A Bumpy Job Ladder Model of Executive Compensation*

Bo Hu†

Abstract

This paper assesses the impact of managerial labor market competition on executive incentive contracts. I develop a dynamic contracting model that features moral hazard, search frictions and poaching offers. The model generates a job ladder along which executives get promoted internally and transit toward larger companies. The ladder is bumpy in that how high the next rung depends on the manager's effort, the realized productivity and size of the poaching firm. The model is applied to two exercises. First, I show that poaching generates a new source of incentives that explains a newly documented empirical puzzle — the firm-size incentive premium. Second, the estimated model is employed to quantitatively account for the trend of executive compensation over decades.

Keywords: Executive Compensation, Managerial Labour Market, Dynamic Moral Hazard, Search Friction, Firm-size Premium

*This article has benefited from the comments and criticisms of participants at the Rotterdam Executive Conference (2018), the 26th Finance Forum at Santander (2018), the Belgrian Financial Research Forum (2019), the search and matching conference at Cambridge (2019), the SUFE IAR Macroeconomics Workshop in Shanghai (2021), and at various seminars at Queen Mary University of London, Vrije Universiteit Amsterdam, HEC Paris, Halle Institute for Economic Research, Tor Vergata University of Rome, University of Gothenburg, Shanghai Jiaotong University. I am grateful to Pieter Gautier, Makoto Watanabe, Sarolta Laczó, Gregory Jolivet, Sebastian Gryglewicz, Shuo Xia, Florian Peters, Ingolf Dittmann, Alvaro Remesal, and Yanos Zylberberg for helpful discussions. I thank Jeffrey Coles, Naveen Daniel, Lalitha Naveen, Cláudia Custódio, Miguel Ferreira, Pedro Matos, Martijn Cremers, and Yaniv Grinstein for generouly sharing data from their respective papers. I acknowledge the financial supports from the Natural Science Foundation of China (72003041), and Shanghai Institute of International Finance and Economics.

†Institute of World Economy, Fudan University, Shanghai Institute of International Finance and Economics (SIIFE); hu_bo@fudan.edu.cn; Room 723, Guoquan Rd. 600, Yangpu District, Shanghai, China.

"The bumpy ladder presents a fun obstacle for children to climb. Each of the sections of the ladder between rungs are on slightly different angles. "

— *from the product description of a bumpy ladder*

1 Introduction

Recent literature has documented a rich set of stylized facts on executive job-to-job mobility. For instance, there has been an increasing trend of transitions since the 1980s. Regarding *who transits,* it is found that transitions happen to all titles, but the transition rate is lower for CEOs and executives in larger firms. On the *moving direction,* the majority of job-to-job transitions have a move towards a larger firm, a promotion in title, and a rise in pay. These facts hint at a job ladder in which more prestigious employers and titles locate on the higher end (Graham et al. 2021, Kaplan and Minton 2012, Murphy and Zabojnik 2007, Huson et al. 2001).[1] Contrary to the abundant empirical findings, a theoretical framework that is compatible with these mobility facts is missing. Indeed, this paper is a step towards such a framework, asking: What are the features of the managerial labor market that generate these job mobility patterns? What differentiates the job ladder of executives from that in a general labor market? What are the impacts on executive contracts, and notably on the hallmark of those contracts — the incentives?

I answer these questions by constructing a job ladder model that features *search frictions* in the executive labor market and *dynamic moral hazard* between the firm and the executive. In the model, firms differ in size, and executives differ in managerial productivity, both of which contribute to production. The firm size is time-invariant, and the executive productivity evolves by a Markov process that depends on effort. While output (equivalently, executive productivity) is observable, effort is not. Thus, *moral hazard* arises. The pair of firm and executive establishes a long-term incentive contract that trade-offs incentives versus costs.

[1]The career path of Brian C. Voegele is a good example of the job ladder: Since his graduation from university, he started as a tax manager in the accounting firm, Ernst & Young LLP. He then was employed by Transocean, Inc. from 1989 to 2005, where he was promoted step by step from a director of tax to the vice president of finance. In June of 2005, he started to serve as the vice president and chief financial officer at Bristow Group Inc. The same year in December, he joined Pride International Inc. and has been a senior vice president and chief financial officer since then. Executive job ladder also exists at the CEO level. The career path of Richard C. Notebaert is as an example of the job ladder as described by Giannetti (2011): "Notebaert led the regional phone company Ameritech Corporation before its 1999 acquisition by SBC Communication Inc.; after, he held the top job at Tellabs Inc., a telecom-equipment maker; finally, in 2002, he became CEO of Qwest Communications International Inc."

During each period, the executive is poached with a probability. The incumbent and the poaching firms then engage in a Bertrand competition, where the executive either renegotiates with the current firm for a higher pay, or transits to the poaching firm. As a result, *job mobility* emerges. The model generates a job ladder along which an executive's compensation grows when other firms poach her. Broadly speaking, my model is a search-frictional version of the assignment models by Tervio (2008) and Gabaix and Landier (2008), or a dynamic moral hazard version of the on-the-job search model by Postel-Vinay and Robin (2002) and Cahuc et al. (2006).

The developed model gives a number of predictions that are confirmed in the data. First, when competing for an executive, a larger firm is capable of bidding higher. The Bertrand competition between two firms then yields that the executive takes the job with the larger employer. That is, job transitions are towards larger firms. Consequently, executives in large companies are less likely to be poached successfully and transit. These predictions are in line with the aforementioned stylized facts. Second, expecting the arrival of poaching offers, firms backload compensation to retain executives. The compensation growth over an executive's tenure is due to bidding by entrants who are smaller and thus unable to poach the executive. Together, the model predicts that, despite a lower job-to-job transition rate, executives in a large firm experience higher pay growth, consistent to the fact that I document in the data.

Overall, the bidding process gives rise to a job ladder where each rung is defined by a compensation level, and each outside offer is a chance to move up. A salient feature of the job ladder is that it is *bumpy*, which distinguishes my model from the classical job ladder models. The meaning of "bumpy" is two-fold. First, executive compensation depends on the realized productivity/output, which is a defining feature of an incentive contract. Because the output is stochastic, so is the compensation. Thus, along the ladder executive compensation can move up and down. Second, in the competition between incumbent and poaching firms, both sides are willing to bid higher if the executive turns out to be more productive. Therefore, it is ex-ante uncertain how high the next rung of the ladder will go — a high realized output leads to fierce bidding, whereas a low realized output leads to mild bidding.

The bumpy job ladder model highlights where executives' incentives originate. The executive puts in effort not only for higher pay from the current firm but also for a better poaching offer. The latter is a new source of incentives that the current firm gets "for free", called the *poaching offer incentives*. Thus, despite that the compensation is backloaded, incentives are not. With poaching offers, a firm may lose an executive to other firms. However, it also benefits because with poaching offer incentives, the current firm only needs to provide lower equity-based pay to motivate the executive. In this sense, the labor market is a double-edged sword — firms save on incentive pay but separate with the manager for a higher probability. This insight guides me in the first application of the model, where I explain an empirical puzzle on incentive pay using poaching offer incentives.

In the first application, I study the impact of firm size on the contractual incentive provision. I document a novel fact that larger firms tend to give executives a higher *proportion* of incentive pay. Specifically, if the firm size doubles, the fraction of incentive-related pay in total compensation increases by 4.74 percentage points (the median fraction is 65%). I refer to this fact the *firm-size incentive premium*, which, to my knowledge, is documented for the first time. Besides using the proportion of incentive pay, in Section 4, I also establish the incentive premium using wealth-performance sensitivities which are better metrics for incentives because the majority of contractual incentives come from the existing holdings of equity and option, rather than the newly granted.

Poaching offer incentives is the key to understand the firm size incentive premium. How strong the poaching incentives are depends not only on the benefits in dollars but also on how the executive perceives the incentives. I show that under a mild condition, poaching offer incentives are lower for executives in larger firms. The intuition is that larger firms are capable of bidding higher, thus those executives are expected to receive higher compensation in the future. Due to a diminishing marginal utility of wealth, the perceived poaching offer incentives are smaller. For this reason, large firms need a higher proportion of contractual incentives to compensate for the weakened poaching offer incentives.

This simple explanation is quantitatively relevant. I estimate the model using simulated method of moments, where I target the first and second order moments on total compensation, firm size, incentives, job turnovers, etc. Notably, the firm-size incentives premium is not targeted. Nevertheless, the estimated model quantitatively captures the premium well. Moreover, I demonstrated that poaching offer incentives account for a large portion of total incentives for small firms. The fraction of poaching offer incentives goes down to around 15% for medium-sized firms, and almost vanishes for top-sized firms.

In a second application, I use counterfactual exercises to quantitatively account for the sharp increases in *executive total and incentive pay*, the rising *inequality across executives*, as well as the stronger correlation *between firm size and compensation* since the mid-1970s (Frydman and Saks 2010). I show that, by increasing the poaching offer arrival rate while preserving all other parameters, the model can match data moments well, including the rising dispersion of executive compensation. The interpretation is that the managerial labor market is much thinner before the 1970s, which is consistent with the evidence provided by Frydman (2005), Murphy and Zabojnik (2007).

The rest of this section reviews the related literature. Section 2 presents stylized facts of executive job mobility that motivate my model. I then set up the model in section 3, where I characterize the optimal contract. Section 4 is the first application of the model where I document the firm-size incentive premium and use the model to explain it. Additionally, I structurally estimate the model and evaluate the quantitative predictions. Section 5 is the second application, and finally Section 6 concludes.

Literature review

This paper contributes to the literature that uses the competitive assignment model to explain the correlation between executive total compensation, performance-based incentives and firm size (Gabaix and Landier 2008, Tervio 2008, Eisfeldt and Kuhnen 2013, Baker and Hall 2004, Edmans et al. 2009, Edmans and Gabaix 2011). Since in these models matching is frictionless, job-to-job transitions are only implicit.

Another strand of literature explains the executive pay differentials in firm size using agency problems, e.g., Gayle and Miller (2009). Gayle et al. (2015) embed a multi-period moral hazard problem into a generalized Roy model and they find that the signal quality is worse in larger firms, which explains most of the pay differentials between small and large firms. The critical difference is that, in their model job-to-job transitions are in general not directed in firm size, whereas in my model there is a hierarchical job ladder from small to large firms. This explains why incentives of labor market competition contribute much more in my framework.

In explaining the rise of executive compensation in recent decades, my paper is an explanation based on job mobility. The literature highlights that the increases in compensation coincide with the increased mobility which is ultimately brought about by the increased importance of general managerial skills rather than firm-specific knowledge (Frydman 2005, Frydman and Saks 2010, Murphy and Zabojnik 2007). Giannetti (2011) develops a model to show that job-hopping opportunities help explain the increase in total pay and the structure of managerial contracts.

Regarding modelling, this paper links two literature strands. One strand is an extensive literature on optimal long-term contracts with private information and commitment frictions.[2] I build on this literature by embedding an optimal contracting problem into an equilibrium search model. In doing so, the outside environment is endogenized which significantly changes the optimal contract. In particular, the outside option of the executive is explicitly modelled as a result of random on-the-job search, and it depends on the executive's previous effort.[3] Another strand uses structural search models to evaluate wage dispersions, e.g., Postel-Vinay and Robin (2002), Cahuc et al. (2006), and Lise et al. (2016).[4] The managerial labor market is a particularly appropriate environment for the sequential auction framework — it happens very often that executives are contacted and "auctioned" by competing firms for promotion (Khurana 2004).

[2]See e.g., Rogerson (1985), Spear and Srivastava (1987), Phelan and Townsend (1991), Thomas and Worrall (1990), Phelan (1995), Edmans et al. (2012), Farhi and Werning (2013), Wang and Yang (2022).

[3]Grochulski and Zhang (2017) study the optimal mix of external and contractual incentives. Their model highlights that market-based incentives reduce the need for contractual incentives. Contrary to my model, they have a frictionless labor market with homogenous firms, and job-to-job transitions never happen in equilibrium. Wang and Yang (2022) study a dynamic principal-agent model and shed light on the interaction between moral hazard and voluntary/involuntary CEO turnovers. They model CEO's market value as an i.i.d. draw each period which captures the change of market conditions, whereas my model imposes explicit structure on the labor market.

[4]Abrahám et al. (2017) combine repeated moral hazard and on-the-job search to explain wage inequality in the general labor market. What distinguishes my model from theirs is that agents' productivity is persistent in my model. This feature gives rise to poaching offer incentives and explains the firm-size incentive premium. See also job ladder models with directed on-the-job search, e.g., Menzio and Shi (2010), Lentz (2014), Tsuyuhara (2016), etc.

2 Three stylized facts of executive job mobility

I construct a novel dataset of executive job mobility and compensation by linking BoardEx with standard data sources, including Execucomp, Compustat, and CRSP. This section first introduces the construction of the data and the definition of job-to-job transitions. Then I document three stylized facts that serve as the first pass for an appropriate theoretical model on executive job mobility.

2.1 Data and the definition of job-to-job transitions

Execucomp contains rich firm-side information and contract-related information for the top five to eight named officers of S&P listed firms. However, there is little information in this database that one can infer about job transition status. To supplement, BoardEx contains a full employment history for each executive. Particularly relevant to this study is that each job episode has the start/end dates, the employer, and the job title. This means that by linking Execucomp to BoardEx, I can trace where and when the executive worked before becoming a named executive in Execucomp, and what her subsequent jobs are.

To merge Execucomp and BoardEx, I use an executive's full name (first, middle and last names), date of birth, and whether the two datasets share the same job episodes, namely, during the period that the executive is a named officer in an Execucomp firm, the same job episodes show up in BoardEx database. When all three pieces of information are consistent, I confirm that the same executive has been identified in both databases and merge relevant observations. I successfully matched a sample of $35,088$ executives. Among these executives, $26,972$ executives had retired before 2016 (the year I collected the data); thus, I have their entire job histories. The total number of executive-fiscal year observations is $218,168$, spanning from 1992 to 2016. See Appendix B for summary statistics.

For each job episode in Execucomp, I define the end-of-job-spell status as either a *job-to-job transition* or an *exit* from the executive labor market. An executive has a job-to-job transition if she starts to work as an executive in another firm within *six months* after her Execucomp spell ends. Otherwise, she exits from the labor market. Based on this definition, the average yearly job-to-job transition rate is 4.98%, and the average exit rate is 6.91%. Relaxing the criterion from *six months* to *a year* necessarily increases the transition rate and lowers the exit rate but does not change the data patterns shown below.

To see that the differentiation between transition and exit does make sense, Figure 1 and 2 show how the two rates change with age. The transition rate increases before age 40, peaks between 42 and 46, and then goes down after age 50. In contrast, the exit rate peaks at age 65. These patterns are consistent with common sense.

Based on this definition of job-to-job transitions, I replicate some patterns documented in previous studies. For example, the transition rate is much lower for CEOs, and transitions can happen within or across industries. Among transitions that industry information is observable in the data[5], 1717 out of 2567 transitions (that is, 67%) are within an industry when the definition of industries is Fama-French 12, and 1407 (55%) when the definition is Fama-French 48. Besides, transitions between private and listed firms are also common, which is worth further examination in future studies.

2.2 The three facts

Fact 1. *Upon transit, executives tend to move to larger firms.*

Firm size is measured by market capitalization. In my sample, there are 9138 job-to-job transitions from a Compustat firm, and 2567 of them have the size information on both the original and targeted firms. Since a transition is included whenever firm size is observed, it is a reasonable sample of job-to-job transitions between publicly listed firms. I find that approximately 60% job-to-job transitions are associated with an increase in firm size. The fraction is stable across age groups and industries, as shown in Table 1. I further explore those transitions towards a smaller firm. It turns out that 20% of those cases are due to a title change from a non-CEO title to a CEO title, while this fraction is only 3.3% in transitions towards a larger firm. Figure 3 portrays the

[5] That is, the transition has to be between *S&P* firms, though the executive needs not to be a named officer in either the original firm or the targeted firm.

distribution of the change in firm size upon a transition. While most of the transitions are between firms of similar sizes, there are a lot of "leap" transitions where the targeted firm is much larger. This pattern lends support to my modelling of the managerial labor market where search is *random*.

Fact 2. *Executives from larger firms are less likely to move.*

As a first pass, figure 4 depicts the transition rates across firm size quantiles with a fitted line. The transition rate decreases from more than 6% at the 5th percentile to less than 3% at the 95th percentile. I then estimate a Cox model to show how firm size affects the hazard of a job-to-job transition. As shown in Table 2 columns (1) to (2), firm size is negatively associated with the hazard rate of job-to-job transitions in two regression settings. In column (1), *age, tenure, year, and industry dummies* are controlled. In column (2), I further add $\log(tdc1)$, *CEO, CFO, director, interlock dummies,* and firm performance metrics including *market book ratio (mbr)* and *operating profitability.* In particular, $tdc1$ is the total compensation, including salary, bonus, and new equity grants. Overall, it is a robust result that executives in larger firms are less likely to have job-to-job transitions.

Fact 3. *Executives in larger firms tend to experience higher pay growth.*

The pay-growth rate is measured by the first-order difference of $\log(tdc1)$. Column (3) in Table 2 presents the regression of $\Delta\log(tdc1)$ on firm size, controlling for the *last period total compensation*, *tenure*, *age*, and *year* $\times$ *industry dummies*. The estimates indicate that starting from the same total compensation level, for a 1% increase in firm size, the compensation growth rate increases by 11.2%. To put the number into context, note that the median pay-growth rate is 7%. Based on the estimates, if firm size doubles, the pay-growth rate increase to 15%. In column (4) estimated coefficient becomes even larger after further controlling for company performance metrics (*operating profitability*, *market-book ratio*), title dummies (*director*, *CEO*, *CFO*).

Table 1: Change of firm size upon job-to-job transitions

Panel A: All executives

Firm size proxy	*Total obs.*	*Size decrease obs.* (%)	*Size increase obs.* (%)
Market Cap	2567	985 (39%)	1582 (61%)
Sales	2617	1051 (40%)	1566 (60%)
Book Assets	2616	1038 (40%)	1578 (60%)

Panel B: Across age groups

Age groups	*Total obs.*	*Size decrease obs.* (%)	*Size increase obs.* (%)
≤ 40	100	34 (34%)	66 (66%)
[40, 45)	381	135 (35%)	246 (65%)
[45, 50)	701	262 (37%)	439 (63%)
[50, 55)	766	304 (40%)	462 (60%)
[55, 60)	420	179 (43%)	241 (57%)
[60, 65)	134	52 (39%)	82 (61%)
[65, 70)	28	7 (25%)	21 (75%)
≥ 70	6	1 (16%)	5 (84%)

Panel C: Across industries

Industries FF-12	*Total obs.*	*Size decrease obs.* (%)	*Size increase obs.* (%)
1 Consumer NonDurables	119	39 (33%)	80 (67%)
2 Consumer Durables	88	33 (38%)	55 (62%)
3 Manufacturing	281	98 (35%)	183 (65%)
4 Energy	120	58 (48%)	62 (52%)
5 Chemicals	71	30 (42%)	41 (58%)
6 Business Equipment	609	229 (38%)	380 (62%)
7 Telcom	60	20 (33%)	40 (67%)
8 Utilities	96	48 (50%)	48 (50%)
9 Wholesale and Retail	381	142 (37%)	239 (63%)
10 Healthcare and Drugs	197	89 (45%)	108 (65%)
11 Finance	314	115 (37%)	199 (63%)
12 Other	231	84 (36%)	147 (64%)

Table 2: Job-to-Job Transitions, Pay-growth and Firm Size

	Job-to-Job transition		$\Delta\log(tdc1)$	
	(1)	(2)	(3)	(4)
log(firm size)	-0.0517***	-.0522***	0.1117***	0.1436***
	(0.0066)	(0.0101)	(0.0090)	(0.0113)
log(tdc1)		-.0306*	-0.2898***	-0.3809***
		(0.0162)	(0.0200)	(0.0255)
CEO		-0.1032**		0.2914***
		(0.0511)		(0.0134)
CFO		-0.0369		0.0313***
		(0.0322)		(0.0054))
director		-0.8811***		0.1070***
		(0.0473)		(0.0113)
interlock		-0.5773***		-0.1297***
		(0.189)		(0.0238)
mbr		-0.0893***		0.0007
		(0.0105)		(.0047)
profitability		-.1191***		0.0172
		(0.0174)		(.0202)
age dummies	x	x	x	x
tenure dummies	x	x	x	x
year dummies	x	x		
indust dummies	x	x		
year $\times$ *indust dummies*			x	x
N	206,676	135,213	129,068	107,593
chi2 / adj. R^2	3262.38	3548.85	0.1573	0.2057

Note: Columns (1) and (2) estimate a Cox proportional hazards model where the event is a job-to-job transition, and the hazard function is stratified on Fama-French 48 industries. I control for age, tenure, and fiscal year dummies. Columns (3) and (4) estimate a linear regression of the first order difference in $\log(tdc1)$ on *lagged log(size)* and *lagged log(tdc1)*. Controls include *operating profitability*, *market-book ratio (mbr)*, *director* (whether the executive served as a director during the fiscal year), *CEO*, *CFO* (whether the executive served as a CEO or a CFO during the fiscal year), and *interlock* (whether the executive is involved in the interlock relationship). Besides, I control for age, tenure, and fiscal year times industry dummies. The standard error (clustered at the firm $\times$ fiscal year level) are shown in parentheses, and I denote symbols of significance by * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, respectively.

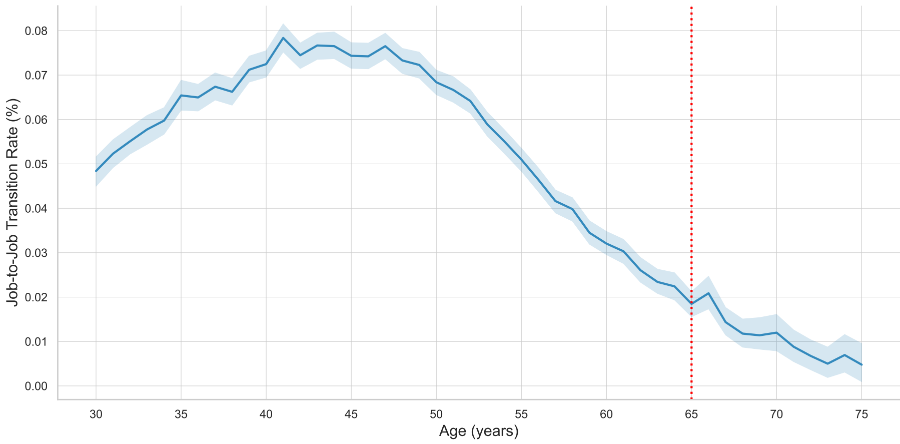

Figure 1: Job-to-job transition rate over age

Note: The figure depicts the estimates of job-to-job transition rates over age with the 95% confidence interval around the estimates. A *job-to-job transition* is defined as an executive leaving the current firm and starting to work in another firm within six months.

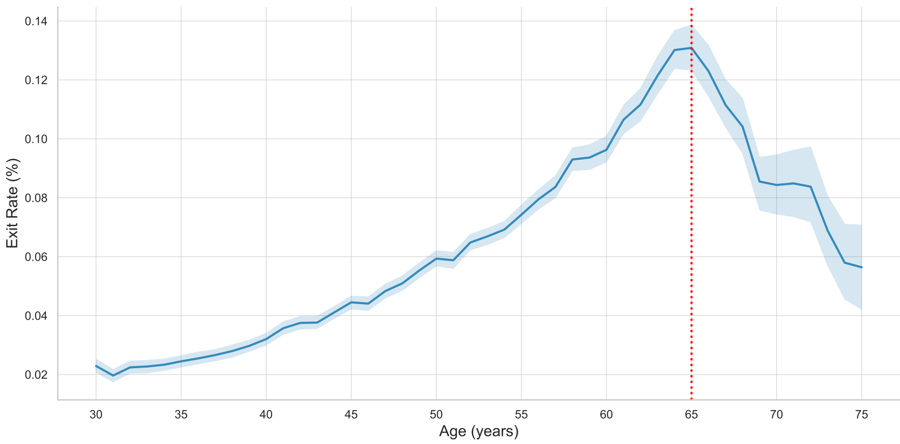

Figure 2: Exit rate over age

Note: The figure depicts the estimates of exit rates over age with the 95% confidence interval around the estimates. An *exit* is defined as an executive leaving the current firm and does not work in another firm for six months.

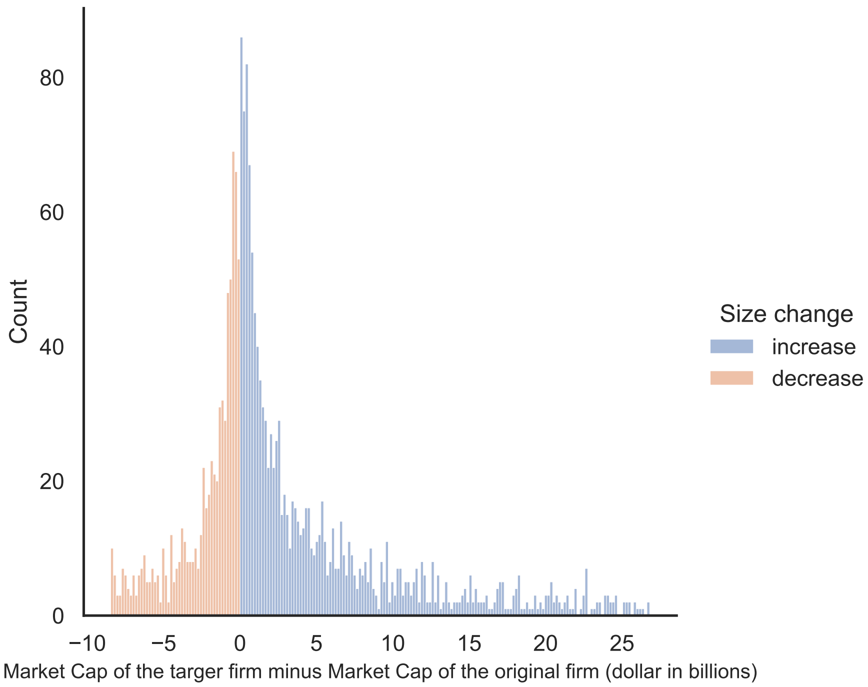

Figure 3: Distribution of changes in firm size upon a job-to-job transition

Note: This bar plot depicts the distribution of change in firm size (measured by market capitalization in billion dollars) in blue (increase in firm size) and red (decrease in firm size).

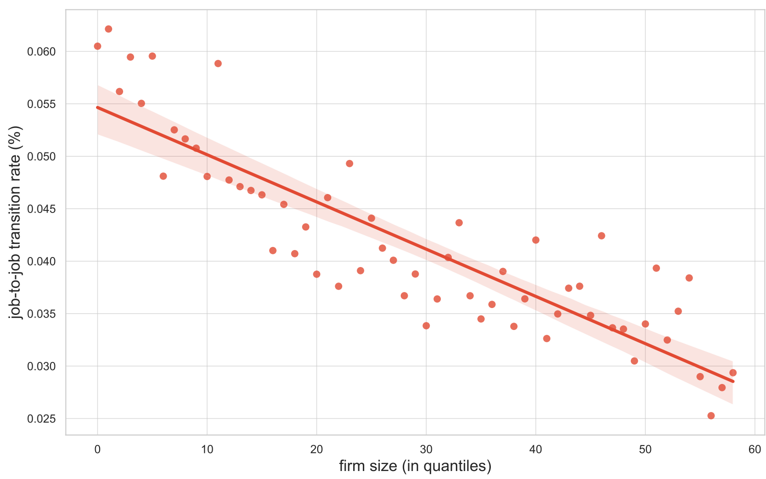

Figure 4: Job-to-job transition rate across firm size

Note: The figure depicts the estimates of job-to-job transition rates across 60 firm size quantiles (scatter points) and a fitted line.

3 The theoretical framework

In this section, I first construct a job ladder model and then characterize the optimal incentive contract. In subsection 3.3, I explain how the model generates a bumpy executive job ladder and how the model gives predictions that are consistent with the three stylized facts of Section 2.

3.1 Set-ups

Executives Time is discrete, indexed by t, and continues forever. There is a continuum of individuals. Each is either employed as an executive or looking for an executive position as a *candidate*. Individuals die with some probability. Once that happens, a "newborn" enters and becomes a candidate searching for an executive position. The model focuses on executives' on-the-job search and its influence on the compensation contract. Death and candidates are added to avoid that all executives are on the top of the job ladder in steady state.

Each individual aims to maximize expected lifetime utility:

$$\mathbb{E}_0 \Sigma_{t=0}^{\infty} (\beta \times (1-\eta))^t (u(w_t) - c(e_t)),$$

where $\tilde{\beta} \in (0,1)$ is the discount factor, $\eta \in (0,1)$ is the death probability, utility of consumption $u : \mathbb{R}_+ \to \mathbb{R}$ is twice differentiable, strictly increasing and concave with $\lim_{w \to 0} u'(w) = \infty$, $c(\cdot)$ is the dis-utility of effort. Effort e_t takes two values, $e_t \in \{0,1\}$ with 1 representing high effort and 0 shirking. Normalize $c(0)$to 0, and denote the cost of taking effort by $c = c(1) > 0$. Let $\beta = \tilde{\beta}(1-\eta)$ be the effective discount factor.

Managerial productivity. Executives are heterogeneous in an observable managerial productivity z, which takes on values in a finite set, $z \in \mathbb{Z} = \{z^{(1)}, z^{(2)}, ..., z^{(n_z)}\}$ with $\underline{z} = z^{(1)} < z^{(2)} < ... < z^{(n_z)} = \bar{z}$. This productivity is not (entirely) firm-specific and thus can be carried through job-to-job transitions between firms. All matches start with productivity $z^{(1)}$, then z evolves according to a Markov process. Formally, given a beginning-of-period productivity z, the executive chooses to take effort or not. At the end of the period, z' is realized and becomes the beginning-of-the-next-period productivity. z' has a probability $\gamma(z'|z)$ if $e = 1$, and $\gamma^s(z')$ if $e = 0$. Note that γ^s is assumed to be independent of z for simplicity. The process of z has two properties that are common in literature:

a. For each executive, z is positively correlated across time; therefore, a competent executive will likely be productive in the next period. This requires that γ is monotone such that for every non-decreasing function $h : \mathbb{Z} \to \mathbb{R}$, $\sum_{z' \in \mathbb{Z}} h(z')\gamma(z'|z)$ is also non-decreasing in z.

b. Making an effort today leads to higher chances of obtaining high productivity. This requires the likelihood ratio defined by $g(z'|z) \equiv \frac{\gamma^s(z'|z)}{\gamma(z'|z)}$ satisfy MLRP, i.e., $g(z'|z)$ is non-increasing in z'.

Firms Firms are different in time-invariant scale of asset which takes on values in a finite set, $s \in \{s^{(1)}, s^{(2)}, ..., s^{(n_s)}\}$ with $\underline{s} = s^{(1)} < s^{(2)} < ... < s^{(n)} = \bar{s}$. A match between an executive of productivity z and a firm of size s generates a flow of output $f(z,s)$ each period. f is strictly increasing and strictly concave. Concavity implies the presence of at least some complementarities between firm size and executive productivity. While managerial productivity z and output $f(z,s)$ are observable, effort e is not. Thus, a moral hazard problem arises.

Managerial labor market The managerial labor market is search frictional, which precludes the optimal assignments assumed in Gabaix and Landier (2008). Search is random. All individuals, employed or not, have a probability λ to sample a poaching firm of size s' from an exogenous distribution with probability $\tilde{p}(s')$. Since the only role of candidates is to maintain an active job ladder, I normalize that all candidates have the same continuation value U. Thus, when a candidate is matched, the firm offers a contract of value U. The candidate enters the next period as an employed executive. U is also the outside option value of an on-the-job executive. When an employed executive meets an outside firm, the incumbent and the outside firms engage in a Bertrand competition à la Postel-Vinay and Robin (2002) where the executive takes the job of the larger firm. Anticipate this subgame, in a renegotiation-proof contract, the firm would specify its counter-offers depending on the manager's productivity and poaching firm size.

Timing Consider an executive with a beginning-of-period productivity z and is currently matched to a firm with size s. Each period has three stages:

1. **Production and pay:** The executive contributes her beginning-of-period productivity z to production $f(z,s)$ and obtains a flow compensation w. With probability η, the executive dies. Otherwise, she proceeds to the next stage.

2. **Update productivity:** The executive chooses an effort level. Then a new productivity z' is drawn from $\gamma(z'|z)$ if she takes effort or from $\gamma^s(z')$ if she shirks. z' is the beginning productivity of the next period.

3. **Poaching offers:** With probability λ, the executive is poached by a firm of size s'. Then the contract is updated based on (z', s').

The compensation w, effort choice e, and the job-to-job transition decisions in each period are stipulated in the contract between the firm and the executive, defined on a proper state of the world, which we now turn to.

3.2 Contracting

Contractual environment To recursively write up the contracting problem, I use the executive's beginning-of-period expected utility, denoted by $V \in \mathbb{V}$, as a co-state variable to summarize the history of productivities and outside offers. This is a standard method in the dynamic contracts literature; see, e.g., Abreu et al. (1990). A dynamic contract defined recursively, is

$$\{e(V), w(V), W(z', s', V)|z' \in \mathbb{Z}, s' \in \mathrm{S} \text{ and } V \in \mathbb{V}\},$$

where e is the effort level suggested by the contract, w is the flow compensation, and W is the promised value given for a given state (z', s'). Since I will focus on publicly listed firms, all of which are relatively large, I impose that $\underline{s}$ is sufficiently high such that:

a. Given the cost of effort c, the benefits of high effort outweigh the cost of incentivizing an executive. This implies that $e = 1$ in any optimal contract.

b. Conditional on giving an executive the value U, the discounted sum of profits of a match is always positive. Under this assumption, dismissal is entirely captured by the exogenous parameter η.[6]

In the following, I first characterize competition for talents between firms, namely, the sequential auction, and then set up the contracting problem.

Sequential auction Let $\Pi(V,z,s)$ denote the discounted sum of profits for a firm with size s, matched to an executive with beginning-of-period productivity z and a promised value to the executive V. The maximum bidding value $\overline{W}(z,s)$ are defined by

$$\overline{W} = \sup\{W \in \mathbb{R} | \Pi(W,z,s) \geq 0\}.$$

The firm will separate with the executive (and the vacancy value is normalized to 0) if the latter demands higher than $\overline{W}$. I call $\overline{W}(z,s)$ the *bidding frontier* of the firm and highlight that it depends on (in fact, increases in) z and s.

The competition between the incumbent and poaching firms works as follows. When an executive from a firm of size s (hereafter firm s) meets a poaching firm of size s' (hereafter firm s'), the two firms enter a Bertrand competition. The best firm s can provide is a promised utility $\overline{W}(z',s)$. When $s' > s$, the executive accepts to move to firm s' which offers $\overline{W}(z',s)$. Any less generous offer by firm s' is successfully countered by firm s. If $s' \leq s$, the executive will stay at her current firm, and be "promoted" to continuation value $\overline{W}(z',s')$ that makes her indifferent between staying and joining firm s'. The above argument defines outside values of the executive contingent on each state (z',s'): $W(z',s') \geq \min\{\overline{W}(z',\tilde{s}), \overline{W}(z',s)\}$. Only when a contract satisfies these constraints that it is renegotiation-proof (Lentz 2014).[7]

The contracting problem In designing the contract, the firm chooses a current period compensation w, a set of promised values $W(z',s')$ for each realization of (z',s'). An executive always

[6]Adding dismissal to the quantitative analysis is straightforward. However, exploring the implications and matching them with facts on executive dismissal is out of the scope of this study.

[7]What distinguishes this model from the original sequential auction framework Postel-Vinay and Robin (2002) is that, in my model pay is not flat. Firms compete on a menu of values (a stream of pay in future) contingent on all possible future histories, which is summarized by $\overline{W}(z,s)$.

holds the outside value U, whether she receives an offer or not. This is equivalent to that during each period she receives an offer from a "virtual" firm. Let's denote the firm size by $s^{(0)}$, which is supposed to be much smaller than $\underline{s}$. Impose that this virtual firm has a bidding frontier $\overline{W}(z, s^{(0)}) \equiv U$ for all $z \in \mathbb{Z}$. Using this virtual firm, the distribution of poaching firms can be written as $p(s) = \mathbb{I}(s = s^{(0)})(1-\lambda) + \mathbb{I}(s \neq s^{(0)})\lambda\tilde{p}(s)$ for $s \in \mathbb{S} \equiv \{s^{(0)}, s^{(1)}, s^{(2)}, ..., s^{(n_s)}\}$. The expected discounted sum of future profits of the firm can be expressed recursively as

$$\Pi(V, z, s) = \max_{w, W(z', s')} \left\{ f(z, s) - w + \tilde{\beta}\mathbb{E}_{z', s' \leq s} \left[\Pi(W(z', s'), z', s) d\tilde{F}(s') | z \right] \right\}. \tag{BE-F}$$

subject to the promise keeping constraint,

$$V = u(w) - c + \tilde{\beta}\mathbb{E}_{z', s'} \left[W(z', s') | z \right], \tag{PKC}$$

the incentive compatibility constraint,

$$\tilde{\beta}\mathbb{E}_{z', s'} \left[W(z', s')(1 - g(z'|z)) | z \right] \geq c, \tag{IC}$$

and for all z' and s', the participation constraints of the firm

$$W(z', s') \leq \overline{W}(z', s), \tag{PC-F}$$

and finally the participation/renegotiation-proof constraints of the executive:

$$W(z', s') \geq \min\{\overline{W}(z', s'), \overline{W}(z', s)\}. \tag{PC-E}$$

I have used $\mathbb{E}$ to denote the expectation with respect to z' and s'. The objective (BE-F) is the Bellman equation of the firm, which includes a flow profit of $f(z, s) - w$ and the continuation value. Note that the continuation value of a firm is normalized to zero if the match separates, either because the executive dies, which happens with probability η, or the executive transits to another firm, which happens if the poaching firm is larger. The promise-keeping constraint (PKC) ensures that the choices of the firm honour the promises made in previous periods to deliver V to the executive. The incentive compatibility constraint (IC) says the continuation value of taking effort is higher than that of shirking. This creates incentives for the executive to pursue the shareholders' interests rather than her own. Note that with the term $1 - g(z'|z)$, the left-hand side is the difference of continuation values between taking effort and shirking. Finally, the participation constraints are stated in (PC-E) and (PC-F). The firm stays in the relationship as long as the promised value is no more than $\overline{W}(z', s)$. The sequential auction pins down the outside value of the executive, i.e., the right-hand side of (PC-E). When there is no poaching firm, $\overline{W}(z', s') = U$. I provide a characterization of the optimal contract in the following proposition.

Proposition 1. *$\Pi(V, z, s)$ is differentiable, strictly decreasing and strictly concave in V, strictly increasing in z and s. Given a beginning-of-period state (V, z, s), the optimal contract follows:*

(i) w is determined by

$$\frac{\partial \Pi(V, z, s)}{\partial V} = -\frac{1}{u'(w)}; \tag{1}$$

(ii) define $W(z')$ as the continuation value determined by

$$\frac{\partial \Pi(W(z'), z', s)}{\partial W(z')} - \frac{\partial \Pi(V, z, s)}{\partial V} = -\mu(1 - g(z'|z)), \tag{2}$$

where μ is the multiplier of IC constraint, then

$$W(z', s') = \begin{cases} \overline{W}(z', s) & \text{if } \; s' \geq s \text{ or } W(z') > \overline{W}(z', s), \\ \overline{W}(z', s') & \text{if } \; s' < s \text{ and } W(z') < \overline{W}(z', s'), \\ W(z') & \text{otherwise.} \end{cases} \tag{3}$$

Proof. See Appendix A. □

Equation (1) says that the current period flow compensation w is directly linked to the promised continuation utility V by equating the principal's and agent's marginal rates of substitution between the present and future compensation. A higher V is associated with a higher flow compensation w. Therefore, if an executive is rewarded with a high continuation value, then in the following periods, she will obtain a series of high flow pay. Suppose there are no participation constraints, then equation (2) says that the continuation utility $W(z')$ only changes to induce high effort. In the extreme case that IC constraint is not binding ($\mu = 0$), $W(z') = V$ keeps constant. Thus, w would also be constant over time. Even when $\mu \neq 0$, according to (2), a higher V tends to induce a higher $W(z')$. Therefore, the optimal dynamic contract has memory, which is a property of the classical infinite repeated moral hazard model (Spear and Srivastava 1987).

When the outside offers are realized such that the participation constraint is binding, the contract dispenses of the history dependence, and the continuation value depends on the current state. This is what Kocherlakota (1996) calls *amnesia*. As stated in (3), when the poaching firm is larger $s' \geq s$, the continuation value equals the bidding frontier of the current firm $W(z', s') = \overline{W}(z', s)$; when the poaching firm is smaller $s' < s$, the continuation value depends on whether the bidding frontier of the outside firm $\overline{W}(z', s')$ can improve upon $W(z')$.

It's worth emphasizing that even when the participation constraint is binding, amnesia of the optimal contract is not "complete" — although $\overline{W}$ does not depend on the previously promised utility V, it does depend on the executive's productivity z' which is stochastically determined by the effort in previous periods. In other words, the boundaries of participation constraints still carry the memory of previous effort. Therefore, the bidding between the incumbent and the poaching firm not only has a level effect, but also generates extra incentives, which plays a crucial role in the following analysis. Finally, given that $\Pi(\cdot)$ is continuous and increasing in z and s, $\overline{W}(z, s)$ is well defined.

3.3 The bumpy job ladder

With the characterization of the optimal contract in hand, we are ready to illustrate how the job ladder works. To be precise, the job ladder is defined on the executive continuation value W. Climbing the job ladder implies that the executive obtains a higher continuation value. According to (1), given z and s, there is corresponding flow pay w. Thus, the job ladder can be equivalently defined by total compensation and firm size.

Predictions on job transitions Consider an executive who works in a firm of size s and has a continuation value W. Divide possible poaching firms into two sets:

$$\mathcal{M}_1(s) \equiv \{s' \in \mathbb{S} | s' > s\},$$

$$\mathcal{M}_2(z, s, W) \equiv \{s' \in \mathbb{S} | s' \leq s, W < \overline{W}(z, s')\}.$$

Then the promise-keeping constraint can be rewritten as follows:

$$V = u(w) - c + \tilde{\beta} \sum_{z' \in \mathbb{Z}} \Bigg[\lambda \sum_{s' \in \mathcal{M}_1} p(s')\overline{W}(z', s) + \lambda \sum_{s' \in \mathcal{M}_2} p(s')\overline{W}(z', s') + \Big(1 - \lambda \sum_{s' \in \mathcal{M}_1 \cup \mathcal{M}_2} p(s')\Big) W(z') \Bigg] \gamma(z'|z), \tag{PKC$'$}$$

(PKC′) explains how, by using poaching offers, an executive climbs the ladder; and how by back-loading compensation, the firm retains her. If the outside firm belongs to $\mathcal{M}_1$, the executive transits and receives the full surplus of her previous employer $\overline{W}(z', s)$. If the outside firm belongs to set $\mathcal{M}_2$, the incumbent firm offers a higher continuation value $\overline{W}(z', s')$ and the executive stays. Other offers are simply dropped since they are not strong enough to improve upon $W(z')$, as shown in the third term in brackets.[8]

The model has following predictions. First, by construction, a job-to-job transition occurs only when the poaching firm $s' \in \mathcal{M}_1$. Thus, executives tend to move to larger firms (Fact 1 in section 2). Second, search is random, and all executives draw poaching firms from the same distribution. Executives in larger firms are less likely to meet an even larger firm and transit (Fact 2). Instead, an outside firm is more likely to belong to $\mathcal{M}_2$. Countering these outside offers lead to higher compensation growth in large firms (Fact 3).

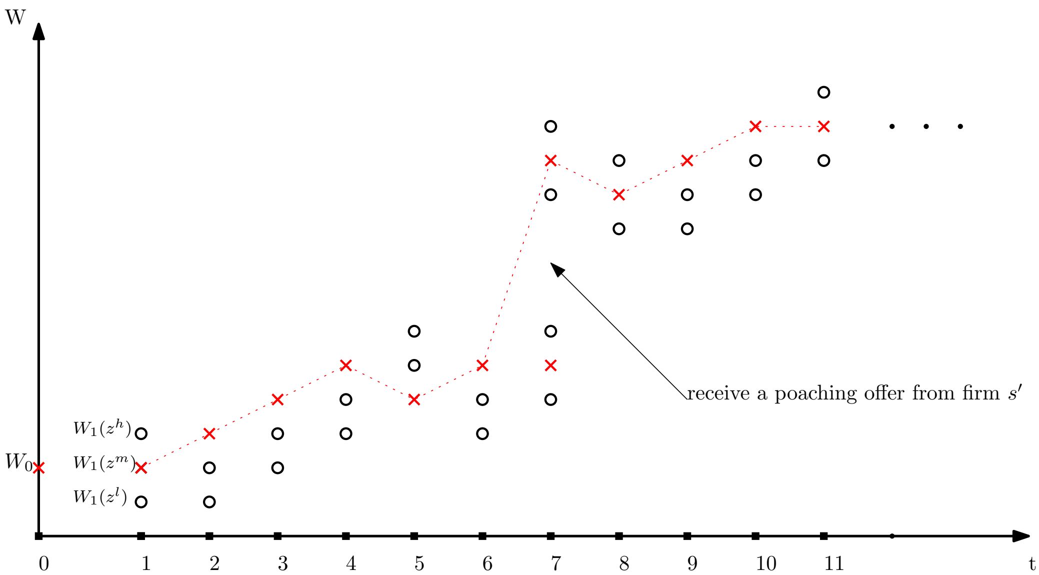

Figure 5: An illustration of the bumpy job ladder

[8]There exist cases where the last two terms can not be separated. Here I omit that for a concise illustration.

A bumpy ladder and poaching offer incentives The ladder is bumpy because, no matter in a promotion or in a transition, how high the next step depends on a stochastic realization z' (which in turn depends on the effort taken by the executive). This is a defining feature of incentive contracts. To illustrate, consider an executive that starts to work in period $t = 0$ with a continuation value W_0. In each period, there are three possible realizations of z: $z^l < z^m < z^h$. Correspondingly, the incentive contract indicates for each outcome the continuation value in period t: $\{W_t(z^h), W_t(z^m), W_t(z^l)\}$. In Figure 5, the menu of continuation value for each z' is represented by a circle/cross, and the cross represents the continuation value of the realized z'. The red dotted line illustrates how the continuation value changes over time.

At $t = 1, z^m$ is realized. This pins down the continuation value $W_1(z^m)$. If there were no moral hazard issue, the optimal contract would indicate $W_2 = W_1(z^m)$, a particular case of Proposition 1 with $\mu = 0$. With moral hazard, the optimal contract implies three possible values for $t = 2$, namely $W_2(z^j), j = h, m, l$. Depending on the outcome, the executive may gradually gain higher values, as shown in figure 5 from $t = 1$ through $t = 4$, or a lower value if the outcome turns out to be low, as shown at $t = 5$. At $t = 7$, there is a poaching offer which increases the continuation value by Bertrand competition between the current and the poaching firms. Depending on whether s' is larger than s or not, there can be a job transition ($s' > s$) or an internal promotion ($s' \leq s$).

Importantly, the poaching offer itself is also a menu of continuation values from an ex-ante perspective: $\{\overline{W}(z^h, \tilde{s}), \overline{W}(z^m, \tilde{s}), \overline{W}(z^l, \tilde{s})\}, \tilde{s} = \max\{s, s'\}$, which features that the offer is more generous if the manager has higher productivity z'. Therefore, despite that compensation is backloaded, incentives from the backloaded compensation remain to exist. Executives anticipate the incentives from poaching offers — taking effort today is not only for higher pay in the current employer but also for higher bids should a poaching firm arrive. In the next section, I use poaching offer incentives to explain the firm-size incentive premium.

4 Application I: Why do larger firms pay more incentives to executives?

In this section, I present and use the model to account for a novel empirical fact, the *firm-size incentive premium*. I first document the premium and show that it is closely related to how active the managerial labor market is. Then I provide a sufficient condition under which the model gives the exact prediction. Finally, I structurally estimate the model and evaluate if the model can capture the premium quantitatively.

4.1 Firm-size incentive premium

The firm-size incentive premium refers to that, executives of larger firms receive *proportionally* higher contractual incentives in total compensation. Here the contractual incentives refer to those brought by holdings of firm stocks and options, which is in contrast to the poaching offer incentives due to labor market, although strictly speaking, both are part of my "theoretical contract". The premium, therefore, reveals the firm-size impact on the *structure* of the contract.

A simple way to proceed is to correlate firm size with the proportion of incentive-related pay (e.g., newly granted shares) in total compensation, which is one way adopted below. However, the vast majority of executive incentives come from the previously granted stocks and options, therefore, I follow the literature and use incentives in executives' firm-related wealth.

The primary measurement of contractual incentives that I use is the wealth-performance sensitivity (denoted by *inc*), which is defined as the dollar change in the executive's firm-related wealth for 100 percentage increase in firm value. In other words, *inc* is the dollar value of the executive's share in the sense that options are converted into stock equivalents according to their delta.[9] For other variables, as in Section 2, firm size is measured by market capitalization, and total compensation is measured by *tdc*1.

One (informal but intuitive) way to think of the previously granted equities is as follows. The expected value of these equities has been included in the utilities of previous periods. Thus, in

[9] *inc* is provided by Coles et al. (2006) who computed based on the method of Core and Guay (2002).

the current period, we may normalize the expected utility of those previous grants to zero. Still, these holdings generate fluctuations in executives' realized income. Therefore, $inc/tdc1$ is the "effective" proportion of incentive pay in the total compensation. In the following, I regress inc on firm size with $tdc1$ controlled. This way, I allow for a more general regression model.

In Table 3, column (2) shows that inc is positively correlated to firm size — if firm size doubles, inc increases by 27%. To put this number into context, take the median 4.05 million dollars of inc in the year 2012. If the firm size doubles, then inc would increase to 5.14 million dollars. This change is substantial given a median annual total pay of 1.32 million dollars.

The size incentive premium also shows up if contractual incentives are calculated in other ways. In column (4), I measure contractual incentives by scaled wealth-performance sensitivity proposed by Edmans et al. (2009). It equals inc divided by $tdc1$, and I denote it by inc^s. Similar as inc, when using inc^s, total compensation is controlled to reflect the impact on the "proportion" of incentives in the contract. In column (5), I use the fraction of incentive pay in $tdc1$ (only incentives in the flow pay) and I denote it by inc^f. Incentive pay includes bonuses, restricted stock grants, option grants, and other long-term incentive payouts. Either way, the analysis identifies a firm-size incentive premium. For example, column (5) says that if the firm size doubles, the percentage of incentive pay in total compensation increases by 4.76 percentage points (the median fraction is 65%).

To connect to the literature, I also present in columns (1) and (3) the estimates when $tdc1$ is not controlled. In these regressions, the coefficients of firm size reflect the impact on the absolute amount of incentives. Interestingly, depending on the way of measuring incentives, the coefficients of firm size differ a lot. Consistent to Edmans et al. (2009), the scaled wealth-performance sensitivity has a lower correlation with firm size than the unscaled inc. However, after controlling for $tdc1$, both size coefficients become moderate.[10]

Firm-size premium and executive labor market The following evidence highlights the link between firm-size premium and the managerial labor market using the variation across industries.

[10] Edmans et al. (2009) showed that inc^s is not correlated with firm size. They use a sample of CEOs from the top 200 or 500 firms, which is guided by their theory, whereas my sample contains all named officers of S&P firms.

Table 3: Size premia

	$\log(inc)$		$\log(inc^s)$		inc^f		$\log(inc)$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(firm size)	0.5824***	0.3451***	0.1368***	0.2554***	6.878***	0.3168***	0.3247***	0.3164***
	(0.0135)	(0.0231)	(0.0089)	(0.018)	(0.1593)	(0.0029)	(0.0036)	(0.0029)
$\log(tdc1)$		0.6401 ***		-0.3042***		0.6919 ***	0.6871 ***	0.6843 ***
		(0.030)		(0.0291)		(0.046)	(0.0055)	(0.0046)
log(firm size) × *JJ rate*						0.7165***		
						(0.1054)		
log(firm size) × *gai*							0.055***	
							(0.0112)	
log(firm size) × *inside CEO*								-0.0866***
								(0.0196)
Obs.	146,747	128,006	127,964	127,964	165,435	128,00	79,476	128,008
adj. R^2	0.4458	0.5236	0.2283	0.2761	0.2663	0.4862	0.4828	0.4854

Note: This table reports firm-size incentive premia (columns 1 to 5) and how they are related to the managerial labor market (columns 6 to 8). Firm size is measured by the market capitalization defined by the common shares outstanding times the fiscal year close price. Executives' compensation level is measured by total annual flow compensation *tdc*1, including the sum of salary and bonus, the value of restricted stocks and options granted, and the value of retirement and long-term compensation schemes. The dependent variable in columns 1 to 2, 6 to 8 is $\log(inc)$ where *inc* is the dollar change in firm-related wealth for a 100 percentage points change in firm value. The dependent variable in columns 3 and 4 is the log of scaled inc, which is calculated by $inc^s = \frac{inc}{tdc1}$. The dependent variable in column 5 is the fraction of incentive pay in total flow compensation: $inc^f = incentive\ pay/tdc1 \times 100$, where incentive pay is the sum of bonus, other annual, restricted stock grants, option grants, and LTIP payouts. All regressions control dummies for age, tenure, and year times industry. The standard errors (clustered at the firm × fiscal year level) are shown in parentheses, and we denote symbols of significance by $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$, respectively.

An industry is an appropriate sub-labor market since more than 60% job-to-job transitions are within industry (see section 2). I use three proxies to measure the activeness of a sub-managerial labor market. The first proxy is the job-to-job transition rate for each industry year (Fama-French 48 industries and fiscal years). A job-to-job transition is defined if the executive starts to work in another firm within six months. The second proxy *gai* is the mean of the general ability index of CEOs at the industry-year level. The general ability index itself is the first principal component of five proxies that measure the generality of the CEO's human capital based on the CEO's lifetime work experience. The last proxy *inside CEO* is the percentage of insider CEOs in the industry in which the firm operates. It counts for all new CEOs between 1993 and 2005 using Fama-French 48-industry groups.[11] In Table 3 columns (6) to (8), I examine how the interaction terms are associated with *inc*. The results are not ambiguous. All interaction terms are significant, and the signs confirm that firm size premia are higher in industry-years where the executive labor market is more active.

As a last piece of evidence, I show that the size incentive premium decreases as executives approach retirement age. Starting from Gibbons and Murphy (1992), *age* has been used as an indicator for career concerns. The older the executive is, the less influential the managerial labor market is on the incentive contract design. If the size incentive premium is caused by the managerial labor market, we expect the incentive premium to decrease with age. This is confirmed, as shown in figure 6, that the size incentive premium starts at 0.652 at age 35 and gradually goes down to around 0.35 after age 50. This pattern holds with or without controls.

4.2 Explaining firm size incentive premium

As demonstrated in Section 3, the possibility of receiving poaching offers generate incentives, referred to as the poaching offer incentives that substitute for the contractual incentives. If poaching offer incentives decrease in firm size, then to offset this loss of poaching incentives, larger

[11] *insider CEO* is provided by Martijn Cremers and Grinstein (2013). *gai* is provided by Custódio et al. (2013). The five proxies to measure general ability of CEO's are: the number of positions that the CEO performed during his/her career, the number of firms where a CEO worked, the number of industries at the four-digit SIC level where a CEO worked, a dummy variable that equals one if a CEO held a CEO position at another firm, and a dummy variable that equals one if a CEO worked for a multi-division firm.

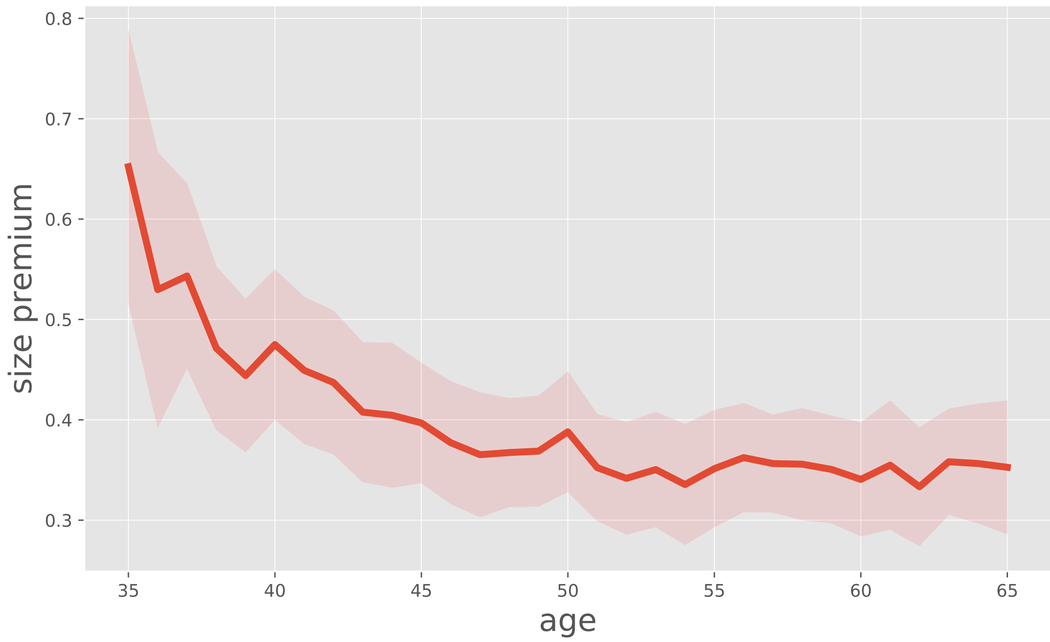

Figure 6: Size premium in performance-based incentives decreases in age

Note: The figure depicts the firm size incentive premium in *inc* at each age from 35 to 65. They are the estimated coefficients of the interaction terms between *age dummies* and *log(firm size)* in the following regression

$$\mathit{log(inc)}_{it} = \Phi' \mathit{age\ dummies}_{it} \times \mathit{log(firm\ size)}_{it} + \Psi' X_{it} + \epsilon_{it},$$

where i denotes an executive, t denotes the fiscal year, *age dummies* is a set of dummies for each age from 35 to 65, *firm size* is measured by the market capitalization, X denotes a vector of control variables and a constant term. We control for total compensations *log(tdc1)*, dummies for executive tenure, age, fiscal year times industries. A 95% confidence interval is plotted using the standard error clustered on firm $\times$ fiscal year.

firms need to offer more contractual incentives using equity-based pay.

In the following, I assume that z' and s' follow continuous distributions on the support $[\underline{z}, \overline{z}]$ and $[\underline{s}, \overline{s}]$, with cumulative density functions denoted by $\Gamma(z, z')$ for z' if $e = 1$, $\Gamma^s(z')$ if $e = 0$, and $\tilde{P}(s')$ for s', respectively. Further denote the cumulative density function of the mixture distribution of firm size by: $P(s') = \mathbb{I}(s = s^{(0)})(1 - \lambda) + \mathbb{I}(s \neq s^{(0)})(1 - \lambda + \lambda \tilde{P}(s))$. It is convenient to think of these distributions as the limits of the discrete distributions in the benchmark when $n_z, n_s \to \infty$, with $\underline{z}, \overline{z}, \underline{s}, \overline{s}$ fixed. I further define an *incentive operator* $\mathcal{I}$ which calculates the incentives an executive receives from any scheme $\{W(z')\}_{z' \in \mathbb{Z}}$, i.e., the difference in utilities between taking effort and shirking:

$$\mathcal{I}\Big(W(z')\Big) \equiv \int_{z'} W(z')\Gamma(z, dz') - \int_{z'} W(z')\Gamma^s(dz').$$

Using $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{I}(\cdot)$, the IC constraint can be rewritten as

$$\underbrace{\lambda \int_{s' \in \mathcal{M}_1} P(ds') \mathcal{I}\Big(\overline{W}(z', s)\Big)}_{\text{poaching incentives I}} + \underbrace{\lambda \int_{s' \in \mathcal{M}_2} \mathcal{I}\Big(\overline{W}(z', s')\Big) P(ds')}_{\text{poaching incentives II}}$$
$$+ \underbrace{\Big(1 - \lambda \int_{s' \in \mathcal{M}_1 \cup \mathcal{M}_2} P(ds')\Big) \mathcal{I}\Big(W(z')\Big)}_{\text{contractual incentives}} \geq c/\beta, \tag{IC$'$}$$

On the left-hand side, I decompose and label incentives by origins. Some incentives are brought by poaching offers when $s' \in \mathcal{M}_1$ or $\mathcal{M}_2$ (poaching incentives I and II). The rest are due to performance-related pay when there is no (strong) poaching offer (contractual incentives).

Suppose that the size of the incumbent firm marginally increases from s to $\tilde{s} > s$. The only change to the left-hand side of (IC$'$) is that in the poaching incentives I, $\mathcal{I}\Big(\overline{W}(z', s)\Big)$ becomes $\mathcal{I}\Big(\overline{W}(z', \tilde{s})\Big)$. Since $\tilde{s} > s$, $\tilde{s}$ has a higher bidding frontier $\overline{W}(z', \tilde{s}) > \overline{W}(z', s)$. By diminishing marginal utility, it can be that the incentives from a higher bidding frontier are lower: $\mathcal{I}\Big(\overline{W}(z', s)\Big) > \mathcal{I}\Big(\overline{W}(z', \tilde{s})\Big)$. This holds if the utility function is sufficiently concave. Then, to satisfy the IC constraint, firm $\tilde{s}$ needs to provide more incentives in the form of performance-based pay than firm s provides. This explains the observed firm-size incentive premium. Proposition 2 gives a sufficient condition for how concave u is required, under the assumption that $\Pi(V, z, s)$ is smooth in z and s.

Proposition 2. *Suppose z' and s' take on values in an interval, $z' \in \mathbb{Z} = [\underline{z}, \overline{z}]$, $s' \in \mathbb{S} = \{s^{(0)}\} \cup [\underline{s}, \overline{s}]$, and $\Pi(V, z, s)$ is strictly increasing and continuously differentiable in z and s. Then poaching offer incentives decrease in the current firm size s if u is sufficiently concave:*

$$-\frac{u''}{u'} \geq \sup_{z \in \mathbb{Z}, s \in \mathbb{S}} \left[\frac{f_{zs} + \kappa'(s)}{f_s(f_z + \kappa(s))} \right]. \tag{4}$$

where $\kappa(s)$ is a positive-valued bounded function that is increasing in s.

Proof. See Appendix A. □

To understand the intuition, first notice that $\mathcal{I}\Big(\overline{W}(z', s)\Big)$ is simply a weighted sum of $\frac{\partial \overline{W}(z', s)}{\partial z'}$ over the domain of z' — the steeper $\overline{W}(z', s)$ is with respect to z', the higher the incentives are to induce effort. So it would be sufficient to show $\frac{\partial \overline{W}(z', s)}{\partial z'}$ decreases in s. It follows from implicit differentiation that

$$\frac{\partial \overline{W}(z,s)}{\partial z} = -\frac{\partial \Pi(z,s,\overline{W})/\partial z}{\partial \Pi(z,s,\overline{W})/\partial \overline{W}} = \frac{f_z(s)+\kappa(s)}{1/u'(\overline{w})}, \tag{5}$$

where $\overline{w}$ is the flow compensation corresponds to $\overline{W}$. The numerator shows that z contributes to the flow profit of this period f_z as well as to future profits captured by $\kappa(s)$, where $\kappa(s)$ adjusts for the possibility that the executive may be separated from the firm. The denominator follows directly from the optimal contract condition (1).

There are two opposing forces shaped by s. On the one hand, s contributes to production, thus the numerator is increasing in s. On the other hand, a larger firm is able to bid higher, then $\overline{w}$ increases in s, making $u'(\overline{w})$ lower; thus, the denominator is also increasing in s.[12] The second force dominates when the utility function has enough concavity, as stated in proposition 2. In numerical exercises, I use a CRRA utility function $u(w) = \frac{w^{1-\sigma}}{1-\sigma}, \sigma > 0$, and a multiplicative production function $f(z,s) = \alpha_0 s^{\alpha_1} z, \alpha_0 \in (0,1), \alpha_1 \in (0,1]$.[13] Given that the production function has a strong firm-size effect, one may expect that u has to be very concave to generate firm-size incentive premia. However, I find that in fact $\sigma > 1$ is sufficient to have poaching offer incentives decrease in firm size.[14]

4.3 Structural estimation

I estimate the model parameters using Simulated Methods of Moments. The moments are partly coefficients from auxiliary regressions, so the approach could alternatively be presented as Indirect Inference. I first describe the numerical methods, the model specifications, and the moments used for identification. Specifically, I do not explicitly target the firm-size incentive premium.

[12] Using that $\Pi(\overline{W}, z, s) = 0$, we immediate have

$$\overline{w}(s) = f(z,s) + \tilde{\beta} \int_{z'} \int_{s' \le s} \Pi(W(z',s'),z',s) P(ds') \Gamma(z,dz'),$$

which is strictly increasing in s.

[13] The multiplicative form is standard in this literature, see, e.g., Gabaix and Landier (2008). The production function allows the executive's effort "roll out" across the entire firm up to a scale of α_0. y has constant return to scale if $\alpha_1 = 1$, and decreasing return to scales if $\alpha_1 < 1$.

[14] This required concavity is not demanding. Indeed, it is consistent with σ values that are calibrated or estimated in related literature. For example, the calibration study on CEO incentive pay by Hall and Murphy (2000) uses σ between 2 and 3; the series of calibration exercises on CEO incentive compensation convexity starting from Dittmann and Maug (2007) are based on $\sigma > 1$; Balke and Lamadon (2022) estimate $\sigma = 1.68$ using the matched employer-employee data from Sweden for the general labor market.

Still the model quantitatively captures it well.

Numerical method To solve the contracting problem, one needs to find the optimal promised values in each state of the world for the next period. This becomes infeasible as soon as reasonable supports are considered for $\mathbb{Z}$ and $\mathbb{S}$. Instead of solving for promised values directly, I use the recursive Lagrangian techniques developed in Marcet and Marimon (2019) and extended by Mele (2014). Under this framework, the optimal contract can be characterized by maximizing a weighted sum of the lifetime utilities of the firm and the executive, where in each period the social planner optimally updates the Pareto weight of the executive to enforce an incentive compatible allocation. This Pareto weight becomes a new state variable that "recursifies" the dynamic agency problem. In particular, this endogenously evolving weight summarizes the contract's promises according to which the executive is rewarded or punished based on performance and poaching offers. Ultimately, solving an optimal contract is to find the sequence of Pareto weights that implements an incentive-compatible allocation. Once these weights are solved, the corresponding utilities can be recovered. This technique improves the speed of computation and makes the estimation feasible.

Model specification and parameters I estimated the model fully parametrically and made several parametric assumptions. As mentioned earlier, I use the constant relative risk aversion utility function $u(w) = \frac{w^{1-\sigma}}{1-\sigma}$, and a production function $f(z,s) = e^{\alpha_0} s^{\alpha_1} z, \alpha_0 < 0$. I model the process of productivity by an $AR(1)$ process,

$$z_t = \rho_0(e) + \rho_z z_{t-1} + \epsilon_t,$$

where ϵ follows a normal distribution $N(0, \sigma_\epsilon)$, and the mean for effort level $e = 0$ is normalized to zero. The process is transformed to a discrete Markov Chain using Tauchen (1986) on a grid of 6 points. The choice of grid points is for speed of estimation. The simulated moments are very robust to this choice. Furthermore, I set the sampling distribution of firm size $\tilde{P}(s)$ a truncated log-normal distribution with an expectation of μ_s and standard deviation of σ_s.[15] Finally, the

[15]The upper and lower bounds of the truncated normal distribution are calibrated to be the 0.99 and 0.01 quantile of market capitalization in data.

discount rate β is set to be 0.9 for the model is solved annually. I set the number of grid points for the Pareto weight to be 50 and for firm size s to be 20. Table 4 lists the complete set of parameters.

Table 4: Parameters

Parameters	*Description*
η	death probability
λ	offer arrival probability
ρ_z	$AR(1)$ coefficient of productivity shocks
μ_z	mean of productivity shocks for $e = 1$
σ_z	standard deviation of productivity shocks
μ_s	mean of $\tilde{P}(s)$
σ_s	standard deviation of $\tilde{P}(s)$
c	cost of effort
σ	relative risk aversion
α_0, α_1	production function parameters

Moments and identifications I now make a heuristic identification argument that justifies the choice of moments used in the estimation. Firstly, for the identification of the productivity process, the exit rate, and the offer arrival rate, there are direct links between the model and the data. The exit rate directly informs η. Likewise, the incidence of job-to-job transitions is monotonically related to λ. The parameters of the productivity process, namely ρ_z, μ_z and σ_ϵ, are informed directly by the estimates of an AR(1) process on the profitability of each firm-executive match,

$$profit_{it} = \beta_0 + \rho_z profit_{it-1} + \epsilon_{it,0},$$

where i represents the executive-firm match, and t represents the year.

Secondly, the two parameters governing the job offer distribution, μ_s and σ_s, are disciplined by the mean and variance of log firm size. Given $\lambda > 0$, the higher μ_s is, the more likely that executives can transit to larger firms, and the larger mean of $\log(size)$. Similarly, the higher σ_s is, the more heterogeneous the poaching firms are, and both the mean and variance of $\log(size)$ increase.

Thirdly, regarding the production function, α_0 is mainly determined by the level of total compensation, and α_1 is determined by the relationship between firm size and total compensation. Therefore, α_0 and α_1 are identified by the mean and variance of $\log(tdc1)$ and $\beta_{tdc1-size}$ in following regression of $\log(tdc1)$ on $\log(size)$:

$$\log(tdc1_{it}) = \beta_1 + \beta_{tdc1-size} \log(size_{it}) + \epsilon_{it,1}.$$

The final part of the identification concerns the parameters σ and c. These parameters govern the level of incentives and how the incentives change with the compensation level. To be consistent with the variable *inc* in the data, I construct in the simulated data an *inc* variable defined by the dollar change in pay for a percentage change in productivity. I use the mean and variance of $\log(inc)$ to inform the effort cost c. To discipline σ, I run regression

$$\log(inc_{it}) = \beta_2 + \beta_{inc-tdc1} \log(tdc1_{it}) + \epsilon_{it,2},$$

and use $\beta_{inc-tdc1}$ to inform σ. The higher σ is, the larger $\beta_{inc-tdc1}$ is.

Table 5: Moments and Estimates

Moment	*Data*	*Model*	*Estimate*	*Standard Error*
Exit rate	0.0691	0.0691	$\eta = 0.0695$	0.0127
J-J transition rate	0.0498	0.0473	$\lambda = 0.3164$	0.0325
$\hat{\rho}_{profit}$	0.7683	0.6299	$\rho_z = 0.8004$	0.0366
$Mean(profit)$	0.1260	0.1144	$\mu_z = 0.0279$	0.0014
$Var(profit)$	0.0144	0.0160	$\sigma_z^2 = 0.1198$	0.0044
$Mean(\log(size))$	7.4515	7.4806	$\mu_s = 1.2356$	0.0365
$Var(\log(size))$	2.3060	2.1610	$\sigma_s = 2.5795$	0.1211
$Mean(\log(tdc1))$	7.2408	7.2665	$\alpha_0 = -1.5534$	0.0147
$Var(log(tdc1))$	1.1846	0.8960	$\alpha_1 = 0.5270$	0.0217
$\beta_{tdc1-size}$	0.3830	0.2822		
$\beta_{inc-tdc1}$	1.1063	1.1997	$\sigma = 1.1038$	0.0030
$Mean(\log(inc))$	8.4994	8.478	$c = 0.0814$	0.0259
$Var(\log(inc))$	3.4438	3.35872		

Estimates Table 5 reports the targeted values of moments in the data and the corresponding values in the estimated model. The last two columns list the parameter estimates and standard errors. While I arranged moments and parameters along the identification argument made in the previous subsection, all parameters are estimated jointly. Overall, the model provides a decent fit to the data.

Looking into the estimates, a job arrival rate $\lambda = 31.64\%$ is required to match the job-to-job transition rate 4.98% in the data. The magnitude of λ indicates that, on average, the executive will receive an outside offer every three years. Most job offers (about 84%) are from poaching firms that are smaller than the current firm and are used to negotiate compensation with the current firm. This is confirmed by a small mean of poaching firms. The magnitude of μ_s indicates that most offers are provided by relatively small firms, though the magnitude of σ_s implies the variation of poaching firms is high. Comparing the data and model-simulated mean and variance of $\log(size)$, I confirm that using a log-normal distribution is sufficient to match the firm size distribution in the data.

The process of productivity is matched reasonably well, given I use only 6 grid points. The mean $\log(tdc1)$ is matched well, but the variance of $log(tdc1)$ and $\beta_{tdc1-size}$ are not. In particular, the variance of $log(tdc1)$ is much lower in the model-generated data. This indicates that the on-the-job search and sequential auction in the model may miss some heterogeneous features of firms and executives. Finally, the optimal dynamic contracting employed by the model provides good matches on the mean and variance of $\log(inc)$, and the slope of inc on total compensation $\beta_{inc-tdc1}$.

4.4 Predicting firm-size incentive premium

I intentionally leave the incentive premium untargeted in the estimation. Instead, I estimate it using the model-simulated data, and compare with the real-world data estimates to examine if the model mechanism can match up with the data. Specifically, I estimate the following regression in both real-world and model-simulated data:

$$\log(inc_{it}) = \beta_5 + \beta_{inc-size} \log(size_{it}) + \beta_6 \log(tdc1_{it}) + \epsilon_{it,3}, \tag{6}$$

and $\beta_{inc-size}$ captures the incentive premium.

Table 6 reports premia estimates. The first row are the incentive premia estimated in re-

Table 6: Predictions on Size premia

	Data	Benchmark	ignore poach inc	more offers	fewer offers
incentive premia	(1)	(2)	(3)	(4)	(5)
w/ tdc1	0.3451	0.3122	-0.0444	0.4299	0.1964
w/o tdc1	0.5824	0.6507	0.4202	0.7093	0.4076

gression (6) with $\log(tdc1)$ controlled, and the second row are the premia without controlling $\log(tdc1)$. Focus on the first two columns, column (1) are the real-world premia (replicates of columns 1 and 2 of Table 3), and column (2) are the model-simulated premia. Comparing these two columns, I find that, even not intentionally targeting these premia, the model captures them well. This reassures that the model mechanism is important in explaining the firm-size premium.

To clarify the role of poaching offers, I simulate a counterfactual scenario where firms mechanically ignore poaching offer incentives when designing incentive contracts. Indeed, column (3) shows that in this case, the incentive premium with $tdc1$ controlled becomes zero. In contrast, the premium with $tdc1$ not controlled becomes smaller: 0.4202, which entirely reflects that the compensation levels are higher in larger firms, the channel discussed by Edmans et al. (2009).

To further evaluate the contribution of poaching offer incentives, in Figure 7 I compare the model generated inc between the benchmark (column 2) where poaching offer incentives are present, and the model variant (column 3) where poaching offer incentives are ignored. The higher inc in the variant reflects the contribution of poaching offers. I divide firms into ten groups based on firm size (group 0 contains the smallest firms). The upper panel of Figure 7 shows the box plot of $\log(inc)$ across size. There are two observations. First, each firm-size group has a wide dispersion of $\log(inc)$, which reflects the large variation in total compensation across executives in the same firm-size group. Second, the contribution of poaching incentives varies across firms. In particular, smaller firms need to provide a higher $\log(inc)$ when poaching offer incentives are ignored (the orange box) compared to the benchmark (the green box). In the lower panel of Figure 7, I calculate the contribution (fraction) of poaching offer incentives for each size group. The proportion of poaching offer incentives is surprisingly high for the smallest firm group: inc would need to be 80% higher when poaching incentives are absent. The fraction quickly goes

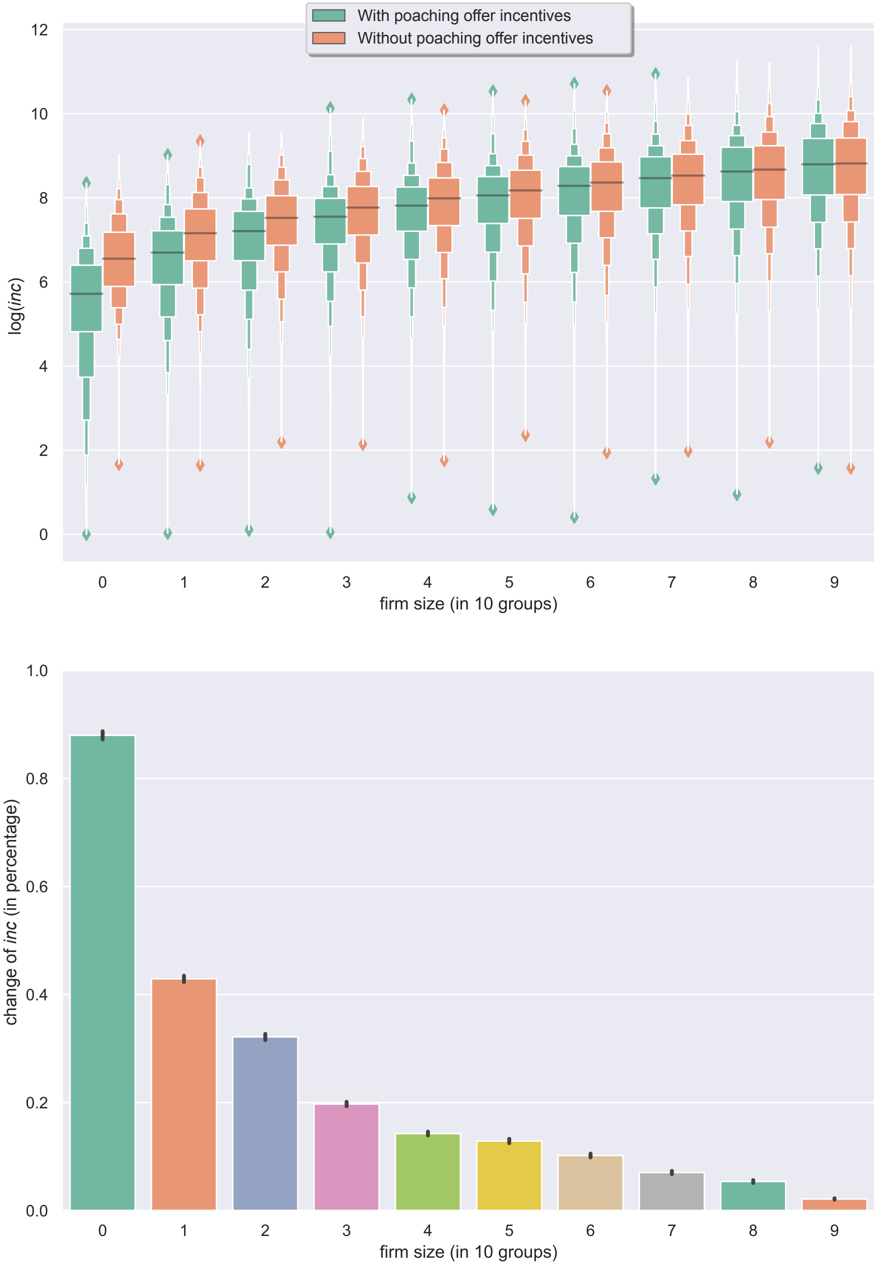

Figure 7: Distribution of $\log(inc)$ and the fraction of poaching offer incentives

Note: The upper panel plots the distribution of $\log(inc)$ for ten firm-size groups. The green boxes are for the benchmark model where poaching offer incentives are considered when designing a contract. The orange boxes are for a variant where poaching offer incentives are mechanically ignored. The increase in median in each group shows that, without poaching offer incentives, firms need to give higher inc. The lower panel calculates the fraction of poaching offer incentives by $\frac{inc^v - inc^b}{inc^b}$, where inc^b and inc^v are the wealth-performance sensitivity in the benchmark and in the variant with poaching offer incentives ignore, respectively.

down to around 15% in medium-sized firms and almost vanishes for top-sized firms.

Finally, I simulate a version with a high job arrival probability $\lambda = 0.6$ and a low job arrival probability $\lambda = 0.1$ in columns (4) and (5), respectively. It shows that when there are more (fewer) job offers, the premium is higher (lower). These exercises inspire the second application of the model, where I use λ as a valve to adjust for the strength of managerial labor market forces. While it is easy to see that poaching offer incentives are monotonically increasing in λ, it is not clear how other moments of the data would change.

5 Application II: The long-run trend of executive compensation

As a second application of the model, I use a counterfactual exercise to quantitatively capture various distributional changes of executive compensation since the 1970s. Frydman and Saks (2010) document that the level and inequality of executive pay were relatively low from the late 1930s to the mid-1970s and had soared since then. Similarly, the correlation between firm size and total compensation is weaker before the 1970s. In table 7, I select two representative periods, 1970 - 1979 and 1990 - 1999, and use the data moments from Frydman and Saks (2010). The mean of total compensation rises from 1090 thousand dollars before 1979 to 4350 thousand dollars after 1990, and the mean of performance-based incentives increases by almost six folds from the 1970s to the 1990s. The interquartile range of third and first quartiles increases from 670 thousand dollars to 3080 thousand dollars. While firm size is closely related to executive pay after 1992, it was weaker in the previous decade. The coefficient mildly increases from 0.199 to 0.264 from the 1970s to 1990s.

All these changes can be accounted for by my model with *an exogenous increase* in job arrival rate λ, preserving all other parameters. Though this paper does not provide an endogenous mechanism for an increase in λ, there has been abundant evidence showing that the executive market had been more active since the mid-1970s. Murphy and Zabojnik (2007) document that an increasing number of CEO openings has been filled through external hires. Huson et al. (2001)

[16] Replicate from Table 3, 4 and 6 of Frydman and Saks (2010), in dollar value of year 2000.

Table 7: The long-run trend in executive compensation

Moments	Data[16]		Model	
	1970 - 1979	1990 - 1999	$\lambda = 0.05$	$\lambda = 0.4$
Mean tdc1 (thousand)	1090	4350	985	4296
Mean size (million)	-	-	2426	5710
Mean inc (million)	2.174	12.034	2.497	12.531
$\beta_{tdc1-size}$	0.199	0.264	0.175	0.240
Percentiles of tdc1 (thousand)				
25th percentile	640	1350	109	1217
50th percentile	930	2360	478	2957
75th percentile	1310	4430	1596	5860

document that the fraction of outsider CEOs increases from 15.3% in the 1970s to 30.0% at the beginning of the 1990s. One explanation for the trend is that executive jobs have increasingly placed greater emphasis on general managerial skills (Murphy and Zabojnik 2004, Frydman 2005). This is also the view taken by this paper. The executive's productivity in the model is general and can be transferred between firms.

I calibrate λ to be 5% for the period 1970-1979 and 40% for 1990 - 1999. These values are chosen to match the data moments under the constraint that all other parameters are equal to the estimated values in Section 4. Since most firms in the sample of Frydman and Saks (2010) are within the rank 500, I only keep the largest 500 firms in the simulated data. The moments calculated by simulated datasets are reported in the last two columns of table 7.

Capturing all the moments of table 7 by only increasing λ is a tough test of the model. Increasing λ essentially amplifies the poaching offer mechanism. We, therefore, require the model is not only "correct" in out-of-sample moment predictions but also "correct" in predicting distributional changes. The results, however, are surprisingly well. As λ increases, executives are more likely to use poaching offers to renegotiate contracts, which leads to higher total compensation $tdc1$ and higher incentives inc. As firms bid for executives, the correlation between pay and firm size increases. Next, since the labor market is frictional, the inequality is amplified with more poaching offers: lucky executives receive many poaching offers, while unlucky ones get few job-hopping opportunities. The simulated moments are mostly close to the data counterparts, with some exceptions. In particular, the model generates much lower *tdc1* in the first two percentiles when $\lambda = 0.05$. This may reflect that the poaching offer distributions of the two periods are different.

This exercise also has other predictions. First, a more active labor market is related to larger average firm size. The mean of firm size doubles as λ increases. This pattern is generally consistent with the considerable increase in market capitalization (measured by the S&P 500 index). Second, the model predicts that firm-size incentive premium increases (not shown in table), a similar pattern as in the last two columns of table 6. This implies that, over time, firm size plays a more prominent role in shaping contract structure.

6 Conclusions

This paper studies the impact of labor market competition on managerial incentive contracts. I develop a dynamic contracting model where executives use poaching offers to renegotiate with their current employer. The model demonstrates that poaching offers have both a level and an incentive effect on executive compensation. In particular, the model generates a bumpy job ladder, which is supported by empirical findings from a newly assembled job-to-job transition dataset. The model explains the firm-size incentive premium, and the model-simulated premia match up well with the data counterparts. Finally, with an exogenous increase in poaching offer arrival rate, the model can quantitatively account for the increases in total and incentive pay, and the rising inequality across executives since the 1970s.

References

Abrahám, Arpád, Fernando Alvarez-Parra, and Susanne Forstner, "The Effects of Moral Hazard on Wage Inequality in a Frictional Labor Market," 2017.

Abreu, Dilip, David Pearce, and Ennio Stacchetti, "Toward a theory of discounted repeated games with imperfect monitoring," *Econometrica*, 1990, pp. 1041–1063.

Baker, George P and Brian J Hall, "CEO incentives and firm size," *Journal of Labor Economics*, 2004, *22* (4), 767–798.

Balke, Neele and Thibaut Lamadon, "Productivity Shocks, Long-Term Contracts, and Earnings Dynamics," *American Economic Review*, July 2022, *112* (7), 2139–77.

Benveniste, L.M and J.A Scheinkman, "Duality theory for dynamic optimization models of economics: The continuous time case," *Journal of Economic Theory*, 1982, *27* (1), 1–19.

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.

Coles, Jeffrey L, Naveen D Daniel, and Lalitha Naveen, "Managerial incentives and risk-taking," *Journal of financial Economics*, 2006, *79* (2), 431–468.

Core, John and Wayne Guay, "Estimating the Value of Employee Stock Option Portfolios and Their Sensitivities to Price and Volatility," *Journal of Accounting Research*, 2002, *40* (3), 613–630.

Cremers, KJ Martijn and Yaniv Grinstein, "Does the market for CEO talent explain controversial CEO pay practices?," *Review of Finance*, 2013, *18* (3), 921–960.

Custódio, Cláudia, Miguel A Ferreira, and Pedro Matos, "Generalists versus specialists: Lifetime work experience and chief executive officer pay," *Journal of Financial Economics*, 2013, *108* (2), 471–492.

Dittmann, Ingolf and Ernst Maug, "Lower salaries and no options? On the optimal structure of executive pay," *The Journal of Finance*, 2007, *62* (1), 303–343.

Edmans, Alex and Xavier Gabaix, "The effect of risk on the CEO market," *The Review of Financial Studies*, 2011, *24* (8), 2822–2863.

— , — , and Augustin Landier, "A Multiplicative Model of Optimal CEO Incentives in Market Equilibrium," *The Review of Financial Studies*, 2009.

— , — , and Dirk Jenter, "Executive Compensation: A Survey of Theory and Evidence," in Benjamin E. Hermalin and Michael S. Weisbach, eds., *The Handbook of the Economics of Corporate Governance*, Vol. 1, North-Holland, 2017, pp. 383–539.

— , — , Tomasz Sadzik, and Yuliy Sannikov, "Dynamic CEO Compensation," *The Journal of Finance*, 2012, *67* (5), 1603–1647.

Eisfeldt, Andrea L and Camelia M Kuhnen, "CEO turnover in a competitive assignment framework," *Journal of Financial Economics*, 2013, *109* (2), 351–372.

Farhi, Emmanuel and Iván Werning, "Insurance and taxation over the life cycle," *Review of Economic Studies*, 2013, *80* (2), 596–635.

Frydman, Carola, "Rising through the ranks: The evolution of the market for corporate executives, 1936-2003," *Working Paper*, 2005.

— 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 and Augustin Landier, "Why has CEO pay increased so much?," *The Quarterly Journal of Economics*, 2008, *123* (1), 49–100.

Gayle, George-Levi and Robert A Miller, "Has moral hazard become a more important factor in managerial compensation?," *American Economic Review*, 2009, *99* (5), 1740–69.

— , Limor Golan, and Robert A Miller, "Promotion, turnover, and compensation in the executive labor market," *Econometrica*, 2015, *83* (6), 2293–2369.

Giannetti, Mariassunta, "Serial CEO incentives and the structure of managerial contracts," *Journal of Financial Intermediation*, 2011, *20* (4), 633–662.

Gibbons, Robert and Kevin J Murphy, "Optimal incentive contracts in the presence of career concerns: Theory and evidence," *Journal of Political Economy*, 1992, *100* (3), 468–505.

Graham, John, Dawoon Kim, and Kim Hyunseob, "Executive Mobility in the United States," *Nanyang Business School Research Paper No. 21-04*, 2021.

Grochulski, Borys and Yuzhe Zhang, "Market-Based Incentives," *International Economic Review*, 2017, *58* (2), 331–382.

Hall, Brian J and Jeffrey B Liebman, "Are CEOs really paid like bureaucrats?," *The Quarterly Journal of Economics*, 1998, *113* (3), 653–691.

__ **and Kevin J Murphy**, "Optimal exercise prices for executive stock options," *American Economic Review*, 2000, *90* (2), 209–214.

Harris, Milton and Bengt Holmstrom, "A theory of wage dynamics," *The Review of Economic Studies*, 1982, *49* (3), 315–333.

Holmström, Bengt, "Managerial incentive problems: A dynamic perspective," *The review of Economic studies*, 1999, *66* (1), 169–182.

Hopenhayn, Hugo A and Juan Pablo Nicolini, "Optimal unemployment insurance," *Journal of political economy*, 1997, *105* (2), 412–438.

Huson, Mark R, Robert Parrino, and Laura T Starks, "Internal monitoring mechanisms and CEO turnover: A long-term perspective," *The Journal of Finance*, 2001, *56* (6), 2265–2297.

Jensen, Michael C and Kevin J Murphy, "Performance pay and top-management incentives," *Journal of Political Economy*, 1990, *98* (2), 225–264.

Kaplan, Steven N and Bernadette A Minton, "How has CEO turnover changed?," *International review of Finance*, 2012, *12* (1), 57–87.

Khurana, Rakesh, *Searching for a corporate savior: The irrational quest for charismatic CEOs*, Princeton University Press, 2004.

Kocherlakota, Narayana R, "Implications of efficient risk sharing without commitment," *The Review of Economic Studies*, 1996, *63* (4), 595–609.

Lentz, Rasmus, "Optimal employment contracts with hidden search," 2014.

Lise, Jeremy, Costas Meghir, and Jean-Marc Robin, "Matching, sorting and wages," *Review of Economic Dynamics*, 2016, *19*, 63–87.

Marcet, Albert and Ramon Marimon, "Recursive Contracts," *Econometrica*, 2019, *87* (5), 1589–1631.

Mele, Antonio, "Repeated moral hazard and recursive Lagrangeans," *Journal of Economic Dynamics and Control*, 2014, *42*, 69–85.

Menzio, Guido and Shouyong Shi, "Block recursive equilibria for stochastic models of search on the job," *Journal of Economic Theory*, 2010, *145* (4), 1453–1494.

Murphy, Kevin J and Jan Zabojnik, "CEO pay and appointments: A market-based explanation for recent trends," *American economic review*, 2004, *94* (2), 192–196.

__ **and** __ , "Managerial capital and the market for CEOs," 2007.

Oyer, Paul, "Why do firms use incentives that have no incentive effects?," *The Journal of Finance*, 2004, *59* (4), 1619–1650.

Phelan, Christopher, "Repeated moral hazard and one-sided commitment," *Journal of Economic Theory*, 1995, *66* (2), 488–506.

__ **and Robert M Townsend**, "Computing multi-period, information-constrained optima," *The Review of Economic Studies*, 1991, *58* (5), 853–881.

Postel-Vinay, Fabien and Jean-Marc Robin, "Equilibrium wage dispersion with worker and employer heterogeneity," *Econometrica*, 2002, *70* (6), 2295–2350.

Rogerson, William P, "Repeated moral hazard," *Econometrica: Journal of the Econometric Society*, 1985, pp. 69–76.

Spear, Stephen E and Sanjay Srivastava, "On repeated moral hazard with discounting," *The Review of Economic Studies*, 1987, *54* (4), 599–617.

Tauchen, George, "Finite state markov-chain approximations to univariate and vector autoregressions," *Economics letters*, 1986, *20* (2), 177–181.

Tervio, Marko, "The difference that CEOs make: An assignment model approach," *American Economic Review*, 2008, *98* (3), 642–68.

Thomas, Jonathan and Tim Worrall, "Income fluctuation and asymmetric information: An example of a repeated principal-agent problem," *Journal of Economic Theory*, 1990, *51* (2), 367–390.

Tsuyuhara, Kunio, "Dynamic contracts with worker mobility via directed on-the-job search," *International Economic Review*, 2016, *57* (4), 1405–1424.

Wang, Cheng and Youzhi Yang, "Optimal CEO turnover," *Journal of Economic Theory*, 2022, *203*, 105475.

Appendix A. Proofs

Proof for proposition 1

Proof. Recall that $\Pi(W,z,s)$ is the expected discounted stream of profits from a match between a firm of size s and an executive of productivity z subject to a utility promise W, $\overline{W}(z,s)$ is the firm's willingness to pay for the executive, and for the virtue firm $s^{(0)}$, $\overline{W}(z,s^{(0)}) = U$. Further denote $\overline{W}(s) = \sup\{\overline{W}(z,s), z \in \mathbb{Z}\}$. Then $\mathbb{X} = [U, \sup_{s\in\tilde{\mathbb{S}}} \overline{W}(s)]$ is set of possible values for the promised utility. In this proof, the functional dependence of Π on s is suppressed to save notations: $\Pi(W,z)$.

Rewrite the problem. The firm chooses a contingent plan of promised continuation utilities for each possible realization $z' \in \mathbb{Z}, s' \in \mathbb{S}$, denoted by $Y \equiv \{W(z',s')\}_{z'\in\mathbb{Z},s'\in\mathbb{S}}$. Y is the control variable. With the assumption of monotone likelihood ratio, I first guess that $W(z',s')$ is nondecreasing in z' and verify the guess after differentiability is established. Let $\mathbb{Y} \equiv \mathbb{X}^{n_z\times n_s}$ be the set of continuation values available. Depending on the realized z' and s', the update to the next period state is given by $\phi : \mathbb{Y}\times\mathbb{Z}\times\mathbb{S} \to \mathbb{X}$: $\phi(Y,z',s') = W(z',s')$. ϕ selects among the menu $\{W(z',s')\}$ the realized promised value $W(z',s')$. Since the promise keeping constraint is always binding (note that $u(\cdot)$ is strictly increasing), I write w as a function of Y:

$$w = u^{-1}\Big(V + c - \tilde{\beta}\sum_{z'}\sum_{s'} W(z',s')p(ds')\gamma(z'|z)\Big), \tag{7}$$

and $u^{-1}(\cdot)$ is a convex function. Then the flow profit is given by $\pi : \mathbb{X}\times\mathbb{Y}\times\mathbb{Z} \to \mathbb{R}$ defined by

$$\pi(V,Y,z) = f(z,s) - u^{-1}\Big(V + c - \tilde{\beta}\sum_{z'}\sum_{s'} W(z',s')p(s')\gamma(z'|z)\Big).$$

It follows that $\pi(V,Y,z)$ is continuous in the first two arguments, and since both $\mathbb{X}$ and $\mathbb{Y}$ are compact-valued, it is also bounded. Moreover, because $u(\cdot)$ is strictly increasing and Γ is monotone, $\pi(V,\cdot,z)$ is strictly decreasing in V and strictly increasing in z. Finally, because $-u^{-1}(\cdot)$ is strictly concave, $\pi(V,Y,\cdot)$ is strictly concave in V and Y.

Given the beginning-of-period executive productivity z, and a promised lifetime utility $V \in [U, \overline{W}(s)]$, the expected discounted sum of profits can be expressed recursively as

$$\Pi(V,z) = \max_{Y\in\Phi(z)} \pi(V,Y,z) + \tilde{\beta}\sum_{z'}\sum_{s'\le s} \Pi\Big[\phi\Big(Y,z',s'\Big),z'\Big]p(s')\gamma(z'|z). \tag{8}$$

where the firm chooses a contingent plan of promised values $Y = \{W(z',s')\}$ from

$$\Omega(z) = \Bigg\{Y \in \mathbb{Y} \,\Bigg|\, \sum_{z'}\sum_{s'} W(z',s')(1-g(z'|z))p(s')\gamma(z'|z) \ge c,$$
$$W(z',s') \in \Big[\min\Big\{\overline{W}(z',s'), \overline{W}(z',s)\Big\}, \overline{W}(z',s)\Big]\Bigg\}.$$

To ensure that $\Omega(z)$ is non-empty, I impose that c satisfy

$$c \leq (1-\lambda) \sup_{\{W(z')\}_{z' \in \mathbb{Z}}} \sum_{z'} W(z') \Big(\gamma(z'|z) - \gamma^s(z') \Big),$$

where in the sup, I require that for all z', $W(z') \in [U, \overline{W}(z', \underline{s})]$. The right-hand side is the incentive that all firm is able to provide. Note that the lower U is, the higher the right-hand is. Therefore, by lowering U, this condition imposes a very weak restriction on c.

$\Omega(z)$ has the following properties. First, $\Omega(z)$ does not depend on V. Second, $\Omega(z)$ is convex because given $\{W_0(z', s')\}, \{W_1(z', s')\} \in \Omega(z)$, the linear combination $\{W_\theta(z', s')\}$ where $W_\theta(z', s') = \theta W_0(z', s') + (1-\theta) W_1(z', s'), \theta \in (0, 1)$ satisfies the constraints of $\Omega(z)$. Finally, $\Omega(z)$ is increasing in the sense that $z_1 \leq z_2$ implies $\Omega(z_1) \subset \Omega(z_2)$. To see this, note that the left-hand side of the incentive compatibility constraint increases in z

$$\sum_{s'} \sum_{z'} W(z', s') \Big(\gamma(z'|z_2) - \gamma^s(z') \Big) p(s') \geq \sum_{s'} \sum_{z'} W(z', s') \Big(\gamma(z'|z_1) - \gamma^s(z') \Big) p(s')$$

due to that γ is monotone.

Existence. Let $\mathbb{O} = \mathbb{X} \times \mathbb{Z}$ be the product space. Denote $C(\mathbb{O})$ be the space of bounded continuous functions $h : \mathbb{O} \to \mathbb{R}$, with the sup norm: $||h|| = \sup_{o \in \mathbb{O}} |h(o)|$. Since here $\mathbb{Z}$ is a finite set, the continuity refers to that for each $z \in \mathbb{Z}$, the z-section of the function is continuous. Consider the right-hand side of (8) as a functional mapping T. Fix $h \in C(\mathbb{O})$, since ϕ is continuous in Y, $\tilde{\beta} \sum_{z'} \sum_{s' \leq s} h\Big[\phi\Big(Y, z', s'\Big), z'\Big] p(s') \gamma(z'|z)$ is continuous in Y. The problem is to maximize a continuous function over the compact set $\Omega(z)$. Hence, the maximum is attained and Th is bounded. By Berge's Maximum Theorem, Th is continuous. That is, $T : C(\mathbb{O}) \to C(\mathbb{O})$. It is easy to check that T satisfies the hypotheses of Blackwell's sufficient conditions for a contraction. Thus, T has a unique fixed point $\Pi \in C(\mathbb{O})$.

Monotonicity. To show that Π is strictly decreasing in V, let's fix z and pick a function $h(V, z)$ that belongs to the set of bounded continuous functions that are nonincreasing in V. Take $V_1, V_2 \in \mathbb{X}$ with $V_1 < V_2$, and let $Y_i \in \Omega(z)$ be the optimal control that attains $Th(V_i, z)$ for $i = 1, 2$. Then

$$\begin{aligned}
(Th)(V_1, z) &= \pi(V_1, Y_1, z) + \tilde{\beta} \sum_{z'} \sum_{s' \leq s} h\Big[\phi\Big(Y_1, z', s'\Big), z'\Big] p(s') \gamma(z'|z) \\
&\geq \pi(V_1, Y_2, z) + \tilde{\beta} \sum_{z'} \sum_{s' \leq s} h\Big[\phi\Big(Y_2, z', s'\Big), z'\Big] p(s') \gamma(z'|z) \\
&\geq \pi(V_2, Y_2, z) + \tilde{\beta} \sum_{z'} \sum_{s' \leq s} h\Big[\phi\Big(Y_2, z', s'\Big), z'\Big] p(s') \gamma(z'|z) \\
&\geq (Th)(V_2, z)
\end{aligned}$$

where the second line uses that $Y_2, Y_1 \in \Omega(z)$, and Y_1 attains $Th(V_1, z)$, and the third line uses that $\pi(V, \cdot, \cdot)$ is strictly decreasing in V. Therefore, the unique fixed point of $\Pi(V, z)$ is strictly decreasing in V.

Next, I show that Π is strictly increasing in z. Fix $V \in \mathbb{X}$; suppose that $h(V, z)$ is non-decreasing in z; and choose $z_1 < z_2$. Let $Y_i \in \Omega(z)$ attain $Th(V, z_i)$ for $i = 1, 2$. Then

$$\begin{aligned}
(Th)(V, z_1) &= \pi(V, Y_1, z_1) + \tilde{\beta} \sum_{z'} \sum_{s' \leq s} h\left[\phi\left(Y_1, z', s'\right), z'\right] p(s')\gamma(z'|z_1) \\
&\leq \pi(V, Y_1, z_2) + \tilde{\beta} \sum_{z'} \sum_{s' \leq s} h\left[\phi\left(Y_1, z', s'\right), z'\right] p(s')\gamma(z'|z_2) \\
&\leq (Th)(V, z_2),
\end{aligned}$$

where second line uses that $\pi(V, Y, z)$ is strictly increasing in z. Hence, the fixed point $\Pi(V, z)$ is strictly increase in z. It follows similarly that $\Pi(V, z, s)$ is strictly increasing in s (to reflect that s is back into the argument list). These monotonicity properties, together with that Π is continuous, imply that $\overline{W}(z, s)$ is well defined.

Concavity. To show that Π is strictly concave in V, let's fix z and let $h(V, z)$ belongs to the set of bounded continuous functions that are weakly concave in V. Then I show that $Th(V, z)$ is strictly concave in V. Take $V_1 \neq V_2 \in \mathbb{X}$ and $V_\theta = \theta V_1 + (1 - \theta)V_2$, $\theta \in (0, 1)$. Let Y_i attain $(Th)(V_i, z), i = 1, 2$. Since $\Omega(z)$ does not depend on V, $Y_\theta = \theta Y_1 + (1 - \theta)Y_2 \in \Omega(z)$.

$$\begin{aligned}
(Th)(V_\theta, z) &= \pi(V_\theta, Y_\theta, z) + \tilde{\beta} \sum_{z'} \sum_{s' \leq s} h\left[\phi\left(Y_\theta, z', s'\right), z'\right] p(s')\gamma(z'|z) \\
&> \theta \left[\pi(V_1, Y_1, z) + \tilde{\beta} \sum_{z'} \sum_{s' \leq s} h\left[\phi\left(Y_1, z', s'\right), z'\right] p(s')\gamma(z'|z)\right] \\
&+ (1 - \theta) \left[\pi(V_2, Y_2, z) + \tilde{\beta} \sum_{z'} \sum_{s' \leq s} h\left[\phi\left(Y_2, z', s'\right), z'\right] p(s')\gamma(z'|z)\right] \\
&= \theta(Th)(V_1, z) + (1 - \theta)(Th)(V_2, z)
\end{aligned}$$

The second line uses that $\pi(V, Y, z)$ is strictly concave in V and Y jointly, and ϕ is concave in Y. Therefore, the fixed point $\Pi(V, z)$ is strictly concave in V.

Differentiability. The concavity also indicates that $\Pi(V, z)$ is differentiable in V almost everywhere. Now I show that $\Pi(V, z)$ is differentiable in V everywhere. Suppose for a fixed (z), $\Pi(\cdot, z)$ is not differentiable at $\tilde{V}$ and denote the firm's optimal choice at that point by $\{\tilde{W}(z', s')\}$ and a corresponding flow pay $\tilde{w}$. In the rest of the paragraph, I suppress the functional dependence of Π on z. Now consider a strategy of the firm that delivers any V around $\tilde{V}$ by changing the flow pay to $w^*(V) \equiv u^{-1}(V - \tilde{V} + u(\tilde{w}))$ while keeping $\{\tilde{W}(z', s')\}$. By construction, this action satisfies all constraints and I denote the firm's value function when taking this action as $\tilde{\Pi}(V)$. It then follows that $\tilde{\Pi}(V) \leq \Pi(V)$ and $\tilde{\Pi}(\tilde{V}) = \Pi(\tilde{V})$. Moreover, V enters $\tilde{\Pi}(V)$ only through

$-w^*(V) = -u^{-1}(V - \tilde{V} + u(\tilde{w}))$. Since $-u^{-1}(\cdot)$ is concave and twice differentiable, $\tilde{\Pi}(V)$ is also concave and differentiable in any point V around $\tilde{V}$ and including $\tilde{V}$. To sum up, $\tilde{\Pi}(V)$ is a function that is concave, twice differentiable, and lower than $\Pi(V)$ and equals to $\Pi(V)$ at $\tilde{V}$. By Benveniste and Scheinkman (1982), we have $\Pi(V)$ is differentiable at $\tilde{V}$. Therefore, $\Pi(V)$ is differentiable in V everywhere.

First-order conditions. To characterize the optimal contract I assign Lagrangian multipliers λ to (PKC), μ to (IC), $\tilde{\beta}\mu_0(z', s')$ to (PC-E) and $\tilde{\beta}\mu_1(z', s')$ to (PC-F). The first order condition w.r.t w gives $u'(w) = 1/\lambda$, and the envelop theorem gives $-\frac{\partial \Pi(z,s,V)}{\partial V} = \lambda$. They together give (1). Participation constraints (PC-E) and (PC-F) can be simplified. If $\overline{W}(z', s') \geq \overline{W}(z', s)$, we have $W(z', s') = \overline{W}(z', s)$. This is the first case in line 1 of (3). If $\overline{W}(z', s') \geq \overline{W}(z', s)$, participation constraints become $\overline{W}(z', s') \leq W(z', s') \leq \overline{W}(z', s)$. Use this to derive the first order condition w.r.t $W(z', s')$:

$$-\frac{\partial \Pi(z', s, W(z', s'))}{\partial W(z', s')} = \lambda + \mu(1 - g(z'|z)) + \mu_0(z', s') - \mu_1(z', s').$$

If $\mu_0(z', s') = \mu_1(z', s') = 0$, $W(z', s') = W(z')$ defined by (2). This is the case in line 3 of (3). If $\mu_0(z', s') > \mu_1(z', s') = 0$, $W(z', s') = \overline{W}(z', s')$. This is the case in line 2 of (3). Finally, if $\mu_1(z', s') > \mu_0(z', s') = 0$, $W_i(z', s') = \overline{W}(z', s)$. This is the second condition in line 1 of (3).

Verify that $W(z', s')$ is non-decreasing in z'. Since $\overline{W}(z', s')$ is non-decreasing in z, $W(z', s')$ is certainly non-decreasing in z' whenever $W(z', s') = \overline{W}(z', s')$; that is, a participation constraint is binding. Next, consider that non-binding $W(z', s')$. Fix s', z, and suppose $z_2' > z_1'$, by first order conditions, we have

$$-\frac{\partial \Pi(W(z_2', s'), z_2')}{\partial W} = \lambda + \mu(1 - g(z_2'|z)) > \lambda + \mu(1 - g(z_1'|z)) = -\frac{\partial \Pi(W(z_1', s'), z_1')}{\partial W}, \quad (9)$$

where I have used $g(z_2'|z) < g(z_1'|z)$. Since (PKC) is less binding as z becomes higher, $-\frac{\partial \Pi(W,z)}{\partial W} = \lambda$ decreases in z, which implies that $-\frac{\partial \Pi(W(z_1',s'),z_1')}{\partial W} > -\frac{\partial \Pi(W(z_1',s'),z_2')}{\partial W}$. Together we have

$$-\frac{\partial \Pi(W(z_2', s'), z_2')}{\partial W} > -\frac{\partial \Pi(W(z_1', s'), z_2')}{\partial W}.$$

By concavity of Π, $W(z_2', s') > W(z_1', s')$. □

Proof for proposition 2

Proof. Take a general incentive scheme $W(z)$ defined on $\mathbb{Z}$, the incentive $\mathcal{I}\Big(W(z)\Big)$ can be written as a weighted sum of $\frac{\partial W(z)}{\partial z}$:

$$\mathcal{I}(W(z)) \equiv \int_{\underline{z}}^{\overline{z}} W(z)(1 - g(z))\Gamma(dz) = \int_{\underline{z}}^{\overline{z}} \omega(z) \frac{\partial W(z)}{\partial z} dz, \quad (10)$$

where $\omega(z) = -\int_{\underline{z}}^{z}(1-g(t))\Gamma(t)dt$. (10) can be easily verified using integration by parts together with $\int_{\underline{z}}^{\overline{z}}(1-g(z))\Gamma(z)dz = 0$. Thus, it is sufficient to show that $\frac{\partial \overline{W}(z,s)}{\partial z}$ decreases in s for all $z \in \mathbb{Z}$. Applying implicit theorem on $\Pi(\overline{W}, z, s) = 0$:

$$\frac{\partial \overline{W}(z,s)}{\partial z} = -\frac{\partial \Pi(\overline{W}, z, s)}{\partial z} \Bigg/ \frac{1}{u'(\overline{w}(z,s))}. \tag{11}$$

where $\overline{w}(z,s)$ is the compensation corresponding to $\overline{W}(z,s)$. Use $\psi_n(s)$ and $\psi_d(s)$ to denote the numerator and denominator of (11), respectively, then $\frac{\partial \overline{W}(z,s)}{\partial z}$ decreases in s if and only if

$$\frac{\psi_n'(s)}{\psi_n(s)} < \frac{\psi_d'(s)}{\psi_d(s)}. \tag{12}$$

I now derive ψ_n and ψ_d. Given that $\Pi(\overline{W}(z,s), z, s) = 0$, I express $\overline{w}$ as a function of s and z:

$$\overline{w}(s,z) = f(z,s) + \tilde{\beta} \int_{z'} \int_{s' \le s} \Pi(W(z',s'), z', s) \tilde{P}(ds') \Gamma(z, dz').$$

Take derivative with respect to s; use that if the poaching firm $s' = s$, then $W(z', s') = \overline{W}(z', s)$, and $\Pi(\overline{W}(z', s), z', s) = 0$, I have $\frac{\partial \overline{w}(z,s)}{\partial s} = f_s(z,s)$. Then

$$\frac{\psi_d'(s)}{\psi_d(s)} = -\frac{u''(\overline{w})}{u'(\overline{w})} f_s(z,s).$$

Turn to ψ_n. Recall the Lagrangian

$$\begin{aligned}
\mathcal{L} = {} & f(z,s) - \overline{w} + \tilde{\beta} \int_{s' \le s} \int_{z'} \Pi(W(z',s'), z', s) \tilde{P}(ds') \Gamma(z, z'), \\
& + \lambda \Big(\overline{w} + \beta \int_{s'} \int_{z'} W(z',s') \tilde{P}(ds') \Gamma(z, dz') - V \Big) \\
& + \mu \Big(\tilde{\beta} \int_{s'} \int_{z'} W(z',s') \Big(\Gamma(z, dz') - \Gamma^0(dz') \Big) \tilde{P}(ds') - c \Big),
\end{aligned}$$

where I have inserted the optimal contingent plan so that the participation constraints are dropped. Using the envelop theorem yields $\psi_n(s) = \frac{\partial \mathcal{L}}{\partial z} = f_z + \kappa(s)$, where

$$\begin{aligned}
\kappa(s) = {} & f_z + \tilde{\beta} \int_{s' \le s} \int_{z'} \Pi(W(z',s'), z', s) \tilde{P}(ds') \Gamma_z(z, dz') \\
& + (\lambda + \mu) \Big(\tilde{\beta} \int_{s'} \int_{z'} W(z',s') \Gamma_z(z, dz') \tilde{P}(ds'),
\end{aligned}$$

is positive and bounded. It follows that $\psi_n'(s) = f_{zs} + \kappa'(s)$, where

$$\kappa'(s) = (\lambda + \mu) \tilde{\beta} \int_{s' \ge s} \tilde{P}(ds') \int_{z'} \frac{\partial \overline{W}(z', s)}{\partial s} \Gamma_z(z, dz')$$

is also positive and bounded. Therefore, a sufficient condition for (12) is

$$-\frac{u''(\overline{w})}{u'(\overline{w})} > \sup_{s \in \mathbb{S}, z \in \mathbb{Z}} \left[\frac{f_{zs} + \kappa'(s)}{f_s(f_z + \kappa(s))} \right].$$

□

Appendix B. Data summary statistics

Here I describe the variables that are used in my analysis. Table 8 reports summary statistics for my sample. Note that all nominal quantities are converted into constant 2004 dollars. Using information from Execucomp, I identify the *gender*, *age* of an executive in each year, the *tenure* in the current executive episode, whether he or she is a *CEO*, *CFO*, or *director* of the board or involved in a *interlock* relationship during the fiscal year. 93% of the executives are male, and the average age is 51. The average length of episodes is 4.71 year. Among all executive-year observations, 18.4% are CEO spells, 9.6% are CFO spells.

In terms of the compensation information, *tdc1* is the total compensation, including salary, bonus, values of stock and option granted, etc. The total compensation has an average of 2,555 thousand dollars, with the 25th percentile of 632 thousand dollars and the 75th percentile of 2,690 thousand dollars. In terms of means, only 16.5% of the total compensation is fixed base salary and the rest are all incentive related. Performance-based incentives not only come from the total compensation each year, but also come from the stocks and options that are granted in previous years. Variable *inc* measures how strong performance-based incentives are in firm-related wealth. It is defined by the dollar change in wealth associated with a 100 percentage points change in the firm's stock price.

For the firm side information, I use market capitalization *mkcap*, the market value of a company's outstanding shares, to measure the firm size. In some robustness checks (not shown in the main text), I also use the book value of assets *at*, and *sales* to measure firm size. They are in million dollars. I use operating profitability, denoted by *profitability*, to measure firm performance. Two alternative measures for firm performance are stock market annualized return, denoted by *annual return*, and market-to-book ratio, denoted by *mbr*.

Execucomp has little information of executives' employment history. The job turnover information comes from BoardEX database. BoardEX contains details of each executive's employment history, including start and end dates, firm names and positions. It also has extra information on education background, social networks, etc. I merge the two databases using three sources of information: the executive's first, middle and last names, the date of birth, and working experiences. If all three aspects are consistent, the executive is identified. By this way, I am able to identify more than 91% of executives in Execucomp, 32,864 executives in total.

Table 8: Summary statistics

Variable	*N*	*mean*	*sd*	*p25*	*p50*	*p75*
age (years)	218168	51.04	6.96	46	51	56
male	218168	0.936	0.244	1	1	1
CEO	218168	0.184	0.387	0	0	0
CFO	218168	0.096	0.295	0	0	0
director	218168	0.339	0.473	0	0	1
interlock	218168	0.013	0.112	0	0	0
tenure (years)	218168	4.71	3.793	2	4	6
tdc1 (thousand dollars)	167,360	2562.245	5726.48	662.892	1292.752	2692.256
inc (million dollars)	146790	33.788	520.537	1.689	5.082	15.667
mkcap (million dollars)	212271	7293.608	23539.41	546.214	1479.479	4714.474
at (million dollars)	216384	14877.73	89106.44	548.0703	1771.067	6385.512
sales (million dollars)	216276	5283.954	16216.22	432.312	1215.325	3852.92
profitability (percentage)	209639	0.119	0.359	0.069	0.121	0.176
annual return (percentage)	211067	0.181	0.802	-0.127	0.106	0.356
mbr	183565	1.669	2.21	0.811	1.198	1.913

Note: The table reports summary sample statistics for the Execucomp/Compustat dataset, which covers named executive officers reported in Execucomp from 1992 to 2016. All dollar values are stated in 2004 dollars. *age* is the executive's age by the end of the fiscal year. The sample episodes with ages lower than 35 or above 70 are dropped. Dummy variables *CEO*, *CFO*, *director* and *interlock* indicate whether the executive serves as a director, CEO, CFO and is involved in the interlock relationship during the fiscal year, respectively. *tenure* (in years) counts the number of fiscal years that the executive works as a named officer. *tdc1* is the total compensation comprised of the following: Salary, Bonus, Other Annual, Total Value of Restricted Stock Granted, Total Value of Stock Options Granted (using BlackScholes), Long-Term Incentive Payouts, and All Other Total. *inc* is the million dollar change in wealth associated with 100 percentage points change in stock price. *mkcap* (in millions) is the market capitalization of the company, calculated by *csho* (Common Shares Outstanding, in millions of shares) multiplied by *prcc_f* (fiscal year end price). *prcc_f* and *csho* are reported in Compustat Fundamentals Annual file. *at* (in millions) is the Total Book Assets as reported by the company. *sales* (in millions) is the Net Annual Sales as reported by the company. *profitability* is the operating profitability, calculated by EBITDA/Assets. *annual return* is the annualized stock return which is compounded based on CRSP MSF (Monthly) returns. MSF returns have been adjusted for splits etc. *mbr* is the Market-to-Book Ratio calculated by the Market Value of Assets divided by Total Book Assets. The market Value of Assets is calculated according to $\textit{Value of Assets (MVA)} = prcc\_f * cshpri + dlc + dltt + pstkl - txditc$.