flavien Moreau, Ben Smith, Robert Ulbricht, Jonathan Vogel, and the seminar participants at UCLA for their insightful comments.

†University of California, Los Angeles. Email: liyanshi@g.ucla.edu. Website: www.liyanshi.com.

1 Introduction

Non-competition employment contracts, agreements that exclude employees from joining competing firms for a duration of time when they leave their employers, are prevalent in the managerial labor market. About 64% of executives employed in public firms in the U.S. have signed contracts that include non-competition clauses. The anti-competitive effects of such contracts are concerning, as restricted labor mobility precludes reallocation of managers to more productive employment.[1] Proponents, however, argue that non-competition protects firms from losing the benefits of their own investment, thereby encouraging a more efficient level of investment. The disagreement over the merits of non-competition contracts manifests itself in the disparate legal landscape across the U.S., with most states taking a permissive stance towards non-competition and one notable exception of California banning non-competition (Bishara, 2011).

This paper assesses the aggregate effect of non-competition contracts in the managerial labor market, considering the beneficial effects of encouraging firm investments and the harmful effects of restricting manager mobility. Despite the two opposing effects being well documented separately in empirical studies (e.g., Garmaise, 2009; Marx et al., 2009; Jeffers, 2017; Lavetti et al., 2017), their net effect is unclear. Studying the tradeoff between investment protection and manager reallocation can inform policymakers in their decision to permit or restrict the use of non-competition contracts.[2] Such a study calls for a model of employment contracts with non-competition clauses to guide quantitative evaluation.

I first make a theoretical contribution to the literature by developing such a model of employment contracts. Motivated by actual contractual practices, I model firms using non-competition to enforce buyout payment when managers are poached, ultimately achieving rent extraction from outside firms. Non-competition creates the following efficiency tradeoff. On one hand, it allows the initial employing firm to partially capture the external payoff and undertake more investment. On the other hand, it distorts manager allocation and creates barriers for outside firms to enter. I show that the privately-optimal contract over-extracts

[1] Reports by the White House (The White House (2016)) and the Department of the Treasury (The Department of the Treasury, 2016) identify non-competition contracts as a likely cause for the declining labor market fluidity, stagnant wage growth, and declining business dynamism observed in the U.S. (see e.g., Davis and Haltiwanger, 2014; Decker et al., 2016).

[2] Recent efforts in non-competition legal reforms have tried to move closer to the California law but been largely unsuccessful. These include perennial legislative proposals in Massachusetts in restricting the use of non-competition. In particular, a bill filed in 2017 in Massachusetts proposes to limit non-competition duration to a maximum of one year. The details can be found at https://malegislature.gov/Bills/190/SD1578. In addition, The White House under The Obama Administration proposes to institute a nation-wide ban. The details can be found at https://obamawhitehouse.archives.gov/sites/default/files/competition/noncompetes-calltoaction-final.pdf.

rent by setting an excessively long non-competition duration. Therefore, restrictions on non-competition duration can improve efficiency.

I also contribute to the empirical literature in two ways: (i) by assembling a new dataset on non-competition arrangements for executives in U.S. public firms, scrapped from contracts disclosed in company filings; and (ii) by providing new empirical evidence from observed uses of non-competition contracts. I find that executives under non-competition are associated with a lower separation rate from their firm and higher firm investment. The evidence also confirms the model's mechanism that non-competition decreases wage-backloading for retention. I then calibrate the model to these data moments. Quantitatively, the calibrated model implies an optimal restriction on non-competition duration close to banning non-competition.

I set the model in a simple production environment and focus on rich contracting features. I model firm-manager matches as production units, where they invest to improve their match productivity. The productive knowledge resides in the manager and is portable to future outside firms, which creates an *investment externality*. The initial match's investment is prone to holdup because they pay the cost while future outside firms can partially appropriate the payoff.

To capture the rich contracting features, I embed the bilateral dynamic contract along the lines of Postel-Vinay and Robin (2002) in the multi-contracting environment of Aghion and Bolton (1987). In contrast to the standard dynamic contract concerning only wage payment, I expand the contracting possibilities in two ways: non-competition restricting the manager's outside employment and buyout payment from the manager to the firm. These two additional possibilities affect outside firms that subsequently contract with the manager, resulting in a *contracting externality*. Further, since the manager cannot commit and will renege on the payment to the initial firm, non-competition is necessary to enforce buyout payment.

The allocative distortion of non-competition buyout results from the information constraint that outside firm productivity is private information. To see why, the contract between the initial firm and the manager first achieves *bilateral efficiency*, as the initial firm aligns the manager's incentive by costlessly backloading wage to retain the manager, given firm commitment and agent risk neutrality. The initial firm and the manager, as a bilateral coalition, then act like a monopolist towards outside firms. To achieve the monopoly price outside firms pay to poach the manager, non-competition enforces the extra buyout payment in addition to wage payment. As a result, the productivity threshold for outside firms to poach the manager is distorted upward.

In the baseline environment, the contract includes a non-competition clause with a buyout

option bunched to a single-price menu contingent on realized current match productivity. That is, the bilateral coalition acts like a non-discriminating monopolist and outside firms that poach the manager fully buyout the non-competition.

The efficiency tradeoff lies in the interaction between the contracting externality and the investment externality. Rent extraction improves investment efficiency. To be precise, more rent extraction leads to higher *marginal bilateral joint value* of investment and thus a more efficient incentive for investment. This interaction generates an *investment-reallocation tradeoff*: a longer duration of non-competition alleviates the holdup due to the investment externality, while aggravating the distortion in manager allocation due to the contracting externality.

The privately-optimal contract, despite being bilaterally efficient, is socially inefficient along the investment-reallocation tradeoff. This social inefficiency can be seen through the operation of the contracting externality: the bilateral coalition of the initial firm and the manager maximizes their bilateral joint value and disregards the value of outside firms. Compared to a planner who aims to maximize *social value*, the bilateral coalition sets excessively long non-competition duration and over-extracts rent. I consider a planner who has at its disposal the policy instrument to cap non-competition duration. The planner desires less rent extraction and caps non-competition duration below the privately-optimal level.

Having established the efficiency implications, I take the model's three predictions of how non-competition contracts affect firm-manager matches to the data. First, non-competition distorts the poaching threshold upward, resulting in less frequent manager separation from the firm. Second, the firm undertakes more investment in response to rent extraction. Third, the wage is less backloaded due to less wage bidding for retention against outside offers. To be precise, the manager starts with a higher wage but experiences lower wage growth over tenure.

I assemble a new dataset on non-competition contracts for executives employed in U.S. public firms. Specifically, I scrap contracts included in company filings in the `SEC Edgar` database. I then use natural language processing and machine learning tools to classify the contracts and extract information on non-competition clauses. The contract data is merged with a rich array of standard data on executives and firms. In the final data sample including 9,758 firm-executive employment relations, 64% of the executives are subject to non-competition.

Using the merged data, I find that observed uses of non-competition indeed have sizable effects on managerial reallocation and firm investment. First, executives with non-competition are 1.2 percentage points per annum less likely to separate from their firms, compared to those without. Second, when the fraction of executives subject to non-competition

increases from 0% to 100%, firms have an investment rate in physical and intangible capital that is 1.4 percentage points higher per annum. In addition, the availability of contract level data allows me to uncover new empirical evidence on how non-competition interacts with wage-backloading. Interestingly, I find that wage is less backloaded for executives with non-competition, confirming the dynamic contracting channel. Specifically, executives with non-competition start with a wage that is \$130k (or 15%) higher in 2010 prices and experience wage growth over the first ten years of tenure that is 1.6 percentage points lower per annum.

To carry out quantitative evaluation, I calibrate the model to match the aforementioned moments and other cross-sectional moments, guided by the model's close link between parameters and moments. In particular, the investment elasticity parameter is pinned down by targeting the investment response to the use of non-competition. The calibrated baseline model suggests that the optimal restriction on non-competition duration is only 30% of the privately-optimal level. The optimal restriction and a ban on non-competition result in welfare gains measured in steady state net output of 6% and 4%, respectively, relative to the laissez-faire outcome.

The welfare gains from restricting non-competition depend on the investment elasticity, that is, how responsive investment is to changes in payoffs. If investment is highly elastic, investment holdup is severe and the gains from alleviating holdup are large. My baseline calibration implies an investment elasticity at the higher end of the range in the literature. Hence, my welfare calculation is conservative. Indeed, fixing the investment elasticity at medium and low levels, the re-calibrated model suggests optimal restrictions on non-competition even closer to banning non-competition.

Finally, I provide two extensions to the baseline model. The first extension introduces business stealing by outside firms from the initial employer and knowledge depreciation during the non-competition period. The bilateral coalition now acts like discriminating monopolist towards outside firms: the contract features a continuum buyout menu, in contrast to the single-price buyout menu in the baseline model. Non-competition is enforced in equilibrium to price discriminate against less productive outside firms, creating a "damaged" version of managerial human capital, whereas in the baseline model non-competition is always fully bought out and never enforced.[3] This extension reconciles with selective non-competition enforcement observed in actual practices.

The second extension recasts the baseline model in a general equilibrium setting to account better for the extent of entry barriers non-competition can create. Specifically, I

[3]This contract is likened to the "damaged goods" phenomenon in industrial organization (Deneckere and McAfee, 1996) where a monopolist intentionally damages goods to achieve price discrimination.

endogenize the arrival rate of outside opportunities by introducing a random labor search market and costly free entry of new firms. The efficiency concerns are more intricate: in addition to the investment and contracting externalities, the labor market induces a *search externality*. Further, the holdup of investment is now two-sided: the entrants' investment to enter also has a positive external effect. The efficiency implications of non-competition in shifting the surplus division between incumbents and entrants hinges on how elastic matching is with respect to firm entry.

1.1 Related Literature

This paper builds on the literature of externalities in bilateral contracts in a multi-contracting environment, in particular exclusionary contracts in vertical contracting (e.g., Aghion and Bolton, 1987; Spier and Whinston, 1995; Segal and Whinston, 2000). The employment contracting setting is analogous to the vertical contracting setting: both involve an incumbent firm restricting its manager's or downstream firm's future trade with outside firms. The efficiency concerns also resemble: rent extraction can enhance investment efficiency but may undermine allocative efficiency. The mechanism leading to allocative inefficiency borrows from Aghion and Bolton (1987) in assuming that firms commit to the contract and act like a monopolist.[4] I make two modifications. First, I modify the stipulated damage payment to buyout payment of non-competition, capturing the features of employment contracts in practice. Second, I extend this framework into dynamic setting to suit quantitative analysis by embedding the bilateral dynamic contract of Postel-Vinay and Robin (2002) and Postel-Vinay and Turon (2010).

This paper contributes to understanding the efficiency implications of employment contractual arrangements for the incentives of firm investment, a topic on which two strands of literature interact. The first stand of literature emphasizes firm-provided investment in general human capital (e.g., Acemoglu, 1997; Acemoglu and Pischke, 1999; Moen and Rosen, 2004; Fella, 2005; Lentz and Roys, 2015).[5] The second strand of literature emphasizes investment externalities through knowledge diffusion along the lines of Lucas and Moll (2014), specifically in the form of employee movement between firms (e.g., Franco and Filson, 2006; Franco and Mitchell, 2008). In relation to both, I study additional contracting features that are widely used in practice yet much disputed in policymaking. Two recent papers, Rauch (2016) and Heggedal et al. (2017), also study the welfare implications of non-competition

[4]This is in contrast to, for example, Spier and Whinston (1995) and Segal and Whinston (2000) in which renegotiation achieves allocative efficiency.

[5]In particular, Acemoglu and Pischke (1999) note that labor market friction induced monopsony can restrict outside opportunity and alleviate bilateral investment holdup. In relation to that, I focus on a setting of sufficiently large friction such that bilateral holdup is absent.

contracts and focus on bilateral inefficiency. In contrast, the efficiency concern in this paper is of a multilateral nature: the contract achieves bilateral efficiency and consequently the bilateral coalition aims to extract rent from third parties.

This paper also contributes to empirical studies on non-competition contracts in encouraging firm investment and restricting labor mobility. In particular, Garmaise (2009), Marx et al. (2009), Starr (2016), and Jeffers (2017) explore exogenous variations in the legal enforceability of non-competition contracts across states or over time. My empirical focus of exploring observed non-competition contracts is similar to Lavetti et al. (2017) who study the physician labor market. Despite potential endogeneity issues, the contract data allows for linking contractual arrangements to wage dynamics. My finding is, however, in contrast to the one in Lavetti et al. (2017). Whereas they find physicians with non-competition experience much faster wage growth, I find that executives with non-competition experience slower wage growth, consistent with wage-backloading in dynamic contracting. The empirical analysis is closely related to Garmaise (2009), who also studies non-competition in the managerial labor market by looking at executives in U.S. public firms. While Garmaise (2009) finds strengthening of non-competition enforcement leading to slower wage growth and attributes it to diminished investment by managers in their human capital, my findings point to instead non-competition lessening wage-backloading.

Lastly, this paper is related to studies on competitive market forces in determining executive compensation (e.g., Frydman and Saks, 2010; Frydman, Forthcoming). For example, Frydman (Forthcoming) documents that the increasing importance of general managerial human capital has led to higher executive mobility and compensation over time. My empirical findings suggest that outside competition for managers affects not only the level of compensation but also the structure of compensation over tenure, confirming retention concerns in dynamic compensation design.

The remainder of the paper is organized as follows. Section 2 introduces the baseline contracting model. Section 3 studies the planner's problem of optimal restriction on non-competition. Section 4 provides two extensions to the baseline model. Section 5 describes the data on executives employed in U.S. public firms. Section 6 presents the empirical findings in the executive data regarding the model's predictions. In Section 7, I calibrate the model and carry out welfare analysis quantitatively. Section 8 concludes.

2 The Model

To motivate the model setting, I first present a brief discussion of the institutional background of non-competition contracts pertinent to modeling contract design. Non-competition

clauses, as literally stated, restrict employees from joining outside competing firms for some period of time after leaving their initial employing firms.[6] However, employment contracts with a non-competition clause also commonly include a buyout clause, which grants the employee an option to buyout the non-competition with a payment. For this reason, a buyout clause is sometimes called a "clawback" or "forfeiture-for-competition" clause.[7] The arrangement of buying out non-competition has obvious advantage over actually enforcing it – the former involves a transfer while the latter is mere "money burning".

The buyout practice is also reflected in law. For example, Texas law requires that a non-competition contract must have a buyout clause attached for certain occupations. The state of New York follows the "employee choice doctrine" – it is the employee's choice to either have the non-competition enforced or have compensation clawed back. Therefore, I focus on non-competition as a means to enforce buyout payment.

Apart from non-competition clauses, managerial compensation design serves retention purposes. A significant portion of compensation awarded is in the form of restricted equity, consisting of unvested stock and un-exercisable options. For the sample of executives included in the empirical part of this study, the fraction of compensation realized through stock vesting and options exercised is over 60% on average. Restricted equity can only be cashed out in a future date conditional on the manager staying with the firm, hence often referred as "golden handcuffs".[8]

These details together motivate a model of a dynamic contract with two additional contracting terms: a non-competition clause and a buyout option. I develop such a model, embedding the bilateral dynamic contract along the lines of Postel-Vinay and Robin (2002) in the multi-contracting environment of Aghion and Bolton (1987).

2.1 The Environment

Time is continuous and infinite, $t \in [0, \infty)$. I fix a probability space $(\Omega, \mathcal{F}, P)$ together with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions. The economy is populated by a measure-one, continuum of over-lapping generations of managers with exponential lifetime, employed by a corresponding continuum of firms.[9] Each manager dies with Poisson intensity δ, upon which event the firm also exits the economy, replaced by a new-born manager matched to

[6] A sample contract for executives employed in U.S. public firms is displayed in Figure 8 in the additional data appendix.

[7] The sample contract in Figure 9 includes a buyout option.

[8] The sample contract in Figure 10 states that the reason for the restricted equity award is to "encourage your continued employment with" the firm.

[9] The firms in the model are single-manager firms. When mapping the model to firm-level data, a firm is considered as a collection of firm-manager matches.

a new-born firm. The agents are risk neutral and discount future at rate ρ. This risk-neutral assumption implies that there is no risk-sharing in bilateral contracting and utility is perfectly transferrable. The effective discount rate is $r = \rho + \delta$.

A firm-manager match at time t is characterized by its idiosyncratic productivity z_t, producing a flow of output $y_t = \exp(z_t)$. A manager is the only input in production.[10] Upon birth, the initial productivity of the new firm-manager match, z_0, is drawn according to the cumulative distribution function $H(\cdot)$. The idiosyncratic productivity $\boldsymbol{z} = \{z_t\}_{t\geq 0}$ evolves stochastically according to a Brownian motion:

$$dz_t = \mu_t\, dt + \sigma\, dB_t,$$

where $\boldsymbol{\mu} = \{\mu_t\}_{t\geq 0}$ is the investment undertaken continuously by the employing firm and $\boldsymbol{B} = \{B_t\}_{t\geq 0}$ is a standard Brownian motion. The investment entails a cost $c(\mu_t)\exp(z_t)$, where the cost function $c(\cdot)$ is strictly increasing, twice continuously differentiable, and convex. Investment is embodied in the manager's human capital that is general and portable. If the manager were to separate from the firm and be employed in outside firms, he takes the accumulated human capital to outside firms, creating an *investment externality*.

The labor market is frictional. Managers are matched to employment opportunities with outside firms at Poisson intensity λ. The new match has productivity $z_t' = z_t + \theta_t$, where the uncertain relative productivity θ_t is drawn according to the cumulative distribution function $F(\cdot)$, defined over $[\theta_m, \infty)$. I assume that $F(\cdot)$ is continuous and satisfies $1 - F(0) > 0$. The employment opportunity with the outside firm is non-durable: it disappears if not taken. The initial employing firm exits when the manager moves to outside firms. The measure of firms in the economy with productivity $z_t \leq z$ at time t is denoted by $G(z,t)$.

2.2 Information Structure and Contracting Possibilities

Information is asymmetric: the firms do not observe each other's productivity.[11] Given the information constraint, the maximum payoff the initial firm and the manager can jointly achieve is by charging the outside firm a monopoly price to poach away the manager. The dynamic contract with wage bidding and non-competition buyout below implements the monopoly pricing.

Firms and managers enter into ex ante bilateral long-term contracts specifying the process through which employment and corresponding transfer are determined ex post. Crucially for

[10] I abstract away from other production inputs but they can be easily incorporated.

[11] It is essential that outside firm productivity is private information. The results also holds when initial firm productivity is also private information.

the externality to be discussed, firms contract with the same manager sequentially, with the initial firm being the Stackelberg leader. When outside firms contract with the manager, they take as given the existing contract the manager has entered with the initial firm. Building on the dynamic contract in Postel-Vinay and Robin (2002), I expand the contracting possibilities in two ways relative to their work.

The first extension introduces the possibility of buyout payment from the manager to the firm at match separation. I assume that firms are deep pocketed while managers are hand-to-mouth. In the absence of outside offers, the transfer from the firm to the manager, i.e. the wage payment, has to be positive. When the manager takes on an outside employment, it opens the possibility for payment from the manager to the initial firm, one that is financed by outside firms.

The other extension is the possibility for non-competition. I assume that firms can commit to the contract but managers cannot commit to the contract.[12] In particular, the manager can renege on the payment at the termination of employment relation, leading to an enforcement problem. To circumvent this enforcement problem, the initial firm instead uses a non-competition clause, excluding the manager from working for outside firms for a duration of π units of time. The manager together with the outside firm can perform buyout payments to avoid non-competition.

The two additional contracting terms adversely affect outside firms that subsequently contract with the manager, resulting in a *contracting externality*. The benevolent external enforcer, or the court of law, if asked by the initial firm, enforces the non-competition as the contract dictates. By permitting and enforcing non-competition, the external enforcer makes non-competition a threat for outside firms to pay. The buyout payment takes place with mutual consent and without external enforcement.[13]

At time t, after initial match productivity z_t is realized and outside firm of match productivity $z'_t = z_t + \theta_t$ arrives, the competition between firms to poach and retain the manager occurs in a two-stage game. In the first stage, the initial and outside firms bid for the manager in an ascending (English) auction. The initial firm commits to bidding up to the entire match value. If the manager is poached by the outside firm, a second stage ensues. In the buyout stage, the initial firm asks the buyout price τ_t and the outside firm then decides whether to buyout the non-competition. The buyout price $\tau_t(\tilde{\pi}) = \tau(z_t, \tilde{\pi}) \geq 0$ is adapted to the filtration generated by the productivity process.

I denote the resulting wage process by $\boldsymbol{w} = \{w_t\}_t$, where $w_t \geq 0$ to be adapted to the

[12] I assume that there is no renegotiation of the contract ex post.

[13] From a property rights perspective, the court of law grants the initial firm the property right to the manager's future employment during the non-competition period which they can sell back to the manager or resell to outside firms.

filtration generated by the productivity process and the history of outside offers.[14] Since wage can be indeterminate under the risk-neutral assumption, I assume a constant wage contract to uniquely pin down the wage. That is, the wage is constant unless the firm needs to raise it in the bidding to retain the manager.[15] Formally, the contract includes the wage process generated by the wage bidding, the non-competition clause, and the buyout menu:[16]

$$\mathcal{C} = (\boldsymbol{w}, \mathcal{M}), \text{ where } \boldsymbol{w} = \{w_t\}_t \text{ and } \mathcal{M} = \{\pi, \{\tau_t(\tilde{\pi}) : \tilde{\pi} \in [0, \pi]\}\}.$$

2.3 The Contracting Problem

Before defining the contracting problem, I introduce three sets of necessary notations. First, the productivity threshold for outside firms to poach the manager at time t is denoted by $\bar{\theta}_t$. The initial match has a stopping time T, which occurs when the outside match productivity is above the poaching threshold, i.e., $\theta_T > \bar{\theta}_t$. Second, the bilateral joint value of a firm-manager match with productivity z_t is denoted by $J(z_t)$. For a firm-manager match excluded by a duration of π, their bilateral joint value is discounted to $e^{-r\pi}J(z_t)$. Third, following the recursive contract approach, the contracts are summarized by the level of promised utility delivered to the manager. When bidding for the manager, the two firms compete in utility terms. The initial and outside firms drop out at reservation utility levels $u_1(z_t) = J(z_t)$ and $u_2(z_t + \theta_t)$, respectively.

First, in the buyout stage, when the manager is poached by the outside firm, they jointly choose from the buyout menu to maximize the value of their match. This leads to the following incentive compatibility constraint:

$$\tilde{\pi}_t(\theta_t) = \arg\max_{\tilde{\pi} \in [0, \pi_t]} e^{-r\tilde{\pi}} J(z_t + \theta_t) - \tau_t(\tilde{\pi}). \tag{IC}$$

Second, since the initial firm commits to bidding up to the entire match value, the outside firm needs to deliver the promised utility $J(z_t)$. This leads to an individual rationality

[14] The non-competition duration in the contract π is some constant value. In fact, due to the proportionality of outside match productivity z_t' to realized current match productivity z_t, the non-competition duration depends only on parameters of the relative productivity distribution $F(\cdot)$ and is independent of realized current match productivity z_t. This result will become clear in subsequent discussion. Therefore, it is without loss of generality to restrict to a contract with a fixed constant non-competition duration.

[15] Such constant wage contract is justified by arbitrarily small amount of risk aversion on the manager's side. Any amount of risk aversion implies that the optimal contract would be a constant wage contract absent outside offer to insure the manager's risk. When risk aversion vanishes as in this setting, the constant wage contract still obtains.

[16] It is essential that the initial firm commits to wage bidding and hence the resulting wage process. However, It is not essential that the initial firm commits to the buyout price ex post. In other words, the contract does not necessarily need to specify the buyout price ex ante.

constraint for the outside firm:

$$e^{-r\tilde{\pi}_t(\theta_t)} J\left(z_t + \theta_t\right) - \tau_t\left(\tilde{\pi}_t\left(\theta_t\right)\right) - J\left(z_t\right) \geq 0, \ \forall \theta_t \geq \bar{\theta}_t. \tag{IR}$$

Given that both the firm and the manager are risk neutral, the bilateral joint value function is given by:

$$J\left(z_t\right) = \max_{\mathcal{C},\boldsymbol{\mu}} \mathbb{E}\left[\int_t^T e^{-rs} e^{z_s} ds + e^{-rT}\left[J\left(z_T\right) + \tau_T\left(\tilde{\pi}_T\left(\theta_T\right)\right)\right] \middle| \mathcal{F}_t\right]. \tag{1}$$

Original contracting problem. The initial firm with productivity z_0 and that promises ex-ante utility U_0 to the manager chooses the contract and investment to maximize its value:[17]

$$V\left(z_0, U_0\right) = \max_{\mathcal{C},\boldsymbol{\mu}} \mathbb{E}\left[\int_0^T e^{-rt}\left(e^{z_t} - w_t\right) dt + e^{-rT} \tau_T\left(\tilde{\pi}_T\left(\theta_T\right)\right) \middle| \mathcal{F}_0\right] \tag{$\mathcal{P}$}$$

subject to the outside firm's incentive compatibility and individual rationality constraints, (IC) and (IR), as well as the promise keeping constraint to the manager:

$$\mathbb{E}\left[\int_0^T e^{-rt} w_t dt + e^{-rT} J\left(z_T\right) \middle| \mathcal{F}_0\right] \geq U_0. \tag{PK}$$

I show that the contract achieves bilateral efficiency in the following Lemma:

Lemma 1 (Bilateral efficiency)**.** *The contract between the initial firm and the manager achieves bilateral efficiency. That is, the contract maximizes their bilateral joint value.*

The bilateral efficiency result is crucial for understanding the contracting problem. Due to firm commitment power and agent risk neutrality, the initial firm aligns the manager's incentive by costlessly backloading wages to retain the manager when outside opportunities arrive. An intuitive illustration of how this is achieved is to incorporate the (PK) constraint into the initial firm's objective in problem ($\mathcal{P}$). A formal proof is provided in Appendix B.1. One minor additional assumption is needed in the model parameters: the Poisson intensity λ for the arrival of outside opportunities is small. Infrequent arrival of outside opportunities ensures that the wage non-negativity constraint, $w_t \geq 0$, never binds.

Simplified contracting problem. With Lemma (1), the contracting problem between the initial firm and the manager becomes a bilateral joint maximization problem. The

[17] The determination of the level of promised utility at time zero, U_0, will be specified when I discuss wage-backloading. For now, some given level of utility is sufficient for the discussion.

bilateral joint value function follows the Hamilton–Jacobi–Bellman (HJB) equation:[18]

$$rJ(z) = \max_{\mathcal{M},\mu} y - c(\mu) e^z + \mu J'(z) + \frac{1}{2}\sigma^2 J''(z) + \lambda \int_{\bar{\theta}}^{\infty} \tau(z, \tilde{\pi}(z,\theta)) \, dF(\theta) \qquad (\mathcal{P}')$$

subject to the outside firm's incentive compatibility and individual rationality constraints:

$$\tilde{\pi}(\theta) = \arg\max_{\tilde{\pi} \in [0,\pi]} e^{-r} J(z+\theta) - \tau(\tilde{\pi}) \qquad \text{(IC)}$$

$$e^{-r\tilde{\pi}(\theta)} J(z+\theta) - \tau(\tilde{\pi}(\theta)) - J(z) \geq 0, \ \forall \theta \geq \bar{\theta}. \qquad \text{(IR)}$$

In the HJB equation, $J'(z)$ and $J''(z)$ denote, respectively, the first and second order derivatives of $J(z)$ with respect to z. From here onwards, I use the recursive formulation of the problem and corresponding notations.

It follows that in problem ($\mathcal{P}'$), the initial firm and the manager, as a bilateral coalition, act like a monopolist towards outside firms. They choose $\mathcal{M}$ to maximize the rent extracted from outside firms and thus their bilateral joint value.

I now define a laissez-faire equilibrium in which the external enforcer enforces the non-competition as contracted.

Definition 1 (Laissez-faire equilibrium)**.** A laissez-faire equilibrium consists of value function $J(z)$, contract $\mathcal{M} = \{\pi, \{\tau(z,\tilde{\pi}) : \tilde{\pi} \in [0,\pi]\}\}$, employment and investment allocation $\{\bar{\theta}, \mu\}$, and productivity distribution $G(z,t)$, given $G(z,0)$, such that:

i. the contract and the investment, together with the value function, solve problem ($\mathcal{P}'$); outside firms are individually rational;

ii. the productivity distribution follows the Kolmogorov Forward (KF) equation:

$$g_t(z,t) = -\mu g_z(z,t) + \frac{\sigma^2}{2} g_{zz}(z,t) + \delta[h(z) - g(z,t)] + \lambda \int_{\bar{\theta}}^{\infty} [g(z-\theta,t) - g(z,t)] \, dF(\theta). \qquad (2)$$

2.4 The Privately-Optimal Contract

In this section, I characterize the privately-optimal contract and demonstrate the economic forces pertinent to efficiency. In the first step, concerning the bilateral coalition's incentive

[18] Time is not a relevant state variable in the value function of individual firms or managers, since there is no aggregate state variable that affects the individuals. That is, the individual value function is time invariant and only depends on the (history of) realized individual idiosyncratic states.

to extract rent, I characterize the privately-optimal contract. In the second step, concerning the incentive for investment, I examine how investment responds to rent extraction.

To start, the simplified contracting problem ($\mathcal{P}'$) allows me to separate the contract decision, $\mathcal{M}$, from the investment decision, μ. The bilateral coalition chooses the contracting terms to maximize expected rent extraction:

$$\max_{\mathcal{M}} \int_{\bar{\theta}}^{\infty} \tau\left(z, \tilde{\pi}\left(z, \theta\right)\right) dF\left(\theta\right). \tag{3}$$

It invests up to the point where the marginal cost of investment is equal to the *marginal bilateral joint value:*

$$c'\left(\mu\right) e^{z} = J'\left(z\right). \tag{4}$$

Lemma 2 (Linearity)**.** *The bilateral joint value function is linear in productivity e^z, $J\left(z\right) = J\left(0\right) e^{z}$.*

The linearity result facilitates subsequent discussions. Intuitively, we can see that the flow payoff of production and investment cost is scaled by e^z and the relative productivity of outside match, $\theta = z' - z$, is independent of z. It follows that all the quantity decisions are independent of z while all the price decisions are linear in e^z. Hence, the bilateral joint value function is also linear in e^z.

Assumption 1 (Monotone hazard rate)**.** *The hazard rate $\frac{f(\theta)}{1-F(\theta)}$ is increasing in θ.*

The monotone hazard rate assumption ensures that there exists a unique non-competition duration or, equivalently, a unique poaching threshold. It is formally stated in the following proposition:

Proposition 1 (Privately-optimal contract)**.** *Under Assumption 1, in the privately-optimal contract, the non-competition duration and the buyout menu are:*

$$\pi = \frac{1}{r}\bar{\theta};\ \tau\left(z, \pi\right) = J\left(0\right) e^{z}\left(e^{r\pi} - 1\right), \tag{5}$$

where the poaching threshold $\bar{\theta} > 0$ is characterized by:

$$e^{\bar{\theta}}\left[1 - \frac{1 - F\left(\bar{\theta}\right)}{f\left(\bar{\theta}\right)}\right] = 1. \tag{6}$$

To illustrate the idea behind these results, I resort again to the intuition that the bilateral coalition acts like a monopolist towards outside firms. By seizing ownership of the manager's future employment during the non-competition π and setting a monopoly price menu (i.e.,

the buyout menu), the coalition maximizes the expected "profit" (i.e., the rent extracted) when reselling that ownership to outside firms. Given the linearity of the bilateral joint value function in Lemma 2, the monopolist charges a constant markup regardless of the realized productivity. That is, the productivity threshold for outside firms to poach the manager is distorted upward, $\bar{\theta} > 0$ independent of the realized productivity. Hence, it is without loss of generality that I have restricted the contract to one with fixed non-competition duration π.

Notice that, in equation (5), the buyout menu bunches to a single price, given the realized productivity. The intuition for this bunching result is that "demand" (the bilateral joint value of the outside match) is linear in "quantity" (the non-competition duration). That is, the bilateral coalition acts like a non-discriminating monopolist. As a result, the bilateral coalition fully extracts rent from the outside firm at the poaching threshold. And all outside firms that can poach the manager fully buyout the non-competition.

Having characterized the privately-optimal contract, the investment response to rent extraction follows immediately. Given the linearity result in Lemma 2 and the contract in Proposition (1), the investment optimality condition in equation (4) simplifies.

Corollary 1 (Investment holdup)**.** *The marginal cost of investment equals the marginal bilateral joint value, i.e., $c'(\mu) = j$, where the marginal bilateral joint value j satisfies:*

$$j = \frac{1 - c(\mu)}{r - \mu - \frac{1}{2}\sigma^2 - \lambda\left(e^{\bar{\theta}} - 1\right)\left(1 - F\left(\bar{\theta}\right)\right)}. \tag{7}$$

The marginal bilateral joint value in equation (7) has a clear economic interpretation: rent extraction, $\left(e^{\bar{\theta}} - 1\right)\left(1 - F\left(\bar{\theta}\right)\right)$, leads to higher marginal bilateral joint value of investment and a more efficient incentive for investment. That is, rent extraction allows the initial firm to partially capture the external payoff to its investment, thus alleviating investment holdup by exactly the amount of extracted rent. This interaction between the contracting externality and the investment externality generates an *investment-reallocation* tradeoff: a longer duration of non-competition alleviates the holdup due to the investment externality, while aggravating the distortion in manager allocation due to the contracting externality.

I conclude this section with three discussions of the contracting result and efficiency implications: (i) the relations to the contracts in closely related papers; (ii) the no-renegotiation assumption; and (iii) the ideas of rivalry and excludability in the use of knowledge.

Contracts in related papers. The non-competition buyout contract differs from Aghion and Bolton (1987) and Postel-Vinay and Robin (2002) in subtle ways. First, the contract in Aghion and Bolton (1987) sets a fixed stipulated damage payment for breach of contract,

creating barriers for outside firms to enter. The non-competition buyout contract achieves the same intended outcome of stipulated payment. While the distortionary effects are the same, non-competition buyout fits the institutional details in employment contracting.

Second, by setting the non-competition during to zero, $\pi = 0$, the outcome in the model reverts back to Postel-Vinay and Robin (2002), in which the poaching threshold is $\bar{\theta} = 0$. However, to model the efficient separation benchmark, their bargaining protocol of firms each making a take-it-or-leave-it offer to the manager under complete information is modified to firms engaging in wage bidding in the form of an ascending (English) auction under asymmetric information.

No renegotiation. Rent extraction leading to distortion in manager allocation hinges on ruling out the possibility of efficient renegotiation ex post. Such efficient renegotiation as in exclusionary contracts of Spier and Whinston (1995) and Segal and Whinston (2000) is made impossible by the asymmetric information assumption.

Rivalry and excludability in the use of knowledge. The model setting that the initial firm exits the economy after losing its manager is innocuous for efficiency analysis. Stark it might be, this setting only implies that the initial firm's outside option is zero, simplifying the accounting. A more general setting is when some fraction of the productive knowledge remains in the initial firm, nesting the other extreme of the initial firm fully retaining the productive knowledge, as in models of knowledge diffusion (see e.g., Lucas and Moll, 2014). Additionally, there can be costs of replacing the manager, as in Heggedal et al. (2017). Further, the firms might engage in duopolistic competition in the product market, as in Franco and Mitchell (2008). In all these settings, the initial firm's outside option would be some fraction of the match value. Crucially, the contract design and resulting externality are exactly unchanged. As long as there is a positive surplus from reallocating the manager and the labor market structure is such that outside firms capture some of the surplus, the bilateral coalition has an incentive to extract that surplus.

The distinction between rivalry and excludability in the use of knowledge, emphasized by Romer (1990) among others, has relevance here. Rivalry refers to the use of knowledge by one precluding the use by others; excludability refers to preventing others from using the knowledge. The key tension here is not the extent of rivalry but rather the extent of excludability. The bilateral coalition's ability to exclude outside firms from employing the manager and hence using the knowledge alters the appropriation of surplus from knowledge diffusion.

2.5 Wage-Backloading

This section examines how non-competition alters the extent of outside competitive pressure and in turn wage-backloading needed for retention. The helps to link the model to observed wage patterns in the data.

I introduce two additional details of the wage setting process, following Postel-Vinay and Turon (2010). First, when new-born managers and firms enter the economy, they engage in Nash bargaining to determine the initial promised utility U_0. The manager's bargaining weight is β and the outside options for both parties are zero. Second, as match productivity is stochastic, the firm (manager) initiates wage resetting if wage becomes too high (low). In particular, when the utility promised to the manager exceeds the bilateral joint value, resulting in negative firm value, the firm initiates wage resetting by reducing the promised utility to the level of bilateral joint value. Conversely, when the promised utility falls below the Nash bargained level, the manager initiates wage resetting, resetting the promised utility to the Nash bargained level.

These additional details are innocuous for efficiency analysis. The firm is still able to achieve bilateral efficiency using the long-term contract. Importantly, regarding manager-initiated wage resetting, the wage non-negativity constraint is slack when the chances of productivity moving downward are large. This condition can be ensured if the calibrated standard deviation for the Brownian motion is large.

I denote the value function of a manager with match productivity z and wage w by $U(z,w)$. Given the wage setting details, I obtain two results. First, when outside offers are unable to poach the manager but exceed the current level of promised utility, the initial firm increases the wage to match the outside offer. It implies a wage-bidding region $\left[\underline{\theta}(z,w), \bar{\theta}\right]$, where the lower threshold $\underline{\theta}(z,w)$ satisfies $u_1(z, \underline{\theta}(z,w)) = U(z,w)$. Second, firm-initiated and manager-initiated wage resetting leads to upper and lower bounds for the wage given the realized productivity, denoted by $\bar{w}(z)$ and $\underline{w}(z)$, respectively.

The manager's value function follows the HJB equation: $\forall w \in [\underline{w}(z), \bar{w}(z)]$,

$$(r+\lambda)U(z,w) = w + \lambda\Bigg\{F\left(\underline{\theta}(z,w)\right)U(z,w) + \int_{\underline{\theta}(z,w)}^{\bar{\theta}} e^{-r\pi}J(z+\theta)\,dF(\theta) + \left[1-F\left(\bar{\theta}\right)\right]J(z)\Bigg\} + \mu U_z(z,w) + \frac{1}{2}\sigma^2 U_{zz}(z,w), \tag{8}$$

with the boundary conditions:

$$U(z,w) = e^{-r\pi}J\left(z+\underline{\theta}(z,w)\right)$$

$$U\left(z,\underline{w}\left(z\right)\right)=\beta J\left(z\right),\ U_{z}\left(z,\underline{w}\left(z\right)\right)=\beta J'\left(z\right),\ U\left(z,\bar{w}\left(z\right)\right)=J\left(z\right),\ \text{and}\ U_{z}\left(z,\bar{w}\left(z\right)\right)=J'\left(z\right).$$

In the HJB equation, $U_{z}\left(z,w\right)$ and $U_{zz}\left(z,w\right)$ denote, respectively, the first and second order derivatives of $U\left(z,w\right)$ with respect to z.

Since outside firms unable to poach the manager have reservation value lowered by non-competition $e^{-r\pi}J\left(z+\theta\right)$, $\forall z\in\left[\underline{\theta}\left(z,w\right),\bar{\theta}\right]$, the firm needs to bid up wage less to retain the manager. In anticipation of that, to deliver the promised utility, the manager starts with a higher wage but experience slower wage growth.

To facilitate solving for wage dynamics over tenure, I perform a change of variable by re-writing the wage setting problem in terms of the wage-productivity ratio, $x\equiv\log\left(\frac{w}{e^{z}}\right)$. The manager's rescaled value function has a single state variable, $u\left(x\right)\equiv U\left(z,w\right)/e^{z}$.[19] Further, the wage bidding threshold $\underline{\theta}\left(z,w\right)$ reduces to $\underline{\theta}\left(x\right)$; the upper and lower bounds of wage, $\bar{w}\left(z\right)$ and $\underline{w}\left(z\right)$, reduce to maximum and minimum levels of wage-productivity ratio, $\bar{x}$ and $\underline{x}$. The HJB equation (8) for the manager's value function simplifies to: $\forall x\in[\underline{x},\bar{x}]$,

$$\left(r-\mu-\frac{1}{2}\sigma^{2}\right)u\left(x\right)=e^{x}+\lambda j\int_{\underline{\theta}(x)}^{\bar{\theta}}\left[1-F\left(\theta\right)\right]de^{\theta}-\left(\mu+\sigma^{2}\right)u'\left(x\right)+\frac{1}{2}\sigma^{2}u''\left(x\right),\qquad(9)$$

with the boundary conditions:

$$u\left(x\right)=e^{-r\pi+\underline{\theta}(x)}j,\ u\left(\underline{x}\right)=\beta j,\ u'\left(\underline{x}\right)=\beta,\ u\left(\bar{x}\right)=j,\ \text{and}\ u'\left(\bar{x}\right)=0.$$

The distribution of wage-productivity ratio at tenure t, $\psi\left(x,t\right)$, conditional on the match continuing, follows the KF equation: $\forall x\in[\underline{x},\bar{x}]$,

$$\psi_{t}\left(x,t\right)=\mu\psi_{x}\left(x,t\right)+\frac{1}{2}\sigma^{2}\psi_{xx}\left(x,t\right)+\lambda\left\{\frac{f\left(\underline{\theta}\left(x\right)\right)}{F\left(\bar{\theta}\right)}\Psi\left(x,t\right)-\left[1-\frac{F\left(\underline{\theta}\left(x\right)\right)}{F\left(\bar{\theta}\right)}\right]\psi\left(x,t\right)\right\}.\quad(10)$$

I focus on the wage dynamics for new-born matches. These matches have initial wage-productivity ratio x_{0} satisfying the Nash bargained result.[20] That is, $\psi\left(x,0\right)$ has a unit mass at $x=\underline{x}$.

The first two terms in the KF equation (10) reflect the evolution due to the productivity process z. The terms in the large bracket reflect the evolution due to wage w jumping

[19] The change of variable transforms the problem from one of solving a partial differential equation to one of solving an ordinary differential equation.

[20] In the calibrated model, majority of new matches are new-born matches and a relatively small fraction of matches occur due to job-to-job transition.

upward in response to outside offers conditional on match survival: the inflow is the fraction with wage-productivity ratio below x, $\Psi(x,t)$, and bid up to x, $f(\underline{\theta}(x))/F(\bar{\theta})$; the outflow is the fraction with wage-productivity ratio at x, $\psi(x,t)$, and bid up to a level above x, $1-F(\underline{\theta}(x))/F(\bar{\theta})$.[21]

2.6 Aggregation

I now characterize the steady state features of the laissez-faire equilibrium. In particular, I characterize the stationary productivity distribution, $g(z)$, and the aggregate net output, defined as output net of investment cost,[22]

$$Y = \int (e^z - c(\mu) e^z) \, dG(z).$$

The distribution from which the new-born matches draw initial productivity, $H(\cdot)$, is specified as a unit mass at zero.

Proposition 2 (Steady state)**.** *In steady state,*

i. the stationary productivity distribution has a double asymptotic Pareto tail:

$$g(z) \sim \begin{cases} e^{-\zeta_+ z}, & z \to +\infty, \\ e^{-\zeta_- z}, & z \to -\infty, \end{cases}$$

where the Pareto indices $\zeta_\pm$ are the roots of the characteristic equation:

$$\frac{1}{2}\sigma^2\zeta^2 - \mu\zeta - \delta + \lambda\left[\int_{\bar{\theta}}^{\infty} e^{\zeta\theta} dF(\theta) - \left(1 - F\left(\bar{\theta}\right)\right)\right] = 0; \tag{11}$$

ii. the aggregate net output is:

$$Y = \frac{\delta(1-c(\mu))}{\delta - \frac{1}{2}\sigma^2 - \mu - \lambda\int_{\bar{\theta}}(e^\theta - 1)\, dF(\theta)}. \tag{12}$$

The first part of Proposition 2 states that, even though the endogenous stationary distribution doesn't have a closed-form expression, the distribution has a double asymptotic Pareto tail. The tail indices are easily characterized according to equation (11). This result will be useful when linking the model implied distribution to the one in the data.

[21] In mathematical terms, the underlying process of wage-productivity ratio x is a jump diffusion process where the jump is state dependent.

[22] This accounting definition for output, by fully expensing investment cost, is the relevant metric for steady state welfare.

In the second part, the aggregate output equation (12) provides some insight into the tradeoff between alleviation in investment holdup and distortion in manager allocation. Rent extraction leads to increased productivity gain from firm investment, μ, while reducing productivity gain from manager reallocation, $\int_{\bar{\theta}}^{\infty}\left(e^{\theta}-1\right)dF\left(\theta\right)$.

One minor detail in this proposition is that the choice of the functional form for $H\left(\cdot\right)$ is a mere normalization, because for net output in steady state the mean value of the new-born match productivity, $\int e^{z}dH\left(z\right)$, is the relevant parameter.

3 Planner's Problem

In this section I study whether restrictions on non-competition contracts can improve efficiency. Specifically, I consider a planner who can cap the duration of non-competition.[23] Before proceeding to the planner's problem, I first examine the first-best allocation in the economy, which is useful for illustrating the efficiency tradeoff.

3.1 First-Best Allocation

The social welfare is defined as the discounted stream of aggregate output net of investment cost. To achieve the first-best, the planner directly chooses the allocation, the level of investment $\boldsymbol{\mu}=\left(\mu_t\right)_{t\geq 0}$ and reallocation threshold $\boldsymbol{\theta}=\left(\theta_t\right)_{t\geq 0}$:

$$\max_{\mu,\theta}\int_0^{\infty}e^{-\rho t}\left[\int\left(e^{z}-c\left(\mu\right)e^{z}\right)dG\left(z,t\right)\right]dt \qquad (\mathcal{P}^{*})$$

subject to the KF equation for productivity distribution (2).

Proposition 3 (First-best)**.** *In the first-best allocation:*

i. the poaching threshold: $\bar{\theta}^{*}=0$;

ii. the marginal cost of investment equals the marginal social value, i.e., $c'\left(\mu^{*}\right)=\gamma^{*}$*, where the marginal social value* γ^{*} *satisfies:*

$$\gamma^{*}=\frac{1-c\left(\mu^{*}\right)}{r-\frac{1}{2}\sigma^{2}-\mu^{*}-\lambda\int_0^{\infty}\left(e^{\theta}-1\right)dF\left(\theta\right)}.$$

The first-best can be achieved as follows. First, managers are allocated to outside firms whenever the outside match productivity exceeds the initial one. Second, investment takes

[23] Apart from the duration of non-competition, policy discussions of restricting non-competition also consider limiting the geographic scope and industry scope.

into account the payoff to the initial firm and the manager, as well as the payoff to outside firms. That is, the marginal cost of investment $c'(\mu^*)$ equals the marginal social value γ^*.

Implementation of the first-best requires that the bilateral coalition can fully capture the external payoff to outside firms and thus investment at the socially efficient level, while manager allocation is undistorted. Under perfect information, this can be achieved by assigning all the bargaining power to initial firms: it makes a take-it-or-leave-it offer to outside firms to fully extract rent without distorting allocation.

3.2 Optimal Restriction on Non-Competition

I now study the planner's optimal policy choice of maximum enforceable non-competition duration π^p binding private contracts $\mathcal{C}$, while leaving the contracting to the private parties. It is without loss of generality to restrict the cap from zero to the privately-optimal level, $\pi^p \in [0, \pi]$. The planner maximizes the social welfare:

$$\max_{\pi^p} \int_0^\infty e^{-\rho t} \left[\int (e^z - c(\mu) e^z) \, dG(z,t) \right] dt \qquad (\mathcal{P}^{**})$$

subject to the private contracts $\mathcal{C}$ following π^p, the firms' investment incentive constraint in equation (4), and the KF equation (2). Since the cap on non-competition duration will always be binding, it is equivalent to consider the planner's problem as one of maximizing social welfare by directly choosing contract $\mathcal{C}$ subject to the firms' investment incentive constraint and the KF equation.

I obtain the following proposition:

Proposition 4 (Optimal restriction on non-competition)**.** *The optimal cap on non-competition duration is below the privately-optimal level, i.e., $\pi^p \leq \pi$.*

The proposition above states that the privately-optimal contract, despite being bilaterally efficient, is socially inefficient along the investment-reallocation tradeoff. This social inefficiency can be seen through the operation of contracting externality. In the contracting problem ($\mathcal{P}'$), the bilateral coalition maximizes their bilateral joint value and disregards the value of outside firms. Compared to a planner who aims to maximize social value, the bilateral coalition sets excessively long non-competition duration and over-extracts rent. Therefore, the planner can improve efficiency by capping non-competition duration.

The optimal cap on non-competition duration can be related to the literature on optimal patent duration starting with Nordhaus (1967). There is an analogous tradeoff between the static and dynamic considerations for the two policies. For the patent policy, increasing patent duration encourages more efficient investment at the expense of static distortion

due to additional incumbent monopoly power. For the non-competition policy, the "static" distortion is due to the bilateral coalition's "monopoly" power over outside firms.

I conclude the efficiency discussion with a comparison of the three allocations, the laissez-faire allocation $\left(\bar{\theta}, \mu\right)$, the optimal cap allocation $\left(\bar{\theta}^p, \mu^p\right)$, and the first-best allocation $\left(\bar{\theta}^*, \mu^*\right)$. They satisfy the following relations: $\bar{\theta} > \bar{\theta}^p > \bar{\theta}^*$ and $\mu^p < \mu < \mu^*$. Both the laissez-faire allocation and the optimal cap allocation are in the interior of the first-best allocation. Compared to the bilateral coalition, the planner would like to restore a more efficient allocation at the expense of less efficient investment.

4 Extensions

This section provides two extensions to the baseline model to account for additional economic forces that affect the contract design or the efficiency results.

4.1 Non-Competition as "Damaged Goods"

In the baseline model, outside firms fully buyout non-competition and, therefore, non-competition is never enforced. To reconcile with observed enforcement of non-competition in actual practices, I extend the baseline model to include business stealing by outside firms from initial firms and knowledge depreciation during the non-competition period. The contract features a continuous buyout menu.[24]

In the baseline environment suppose that when a manager leaves to join the outside firm, there is a business stealing effect inflicted upon the initial employer. In addition to the lost marginal productive value of the manager, the initial firm also suffers stolen business of amount $e^{-\eta t}\nu y$ at subsequent time $t \geq 0$ while it survives. The knowledge depreciation rate $\eta \geq 0$ captures the idea that the manager's inside knowledge about the initial firm becomes less relevant and integral over time. The total value of stolen business after π duration of non-competition is:

$$\Upsilon\left(z, \pi\right) = \int_{\pi}^{\infty} e^{-(r+\eta)t+z}\nu dt = e^{-(r+\eta)\pi+z}\frac{\nu}{r+\eta}.$$

The social surplus from moving to the outside firm θ subject to non-competition π, conditional on current realized productivity z, is denoted by $S\left(\theta, \pi \middle| z\right) \equiv e^{-r\pi}J\left(z+\theta\right) - \Upsilon\left(z, \pi\right) - J\left(z\right)$.

[24]The contract example in Figure 9 in the additional data appendix has two part buyout menu

Assumption 2 (Log-submodularity). *The social surplus $S(\theta, \pi|\, z)$ is log-submodular in relative productivity θ and non-competition duration π, i.e., $\frac{\partial^2 \log S(\theta,\pi|z)}{\partial\theta\partial\pi} < 0$.*

Assumption 2 is satisfied when ν and η are sufficiently large. When the assumption does not hold, for example in the absence of business stealing $\nu = 0$ or knowledge depreciation $\eta = 0$, the model environment and the contract revert back to the baseline one.

Proposition 5. *Under Assumption 1 and 2, for firms that poach the manager, their non-competition buyout decision:*

$$\tilde{\pi}(\theta) = \max\left\{\frac{1}{\eta}\left[\log\left(\frac{r+\eta}{r}\nu\right) - \log\left[1 - \frac{1-F(\theta)}{f(\theta)}\right] - \theta\right], 0\right\},\ \forall\theta > \bar{\theta},$$

where the poaching threshold $\bar{\theta}$ is characterized by:

$$e^{\bar{\theta}}\left[1 - \frac{1-F\left(\bar{\theta}\right)}{f\left(\bar{\theta}\right)}\right] = \frac{(r+\eta)}{\eta^{\frac{\eta}{r+\eta}} r^{\frac{r}{r+\eta}}}\nu^{\frac{r}{r+\eta}}.$$

The contract features a continuum buyout menu, in contrast to the single buyout price in the baseline model. Non-competition is enforced in equilibrium to price discriminate against less productive outside firms, whereas in the baseline model non-competition is always fully bought out and never enforced. This extension reconciles the model with selective non-competition enforcement observed in actual practices.

Non-competition enforcement is likened to the "damaged goods" phenomenon in industrial organization (Deneckere and McAfee, 1996) where a monopolist intentionally damages goods to achieve price discrimination. In this setting, the initial firm as the monopolist uses non-competition to create a damaged version of managerial human capital to achieve price discrimination against outside firms and maximize rent extracted. Assumption 2, along with Assumption 1, is necessary for price discrimination to be optimal for a monopolist (see Anderson and Dana, 2009).

4.2 Free Entry of New Firms

The baseline model captures the barriers to entry by distorting manager allocation, taking as given the outside opportunities at new firms. It abstracts away from the incentives for new firms to enter in the first place. This limitation is due to the assumption that outside opportunities arrive exogenously. To better account for the entry channel, I recast the baseline model in a general equilibrium setting. Specifically, I endogenize the arrival rate of outside opportunity by introducing a random labor search market and costly free entry of new firms.

Consider, in the baseline environment, a measure one of managers search on the job and a measure v of ex-ante identical entrant firms post vacancies. The managers and entrants are matched with technology $\lambda(v) \equiv M(1, v)$. I assume that $M(1, v)$ is increasing in the measure of entrants v and satisfies constant return to scale. The vacancy filling rate for entrants is $\lambda(v)/v$ and the job finding rate for managers is $\lambda(v)$. Upon being matched with a manager with existing match productivity z, the entrant firm draws a relative productivity θ from distribution $F(\theta)$. New firms incur a flow cost of κ maintaining an open vacancy. Building on the results in the baseline model, I obtain the result that a new firm that draws a productivity θ and poaches away a manager has a value $J(z+\theta) - J(z) - \tau(z, \pi)$. In the steady state equilibrium, the free entry condition is:

$$\frac{\lambda(v)}{v} \int \int_{\bar{\theta}}^{\infty} [J(z+\theta) - J(z) - \tau(z, \pi)] \, dF(\theta) \, dG(z) \leq \kappa \text{ with } \text{“}=\text{”} \text{ if } v > 0. \tag{13}$$

Given the linearity of the value function in Lemma (2), the free entry condition (13) simplifies to:

$$\frac{\lambda(v)}{v} \left[\int_{\bar{\theta}}^{\infty} \left(e^{\theta} - e^{\bar{\theta}} \right) dF(\theta) \right] \left[\int J(z) \, dG(z) \right] \leq \kappa \text{ with } \text{“}=\text{”} \text{ if } v > 0. \tag{14}$$

The free entry condition in equation (14) shows how entry will respond to the incumbents' rent extraction achieved by non-competition. Since entrants' ex post surplus, $\int_{\bar{\theta}}^{\infty} \left(e^{\theta} - e^{\bar{\theta}} \right) dF(\theta)$, is diminished, the measure of entrants v decreases and so does the arrival rate of outside opportunity for managers $\lambda(v)$.

The efficiency concerns are more complicated: in addition to the investment externality and the contracting externality, the labor market introduces a *search externality*.[25] Without formally solving optimal policies, I discuss intuitively how the efficiency results might change. The holdup of investment is now two-sided: in addition to the positive external effect of incumbent firms' continuous investment, there is also a positive external effect of new firms' investment to enter. In other words, there is some complementary between the investments on the two sides. The insight of Hosios (1990) has some relevance here. According to the Hosios condition, the constrained efficient outcome is obtained when the surplus division between the two sides equals their contributions to matching, i.e. the elasticity of the matching function to the measures of mangers and entrants. Non-competition shifts that surplus division. If matching is highly elastic to the measure of entrants, then it is desirable to restriction non-competition and shift the surplus division in favor of entrants. I leave a more complete analysis to future exploration.

[25] Diamond and Maskin (1979) also consider buyout payment of employment contract in a general equilibrium random search model and examines the interaction of contracting externality and search externality.

5 The Data

To quantitatively evaluate the theory, I use the data on executives employed in U.S. public firms. The close scrutiny of the managerial labor market allows me to put together a rich array of data from various sources. Specifically, I assemble a new dataset on non-competition contracts from company filings in `SEC Edgar` and merge the contract data with three sets of standard data. These standard data include executive compensation from `Execucomp`, executive biographies from `Capital IQ People Intelligence`, and firm-level information from `Compustat`. In the following sections, I provide a brief description of the relevant data features.

5.1 Executive Compensation and Movements

The `Execucomp` dataset is the basis for the firm-executive matches in the sample.[26] I make use of the data in two regards. First, it provides information on the level and composition of executive compensation, including cash compensation such as salary, bonus, and non-equity incentive payment, as well as equity compensation in the form of stock and option grants. This information is well suited for examining compensation design and wage-backloading. Second, it also provides information on employment history, allowing me to keep track of firm-executive match separation and executive movements across firms and measure executive tenure with the firm. However, the employment history information in `Execucomp` is less than ideal, as some observations are missing starting and ending dates for employment. To improve measurement, I supplement with available employment history data from `Capital IQ People Intelligence`.

5.2 Employment Contracts

To collect data on executive contractual arrangements, I conduct textual analysis of employment contracts included in company filings in the `SEC Edgar` database from 1994 to 2015. Specifically, I apply natural language processing and machine learning tools to classify contracts. I then extract information on contractual terms, including (1) whether an employment contract includes a non-competition clause; (2) and if so, the duration of the non-competition period, most commonly one year, eighteen months, or two years, but in some cases as long as five years. Further details on how the contracts are classified and processed are provided in the additional data appendix D.2.

[26] The entire sample in `Execucomp` includes 45,287 executives employed at 3,557 firms for the period from 1992 to 2015.

A total of 45,446 contracts were merged with the sample of firms-executive matches included in the `Execucomp` dataset.[27] I keep the firm-executive matches in `Execucomp` linked to at least one contract. The merged sample includes 17,928 executives employed at 2,916 firms, a total of 19,035 firm-executive matches. I define an executive being subject to non-competition if at least one non-competition clause is found among his contracts with the firm; otherwise, the executive is not subject to a non-competition clause.

5.3 Sample Selection

The merged sample is filtered in five steps following the standard procedures in the literature. First, I exclude firms operating in regulated industries (SIC Codes 4900-4999) and financial industries (6000-6999). Second, I exclude firm observations with missing or non-positive book value of assets or sales, as well as firm observations with less than $5 million in physical capital in 2010 dollars. Third, to avoid bias due to merger and acquisition activities, I exclude firm observations with annual asset or sales growth over 100%. Fourth, I drop firm-executive matches for which tenure cannot be reliably determined. An accurate measure of tenure is needed because match separation and executive compensation are tenure dependent.[28] Finally, I restrict the observations to executives in the age range between 25 and 65. The final sample includes 9,204 executives, 2,009 firms, 9,758 firm-executive matches, and 54,922 firm-executive-year observations. The first and fourth steps lead to the largest reductions in sample size. Specifically, 4,469 firm-executive matches (or 23%) are dropped when filtering by industry; 3,487 firm-executive matches (or 18%) are dropped due to missing tenure.

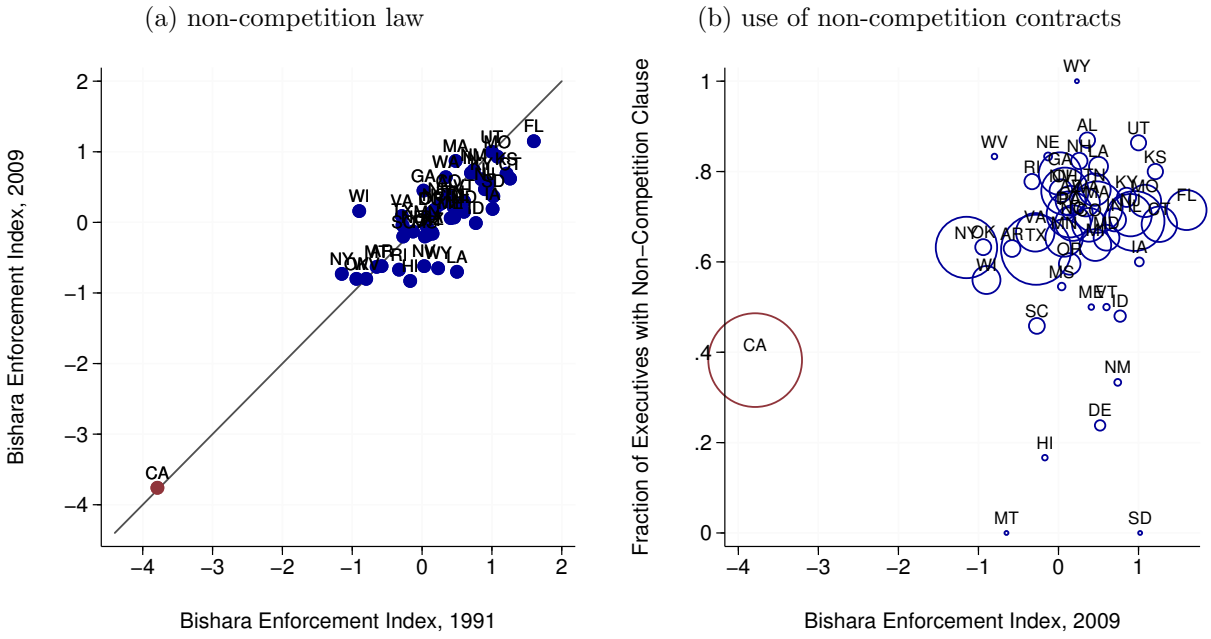
5.4 Non-Competition Legal Regime

Non-competition law varies across states and some measure of state non-competition legal regime is useful. Following empirical studies on non-competition contracts (e.g., Prescott et al., 2016; Lavetti et al., 2017), I use the Bishara enforcement index as a proxy for state legal regime. Some explanation of the index is in order. Bishara (2011) scores the enforceability of non-competition contracts based on legislation and case law across jurisdictions along the following dimensions: whether a state statute of general enforceability exists, scope of employer's protectable interest, plaintiff's burden of proof, consideration provision, modification of overly-broad contracts, and enforceability upon firing. Building on that, Starr (2016)

[27] Gillan et al. (2009) and Bishara et al. (2015) respectively hand-collect around 500 and 1000 executive employment contracts.

[28] Although the model doesn't capture any tenure dependence of firm-executive match separation, it is well documented in empirical studies on executive turnover that the turnover rate tends to decrease with tenure (e.g., Taylor (2010)). Therefore, in my empirical analyses, I need an appropriate control for tenure.

Figure 1: Differences in non-competition law and contracts across states



Notes: The figure in Panel (a) plots the Bishara enforcement index for the year 1991 and 2009. The two years roughly cover the beginning and end years of the executive data sample (1992 to 2015). The figure in Panel (b) plots the fraction of executives with non-competition clauses against the Bishara enforcement index in 2009. The size of the circles represents the total number of firm-executive matches in the state where the company headquarter is located.

constructs state-level weighted indices for 1991 and 2009 which I borrow. The indices are plotted in panel (a) of Figure 1.

I group the states into two legal regimes: (i) California in the regime that bans non-competition; (ii) non-California states in the laissez-faire regime that permits and enforces the non-competition terms as contracted. The reasons for this grouping are three-fold. First, California has a distinctively different non-competition law – a statutory ban[29]. Its enforcement index is much lower than the rest.[30] Second, non-California states have very similar non-competition laws. Their enforcement indices are very close. Third, although state non-competition law has changed over time, these changes are relatively minor. As panel (a) of Figure 1 shows, the 1991 and 2009 index levels are broadly on the 45 degree line.

[29]One exceptional situation to the statutory ban on non-competition contracts in California is sale of business ownership. In situations of selling business ownership, non-competition contracts can be enforced.

[30]North Dakota in fact has a lower Bishara enforcement index than California. Since all 58 North Dakota observations in the merged sample were filtered out in the final sample, I leave out North Dakota in the discussion.

5.5 Data Summary

The summary statistics are presented in Table 6 for the entire sample, sorted by whether there is a non-competition clause, and if so by the length of non-competition duration. Nominal values are deflated to year 2010 dollars using the CPI.

The data shows that 64% of executives are subject to non-competition, confirming the wide prevalence of non-competition contracts for executives.[31] Panel (b) of Figure 1 plots the fraction of executives under non-competition contracts by state. As expected, California employers use non-competition least frequently at 40%. The fraction in non-California states is close to 70%.

I note that, despite the statutory ban in California, non-competition contracts are still used. There are many possible reasons for this behavior. An important one is that employers might be able to enforce the non-competition in another state by engaging in jurisdictional arbitrage. I abstract away from modeling and accounting for these intricacies.

6 Empirical Patterns: Effects of Non-Competition

In this section, I take the model's three predictions of how non-competition contracts affect firm-executive matches to the data. Specifically, I look at how observed uses of non-competition contracts affects match separation, firm investment, and executive wage-backloading. I also look at whether these effects differ between California and non-California states, representing two different legal regimes.

I note one important caveat on endogeneity when interpreting these magnitudes. While the results control for observable firm and manager characteristics, there can be selection into non-competition contracts due to unobservable characteristics.

6.1 Executive Mobility and Reallocation

To examine the restrictive effect of non-competition on executive mobility, I estimate the Cox proportional hazard model with the following specification:

$$\log H_{ijt} = \beta NC_{ij} + \gamma X_{ijt} + \varepsilon_{ijt},$$

[31] Bishara et al. (2015) also study the use of restrictive covenants for CEOs, using around 1,000 manually collected employment contracts for a sample of randomly selected public firms. They find that 78.7% of CEOs have signed non-competition. The corresponding number in my sample is 67.8%. Errors abound in both manual and automated employment contract collecting. If one were to lend all the confidence to Bishara et al. (2015) and attribute the discrepancy to the automated approach, it would imply that in my sample a fraction of matches with non-competition are misclassified as without. The effects of non-competition contracts in the following subsections would be underestimated.

where the separation hazard for executive i at firm j in period t, H_{ijt}, depends on whether the executive is under non-competition with the firm, NC_{ij}, and other observable characteristics of the executive and the firm, X_{ijt}.

Table 1 reports the regression results for separation events. The baseline regression in column (1) controls for industry, year, and state fixed effects for the sub-sample of non-California observations. It shows that executives with non-competition are associated with a separation hazard rate that is 86% of those without non-competition. This ratio is obtained by taking the exponential of the coefficient for non-competition, -0.154. Column (2) shows that, when including California, the magnitude is slightly smaller. A closer look at whether it matters to be located in California is reported in column (3). As expected, the effect of non-competition is almost zero in California and close to the baseline outside California. Column (4) shows that the magnitude is slightly larger when controlling for firm fixed effects. I use a Poisson regression instead in this specification, as the Cox regression is not capable of controlling for a large number of fixed effects. It reassures that the result is not sensitive to unobservable firm fixed effects.

I also report the regression results for job-to-job transition events, a more direct measure of executive mobility. I note beforehand that I do not use the job-to-job transition results for subsequent quantitative analysis for two reasons. First, the data doesn't permit accurate measurement of job-to-job transition, since the sample includes only top executive jobs in Compustat firms satisfying regulatory disclosure requirement. In fact, it tends to under-measure actual job transitions. In addition, the events as observed are low frequency, and the regression methods tend to have difficulty in obtaining accurate estimation of the effects. Despite the issues, the regression results, as presented in Table 7 in Appendix C, still offer corroborative evidence. Executives under non-competition are associated with lower overall job-to-job transition rates. Despite both being insignificant, the effect on within industry transition is of a larger magnitude than the effect on between-industry transition.

6.2 Firm Investment

Firm investment response to the use of non-competition contracts motivates the investment regression equation below:

$$INV_{jt} = \beta \bar{NC}_j + \gamma X_{jt} + \varepsilon_{jt},$$

where firm j's investment expenditure in period t, INV_{jt}, depends on the fraction of executives under non-competition with the firm, $\bar{NC}_j$. The equation is at firm-level because investment is reported at firm level. Standard control variables for investment such as Tobin's Q and cash are included. The definition for investment is total investment, inclusive of

Table 1: Effect of non-competition on firm-executive match separation

	Y/N				Duration
	(1)	(2)	(3)	(4)	(5)
Non-Competition	-0.154***	-0.120***		-0.179***	-0.053***
	(0.031)	(0.038)		(0.044)	(0.017)
Non-Competition × Non-CA			-0.148***		
			(0.032)		
Non-Competition × CA			-0.001		
			(0.010)		
Sample	Non-CA	All	All	Non-CA	Non-CA
Regression	Cox	Cox	Cox	Poisson	Cox
Industry FEs	Y	Y	Y	N	Y
Year FEs	Y	Y	Y	Y	Y
State FEs	Y	Y	Y	N	Y
Firm FEs	N	N	N	Y	N
Observations	44,927	52,604	52,604	43,534	40,118

*Notes: All specifications control for executive age, firm asset, total Tobin's Q, and return on asset. The hazards in column (1) to (3) and (5) are stratified by whether the executive holds the role of CEO, whether the executive is interlocked, and the gender of the executive. The specification in column (4) also control for tenure, the square of tenure, whether the executive holds the role of CEO, whether the executive is interlocked, and the gender of the executive. Standard errors clustered by state are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.*

Table 2: Effect of non-competition on firm investment

	(1)	(2)	(3)	(4)
Non-Competition	0.014***	0.011**		0.002
	(0.003)	(0.005)		(0.004)
Non-Competition × Non-CA			0.017***	
			(0.003)	
Non-Competition × CA			-0.018***	
			(0.001)	
Sample	Non-CA	All	All	Non-CA
Industry FEs	Y	Y	Y	N
Year FEs	Y	Y	Y	Y
State FEs	Y	Y	Y	N
Firm FEs	N	N	N	Y
Observations	18,053	21,124	21,124	18,053
Adjusted R^2	0.47	0.45	0.45	0.43

*Notes: All specifications control for total Tobin's Q and cash. Standard errors clustered by state are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.*

physical and intangible investments, following Peters and Taylor (2017) among others.[32]

Table 2 shows the investment regression results. As in the previous section on executive mobility, I control for industry, year, and state fixed effects and carry out robustness checks dividing the sample by whether in California and controlling for firm fixed effects. For the sample of non-California firms, more usage of non-competition is associated with more investment. Specifically, in the baseline regression in Column (1), when the fraction of executives subject to non-competition increases by a magnitude of 100%, firms have an investment rate that is 1.4 percentage points (or 9%) higher per annum. I note that the regression in Column (3) indicates the opposite for California firms. The model cannot capture or account for this result.

I also check the regression for physical investment and intangible investment separately, reported in Table 8 in Appendix C. The pattern of higher fraction of non-competition associated with more investment holds not only for intangible investment but also physical investment. In fact, the magnitude is larger and more significant for physical investment. This is consistent with previous empirical studies which have also found physical investment responds to non-competition (see e.g., Garmaise, 2009; Jeffers, 2017). Therefore, I use total investment as the measure of investment.

6.3 Wage-Backloading

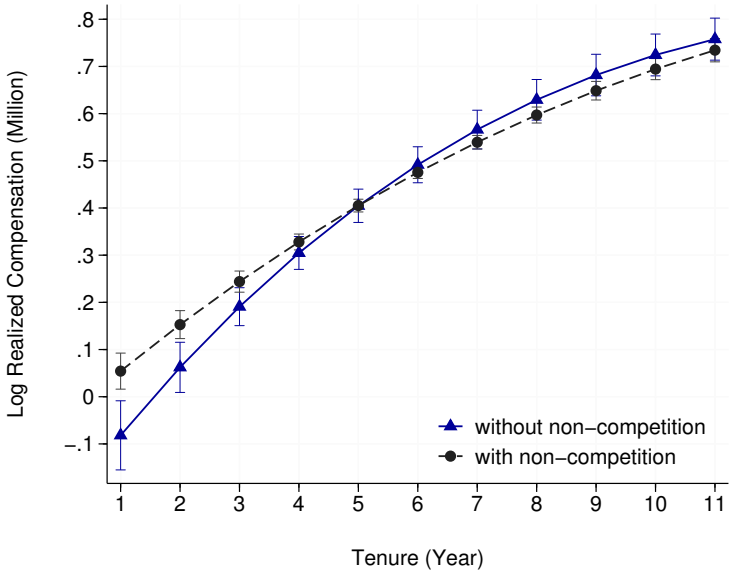
To examine how non-competition contract interacts with wage-backloading, I use the wage regression equation specified as follows:

$$W_{ijt} = \beta_1 NC_{ij} + \sum_{k=1}^{3} \beta_{2,k} T_{ijt}^k + \sum_{k=1}^{3} \beta_{3,k} \cdot T_{ijt}^k \times NC_{ij} + \gamma X_{ijt} + \varepsilon_{ijt},$$

where the wage for executive i at firm j in period t, W_{ijt}, depends on whether the executive is under non-competition with the firm, NC_{ij}, the tenure of the executive at the firm, T_{ijt}, and other observable characteristics of the executive and the firm, X_{ijt}. To allow for the tenure effect to depend on non-competition contract, I include the interaction of tenure with non-competition, $T_{ijt} \times NC_{ij}$. To allow for a non-linear tenure effect, as the wage-bidding channel in the model predicts, I also include higher order polynomials of tenure, T_{ijt}^2, T_{ijt}^3, and their interactions with non-competition, $T_{ijt}^2 \times NC_{ij}$, $T_{ijt}^3 \times NC_{ij}$. The control variables X_{ijt} include the standard ones in the executive compensation literatures: firm asset, Tobin's

[32] Total investment is defined as the sum of physical investment and intangible investment, normalized by the sum of lagged physical capital and intangible capital. Intangible investment is defined as R&D expense plus 30% of selling, general, and administrative expense. Intangible capital is the estimated replacement cost of the firm's intangible capital calculated by Peters and Taylor (2017).

Figure 2: Wage-backloading by whether under non-competition



Notes: This figure plots wage over tenure by whether the executive is subject to non-competition, based on the marginal effects at means in the baseline regression in column (1) of Table 3. The bars display 95% confidence interval.

Q, return on asset, whether the executive holds the role of CEO, whether the executive is interlocked, and the gender of the executive.

Two types of compensation measure are reported by public firms for their top executives per SEC regulations – awarded compensation and realized compensation. A large part of awarded compensation is in the form of restricted equity, which are deferred to future dates contingent on the executive staying with the firm. Deferred compensation is exactly the means to achieve wage-backloading for retention purposes in compensation design.[33] Therefore realized compensation is more pertinent to gauging wage-backloading than awarded compensation.[34]

Table 3 shows the tenure effect and its interaction effect with non-competition on wage. Column (1) reports the baseline regression, using realized compensation as the wage measure

[33]Much discussion in the executive compensation literature revolves around the moral hazard aspect of agency problem, as opposed to retention due to limited commitment. Distinguishing between limited commitment and moral hazard is difficult, as noted by Gopalan et al. (2014). My results suggest that retention is indeed an important consideration in contract and compensation design. The narratives of the contracts, for example the restricted stock award agreement between Amazon and its key employees in Figure 10 in the additional data appendix, also suggest retention concerns.

[34]I use the realized compensation (tdc_total2) according to the post-2006 definition as the wage measure. This measure values stock and options grants at the market value rather than the book value. Further details are provided in the additional data appendix on how regulatory disclosure requirement relates to the discussion and robustness checks with alternative compensation measures.

Table 3: Effect of non-competition on wage-backloading

	Realized				Awarded
	Y/N			Duration	Y/N
	(1)	(2)	(3)	(4)	(5)
Non-Competition	0.222***	0.228***	0.134***	0.110***	0.158***
	(0.051)	(0.050)	(0.043)	(0.030)	(0.049)
Tenure/10	1.688***	1.648***	1.430***	1.527***	0.371***
	(0.159)	(0.157)	(0.133)	(0.136)	(0.136)
Tenure/10 × Non-Competition	-0.589***	-0.597***	-0.294**	-0.255***	-0.238*
	(0.153)	(0.154)	(0.126)	(0.077)	(0.130)
$(\text{Tenure}/10)^2$	-0.856***	-0.861***	-0.654***	-0.738***	-0.181*
	(0.125)	(0.126)	(0.116)	(0.102)	(0.102)
$(\text{Tenure}/10)^2$ × Non-Competition	0.466***	0.477***	0.285**	0.195***	0.143
	(0.121)	(0.123)	(0.111)	(0.051)	(0.100)
$(\text{Tenure}/10)^3$	0.134***	0.137***	0.099***	0.109***	0.030
	(0.024)	(0.025)	(0.024)	(0.020)	(0.020)
$(\text{Tenure}/10)^3$ × Non-Competition	-0.098***	-0.101***	-0.065***	-0.038***	-0.027
	(0.024)	(0.024)	(0.023)	(0.010)	(0.020)
Age	N	Y	N	N	N
Industry FEs	Y	Y	N	Y	Y
Year FEs	Y	Y	Y	Y	Y
State FEs	Y	Y	N	Y	Y
Firm FEs	N	N	Y	N	N
Observations	17,948	17,948	17,948	15,998	17,970
Adjusted R^2	0.55	0.56	0.34	0.55	0.60

*Notes: Non-competition in the specifications in column (1)-(3) and (5) is a binary variable indicating whether the executive is subject to non-competition. Non-competition in the specification in column (4) is a continuous variable indicating the duration of non-competition, which equals zero if there is no non-competition. All specifications control for firm asset, total Tobin's Q, return on asset, whether the executive holds the role of CEO, whether the executive is interlocked, and the gender of the executive. The specification in column (2) also controls for age, age squared, and age cubic. Tenure is rescaled, dividing by 10, for the purpose of displaying coefficient scale properly. The standard errors clustered by state are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.*

and controlling for industry, year, and state fixed effects. It confirms the model prediction on wage dynamics over tenure. The positive coefficient for tenure and negative coefficient for tenure squared imply that wage increases non-linearly with tenure. Further, the positive coefficient for non-competition and negative coefficient for the interaction between tenure and non-competition imply that non-competition increases the starting wage and lowers wage growth over tenure.

I carry out a few robustness checks. First, column (2) also controls for age and the result is unchanged. I do not find significant age effect in fact.[35] Second, column (3) shows that, when controlling for firm fixed effect, the result is unchanged although smaller in magnitude. Third, the result in column (4) shows that the extent of wage-backloading decreases with the duration of non-competition. Finally, column (5) shows that, when using awarded compensation as wage measure, the tenure effect and its interaction with non-competition are much smaller.

To give a more clear sense of the magnitudes, Figure 2 plots compensation over tenure by whether the executive is subject to non-competition, according to the marginal effects at means in the baseline regression in column (1) of Table 3. First, an executive with non-competition is associated with a starting wage that is \$130k higher, or equivalently 15%. Second, an executive with non-competition is associated with an average wage growth over the first ten years of tenure that is 1.6 percentage points lower per annum. This number is the difference between 8.4 percentage points for executives without non-competition and 6.8 percentage points for executives with non-competition. The two wage-tenure lines cross at around tenure of five years. Figure 5 in Appendix C plots wage over tenure by the duration of non-competition, according to the marginal effects at means in the regression in column (4) of Table 3. An increase of one-year in non-competition duration is associated with a starting wage that is \$90k higher and an average wage growth over the first ten years of tenure that is 0.9 percentage points lower per annum.

The composition of compensation further confirms that wage is less backloaded when non-competition is used. Figure 6 in Appendix C shows that the fraction of compensation in the form of cash and deferred equity over tenure. Two patterns are noteworthy. First, the decrease in the fraction of cash compensation and increase in the fraction of deferred equity compensation over tenure is consist with firms using restricted equity – "golden handcuff" – to backload wage. Second, executives with non-competition have a higher fraction of compensation in cash and a lower fraction in deferred equity after tenure of three years.

[35]The executives in the sample have an average age of 51 and standard deviation of 7. The average age is similar between the executives with and without non-competition. This is the part of life cycle during which the age effect on earnings is found to be small.

Table 4: Parameter values

Parameter	Value	Calibration Target
Discount rate	$\rho = 0.05$	A priori information
Manager bargaining weight	$\beta = 0.5$	A priori information
Exogenous separation rate	$\delta = 0.07$	Separation hazard rate
Outside opportunity arrival rate	$\lambda = 0.09$	Separate hazard rate ratio
Outside opportunity dist. lower bound	$\theta_m = 0.65$	Wage growth over tenure
Outside opportunity dist. shape	$\alpha = 4$	Wage growth diff. over tenure
Standard deviation of Brownian motion	$\sigma = 0.24$	Pareto right tail
Investment cost function, level	$\phi = 82$	Investment rate
Investment cost function, elasticity	$\varphi = 2$	Investment response

This result points to non-competition lessening wage backloading by using more cash and less restricted equity.

7 Quantitative Analysis

This section quantitatively assesses optimal restrictions on non-competition using the calibrated model.

7.1 Model Calibration

I calibrate the model to match the following moments in the data: the effect of non-competition according to the reduced form estimates and other cross-sectional moments. The model is calibrated at annual frequency, with one unit of time in the model corresponding to one year in the data.

Two functional forms are specified. First, the distribution of outside match productivity is a Pareto distribution, $F(\theta) = 1 - \exp(\alpha(\theta_m - \theta))$, $\forall \theta \in [\theta_m, \infty)$. Second, the investment cost function is $c(\mu) = \frac{\phi}{1+1/\varphi}\mu^{1+\frac{1}{\varphi}}$, where φ represents the investment elasticity. The set of model parameters is $\{\rho,\ \beta, \delta, \lambda, \theta_m, \alpha, \sigma, \phi,\ \varphi\}$. The discount rate ρ is preset at 0.05 following the standard in the literature to match the interest rate. The manager's bargaining weight is 0.5. The remaining seven parameters are calibrated and their numerical values are displayed in Table 4. I discuss below in detail how the parameters are linked to the moments.

Outside opportunities. The separation rate for executives without non-competition reveals the arrival rate of competitive outside opportunities, $\delta + \lambda[1 - F(0)] = 8.7\%$. The response of separation rate to non-competition corresponds to the extent of excluded but

Table 5: Calibration targets

	Data	Model
Separation hazard rate (w.o. non-competition)	8.7%	8.6%
Separation hazard rate ratio (w./w.o. non-competition)	0.86	0.87
Investment rate (w.o. non-competition)	0.16	0.17
Investment response to non-competition	0.014	0.013
Avg wage growth, tenure 0-10 yrs (w.o. non-competition)	8.4%	8.5%
Avg wage growth, tenure 0-10 yrs (w.o.−w. non-competition)	1.6%	1.6%
Pareto right tail	1.16	1.16

otherwise competitive outside opportunities, $\left(\delta + \lambda\left[1 - F\left(\bar{\theta}\right)\right]\right) / \left(\delta + \lambda\left[1 - F\left(0\right)\right]\right) = 0.86$. The differences in wage-tenure profiles between executives with and without non-competition relates to the interval of the excluded but otherwise competitive outside opportunity, $F\left(\bar{z}\right) - F\left(0\right)$. The four moments are used to jointly calibrate the four parameters for exogenous separation rate, δ, arrival rate of outside opportunities, λ, and the two distribution parameters. Figure 3 shows the fit of the model generated wage-tenure profiles with the data.

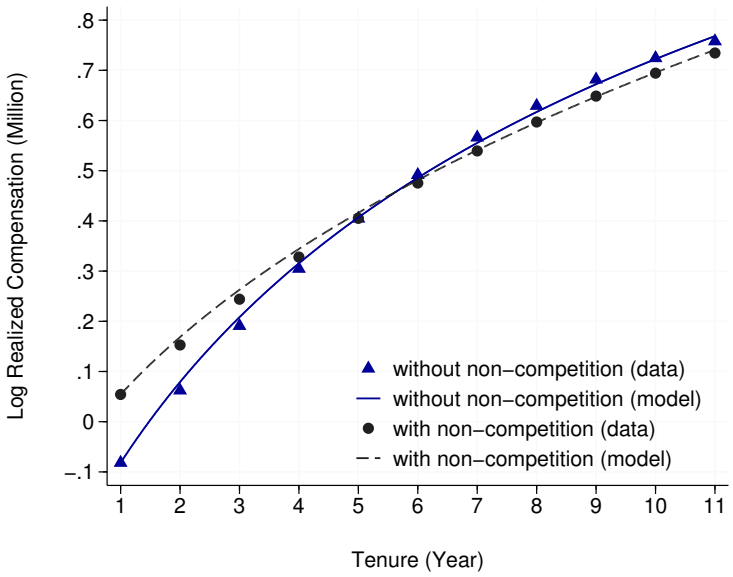
Productivity stochastic process. The first part of Proposition 2, equation (11) for the endogenous stationary productivity distribution, shows that the Pareto right tail index can reveal the standard deviation of the Brownian motion σ. This calibration strategy of relating the stochastic component of investment outcome to the cross sectional firm distribution follows Luttmer (2007) and Atkeson and Burstein (2010). Using Compustat data, I fit an empirical distribution of firm size measured in terms of employment in a given year. I obtain an average shape parameter of 1.16 during the period between 1992 and 2015. This moment implies a standard deviation of 0.24.

Investment cost function. Let $\mu^{\mathcal{M}}$ denote the level of investment with non-competition and $\mu^{\emptyset}$ the one without non-competition. The investment response to non-competition, at first-order log-linear approximation, is:

$$\log c\left(\mu^{\mathcal{M}}\right) - \log c\left(\mu^{\emptyset}\right) \approx \left(1 + \varphi\right) \frac{\lambda}{r - \mu^{\mathcal{M}} - \frac{1}{2}\sigma^2} \left(e^{\bar{\theta}} - 1\right) \left(1 - F\left(\bar{\theta}\right)\right). \tag{15}$$

The investment response equation (15) shows how investment expense responds to changes in marginal bilateral joint payoff. Given the other parameters, matching firm investment response in the data recovers the investment elasticity. Together matching the level parameter

Figure 3: Wage-backloading by non-competition, model prediction



of the cost function ϕ to firm investment rate, I obtain an investment elasticity of 2.

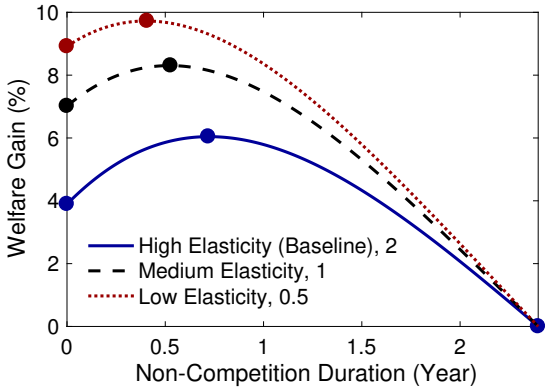
I conclude this section with a discussion of non-targeted moments. The average non-competition duration in the sample is around 1.6 years. The calibrated model suggests that the privately-optimal non-competition duration is 2.4 years, using the non-competition duration formula, $\pi = \frac{1}{r}\log\frac{\alpha}{\alpha-1}$, in Proposition 1. The corresponding buyout payment is according to the formula, $\tau(z) = \frac{1}{\alpha-1}J(z)$, is equivalent to around 10 times of the manager's starting wage, or 12 millions in 2010 dollars on average. As a sanity check, I cross examine the number with the amount of buyout payment in a few non-competition cases listed in Table 9 and find it lies in a reasonable range.

7.2 Policy Evaluation

This section quantitatively assesses the optimal restriction on non-competition in Proposition 4. I first carry out the assessment using the baseline calibrated model and then discuss the sensitivity of the results. Figure 4 plots the welfare gains at a range of non-competition duration from zero to the privately-optimal level.

The calibrated baseline model suggests that the optimal cap on non-competition duration is 0.7 years, which is only 30% of the privately-optimal level. To put the number in perspective, it is in general considered easy in many states to enforce a non-competition if the duration is not over two years. Additionally, a legislative bill in Massachusetts in 2017

Figure 4: Welfare gain from capping non-competition duration



proposes to restrict the non-competition duration to a maximum of one year.[36] The optimal restriction and a ban on non-competition result in steady-state welfare gains of 6% and 4%, respectively, relative to the laissez-faire outcome.

Investment elasticity. The welfare gains from restricting non-competition depends on investment elasticity, that is, how responsive investment is to changes in the marginal bilateral joint value. If investment is highly elastic, the investment holdup is severe and the benefits from alleviating holdup are large. Therefore, reducing the cap on non-competition duration does incur more loss associated with lower investment, leading to a higher optimal cap on non-competition duration.

I note that the investment elasticity parameter in the baseline calibration, $\varphi = 2$, is at the higher end of the range in the literature which center around a unity. Hence, my welfare calculation is conservative. To see exactly how sensitive the welfare gains are, I fix the investment elasticity at medium ($\varphi = 1$) and low ($\varphi = .5$) levels and re-calibrate the model to match the data moments except the investment response. Unsurprisingly, the medium and low levels of investment elasticities imply an even lower optimal cap on non-competition duration and larger welfare gains.

8 Conclusion

In this paper, I studied the aggregate effect of non-competition contracts in the managerial labor market, considering the beneficial effects of encouraging firm investment and harmful effects of restricting labor mobility. I developed a dynamic contracting model in which non-competition is used by initial employing firms to extract rent from outside firms. The

[36]The details can be found at https://malegislature.gov/Bills/190/SD1578.

model captures the tradeoff between alleviating investment holdup and distorting manager allocation. Empirical evidence from new contract-level data supported the model's predictions. I assessed the model's welfare implications quantitatively, reaching the conclusion that capping non-competition duration, close to banning non-competition completely, is socially optimal.

There are other potential channels of non-competition contracts that I abstract away from. One channel is risk-sharing between managers and firms, which is absent due to the risk-neutral assumption in my model. Non-competition contracts can improve risk-sharing by restricting managers' outside opportunities. Another channel is the agglomeration effects of industry clusters. Non-competition contracts prevent the formation of industry clusters by limiting technology spillover and discouraging entrepreneurship.[37] Incorporating these additional channels in future work are useful. The agglomeration channel would further reinforce the conclusion reached here, while the risk-sharing channel would attenuate it.

Finally, non-competition contracts have permeated into broader labor markets beyond the managerial one. A survey conducted by Prescott et al. (2016) indicates that about 30 million or, equivalently, 18% of U.S. workers are subject to such contracts. The quantitative evaluation focused on the managerial labor market in this paper has relevance in a broader context, particularly the high-skilled segments of labor markets in which the same economic forces of similar magnitudes operate.

[37] Gilson (1999) suggests that the ban on non-competition in California was conducive to the emergence of Silicon Valley and its surpassing of Boston's Route 128 tech district. Several studies point to the importance of employee mobility in the formation of industry clusters through job-hopping between firms and employee spin-offs into entrepreneurship (e.g., Franco and Filson, 2006) and the adverse effects non-competition contracts bring about (Franco and Mitchell, 2008; Samila and Sorenson, 2011).

References

Acemoglu, Daron, "Training and Innovation in an Imperfect Labour Market," *The Review of Economic Studies*, 1997, *64* (3), 445–464.

_ **and Jorn Steffen Pischke**, "The Structure of Wages and Investment in General Training," *Journal of Political Economy*, 1999, *107* (3), 539–572.

Aghion, Philippe and Patrick Bolton, "Contracts as a Barrier to Entry," *American Economic Review*, 1987, *77* (3), 388–401.

Anderson, Eric T. and James D. Dana, "When Is Price Discrimination Profitable?," *Management Science*, 2009, *55* (6), 980–989.

Atkeson, Andrew and Ariel Tomas Burstein, "Innovation, Firm Dynamics, and International Trade," *Journal of Political Economy*, 2010, *118* (3), 433–484.

Bishara, Norman D., "Fifty Ways to Leave Your Employer: Relative Enforcement of Covenants Not to Compete, Trends, and Implications for Employee Mobility Policy," *University of Pennsylvania Journal of Business Law*, 2011, *13* (3), 751–795.

_ **, Kenneth J. Martin, and Randall S. Thomas**, "An Empirical Analysis of Noncompetition Clauses and Other Restrictive Postemployment Covenants," *Vanderbilt Law Review*, 2015, *68.*

Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin, "Wage Bargaining with On-the-Job Search: Theory and Evidence," *Econometrica*, 2006, *74* (2), 323–364.

Clementi, Gian Luca, Thomas F. Cooley, and Cheng Wang, "Stock Grants as a Commitment Device," *Journal of Economic Dynamics and Control*, 2006, *30* (11), 2191 – 2216.

Davis, Steven J. and John Haltiwanger, "Labor Market Fluidity and Economic Performance," September 2014, (20479).

Decker, Ryan A., John Haltiwanger, Ron S. Jarmin, and Javier Miranda, "Declining Business Dynamism: What We Know and the Way Forward," *American Economic Review*, 2016, *106* (5), 203–07.

Deneckere, Raymond J. and R. Preston McAfee, "Damaged Goods," *Journal of Economics and Management Strategy*, 1996, *5* (2), 149–174.

Diamond, Peter A. and Eric Maskin, "An Equilibrium Analysis of Search and Breach of Contract, I: Steady States," *The Bell Journal of Economics*, 1979, *10* (1), 282–316.

Fella, Giulio, "Termination Restrictions and Investment in General Training," *European Economic Review*, 2005, *49* (6), 1479 – 1499.

Franco, April M. and Matthew F. Mitchell, "Covenants Not to Compete, Labor Mobility, and Industry Dynamics," *Journal of Economics & Management Strategy*, 2008, *17* (3), 581–606.

Franco, April Mitchell and Darren Filson, "Spin-Outs: Knowledge Diffusion Through Employee Mobility," *The RAND Journal of Economics*, 2006, *37* (4), 841–860.

Frydman, Carola, "Rising Through the Ranks. The Evolution of the Market for Corporate Executives, 1936-2003," *Management Science*, Forthcoming.

_ **and Raven E. Saks**, "Executive Compensation: A New View from a Long-Term Perspective, 1936-2005," *The Review of Financial Studies*, 2010, *23* (5), 2099–2138.

Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "The Dynamics of Inequality," *Econometrica*, 2016, *84* (6), 2071–2111.

Garmaise, Mark J., "Ties that Truly Bind: Noncompetition Agreements, Executive Compensation, and Firm Investment," *Journal of Law, Economics, and Organization*, 2009.

Gillan, Stuart L., Jay C. Hartzell, and Robert Parrino, "Explicit versus Implicit Contracts: Evidence from CEO Employment Agreements," *The Journal of Finance*, 2009, *64* (4), 1629–1655.

Gilson, Ronald J., "The Legal Infrastructure of High Technology Industrial Districts: Silicon Valley, Route 128, and Covenants Not to Compete.," *New York University Law Review*, 1999, *74*, 575–629.

Gopalan, Radhakrishnan, Todd Milbourn, Fenghua Song, and Anjan V. Thakor, "Duration of Executive Compensation," *The Journal of Finance*, 2014, *69* (6), 2777–2817.

Heggedal, Tom-Reiel, Espen R. Moen, and Edgar Preugschat, "Productivity Spillovers through Labor Mobility in Search Equilibrium," *Journal of Economic Theory*, 2017, *169* (Supplement C), 551 – 602.

Hopenhayn, Hugo and Liyan Shi, "Knowledge Creation and Diffusion with Limited Appropriation," *Working Paper*, 2017.

Hosios, Arthur J., "On the Efficiency of Matching and Related Models of Search and Unemployment," *The Review of Economic Studies*, 1990, *57* (2), 279–298.

Jeffers, Jessica S., "The Impact of Restricting Labor Mobility on Corporate Investment and Entrepreneurship," *Working Paper*, 2017.

Lavetti, Kurt, Carol Simon, and William D. White, "The Impacts of Restricting Mobility of Skilled Service Workers: Evidence from Physicians," *Working Paper*, 2017.

Lentz, Rasmus and Nicolas Roys, "Training and Search On the Job," *Working Paper*, 2015.

Lucas, Robert E. and Benjamin Moll, "Knowledge Growth and the Allocation of Time," *Journal of Political Economy*, 2014, *122* (1), 1–51.

Luttmer, Erzo G. J., "Selection, Growth, and the Size Distribution of Firms," *The Quarterly Journal of Economics*, 2007, *122* (3), 1103–1144.

Marx, Matt, Deborah Strumsky, and Lee Fleming, "Mobility, Skills, and the Michigan Non-Compete Experiment," *Management Science*, 2009, *55* (6), 875–889.

Moen, Espen R. and Asa Rosen, "Does Poaching Distort Training?," *The Review of Economic Studies*, 2004, *71* (4), 1143–1162.

Naidu, Suresh, "Recruitment Restrictions and Labor Markets: Evidence from the Postbellum U.S. South," *Journal of Labor Economics*, 2010, *28* (2), 413–445.

Nordhaus, William D., "The Optimal Life of a Patent," Cowles Foundation Discussion Papers 241, Cowles Foundation for Research in Economics, Yale University 1967.

Nuno, Galo and Benjamin Moll, "Social Optima in Economies with Heterogeneous Agents," *Review of Economic Dynamics*, 2017.

Peters, Ryan H. and Lucian A. Taylor, "Intangible Capital and the Investment-Q Relation," *Journal of Financial Economics*, 2017, *123* (2), 251 – 272.

Postel-Vinay, Fabien and Helene Turon, "On-the-Job Search, Productivity Shocks, and the Individual Earnings Process," *International Economic Review*, 2010, *51* (3), 599–629.

_ **and Jean-Marc Robin**, "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 2002, *70* (6), 2295–2350.

Prescott, J.J., Norman Bishara, and Evan P Starr, "Understanding Noncompetition Agreements: The 2014 Noncompete Survey Project," *Michigan State Law Review*, 2016, *2*, 369–464.

Rauch, James E., "Dynastic Entrepreneurship, Entry, and Non-Compete Enforcement," *European Economic Review*, 2016, *86* (Supplement C), 188 – 201.

Romer, Paul M., "Endogenous Technological Change," *Journal of Political Economy*, 1990, *98* (5), S71–S102.

Samila, Sampsa and Olav Sorenson, "Noncompete Covenants: Incentives to Innovate or Impediments to Growth," *Management Science*, 2011, *57* (3), 425–438.

Segal, Ilya R. and Michael D. Whinston, "Exclusive Contracts and Protection of Investments," *The RAND Journal of Economics*, 2000, *31* (4), 603–633.

Spier, Kathryn E. and Michael D. Whinston, "On the Efficiency of Privately Stipulated Damages for Breach of Contract: Entry Barriers, Reliance, and Renegotiation," *The RAND Journal of Economics*, 1995, *26* (2), 180–202.

Starr, Evan P, "Consider This: Training, Wages and the Enforceability of Covenants Not to Compete," *Working Paper*, 2016.

Taylor, Lucian A., "Why Are CEOs Rarely Fired? Evidence from Structural Estimation," *The Journal of Finance*, 2010, *65* (6), 2051–2087.

The Department of the Treasury, "Non-Compete Contracts: Economic Effects and Policy Implications," 2016. https://www.treasury.gov/resource-center/economic-policy/Documents/UST%20Non-competes%20Report.pdf.

The White House, "Non-Compete Agreements: Analysis of the Usage, Potential Issues, and State Responses," 2016. https://obamawhitehouse.archives.gov/sites/default/files/non-competes_report_final2.pdf.

Appendices

A A One-Period Example

This section provides a simple one-period example which encapsulates the essential features of the full model in Section 2 to illustrate the key insights. In particular, it supplements the full model with the details of the game played by the agents.

Technology. The economy lasts for one period. The events occur as follows. There is a manager matched to an initial firm. They can undertake investment in their match productivity $z \in \mathcal{Z}$ at a cost of $c(z)$. Employment opportunity with outside firms for the manager arrives with probability λ. The outside match productivity $z' = z\theta \in \mathcal{Z}'$, where the relative productivity $\theta \in \Theta = [\theta_m, \infty)$ is drawn according to the cumulative distribution function $F(\cdot)$. Employment and production take place. The agents are risk neutral. The usual assumptions in Section 2 are retained.

Definition 2 (Allocation)**.** An allocation $\left(\bar{\theta}, z\right)$ consists of: (i) the poaching threshold $\bar{\theta}$ such that manager to be employed at the outside firm with relative productivity $\theta > \bar{\theta}$; and (ii) the level of investment z.

Information structure. Information is asymmetric: the firms do not observe each other's productivity. Given the information constraint, the maximum payoff the initial firm and the manager can jointly achieve is by charging the outside firm a monopoly price to poach away the manager. The contract with non-competition buyout and wage-bidding below implements the monopoly price.

Contract. The initial firm and the manager enter into a contract ex-ante. The firm can commit to the contract, which delivers a promised level of utility U_0 to the manager. The contract is publicly observable and includes the following terms: (i) the default wage payment from the firm to the manager w_0; (ii) the firm's wage bidding strategy when in competition with the outside firm, $w : \mathcal{Z} \to \mathbb{R}_+$. The firm has limited liability in delivering wage payment. Hence, the maximum wage bidding that it can commit to equals its productivity. That is, $w(z) \in [w_0, z]$; and (iii) the non-competition clause which prevents the manager from working at the outside firm for $\pi \in [0, 1]$ fraction of production time, reducing the outside match value to $(1 - \pi)$ fraction. To summarize, the contract is denoted by $\mathcal{C} = \{w_0, w, \pi\}$.

The firms and the manager play a two-stage game to determine the manager's employment before production takes place: (i) bidding for the manager in an ascending (English)

auction; and (ii) buyout of non-competition. The details of the game is as follows:

i. *Bidding*: The initial and outside firms bid for the manager in an ascending (English) auction: the wage is raised continuously from the current level w_0 until one firm drops out. The wages at which the initial and outside firms would drop out are denoted by w and $w' : \mathcal{Z}' \to \mathbb{R}_+$, respectively. If the initial firm drops out first, the manager moves to the outside firm. A second stage ensues.

ii. *Buyout*: The initial firm chooses the buyout price $\tau : \mathcal{Z} \to \mathbb{R}_+$.[38] The outside firm chooses whether to buyout the non-competition.

The initial firm and outside firm's prior beliefs about each other's productivity are denoted by $G(z'|z) = F(z'/z)$ and $G'(z|z') = F(z/z')$. If the manager is poached by the outside firm, the initial firm updates its posterior belief of the outside firm's productivity, denoted by $P(z'|w'(z') > w(z))$; the outside firm learns perfectly the initial firm's productivity.

Equilibrium definition. In the bidding stage, the expected payoffs for the initial firm, the manager, and the outside firm are, respectively:

$$V(\pi, w, w', \tau|z) = \int \left[(z - w_0)\mathbf{1}_{\{w'(z')<w_0\}} + (z - w'(z'))\mathbf{1}_{\{w_0<w'(z')<w(z)\}} + \tau(z)\mathbf{1}_{\{w'(z')\geq w(z)\cap\tau(z)\leq\pi z'\}}\right] dG(z'|z), \tag{16}$$

$$U(\pi, w, w', \tau|z) = \int \left[w_0\mathbf{1}_{\{w'(z')<w_0\}} + w'(z')\mathbf{1}_{\{w_0<w'(z')<w(z)\}} + w(z)\mathbf{1}_{\{w'(z')\geq w(z)\}}\right] dG(z'|z), \tag{17}$$

$$V'(\pi, w, w', \tau|z') = \int \left[\max\{(1-\pi)z', z' - \tau(z)\} - w(z)\right]\mathbf{1}_{\{w'(z')\geq w(z)\}} dG'(z|z'). \tag{18}$$

In the buyout stage, the initial firm's expected payoff is:

$$\chi(\pi, w, w', \tau|z) = \int \tau(z)\mathbf{1}_{\{\tau(z)\leq\pi z'\}} dP(z'|\{w'(z') \geq w(z)\}). \tag{19}$$

These payoff functions have taken into account the outside firm's optimal buyout decision.

Definition 3 (Perfect Bayesian Equilibrium)**.** A Perfect Bayesian Equilibrium consists of strategies $\{w_0^*, w^*, \pi\}$, z^*, w'^*, and τ^* and posterior belief P such that:

[38] It is without loss of generality to restrict the buyout price to one as a function of the initial firm's productivity only. This is because the initial firm cannot improve its payoff by price discriminating towards outside firms.

- the contract and investment is optimal:

$$\{w_0^*, w^*, \pi^*\}, z^* = \underset{w_0, w(z)\in[w_0,z],\pi,z}{\operatorname{argmax}} -c(z) + V(\pi, w, w'^*, \tau^* | z) \tag{20}$$

subject to the promise keeping constraint

$$U(\pi, w, w'^*, \tau^* | z) = U_0; \tag{21}$$

- the outside firm's bid is optimal in the bidding stage:

$$w'^* \in \underset{w'}{\operatorname{argmax}} V'(\pi^*, w^*, w', \tau^* | z'), \ \forall z'; \tag{22}$$

- the initial firm's buyout price is optimal in the buyout stage:

$$\tau^* \in \underset{\tau}{\operatorname{argmax}} \chi(\pi^*, w^*, w'^*, \tau | z), \forall z; \tag{23}$$

- the initial firm's posterior belief is updated according to:

$$P(z' | w'^*(z') \geq w^*(z)) = \frac{\int_{z\theta_m}^{z'} \mathbf{1}_{\{w'^*(\tilde{z}')\geq w^*(z)\}} dG_1(\tilde{z}'|z)}{\int_{z\theta_m}^{\infty} \mathbf{1}_{\{w'^*(z')\geq w^*(z)\}} dG_1(z'|z)}. \tag{24}$$

Bilateral efficiency. Given the assumptions of firm commitment and agent risk neutrality, the bilateral efficiency result applies here. This result is easily obtained by integrating the promise keeping constraint in equation (21) to the initial firm's objective function in equation (20), leading to the bilateral joint payoff:

$$\begin{aligned} J(\pi, w, w'^*, \tau^* | z) &\equiv V(\pi, w, w'^*, \tau^* | z) + U(\pi, w, w'^*, \tau^* | z) \\ &= \int \left[z\mathbf{1}_{w'(z')<w(z)} + \left(\tau(z)\mathbf{1}_{\{\tau(z)\leq \pi z'\}} + w(z)\right)\mathbf{1}_{\{w'(z')\geq w(z)\}} \right] dG(z' | z). \end{aligned} \tag{25}$$

Lemma 3 (Bilateral efficiency)**.** *The contract maximizes their bilateral joint value for the initial firm and the manager in the bidding stage:*

$$\{w_0^*, w^*, \pi^*\} = \underset{w_0, w(z)\in[w_0,z],\pi}{\operatorname{argmax}} J(\pi, w, w'^*, \tau^* | z).$$

Equilibrium contract. Building on Lemma (3), I now solve for the equilibrium.

Proposition 6. *Under Assumption 1, the privately-optimal allocation is characterized by*

$$\bar{\theta} - \frac{1 - F\left(\bar{\theta}\right)}{f\left(\bar{\theta}\right)} = 1 \tag{26}$$

$$c'\left(z\right) = 1 + \left(\bar{\theta} - 1\right)\left(1 - F\left(\bar{\theta}\right)\right); \tag{27}$$

it is implemented by the contract with non-competition duration and wage bidding:

$$\pi^* = 1 - \frac{1}{\bar{\theta}} \text{ and } w^*\left(z\right) = z.$$

Proof. To start with, the bidding strategy $w'^*\left(z'\right)$ is strictly increasing in z'. Therefore, there exists a unique poaching threshold $\bar{z} = z\bar{\theta}$ such that $w'^*\left(z\bar{\theta}\right) = w^*\left(z\right)$, $\forall z$. Performing a change of variable from z' to θ, the Bayes rule for posterior belief in equation (24) simplifies to:

$$P\left(\theta | \, \theta \geq \bar{\theta}\right) = \frac{F\left(\theta\right) - F\left(\bar{\theta}\right)}{1 - F\left(\bar{\theta}\right)}, \forall \theta \geq \bar{\theta}. \tag{28}$$

Given the posterior belief in equation (28), the initial firm's problem of choosing buyout price $\tau\left(z\right)$ in equation (23) is equivalent to choosing a buyout threshold θ_0, which satisfies $\tau\left(z\right) = \pi z \theta_0$. That is,

$$\theta_0 = \underset{\theta \geq \bar{\theta}}{\text{argmax}} \left(1 - F\left(\theta\right)\right)\theta.$$

I can restrict the buyout threshold to at least above the poaching threshold $\theta_0 \geq \bar{\theta}$. This is because the initial firm can at least charge a buyout price such that the outside firm at the poaching threshold has binding outside option, which is to have the non-competition enforced. Accordingly, the outside firm's bid at the poaching threshold is $w'^*\left(z\bar{\theta}\right) = \left(1 - \pi^*\right) z\bar{\theta}$.

Combining the outside firm's bidding strategy and initial firm's buyout price strategy, the bilateral joint value in equation (25) is:

$$F\left(\bar{\theta}\right) z + \left(1 - F\left(\bar{\theta}\right)\right)\left(1 - \pi\right) z\bar{\theta} + \left(1 - F\left(\theta_0\right)\right)\pi z \theta_0 \leq \left[F\left(\theta_0\right) + \left(1 - F\left(\theta_0\right)\right)\theta_0\right] z. \tag{29}$$

The inequality in expression (29) follows from two relations: $\left(1 - F\left(\bar{\theta}\right)\right)\bar{\theta} \leq \left(1 - F\left(\theta_0\right)\right)\theta_0$ and $F\left(\bar{\theta}\right) \leq F\left(\theta_0\right)$. The maximum of the right-hand side is obtained when $\theta_0 = \bar{\theta}$, where $\bar{\theta}$ satisfies equation (26). This outcome can be ensured by contracting the initial firm's bidding strategy as $w^*\left(z\right) = \left(1 - \pi^*\right) z\bar{\theta}$. The limited liability constraint for the initial firm, $\left(1 - \pi^*\right)\bar{\theta} \leq 1$, implies a minimum duration of non-competition at $1 - 1/\bar{\theta}$.

There exists a continuum of Perfect Bayesian Equilibria which all achieve the same privately allocation and payoff, indexed by the level of non-competition duration, $\pi^* \in$

$\left[1-\frac{1}{\bar{\theta}}, 1\right]$. The corresponding wage bidding and buyout price strategies are

$$w^*(z) = (1-\pi^*) z\bar{\theta};\ w'^*(z') = (1-\pi^*) z', \forall z' \leq z\bar{\theta};\ \tau^*(z) = \pi^* z\bar{\theta}.$$

I select the one with the minimum duration of non-competition and maximum wage bidding, given that non-competition relies on external enforcement.

Finally, investment is such that marginal cost equals marginal bilateral payoff as in equation (27). □

The implementation through non-competition buyout above shows that it is not necessary to contract on the amount of buyout payment τ ex ante. The result also holds when the payment is determined through ex post bargaining.

Discussion of firm commitment. I conclude the one-period example with a brief discussion of the role of firm commitment. It is not essential that the initial firm commits to the buyout price. In other words, the contract does not need to explicitly specify the buyout price ex ante.

However, it is essential that the firm commits to wage bidding, which ensures bilateral efficiency. If, to the contrary, the initial firm cannot commit to wage bidding, it maximizes its own payoff when bidding:

$$\max_{\tilde{\theta}} \int_0^{\tilde{\theta}} \left[1-(1-\pi)\theta\right] dF(\theta) + \left[1-F\left(\tilde{\theta}\right)\right] \pi\tilde{\theta}.$$

The optimality condition with respect to the poaching threshold $\tilde{\theta}$ implies that $\tilde{\theta} - \pi\left(1-F\left(\tilde{\theta}\right)\right)/f\left(\tilde{\theta}\right) = 1$. Comparing with equation (26), it implies that, as long as $\pi < 1$, the poaching threshold is below the one in the commitment case, i.e. $\tilde{\theta} < \bar{\theta}$. Since the contract no longer achieves bilateral efficiency, there is disagreement over the non-competition duration: the firm prefers a longer duration while the manager prefers a shorter one.

B Derivations and Proofs

B.1 Proof of Lemma 1

Step 1. Original formulation of contracting problem

The contract $\mathcal{C}$ must deliver the initial level of promised utility to the manager according to:

$$U_0(\boldsymbol{z}, \boldsymbol{\theta}, \mathcal{C}) = \mathbb{E}\left[\int_0^T e^{-rt} w_t dt + e^{-rT} J(z_T) \middle| \mathcal{F}_0\right] \geq U_0, \tag{30}$$

The initial firm derives utility according to:

$$V_0(\boldsymbol{z}, \boldsymbol{\theta}, \mathcal{C}) = \mathbb{E}\left[\int_0^T e^{-rt}\left(e^{z_t} - c(\mu_t) e^{z_t} - w_t\right) dt + e^{-rT} \tau_T\left(\tilde{\pi}_T(\theta_T)\right) \middle| \mathcal{F}_0\right] \tag{31}$$

The initial firm's contracting problem is:

$$\max_{\mathcal{C}, \boldsymbol{\mu}} V_0(\boldsymbol{z}, \boldsymbol{\theta}, \mathcal{C})$$

subject to the (IC), (IR) and (30) constraints.

Equation (30) and (31) can be re-written as:

$$U_0(\boldsymbol{z}, \boldsymbol{\theta}, \mathcal{C}) = \mathbb{E}\left[\int_0^\infty R(0,t)\left(w_t + \lambda \int_{\bar{\theta}_t} J(z_t) dF(\theta_t)\right) dt \middle| \mathcal{F}_0\right] \geq U_0,$$
$$V_0(\boldsymbol{z}, \boldsymbol{\theta}, \mathcal{C}) = \mathbb{E}\left[\int_0^\infty R(0,t)\left(e^{z_t} - c(\mu_t) e^{z_t} - w_t + \lambda \int_{\bar{\theta}_t} \tau_t\left(\tilde{\pi}_t(\theta_t)\right) dF(\theta_t)\right) dt \middle| \mathcal{F}_0\right],$$

where the effective discounting, adjusted for job-to-job transition, is:

$$R(0,t) = \exp\left(-\int_0^t \left(r + p_{t'}\right) dt'\right), \text{ where } p_t = \lambda\left[1 - F\left(\bar{\theta}_t\right)\right].$$

Step 2. Dynamic programming

The manager's continuation utility at time t is given by:

$$U_t(\boldsymbol{z}, \mathcal{C}) = \mathbb{E}\left[\int_t^\infty R(t,s)\left(w_s + \lambda \int_{\bar{\theta}_s} J(z_s) dF(\theta_s)\right) ds \middle| \mathcal{F}_t\right].$$

By the Martingale Representation Theorem, there exists a process $\left\{\sigma_t^U\right\}_{t\geq 0}$ such that $\{U_t\}_{t\geq 0}$ satisfies the following stochastic differential equation:

$$dU_t = \left((r+p_t)\, U_t - w_t - \lambda \int_{\bar{\theta}_t} J(z_t)\, dF(\theta_t) \right) dt + \sigma_t^U dB_t. \tag{32}$$

The initial firm's continuation utility at time t is given by:

$$V_t(\boldsymbol{z}, \boldsymbol{\theta}, \mathcal{C}) = \mathbb{E}\left[\int_t^\infty R(t,s) \left(e^{z_s} - c(\mu_s)\, e^{z_s} - w_s + p_s \int_{\bar{\theta}_s} \tau_s(\tilde{\pi}_s(\theta_s))\, dF(\theta_s) \right) ds \middle| \mathcal{F}_t \right].$$

I apply dynamic programming to the initial firm's contracting problem and use the manager's continuation value as a state variable:

$$V(z_t, U_t) = \max_{\mathcal{C}, \boldsymbol{\mu}} V_t(\boldsymbol{z}, \boldsymbol{\theta}, \mathcal{C}).$$

The initial firm's contracting problem can be written as the following HJB equation:

$$\begin{aligned}
(r+p_t)\, V(z_t, U_t) = \max_{(w_t, \pi_t, \tau_t(\cdot)), \mu_t, \sigma_t^U} \; & e^{z_t} - c(\mu_t)\, e^{z_t} - w_t + p_t \int_{\bar{\theta}_t} \tau_t(\tilde{\pi}_t(\theta_t))\, dF(\theta_t) \\
& + \mu_t V_z(z_t, U_t) + \frac{1}{2}\sigma^2 V_{zz}(z_t, U_t) \\
& + V_U(z_t, U_t) \left((r+p_t)\, U_t - w_t - \lambda \int_{\bar{\theta}_t} J(z_t)\, dF(\theta_t) \right) \\
& + \frac{1}{2}\sigma_t^U V_{UU}(z_t, U_t) + \sigma \sigma_t^U V_{zU}(z_t, U_t)
\end{aligned} \tag{33}$$

subject to the (IC) and (IR) constraints.

Step 3. Bilateral efficiency

Taking derivative with respect to w_t, I obtain:

$$V_U(z_t, U_t) \geq -1 \text{ with "=" if } w_t > 0. \tag{34}$$

If λ is sufficiently small, according to equation (32), the wage non-negative constraint will never bind. That is, equation (34) becomes $V_U(z_t, U_t) = -1$. This in turn implies that:

$$V_{UU}(z_t, U_t) = 0 \text{ and } V_{zU}(z_t, U_t) = 0. \tag{35}$$

Since $J(z_t) = V(z_t, U_t) - V_U(z_t, U_t) U_t$, I also obtain:

$$J(z_t) = V(z_t, U_t) + U_t. \tag{36}$$

Substituting equation (34), (35) and (36) into the HJB equation (33):

$$rJ(z_t) = \max_{(\pi_t, \tau_t(\cdot)), \mu_t} e^{z_t} - c(\mu_t) e^{z_t} + \lambda \int_{\bar{\theta}_t} \tau_t(\tilde{\pi}_t(\theta_t)) dF(\theta_t) + \mu_t J'(z_t) + \frac{1}{2}\sigma^2 J''(z_t).$$

B.2 Proof of Lemma 2, Proposition 1, and Corollary 1

Proof. The contract design of the additional clauses is reduced to maximizing the rent extracted from outside firms:

$$\chi(z) = \max_{\mathcal{M}} \int_{\bar{\theta}}^{\infty} \tau(z, \tilde{\pi}(z, \theta)) dF(\theta)$$

subject to the (IC) and (IR) constraints.

Taking advantage of the envelope condition for the (IC) constraint, together with the binding (IR) constraint at the poaching threshold, the buyout payment satisfies:

$$\tau(z, \tilde{\pi}(z, \theta)) = e^{-r\tilde{\pi}(z,\theta)} J(z + \theta) - \int_{\bar{\theta}}^{\theta} e^{-r\tilde{\pi}(z,\hat{\theta})} J'\left(z + \hat{\theta}\right) d\hat{\theta} - J(z). \tag{37}$$

The problem of maximizing the rent extracted becomes:

$$\begin{aligned} &\max_{\mathcal{M}} \int_{\bar{\theta}}^{\infty} \left[e^{-r\tilde{\pi}(z,\theta)} J(z + \theta) - \int_{\bar{\theta}}^{\theta} e^{-r\tilde{\pi}(z,\hat{\theta})} J'\left(z + \hat{\theta}\right) d\hat{\theta} - J(z) \right] F(\theta) \\ =&\max_{\mathcal{M}} \int_{\bar{\theta}}^{\infty} \left[e^{-r\tilde{\pi}(z,\theta)} J(z + \theta) - \frac{1 - F(\theta)}{f(\theta)} e^{-r\tilde{\pi}(z,\theta)} J'(z + \theta) - J(z) \right] F(\theta). \end{aligned}$$

The first order condition with respect to $\tilde{\pi}(z, \theta)$ is:

$$e^{-r\tilde{\pi}(z,\theta)} \left[J(z + \theta) - \frac{1 - F(\theta)}{f(\theta)} J'(z + \theta) \right] \geq 0 \text{ with } \text{“}=\text{”} \text{ if } \tilde{\pi}(z, \theta) > 0, \ \forall \theta \geq \bar{\theta}. \tag{38}$$

It implies that the equilibrium non-competition buyout is:

$$\tilde{\pi}(z, \theta) = 0, \ \forall \theta \geq \bar{\theta}. \tag{39}$$

That is, all outside firms that are competitive enough to poach the manager fully buyout the non-competition. The buyout payment obtained by replacing the equilibrium non-

competition buyout (39) into equation (37) is:

$$\tau(z,\pi) = J\left(z+\bar{\theta}\right) - J(z).$$

The first order condition with respect to $\bar{\theta}$ is:

$$e^{-r\tilde{\pi}\left(z,\bar{\theta}\right)} J\left(z+\bar{\theta}\right) - \frac{1-F\left(\bar{\theta}\right)}{f\left(\bar{\theta}\right)} e^{-r\tilde{\pi}\left(z,\bar{\theta}\right)} J'\left(z+\bar{\theta}\right) - J(z) = 0. \tag{40}$$

Replacing the equilibrium non-competition buyout (39) into in equation (40), I obtain:

$$L\left(\bar{\theta};z\right) \coloneqq J\left(z+\bar{\theta}\right) - \frac{1-F\left(\bar{\theta}\right)}{f\left(\bar{\theta}\right)} J'\left(z+\bar{\theta}\right) - J(z) = 0.$$

Assumption 1 guarantees that $L\left(\bar{\theta};z\right)$ is strictly increasing in $\bar{\theta}$ given z. Additionally, $L(0;z) < 0$ and $L(\infty;z) > 0$. There exists a unique solution $\bar{\theta} > 0$.

Finally, the poaching threshold also satisfies:

$$e^{-r\pi} J\left(z+\bar{\theta}\right) = J(z),$$

which implies that the non-competition duration is:

$$\pi = \frac{1}{r}\left[\log J\left(z+\bar{\theta}\right) - \log J(z)\right].$$

I guess and verify that the bilateral joint value function is linear in e^z, i.e. $J(z) = je^z$. The linear guess implies three results. First, the equation characterizing the poaching threshold reduces to:

$$L\left(\bar{\theta}\right) \coloneqq e^{\bar{\theta}}\left[1 - \frac{1-F\left(\bar{\theta}\right)}{f\left(\bar{\theta}\right)}\right] - 1 = 0.$$

The poaching threshold is a constant markup, $\bar{\theta} > 0$, of the initial firm's productivity. Second, the buyout payment is proportional productivity, $\tau(z,\pi) = je^z\left(e^{\bar{\theta}} - 1\right)$. Finally, the investment decision reduces to $\mu = (c')^{-1}(j)$, which is independent of productivity. Combining the three results above and replacing them in the HJB equation, I obtain equation (7) which the marginal bilateral joint value j satisfies.

Having obtained that $J(z) = je^z$, it immediately follows that $\pi = \frac{\bar{\theta}}{r}$. □

B.3 Derivation of Simplified Wage Process

To study properties of wage dynamics over tenure, I re-write the manager's HJB equation (8) by simplifying the state space. The manager's value function can be reduced to one with a single state variable – the log wage-productivity ratio, $x = \log\left(\frac{w}{e^z}\right)$. I define the rescaled value function, $U(z,w) \equiv u(x)\, e^z$. The derivatives of the value functions have the following relations:

$$U_z(z,w) = [u(x) - u'(x)]\, e^z \tag{41}$$

$$U_{zz}(z,w) = [u(x) - 2u'(x) + u''(x)]\, e^z. \tag{42}$$

The wage bidding threshold $\underline{\theta}(z,w)$ reduces to $\underline{\theta}(x)$, which satisfies:

$$u(x) = e^{-r\pi + \underline{\theta}(x)} j. \tag{43}$$

Substituting equation (41), (42) and (43) into the original HJB equation (8), dividing both sides by e^z, and re-arranging:

$$\begin{aligned}\left(r - \mu - \frac{1}{2}\sigma^2 + \lambda\right) u(x) =& e^w + \lambda\left\{F(\underline{\theta}(x))\, u(x) + j e^{-r\pi}\int_{\underline{\theta}(x)}^{\bar{\theta}} e^{\theta} dF(\theta) + \left[1 - F\left(\bar{\theta}\right)\right] j\right\} \\ & - \left(\mu + \sigma^2\right) u'(x) + \frac{1}{2}\sigma^2 u''(x).\end{aligned}$$

After an integration by parts for $\int_{\underline{\theta}(x)}^{\bar{\theta}} e^{\theta} dF(\theta)$, I obtain the new HJB equation (9). Transforming the HJB equation from a partial differential equation to an ordinary differential equation simplifies the problem.

The original KF equation for the joint distribution of productivity and wage over tenure, conditional on match surviving, is as follows: $\forall w \in [\underline{w}(z), \bar{w}(z)]$,

$$\begin{aligned}\psi_t(z,w,t) =& -\mu\psi_z(z,w,t) + \frac{1}{2}\sigma^2\psi_{zz}(z,w,t) \\ & + \lambda\left\{\frac{f(\underline{\theta}(z,w))}{F\left(\bar{\theta}\right)}\Psi(z,w,t) - \left[1 - \frac{F(\underline{\theta}(z,w))}{F\left(\bar{\theta}\right)}\right]\Psi(z,w,t)\right\}.\end{aligned} \tag{44}$$

The first two terms in equation (44) capture the diffusion process for firm productivity. The terms within the large bracket capture the jump process for wage. The outflow due to wage bidding is when the outside opportunity is above the poaching threshold, $1 - \frac{F(\underline{\theta}(z,w))}{F\left(\bar{\theta}\right)}$. The inflow is the measure with wage below w, $\Psi(z,w,t)$, and the outside opportunity bid up wage to exactly w.

Similarly, the distribution function can be reduced to a single state variable, $\psi(z,w,t) = \psi(x,t)$. The derivatives have the following relations:

$$\psi_z(z,w,t) = -\psi_x(x,t) \tag{45}$$

$$\psi_{zz}(z,w,t) = \psi_{xx}(x,t). \tag{46}$$

Substituting equation (45) and (46) into the KF equation (44) transforms it into equation (10).

B.4 Proof of Proposition 2

Proof. To derive the stationary productivity distribution, I apply a bilateral Laplace transform of the steady state version of KF equation (2):

$$-\mu g_z(z) + \frac{1}{2}\sigma^2 g_{zz}(z,t) + \delta(h(z) - g(z)) + \lambda \int_{\bar{\theta}}^{\infty} (g(z-\theta) - g(z))\, dF(\theta) = 0. \tag{47}$$

The bilateral Laplace transform is defined as $\hat{g}(\zeta) = \int_{-\infty}^{\infty} e^{-\zeta z} g(z)\, dz$. For the part of the evolution of productivity due to diffusion, I have $\hat{g}_z(\zeta) = \zeta \hat{g}(\zeta)$, $\hat{g}_{zz}(\zeta) = \zeta^2 \hat{g}(\zeta)$. The part of the evolution due to jump:

$$\int_{\bar{\theta}}^{\infty} (g(z-\theta) - g(z))\, dF(\theta) = \int_{\bar{\theta}}^{\infty} \left(e^{-\zeta\theta} - 1\right) dF(\theta)\ \hat{g}(\zeta).$$

The Laplace transformation of (47) yields:

$$\left[\frac{1}{2}\sigma^2\zeta^2 - \mu\zeta - \delta + \lambda \int_{\bar{\theta}}^{\infty} \left(e^{-\zeta\theta} - 1\right) dF(\theta)\right] \hat{g}(\zeta) + \delta \hat{h}(\zeta) = 0. \tag{48}$$

As long as the distribution of jumps is as not as fat as endogenous productivity distribution absent jumps, there is a unique endogenous stationary productivity distribution with double asymptotic Pareto tails.[39] The Pareto indices are the roots of the characteristic equation:

$$\frac{1}{2}\sigma^2\zeta^2 - \mu\zeta - \delta + \lambda \int_{\bar{\theta}}^{\infty} \left(e^{-\zeta\theta} - 1\right) dF(\theta) = 0.$$

The Laplace transform can also serve as the moment generating function of the underlying variable. For negative integers of ζ, $\hat{g}(\zeta) = \mathbb{E}\left[e^{-\zeta z}\right]$ is the $-\zeta$th moment of productivity.

[39] A formal discussion can be found in Gabaix et al. (2016).

The aggregate output is the first moment, which according to equation (48):

$$Y = \hat{g}(-1)(1 - c(\mu)) = \frac{\delta \hat{h}(-1)(1 - c(\mu))}{\delta - \frac{1}{2}\sigma^2 - \mu - \lambda \int_{\bar{\theta}}^{\infty} (e^{\theta} - 1) dF(\theta)}.$$

Given that $h(\cdot)$ is a unit mass at zero, $\hat{h}(-1) = 1$. I obtain the expression for aggregate net output in equation (12). This step also shows that the mean value of the new-born match productivity, $\int e^z dH(z)$, is the relevant parameter. The choice of the functional form for $H(\cdot)$ is a mere normalization. □

B.5 Proof of Proposition 3

Proof. I apply the techniques and tools discussed formally in Nuno and Moll (2017) for optimal control problems with a continuum of heterogeneous agents in continuous time.

For the ease of exposition, I introduce the following compact notations and re-write the equations. Let $L^2(\Phi)$ be the space of functions with a square that is Lebesque-integrable over Φ. The inner product $\langle u, f \rangle_{\Phi} = \int_{\Phi} ufdx$, $\forall u, f \in L^2(\Phi)$, which is used throughout the remaining of this section, helps to keep track of the equations. In our environment, $z \in \mathbb{Z} = \mathbb{R}$, $\theta \in \Theta = \left[\bar{\theta}, \infty\right)$, $t \in \mathbb{T} = \mathbb{R}^+$, and $\Phi = \mathbb{Z} \times \mathbb{T}$.

The KF equation (2) for productivity distribution $g(z,t)$ can be rewritten as:

$$g_t = \mathcal{A}^* g + \delta h + \lambda \langle g(z - \theta, t) - g, f \rangle_{\Theta}, \ \forall z \in \mathbb{Z}, \tag{49}$$

where $\mathcal{A}^*$ is the adjoint operator of $\mathcal{A}$:

$$\mathcal{A}^* g = -\delta g - \mu g_z + \frac{1}{2}\sigma^2 g_{zz}.$$

The problem ($\mathcal{P}^*$) of solving first-best allocation $\left\{\bar{\theta}, \mu\right\}$ in maximizing the social welfare becomes:

$$\max_{\left\{\bar{\theta}, \mu\right\}} \left\langle e^{-\rho t}, \langle e^z - c(\mu) e^z, g \rangle_{\mathbb{Z}} \right\rangle_{\mathbb{T}} \tag{$\mathcal{P}^{*\prime}$}$$

subject to the KF equation (49).

The Lagrangian for problem ($\mathcal{P}^{*\prime}$) is:

$$\mathcal{L} = \left\langle e^{-\rho t}, \langle e^z - c(\mu) e^z, g \rangle_{\mathbb{Z}} \right\rangle_{\mathbb{T}} + \left\langle \Gamma, e^{-\rho t} \left(-g_t + \delta h + \mathcal{A}^* g + \lambda \langle g(z - \theta, t) - g, f \rangle_{\Theta}\right) \right\rangle_{\mathbb{Z} \times \mathbb{T}},$$

where $\Gamma = \Gamma(z,t)$ is the Lagrange multiplier associated with equation (49). Modifying the

second line in the Lagrangian, I obtain:

$$\left\langle \Gamma, e^{-\rho t}\left(-g_t + \lambda \left\langle g\left(z-\theta,t\right)-g, f\right\rangle_{\not\leq} + \delta h + \mathcal{A}^* g\right)\right\rangle_{\mathbb{Z}\times\mathbb{T}}$$
$$= \left\langle e^{-\rho t}\left(\Gamma_t - \rho\Gamma + \mathcal{A}\Gamma + \lambda\left\langle f, \Gamma\left(z+\theta\right)-\Gamma\right\rangle_{\Theta}\right), g\right\rangle_{\mathbb{Z}\times\mathbb{T}} + \delta\left\langle e^{-\rho t}\Gamma, h\right\rangle_{\mathbb{Z}\times\mathbb{T}} + \left\langle \Gamma\left(z,0\right), g\left(z,0\right)\right\rangle,$$

where the infinitesimal operator $\mathcal{A}$ is defined as:

$$\mathcal{A}\Gamma = -\delta\Gamma + \mu\Gamma' + \frac{1}{2}\sigma^2\Gamma''.$$

The modified and re-arranged Lagrangian is:

$$\begin{aligned}\mathcal{L} = &\left\langle e^{-\rho t}\left(e^z - c\left(\mu\right)e^z + \Gamma_t - \rho\Gamma + \mathcal{A}\Gamma + \lambda\left\langle f, \Gamma\left(z+\theta\right)-\Gamma\right\rangle_{\Theta}\right), g\right\rangle_{\mathbb{Z}\times\mathbb{T}} \\ &+ \delta\left\langle e^{-\rho t}\Gamma, h\right\rangle_{\mathbb{Z}\times\mathbb{T}} + \left\langle \Gamma\left(z,0\right), g\left(z,0\right)\right\rangle.\end{aligned}$$

Therefore, I obtain:

$$\rho\Gamma = \max_{\left\{\bar{\theta},\mu\right\}}\left\{e^z - c\left(\mu\right)e^z + \mathcal{A}\Gamma + \lambda\left\langle f, \Gamma\left(z+\theta\right)-\Gamma\right\rangle_{\Theta} + \Gamma_t\right\}. \tag{50}$$

In the recursive formulation in equation (50), there is no aggregate state variables. Therefore, $\Gamma\left(z,t\right) = \Gamma\left(z\right)$ and $\Gamma_t = 0$. Rewriting equation (50):

$$r\Gamma\left(z\right) = e^z - c\left(\mu\right)e^z + \mu\Gamma'\left(z\right) + \frac{1}{2}\sigma^2\Gamma''\left(z\right) + \lambda\int_{\bar{\theta}}\left[\Gamma\left(z+\theta\right)-\Gamma\right]dF\left(\theta\right). \tag{51}$$

The Lagrange multipliers $\Gamma\left(z\right)$ is the shadow social value function associated with a manager with current productivity z. I guess and verify that the shadow social value function is linear in e^z, i.e. $\Gamma\left(z\right) = \gamma^* e^z$. The linear value function implies the following. First, the poaching threshold is $\bar{\theta}^* = 0$ since $\Gamma\left(\cdot\right)$ is increasing. Second, first order condition with respect to μ is $c'\left(\mu^*\right) = \gamma^*$, which implies that investment is independent of productivity. Replacing the two results above in the HJB equation (51), I obtain that:

$$\gamma^* = \frac{1 - c\left(\mu^*\right)}{r - \frac{1}{2}\sigma^2 - \mu^* - \lambda\int_0\left[e^\theta - 1\right]dF\left(\theta\right)}.$$

The steady state welfare is the value of discounted value of the stream of steady state net output,

$$\int\Gamma\left(z\right)dG\left(z\right) + \frac{\delta}{\rho}\int\Gamma\left(z\right)dH\left(z\right) = \frac{1}{\rho}\frac{\delta\left[1-c\left(\mu\right)\right]}{\delta - \frac{1}{2}\sigma^2 - \mu - \lambda\int_{\bar{\theta}}^{\infty}\left[e^\theta - 1\right]dF\left(\theta\right)}.$$

□

B.6 Proof of Proposition 4

Proof. I consider the modified version of the planner's problem. It chooses the contract directly to maximize the social welfare:

$$\left\langle e^{-\rho t}, \left\langle e^{z} - c\left(\mu\right) e^{z}, g\right\rangle_{\mathbb{Z}}\right\rangle_{\mathbb{T}}$$

subject to the firms' investment incentive constraint in equation 4, and the KF equation (2).

The Lagrangian for this problem is

$$\begin{aligned}\mathcal{L} = & \left\langle e^{-\rho t}, \left\langle e^{z} - c\left(\mu\right) e^{z}, g\right\rangle_{\mathbb{Z}}\right\rangle_{\mathbb{T}} \\ & + \left\langle \Gamma, e^{-\rho t}\left(-g_t + \delta h + \mathcal{A}^{*} g + \lambda \left\langle g\left(z-\theta, t\right) - g, f\right\rangle_{\Theta}\right)\right\rangle_{\mathbb{Z}\times\mathbb{T}} \\ & + \left\langle \xi, e^{-\rho t}\left\langle J'\left(z\right) - c'\left(\mu\right) e^{z}, g\right\rangle_{\mathbb{Z}}\right\rangle_{\mathbb{T}}\end{aligned}$$

where $\Gamma = \Gamma\left(z,t\right)$, $\xi = \xi\left(z,t\right)$ are the Lagrange multipliers associated with equation (49) and 4. The modified and re-arranged Lagrangian is:

$$\begin{aligned}\mathcal{L} = & \left\langle e^{-\rho t}\left(e^{z} - c\left(\mu\right) e^{z} + \Gamma_t - \rho\Gamma + \mathcal{A}\Gamma + \lambda\left\langle f, \Gamma\left(z+\theta\right) - \Gamma\right\rangle_{\Theta} + \xi\left(J'\left(z\right) - c'\left(\mu\right) e^{z}\right)\right), g\right\rangle_{\mathbb{Z}\times\mathbb{T}} \\ & + \delta\left\langle e^{-\rho t}\Gamma, h\right\rangle_{\mathbb{Z}\times\mathbb{T}} + \left\langle \Gamma\left(z,0\right), g\left(z,0\right)\right\rangle\end{aligned}$$

Therefore,

$$\begin{aligned}r\Gamma\left(z\right) = \max_{\pi^{p},\tau(\cdot),\mu} & e^{z} - c\left(\mu\right) e^{z} + \mu\Gamma_z\left(z\right) + \frac{1}{2}\sigma^{2}\Gamma_{zz}\left(z\right) + \lambda\int_{\bar{\theta}}^{\infty}\left[\Gamma\left(z+\theta\right) - \Gamma\left(z\right)\right] dF\left(\theta\right) \\ & + \xi\left(J'\left(z\right) - c'\left(\mu\right) e^{z}\right).\end{aligned} \tag{52}$$

Comparing the HJB equation in problem ($\mathcal{P}'$) with equation (52), the right-hand side of equation (52) differs by taking into account the value of the outside firms. Therefore, $\pi^{p} < \pi$. □

B.7 Proof of Proposition 5

Proof. The proof in this extension follows the same steps as the one for Proposition 1 in the baseline. I note two differences that lead to a continuum buyout menu.

The first difference is that the bidding strategies take into account the business stealing effect. In order to poach away the manager, the outside firm would need to bid as much as

the initial firm's maximum willingness to pay, which includes the total discounted value of business stealing, $\Upsilon\left(z,\tilde{\pi}\left(z,\theta\right)\right)$.

For the problem of maximizing rent extracted, the (IC) and (IR) constraints are modified to:

$$\tilde{\pi}\left(z,\theta\right)\in\arg\max_{\tilde{\pi}\in[0,\pi]}S\left(\theta,\tilde{\pi}\middle|\,z\right), \tag{53}$$

$$S\left(\bar{z},\pi\left(z,\bar{z}\right)\middle|\,z\right)\geq 0,\ \forall\theta\geq\bar{\theta}. \tag{54}$$

Taking advantage of the envelope condition for the (53) constraint, together with the binding (54) constraint at the poaching threshold, the buyout payment satisfies:

$$\tau\left(z,\tilde{\pi}\left(z,\theta\right)\right)=S\left(\theta,\tilde{\pi}\left(z,\theta\right)\middle|\,z\right)-\int_{\bar{\theta}}^{\theta}S_{\theta}\left(\tilde{\theta},\tilde{\pi}\left(z,\tilde{\theta}\right)\middle|\,z\right)d\tilde{\theta}. \tag{55}$$

The problem of maximizing rent extracted becomes:

$$\begin{aligned}&\max_{\mathcal{M}}\int_{\bar{\theta}}^{\infty}\left[S\left(\theta,\tilde{\pi}\left(z,\theta\right)\middle|\,z\right)-\int_{\bar{\theta}}^{\theta}S_{\theta}\left(\tilde{\theta},\tilde{\pi}\left(z,\tilde{\theta}\right)\middle|\,z\right)d\tilde{\theta}\right]F\left(\theta\right)\\=&\max_{\mathcal{M}}\int_{\bar{\theta}}^{\infty}\left[S\left(\theta,\tilde{\pi}\left(z,\theta\right)\middle|\,z\right)-\frac{1-F\left(\theta\right)}{f\left(\theta\right)}S_{\theta}\left(\theta,\tilde{\pi}\left(z,\theta\right)\middle|\,z\right)\right]F\left(\theta\right).\end{aligned}$$

The first order condition with respect to $\tilde{\pi}\left(z,\theta\right)$ is:

$$S_{\pi}\left(\theta,\tilde{\pi}\left(z,\theta\right)\middle|\,z\right)-\frac{1-F\left(\theta\right)}{f\left(\theta\right)}S_{\theta\pi}\left(\theta,\tilde{\pi}\left(z,\theta\right)\middle|\,z\right)\geq 0\text{ with }\text{“}=\text{”}\text{ if }\tilde{\pi}\left(z,\theta\right)>0,\ \forall\theta\geq\bar{\theta}. \tag{56}$$

The second difference arises in equilibrium non-competition buyout according to the first order condition (56). When Assumption 2 holds, non-competition buyout doesn't need to be at the corner solution.

The first order condition with respect to $\bar{\theta}$ is

$$S\left(\bar{\theta},\tilde{\pi}\left(z,\bar{\theta}\right)\middle|\,z\right)-\frac{1-F\left(\bar{\theta}\right)}{f\left(\bar{\theta}\right)}S_{\theta}\left(\bar{\theta},\tilde{\pi}\left(z,\bar{\theta}\right)\middle|\,z\right)=0. \tag{57}$$

Replacing the equilibrium non-competition buyout into equation (57), I obtain:

At the poaching threshold:

$$e^{-r\pi}J\left(z+\bar{\theta}\right)=J\left(z\right)+\Upsilon\left(z,\tilde{\pi}\left(z,\theta\right)\right),$$

which implies that

$$\pi = \frac{1}{r}\left[\log J\left(z+\bar{\theta}\right) - \log\left[je^{z} + e^{-(r+\eta)\tilde{\pi}(z,\theta)+z}\frac{\upsilon}{r+\eta}\right]\right].$$

$$e^{\bar{\theta}}\left[1 - \frac{1-F\left(\bar{\theta}\right)}{f\left(\bar{\theta}\right)}\right] = \frac{(r+\eta)}{\eta^{\frac{\eta}{r+\eta}} r^{\frac{r}{r+\eta}}}\nu^{\frac{r}{r+\eta}};$$

□

B.8 Derivation of Investment Response

Given the functional form of investment cost, the optimality condition for investment becomes $c'(\mu) = \mu^{\frac{1}{\varphi}} = j$. Taking log-difference:

$$d\log\mu = \varphi d\log j. \tag{58}$$

The log-difference of the expression for the marginal bilateral joint value in equation (7),

$$d\log j \approx -d\log\left(r - \mu - \frac{1}{2}\sigma^2 - \lambda\left(e^{\bar{\theta}}-1\right)\left(1-F\left(\bar{\theta}\right)\right)\right) \approx \frac{\lambda}{r-\mu-\frac{1}{2}\sigma^2}\left(e^{\bar{\theta}}-1\right)\left(1-F\left(\bar{\theta}\right)\right). \tag{59}$$

Combining equation (58) and (59), I obtain the investment response equation:

$$d\log\mu = \varphi\frac{\lambda}{r-\mu-\frac{1}{2}\sigma^2}\left(e^{\bar{\theta}}-1\right)\left(1-F\left(\bar{\theta}\right)\right). \tag{60}$$

Since $d\log c(\mu) = \left(1+\frac{1}{\varphi}\right)d\log\mu$, I obtain the investment expense response equation (15).

C Additional Figures and Tables

C.1 Summary Statistics of Data Sample

Table 6: Sample summary statistics

	All	Non-Compete		Duration (Year)		
		No	Yes	(0,1]	(1,2]	(2,5]
Fraction (%)	100	36	64	50	39	10
Manager Characteristics						
Manager Age	45.9	46.2	45.8	45.7	45.7	45.9
Tenure	8.6	8.4	8.6	8.0	9.1	10.5
Job-to-Job Transition Rate (%)	1.16	1.24	1.12	1.18	1.07	1.02
– Within Industry	0.39	0.45	0.36	0.43	0.30	0.30
– Between Industry	0.77	0.79	0.76	0.75	0.77	0.72
Separation Rate (%)	8.49	8.68	8.38	8.83	8.11	7.11
Firm Characteristics						
Firm Age	18.5	18.6	18.4	17.4	19.2	19.2
Asset (mn)	5,448	5,455	5,444	4,754	4,906	6,767
Sales (mn)	5,228	4,653	5,545	4,141	5,531	7,774
Emploment (thousands)	20	16	22	19	21	31
Total Investment	0.181	0.189	0.177	0.185	0.170	0.163
– Phyiscal Investment	0.062	0.062	0.062	0.064	0.061	0.060
– Intangible Investment	0.121	0.129	0.116	0.122	0.110	0.104
Tobin's Q	1.44	1.44	1.44	1.47	1.42	1.40
Cash	0.28	0.25	0.30	0.29	0.33	0.28
ROA	0.022	0.018	0.024	0.018	0.026	0.039
Compensation						
Awarded Compensation (mn)	3.15	3.13	3.17	2.82	3.17	4.18
Realized Compensation (mn)	2.98	2.93	3.01	2.76	2.87	4.08
– Cash	0.82	0.81	0.82	0.76	0.82	1.08
– Deferred Equity	1.84	1.88	1.82	1.64	1.68	2.61
Unvest	2.90	2.97	2.86	2.50	2.95	4.02
Vest	4.68	4.74	4.65	4.20	4.40	8.10

Notes: Manager age and firm age refer to the respective age at the beginning of the match. Job-to-job transition rates are defined as movements between jobs observed in the sample of Compustat firms. Industry definition is based on two-digit SIC codes. Physical investment is defined as capital expenditure (capx) normalized by lagged property, plant and equipment (ppegt). Intangible investment is defined as R&D expense (xrd) plus 30% of selling, general and administrative expense (xsga) normalized by lagged total capital (ppegt+k_int). R&D rate is zero whenever missing. Total investment is the sum of physical and intangible investments. Tobin's Q is defined as book assets (at) plus market value of equity (prcc_f×csho) minus common equity (ceq) and deferred taxes (txdb) normalized by property, plant and equipment (ppegt). Cash flow is defined as income before extraordinary items (ib) plus depreciation and amortization (dp) normalized by lagged property, plant and equipment (ppegt). Return on asset (ROA) is defined as net income (ni) normalized by book assets (at). Nominal values are deflated to year 2010 dollars using the CPI.

C.2 Additional Regression Results

Table 7: Effect of non-competition on job-to-job transition

	All	Within Industry	Between Industry
	(1)	(2)	(3)
Non-Competition	-0.122**	-0.168	-0.100
	(0.055)	(0.132)	(0.075)
Industry FEs	Y	Y	Y
Year FEs	Y	Y	Y
State FEs	Y	Y	Y
Observations	52,966	52,966	52,966

*Notes: All specifications control for manager age, firm asset, total Tobin's Q, and return on asset. Hazards are stratified by whether the executive holds the role of CEO, whether the executive is interlocked, and the gender of the executive. Standard errors clustered by state are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.*

Table 8: Effect of non-competition on firm investment, alternative measures

	Intangible		Physical	
	(1)	(2)	(3)	(4)
Non-Competition	0.005		0.013***	
	(0.003)		(0.004)	
Non-Competition × Non-CA		0.006*		0.013***
		(0.003)		(0.004)
Non-Competition × CA		-0.013***		-0.018***
		(0.001)		(0.002)
Sample	Non-CA	All	Non-CA	All
Observations	18,147	21,226	17,918	20,739
Adjusted R^2	0.51	0.52	0.31	0.30

*Notes: All specifications control for total Tobin's Q and cash. Standard errors clustered by state are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.*

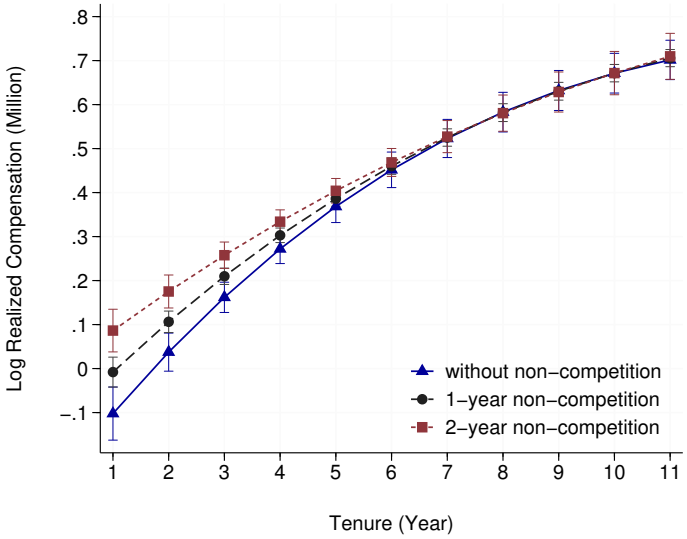
Table 9: Non-competition buyout cases

Case	Year	Buyout (mn)	
		Original	2010$
HP former CEO Mark Hurd moving to Oracle	2010	14	14
Contract between Engility and Anthony Smeraglinolo	2016	4.5	4.1
Tasciyan v. Marsh USA Inc.	2006	0.57	0.62

Notes: Mark Hurd, former CEO of HP, didn't in fact have a non-competition agreement, which wouldn't be enforceable in California. He instead entered into non-disclosure agreements with HP, who then clawed back some compensation for "inevitable disclosure".

C.3 Additional Figures

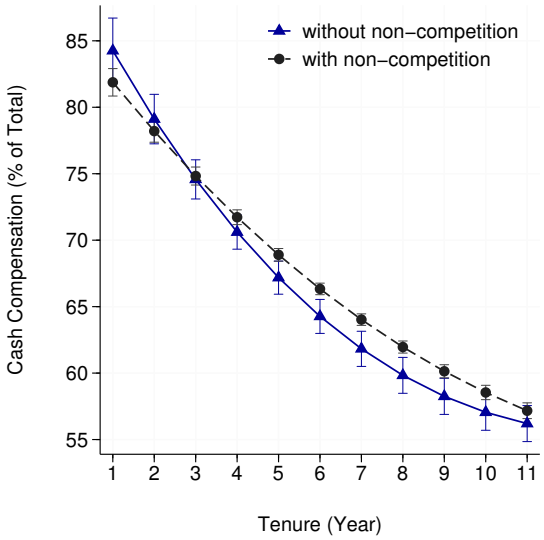
Figure 5: Wage-backloading by duration of non-competition



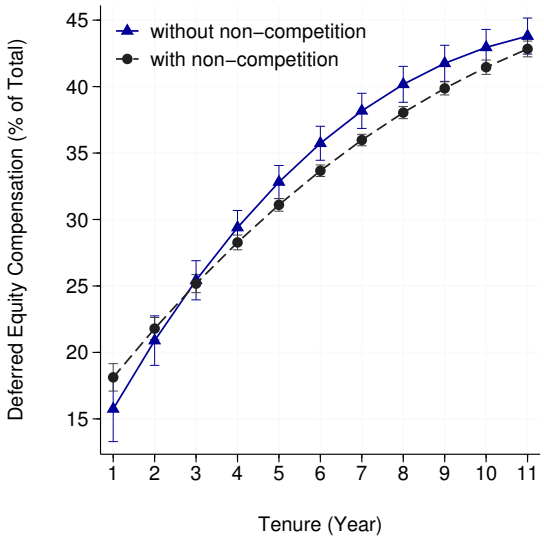
Notes: This figure plots wage over tenure by the duration of non-competition, based on the marginal effects at means in the regression in column (4) of Table 3. The bars display 95% confidence interval.

Figure 6: Composition of compensation by whether under non-competition

(a) fraction in cash compensation



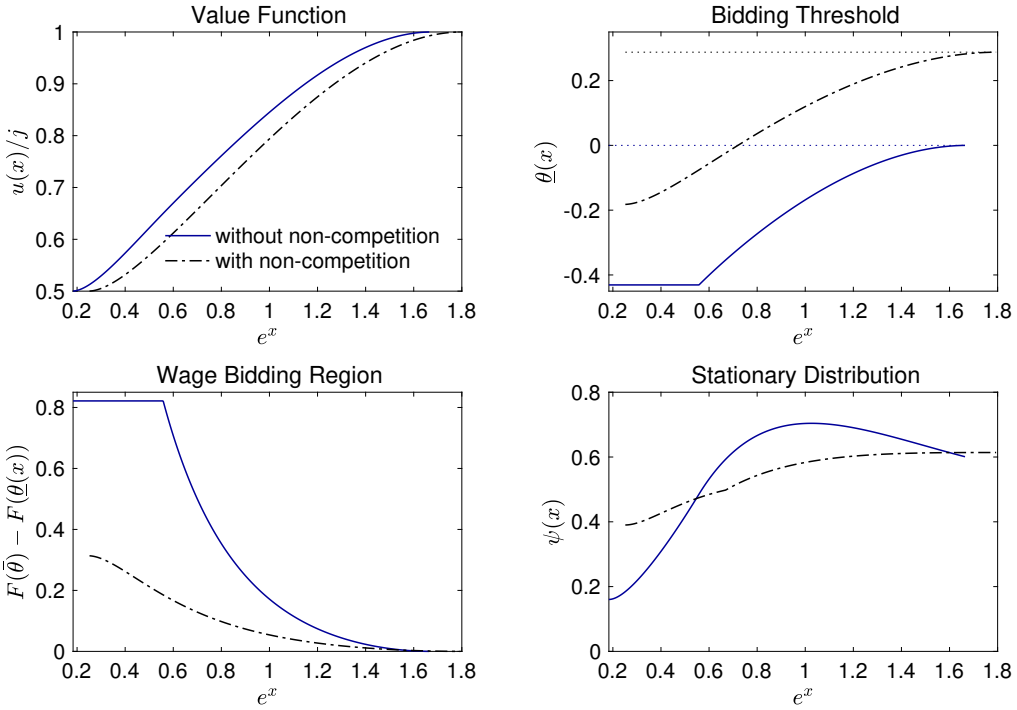
(b) fraction in deferred equity compensation



Notes: The figures plot the fraction of compensation in cash and deferred equity over tenure, by whether the executive is subject to non-competition. The estimates are based on the marginal effects at means. Cash compensation is defined as the sum of salary, bonus, and non-equity incentives. Deferred equity compensation is defined as the sum of the value of shares vested and the value of options exercised. The bars display 95% confidence interval.

C.4 Numerical Solutions

Figure 7: Numerical solution of wage-backloading



Notes: This figure plots the numerical solution of the wage bidding thresholds and corresponding value function and distribution in the wage-backloading section.

D Additional Data Appendix

D.1 Contract Examples

Figure 8: Example of non-competition agreement

Exhibit 10.23

COVENANT NOT TO COMPETE
AND NON-DISCLOSURE AGREEMENT

PARTIES:

Eric Dean Sprunk ("EMPLOYEE")

and

NIKE, Inc., divisions, subsidiaries
and affiliates. ("NIKE"):

AGREEMENT:

In consideration of the foregoing, and the terms and conditions set forth below, the parties agree as follows:

1. **Covenant Not to Compete**.

(a) **Competition Restriction**. During EMPLOYEE's employment by NIKE, under the terms of any employment contract or otherwise, and for one year thereafter, (the "Restriction Period"), EMPLOYEE will not directly or indirectly, own, manage, control, or participate in the ownership,
management or control of, or be employed by, consult for, or be connected in any manner with, any business engaged anywhere in the world in the athletic footwear, athletic apparel or sports equipment and accessories business, or any other business which directly competes with NIKE or any of its parent, subsidiaries or affiliated corporations ("Competitor"). ***By way of illustration only***, examples of NIKE competitors include, but are not limited to: Adidas, FILA, Reebok, Puma, Champion, Oakley, DKNY, Converse, Asics, Saucony, New Balance, Ralph Lauren/Polo Sport, B.U.M, FUBU, The Gap, Tommy Hilfiger, Umbro, Northface, Venator (Foot lockers), Sports Authority, Columbia Sportswear, Wilson, Mizuno, Callaway Golf and Titleist. This provision is subject to NIKE's option to waive all or any portion of the Restriction Period as more specifically provided below.

Notes: The figure displays snapshots of relevant contractual details. The full text of the contract can be found here: https://www.sec.gov/Archives/edgar/data/320187/000119312510161874/dex1023.htm.

Figure 9: Example of non-competition with forfeiture-for-competition clause

Exhibit 10.3

EXECUTION VERSION

NON-COMPETITION AGREEMENT

This Non-Competiton Agreement (the "Agreement") is entered into as of June 29, 2016 by and between Engility Holdings, Inc. (the "Company") and Anthony Smeraglinolo (the "Executive").

WHEREAS, pursuant to the Separation Agreement, the Executive agreed to be bound by a Confidentiality and Non-Competition Restrictive Covenants agreement (the "Non-Competition Agreement") provided for under the CIC Severance Plan, which includes a covenant to not compete with the Company for a minimum period of twelve months following the date of the Executive's separation from employment (the "First Non-Competition Period");

WHEREAS, the Non-Competition Agreement provides that the duration of the covenants thereunder may, at Executive's discretion, be extended beyond the First Non-Competition Period for a period of up to three years;

V. Potential Forfeiture of Payment

In addition to the remedies provided in Section IV of this Agreement, in the event that the Executive has breached the non-competition covenants contained in this Agreement or in the Separation Agreement (i) during the First Non-Competition Period, the Executive shall forfeit all right and interest to $3,000,000, or (ii) during the Second Non-Competition Period, $1,500,000. The Executive shall be required to pay to the Company the applicable forfeiture amount, in cash, within fifteen (15) days after demand is made therefore by the Company, as liquidated damages for the breach of such restrictive covenants. The provisions of this Section V shall constitute an amendment of Section I of the Separation Agreement and Exhibit A of the CIC Severance Plan.

Notes: Link: https://www.sec.gov/Archives/edgar/data/1544229/000156459016021836/egl-ex103_275.htm.

Figure 10: Example of restricted stock option awards

Exhibit 10.12

THE SHARES ISSUABLE UPON VESTING OF THIS AWARD WILL NOT BE RELEASED TO YOU UNTIL ALL APPLICABLE WITHHOLDING TAXES HAVE BEEN COLLECTED FROM YOU OR HAVE OTHERWISE BEEN PROVIDED FOR.

AMAZON.COM, INC.

RESTRICTED STOCK UNIT AWARD AGREEMENT

TO: <<Participant>>

To encourage your continued employment with Amazon.com, Inc. (the "Company") or its Subsidiaries, you have been granted this restricted stock unit award (the "Award") pursuant to the Company's 1997 Stock Incentive Plan (the "Plan"). The Award represents the right to receive shares of Common Stock of the Company subject to the fulfillment of the vesting conditions set forth in this agreement (this "Agreement").

6. Termination of Employment. The unvested portion of the Award will terminate automatically and be forfeited to the Company immediately and without further notice upon the voluntary or involuntary termination of your employment for any reason with the Company or any Subsidiary (including as a result of death or disability). A transfer of employment or services between or among the Company and its Subsidiaries shall not be considered a termination of employment. Unless the Plan Administrator determines otherwise, and except as otherwise required by local law, for purposes of this Award only, any reduction in your regular hours of employment to less than 30 hours per week is deemed a termination of your employment with the Company or any Subsidiary. In case of termination of your employment for Cause, the Award shall automatically terminate upon first notification to you of such termination, unless the Plan Administrator determines otherwise. If your employment is suspended pending an investigation of whether you should be terminated for Cause, all of your rights under the Award likewise shall be suspended during the period of investigation. No Shares shall be issued or issuable with respect to any portion of the Award that terminates unvested and is forfeited.

Notes: Link: https://www.sec.gov/Archives/edgar/data/1018724/000095014903000355/v87419orexv10w12.htm.

D.2 Collecting Employment Contract Data

I collect executive employment contracts included in company filings from 1994 onwards in the U.S. Securities and Exchange Commission (SEC) Edgar database. The SEC requires that public companies disclose contracts material to their business. Management employment contracts and compensatory plans involving directors or executive officers are deemed material and therefore filed in the Edgar database. Among them are forms of contracts including initial employment agreements, letters of employment, amendments to existing employment agreements, stand-alone non-competition agreements, retention agreements, and separation and severance agreements. These employment contracts provide the source of information on executive non-competition arrangements used for this study.

D.2.1 Contract Classification

To gather the contracts, I search with an automated crawler across the SEC `Edgar` database. The contracts are appended as exhibits under "exhibit 10" designation in annual and quarterly reports (10K and 10Q forms, respectively) and current reports for major events (8K forms).

One issue to overcome in collecting contracts is that various other types of contracts can resemble employment contracts. For example. supplier-buyer purchase agreements and joint venture agreements have similar legal concerns as employment contracts. Many companies do not clearly indicate the type of the contract when filing it. To filter out irrelevant contracts, I use natural language processing tools and supervised machine learning algorithms to classify whether a document is an employment agreement. Specifically, I use a subset of contracts with sufficient information in the document title as the dataset. This enables me to label these contract into employment or non-employment type using the information in the document title.[40] The dataset includes 18,904 employment contracts and 11,932 non-employment contracts. I split the dataset with 75% as the training set and 25% as the test set. I then use the word dictionary of the training set as word features and train a logistic regression classification algorithm. The classification algorithm yields an accuracy rate of over 97% in the test set. The key words that classify a document as an employment contract are shown in Figure 11. In particular, word features such as "non-competition", "compete", and "retention" contribute to a document being classified as an employment contract. Using the classification algorithm, I obtain a total of 68,267 employment contracts.

[40] Some contracts were filed with an informative document title clearly indicating the contract type, while others were filed with uninformative document title. One such uninformative document title is "exhibit".

Figure 11: Key words for classifying a document as an employment contract

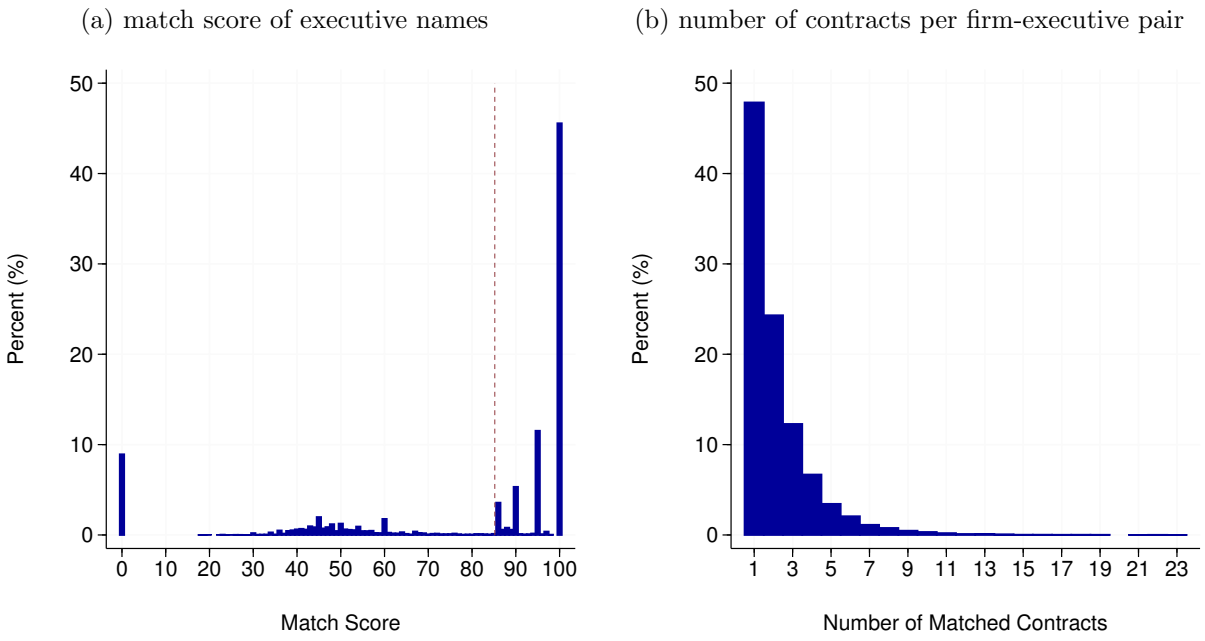
D.2.2 Textual Analysis

I perform textual analysis to extract relevant details on non-competition in the contracts. Of particular interest are contractual terms including (1) whether an employment contract includes a non-competition covenant; (2) if so, the duration of post-separation non-competition period, most commonly one year, eighteen months, or two years. The challenge is that the key information is buried in lengthy discussions of varying legal formats. I address this issue by carefully observing and applying the common features below.

To accurately identify the name of the executive that a contract binds, I apply four methods. First, I parse the content explicitly specifying the contracting parties and extract the texts that refers to the name of the employee. This is achieved by taking advantage of key word marks such as ("employee") as in contract example 8 and ("executive") as in contract example 9 which immediately follow the executive name. It is able to identify names for most contracts and is also the most accurate method. The second method is parsing out the employee name in the title of the contract if it is included. The third method, applicable to cases involving letters of employment, is to capture the name of the person to whom the letter is addressed. Finally, if all of the above fails, I find the name in the signature portion of the contract which usually has a key mark "/s/" preceding it. The last one is the least reliable approach due to instances of a signature representing the firm appearing alongside the one for the employee.

It is straightforward to determine whether a contract includes a non-competition clause. An employment relation has a non-competition clause if at least one contract has at least

Figure 12: Matching executive names between the contracts and Execucomp



one count of words that are some variation of "non-competition".[41] If a contract includes non-competition clause, it has on average 2.9 word counts related to "non-competition". Determining the duration of non-competition is slightly more difficult. I do so by extracting time in the content on details of the non-competition. Key words such as "restriction period" and "non-competition period" as in contract example 8 help to improve accuracy.

Most contracts also specify the legal jurisdiction under which it is to be governed. This information can be reliably gathered, which largely coincides with the state where the company headquarter is located.

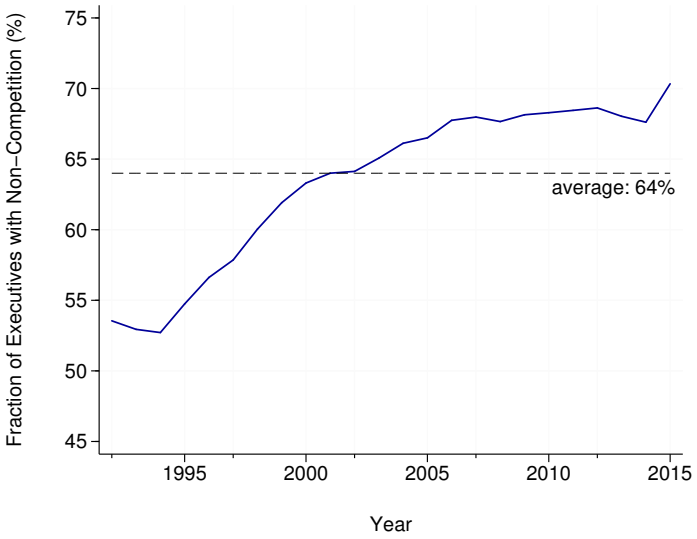
D.2.3 Matching Names

I match the executive names in the contracts to the set of executive names in `Execucomp`. Since company filings have unique company identification, matching among the set of executive names within the identified company leads to high accuracy. I use the string matching package in python `FuzzyWuzzy` and obtain the score for each contract matched to the closest executive name in `Execucomp`.[42] Panel (a) of Figure 12 shows the distribution of match scores. Almost half of the contracts have match score of 100. Of the 68,267 employment

[41] Finally, instances of multiple contracts or amendments to existing contract are observed. There are finer details on when the contract is signed but this information hard to extract. It also happens that contracts are filed with delay.

[42] This package uses Levenshtein Distance to calculate the differences between sequences of strings

Figure 13: Uses of non-competition contracts over time



Notes: This figure plots the fraction of executives with non-competition in a given year.

contracts gathered, I keep 45,446 of them with a score of 86 and above.

I keep in the firm-executive employment relations in `Execucomp` linked to at least one contract in `Edgar`. Panel (b) of Figure 12 shows the distribution of the number of contracts for the firm-executive pair. The average number of contracts is 2.3 per employment relation. At the maximum 23 documents of contracts are found for John J. Dooner, Jr., who held various senior executive roles at the advertising company Interpublic Group of Companies, Inc.

D.3 Contract Data

The contract data shows that there is an overall upward trend in the use of non-competition over time. As Figure 13 shows, during the earlier years of the data sample, the fraction of executives subject to non-competition is below 55%; during the later years, the fraction is close to 70%. The trend confirms the increasing usage of non-competition contracts.

D.4 Employment History Data

The `Execucomp` dataset reports the dates the executive joined the company (joined_co, rejoin) and left the company (leftco, releft). In the case of CEOs, the dataset also reports the dates the executive became CEO (becameceo) and left as CEO (leftofc). However, this information is less than ideal with some missing ones. I supplement it with available employment history from `Capital IQ People Intelligence` (startyear, endyear).

The exact employment dates allow me to measure tenure and separation properly. The tenure variable is defined as one for the first year of employment. The dummy for separation event is defined as one for the last year of employment. I define the dummy for job-to-job transition events as one for the last year of employment if the executive is subsequently employed at another firm. The within-industry and between-industry job-to-job transitions are defined using industry definition based on two-digit SIC codes. I drop the observations for which tenure cannot be reliably measured. This is the step 4 of data sample filtering in subsection 5.3.

D.5 Compensation Data

The detailed compensation data in `Execucomp` is used to assess how non-competition interacts with compensation design. I note the nuances of different compensation measures in two regards. First, public firms are required to disclose compensation for their top executives per SEC regulation. Two types of total compensation measure are reported – awarded compensation and actual realized compensation. A large part of awarded compensation are in the form of restricted equity deferred to future dates contingent on the executive staying with the firm. Therefore, realized compensation is more pertinent than awarded compensation to gauging wage-backloading. Second, the exact disclosure requirement has gone through regulatory changes. In particular, prior to 2006, the two measures are tdc1 for awarded compensation and tdc2 for realized compensation. Starting in 2006, two alternative compensation measures, total_alt1 for awarded and total_alt2 for realized, are reported in compliance with the 2006 financial accounting standard for equity compensation (FAS 123R). The main distinction between pre- and post- 2006 measures is that for the latter stock and option awards reflect the estimated fair value at grant date and exercise or vest date. For these two reasons, the post-2006 measure for realized compensation (total_alt2) is the most relevant one that I use.

Formally, realized compensation (total_alt2) includes salary, bonus, value of shares vested, value of options exercised, non-equity incentives, change in pension, and other compensation. For compensation composition, I define cash compensation as the sum of salary, bonus, and non-equity incentives; deferred equity compensation as the sum of value of shares vested and value of options exercised.

Nevertheless it is reassuring that the wage-tenure profile and its interaction with non-competition are robust to choices of pre- and post- 2006 measures. The baseline compensation regression in Table 3 using the four compensation measures is reported in Table 10. I include the entire final sample in this regression, both California and non-California obser-

Table 10: Wage-backloading based on alternative compensation measures

	Realized		Awarded	
	Post 2006	Pre 2006	Post 2006	Pre 2006
	(1)	(2)	(3)	(4)
Non-Competition	0.220***	0.125***	0.103*	0.043
	(0.044)	(0.035)	(0.057)	(0.040)
Tenure/10	1.730***	1.168***	0.206	0.043
	(0.137)	(0.074)	(0.175)	(0.114)
Tenure/10 × Non-Competition	-0.647***	-0.355***	-0.140	0.017
	(0.122)	(0.080)	(0.157)	(0.104)
$(\text{Tenure}/10)^2$	-0.931***	-0.614***	-0.134	-0.018
	(0.120)	(0.063)	(0.105)	(0.064)
$(\text{Tenure}/10)^2$ × Non-Competition	0.549***	0.310***	0.128	0.003
	(0.108)	(0.062)	(0.100)	(0.065)
$(\text{Tenure}/10)^3$	0.150***	0.093***	0.027	0.003
	(0.023)	(0.014)	(0.019)	(0.012)
$(\text{Tenure}/10)^3$ × Non-Competition	-0.115***	-0.063***	-0.026	-0.001
	(0.022)	(0.014)	(0.019)	(0.012)
Industry FEs	Y	Y	Y	Y
Year FEs	Y	Y	Y	Y
State FEs	Y	Y	Y	Y
Observations	19,269	47,976	19,300	44,459
Adjusted R^2	0.51	0.47	0.54	0.48

*Notes: All specifications control for firm asset, total Tobin's Q, return on asset, whether the executive holds the role of CEO, whether the executive is interlocked, and the gender of the executive. The Standard errors clustered by state are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.*

vations. Column (1) and (2) show that, with both measures of realized compensation, wage grows over tenure; executives with non-competition are associated with a higher starting wage and a lower wage growth over tenure. However, the effect of non-competition on wage-backloading is much larger when using the post-2006 measure, almost the double of the one with pre-2006 measure. This is sensible since equity compensation constitutes around 60% of overall compensation and the book value tends to be lower than the fair value. Therefore, the pre-2006 measure under-estimates the extent of wage-backloading. Column (3) and (4) show that these effects are largely absent for awarded compensation.