A Job Ladder Model of Executive Compensation*

Bo Hu†

Abstract

This paper examines the impact of managerial labor market competition on executive incentive contracts. I develop a dynamic contracting model that incorporates moral hazard, search frictions, and poaching offers. The model generates a job ladder along which executives can either use outside offers to renegotiate with the current firm or transition to outside firms. I show that poaching offers generate a new source of incentives, which explains a novel empirical finding whereby larger firms give executives a higher proportion of incentive compensation.

Keywords: Executive Compensation, Managerial Labor Market, Dynamic Moral Hazard, Search Frictions, Firm-size Incentive Premium

*The author thanks the editor Loukas Karabarbounis, the associate editor of this journal, and two anonymous referees for extremely valuable suggestions that have significantly improved the paper. This article has greatly benefited from the comments from Pieter Gautier, Makoto Watanabe, Gregory Jolivet, Shuo Xia, Hélène Turon, Satoshi Tanaka, Sarolta Laczó, Alvaro Remesal, Yanos Zylberberg, as well as participants at the Rotterdam Executive Conference (2018), the 26th Finance Forum in Santander (2019), the Belgian Financial Research Forum (2019), the SaM Annual Conference at Cambridge (2019), the SUFE IAR Macroeconomics Workshop (2022), the Inaugural Search and Matching Workshop in Asia-Pacific (2022) and various seminars at Tinbergen Institute, HEC Paris, IWH Halle, Tor Vergata University of Rome, University of Gothenburg, and Shanghai Jiao Tong University. The author also thanks Jeffrey Coles, Naveen Daniel, Lalitha Naveen, Cláudia Custódio, Miguel Ferreira, Pedro Matos, Martijn Cremers, and Yaniv Grinstein for generously sharing data from their respective papers. Hu acknowledges financial support from the National Natural Science Foundation of China (72003041), the Shanghai Pujiang Program (21PJC011), the 2024 Fudan University Gaofeng Program in Theoretical Economics, and the Shanghai Institute of International Finance and Economics. Part of the research was done when the author was a PhD student at Tinbergen Institute and when the author was a visiting scholar at Erasmus University Rotterdam. The author is responsible for all errors.

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

1 Introduction

Executives are incentivized by tying their compensation closely to firm performance through bonuses, stocks, options, etc. Traditionally, it is believed that incentive contracts are designed to align the interests of executives with those of shareholders. However, in recent decades, competition for executives has become increasingly influential in shaping incentive contracts. Companies are more frequently looking outside for executive talent. The fraction of outsider CEOs increased from 15.3% in the 1970s to 30.0% at the beginning of the 1990s (Huson et al. 2001). At the same time, more companies are considering ”competitiveness” as a key principle in the design of executive contracts. As stated in IBM’s annual report, compensation for each executive is based on ”the skills and experience of senior executives that are highly sought after by other companies and, in particular, by our [IBM’s] competitors.”

Executives, in turn, recognize that their first appointment leading a firm can pave the way to a series of more prestigious executive jobs. Recent studies have revealed a discernible increase in executive job-to-job transitions since the 1980s. These transitions occur across all executive levels and are often accompanied by a promotion, reflected through a title change and pay increase.[1] To supplement these findings, I merge two data sources, ExecuComp and BoardEx, and thereby connect information on executive compensation to the executives’ résumés. Using this new dataset, I document that the majority of executive job transitions involve a move to a larger firm and that the transition rate is notably lower for executives in larger firms. This evidence suggests the existence of a hierarchical job ladder, with more prestigious positions and employers situated at higher levels.

Despite extensive empirical findings, there is a lack of a theoretical framework that aligns with the facts of executive mobility and, at the same time, connects to the design of incentive contracts. This paper aims to bridge this gap by asking several questions: What characteristics of the managerial labor market give rise to these job mobility patterns? How do executive job transitions differ from those seen in the general labor market? And, crucially, what are the implications of these dynamics for executive contracts, particularly in light of the distinctive incentive structures they often feature?

To address these questions, I develop a search-matching model that integrates elements from two strands of the literature. First, building on the standard repeated agency framework, e.g., Spear and Srivastava (1987), I introduce moral hazard as a key driver of incentive pay. In this setup, an executive’s managerial productivity evolves stochastically, influenced by both her current and past efforts. Firms have different time-invariant asset scales. Once a firm and an executive are matched, the pair produces according to a production function that increases with firm asset scales (firm size) and executive productivity. Although output is observable (from

[1] See, e.g., Clementi and Cooley (2023), Graham et al. (2021), Kaplan and Minton (2012), Frydman (2005), Huson et al. (2001).

which the executive's productivity is also observed), the effort remains hidden, giving rise to the moral hazard problem. To resolve this problem, each firm-executive pair establishes an optimal long-term contract that balances the executive's incentives with the associated costs to the firm.

Second, drawing on the seminal work by Postel-Vinay and Robin (2002), I endogenize the executive's outside options by modeling a frictional labor market with on-the-job search. The managerial labor market is particularly well-suited to this approach, as executives are frequently "auctioned" by competing firms seeking to lure talent with promises of promotion and better terms (Khurana 2004). Specifically, an executive is randomly approached by outside firms, initiating a competitive scenario. If an executive receives an offer from another firm, she faces a decision: Does she renegotiate the current contract for better terms or move to the new firm? Since larger firms can afford higher bids for executives, the model predicts that job transitions will be toward larger firms, and that executives in large firms are less likely to transition, a prediction that is consistent with the aforementioned facts.

My model offers two novel insights. First, the wages determined by the optimal contract are non-monotonic in tenure. Specifically, due to moral hazard, the wage follows a stochastic process with a threshold; if productivity surpasses this threshold, the wage increases. If not, it decreases. This wage process is further constrained by an upper bound imposed by the incumbent firm (namely, the maximum wage the incumbent firm is willing to pay) and a lower bound imposed by the poaching firm (the maximum wage the outside, smaller firm is willing to pay).

These bounds play a crucial role in shaping the wage process. If a poaching firm (assuming it's smaller than the incumbent) is sufficiently large, it can push the lower wage bound upward, leading to back-loaded wages, as highlighted by Postel-Vinay and Robin (2002). Conversely, a significant drop in productivity may force the incumbent firm to revise the wage downward—a mechanism introduced by Postel-Vinay and Turon (2010) and subsequently incorporated by Lise et al. (2016). The interaction of the agency problem, poaching offers, and productivity shocks creates rich wage dynamics, which feature relatively modest ups and downs driven by executive performances and more significant adjustments triggered by extremely low realizations of firm output and sufficiently competitive poaching offers.

The second insight is that when a firm attempts to poach an executive, it is willing to bid more for an executive with higher productivity; consequently, an executive is incentivized not only by the compensation she receives from her current firm but also by the prospect of receiving more attractive offers from external firms. This potential for outside offers introduces a novel source of incentives, which I refer to as *poaching-offer incentives*. To capture this mechanism, my model incorporates a persistent effect of effort; specifically, effort functions as an investment in human capital, increasing the likelihood of higher productivity in all future periods, and the executive's productivity can be carried across executive jobs. This approach contrasts with the standard repeated moral hazard model of Spear and Srivastava (1987), according to which exerting effort enhances output today but has no effect on future output.

Poaching offers, therefore, play dual roles in my model: on the one hand, a firm risks losing an executive to an outside firm, leading to Bertrand competition with the outside firm; on the other hand, the incumbent firm benefits from poaching-offer incentives because, with these incentives, the incumbent firm can reduce the share of incentive pay in the compensation package while still motivating the executive to exert effort. In effect, poaching offers lead to back-loaded compensation while simultaneously providing immediate incentives.

Using the concept of poaching-offer incentives, I explore the impact of firm size on contractual incentives. Firm size is one of the most pivotal characteristics shaping executive compensation. I document a novel fact: larger firms tend to allocate a higher proportion of incentive pay in their executive compensation packages. Specifically, if the firm size doubles, the fraction of incentive-related pay in the executive's total compensation increases by 4.59 percentage points (with the median fraction being 65%). I refer to this as the *firm-size incentive premium*; to the best of my knowledge, this is the first time this premium has been documented. This premium is also confirmed when using wealth–performance sensitivities, which are considered to be more accurate measures of incentives (Edmans et al., 2017).

The firm-size incentive premium is empirically related to job transitions, and I show that this premium is more pronounced in industries where executives exhibit higher job-transition rates. Furthermore, when executives transition to new positions, the proportion of incentive pay in their total compensation increases; this increase is more pronounced when the gap between the size of the target firm and the original firm is larger. These findings point to a theoretical explanation for the premium that is closely linked to job transitions.

Poaching-offer incentives are crucial to understanding the firm-size incentive premium. I show that there are two reasons why poaching incentives decrease as the size of the incumbent firm increases. First, there is a job ladder effect: a larger incumbent firm is positioned higher on the job ladder, reducing the likelihood that an executive will encounter an even larger competing firm. Second, there is a wealth effect: larger firms possess the capacity to bid higher, thus, executives at these firms expect higher future compensation. As a result, the perceived incentives from potential poaching offers diminish. The job ladder and the wealth effects compel larger firms to give more contractual incentives to compensate for the weakened poaching-offer incentives.

This simple explanation is quantitatively relevant. I calibrate the model by targeting the first and second-order moments of total compensation, firm size, incentives, job turnovers, and selected regression coefficients. Notably, the firm-size incentive premium is not targeted in the calibration. Yet, the calibrated model captures the premium well. The model-generated premium aligns closely with the premium estimated from real data, both conditional and unconditional on total compensation. Furthermore, I demonstrate that poaching-offer incentives constitute a significant portion of total incentives for small firms. As firm size increases, the fraction of poaching-offer incentives decreases; it is around 15% for medium-sized firms and nearly vanishes for the largest firms.

The remainder of this section reviews the related literature. In Section 2, I describe the data and document the stylized facts about executive job mobility that motivate my model. I then present the firm-size incentive premium, which is the central phenomenon that this paper seeks to explain. Section 3 is devoted to the theory; in this section, I characterize the optimal contract, describe the wage dynamics, and explain the firm-size incentive premium through the lens of poaching-offer incentives. In Section 4, I calibrate the model and assess its quantitative implications. Section 5 discusses alternative explanations of the firm-size incentive premium, and provides further evidence in support of my explanation. Section 6 concludes the paper.

Related Literature

My approach to modeling executive on-the-job search and compensation negotiation is based on the sequential auction framework pioneered by Postel-Vinay and Robin (2002), Dey and Flinn (2005), and Cahuc et al. (2006), and further developed by Postel-Vinay and Turon (2010), Lise et al. (2016), among others. The dynamic moral hazard component builds on the literature relating to optimal long-term contracts in the presence of private information and commitment frictions, 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), and Wang and Yang (2022).

Connecting these studies, Lentz (2014) explores the optimal employment contract in the presence of hidden search efforts and poaching firm competition. Both his model and mine have moral hazard as a central element. In Lentz's setting, moral hazard arises from the employee's private decision on search intensity. His model generates insightful results on wage back-loading, worker flow patterns, and wage dispersion. As in the original model by Postel-Vinay and Robin (2002), the wage process generated in Lentz's model exhibits downward rigidity. In contrast, the moral hazard in my model stems from the executive's effort choices. As a result, in addition to the impact of poaching offers, the wages in my model can increase as a reward for high performance and decrease as a punishment for low performance.[2]

As an addition to the literature on executive compensation, my paper provides an explana-

[2]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 homogeneous 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 the 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. 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), Tsuyuhara (2016), etc.

tion for the firm-size incentive premium, focusing on the role of executive job mobility. Empirical evidence from Frydman (2005), Frydman and Saks (2010), and Murphy and Zabojnik (2006) suggests that the rise in executive compensation is closely linked to increased mobility and driven by the growing importance of general managerial skills as opposed to firm-specific knowledge. On the theoretical side, Gabaix and Landier (2008) and Tervio (2008) employ competitive assignment models to explain the correlation observed between executive compensation and firm size.[3] Giannetti (2011) presents a model that shows how job-hopping opportunities can explain both the increase in total pay and the structure of managerial contracts. More recently, Shi (2023) examines the optimal regulation of non-compete clauses, while Chemla et al. (2023) consider the outside value of executives within a general equilibrium framework. However, to the best of my knowledge, none of the existing research has examined executive compensation and incentive packages within the context of job ladders and poaching-offer incentives. This paper addresses this gap.

2 Motivating Facts: Job Transitions and Firm-size Incentive Premium

This section introduces the data and the key empirical facts that motivate my theoretical analysis. I construct a new dataset that combines executive job mobility, compensation contracts, and firm-side information by merging ExecuComp, Compustat, and BoardEx. Using this new dataset, I uncover several empirical regularities in executive job transitions, particularly concerning firm size and industry variations. I then document a novel empirical finding: executives at larger firms receive a greater proportion of their compensation in incentive pay. I term this the *firm-size incentive premium*. I demonstrate that the firm-size incentive premium is more pronounced in industries in which executive job-transition rates are higher. These facts combined motivate the model in the following section.

2.1 Data

Compustat and ExecuComp are the standard sources for studies of executive compensation. They provide comprehensive firm-level variables such as financial statements and industry classifications, along with compensation contract variables (including salary, bonus, total compensation, and incentive pay) for up to nine named officers of S&P-listed firms. The variables that are particularly relevant to this study are measures of total executive compensation and firm size. I use the $TDC1$ variable in ExecuComp to measure total compensation, denoted as $total\,pay$.

[3] See also Eisfeldt and Kuhnen (2013), Baker and Hall (2004), Edmans et al. (2009), Edmans and Gabaix (2011). In a different vein, Gayle and Miller (2009) and Gayle et al. (2015) attribute the pay differentials across firms to variations in principal-agent problems.

This metric encompasses all components of the compensation package, including salary, bonus, other annual compensation, and the values of stock and options granted during the fiscal year (hereafter, year). Importantly, the values of granted stock and options are estimated before the realization of company earnings and thus reflect their ex ante values as awarded by the board.[4] I use market capitalization (*mktcap*) to measure firm size. I should emphasize that the findings documented below are robust to alternative measures of firm size, including book values of assets and net annual sales.[5]

BoardEx offers detailed employment histories for executives, including start/end dates, employers, and job titles for each position held. I merge this data with ExecuComp using an executive's full name, date of birth, and overlapping employment periods in both datasets. This allows me to trace an executive's career before their inclusion in ExecuComp as a named officer and track their subsequent employment and job titles.

For the empirical analysis, I use a sample of 47,716 executives from ExecuComp, corresponding to 275,611 executive-fiscal-year observations from 1992 to 2016. On average, each executive has 5.78 fiscal-year records. I successfully matched 34,089 executives between ExecuComp and BoardEx, corresponding to 217,588 executive-year records, which represent 71.44% of the executives and 78.95% of the executive-year observations in ExecuComp. Further details on dataset merging and key variable summary statistics are available in the Online Appendix.

2.2 Executive Job Transitions

Using the matched dataset, I define the end-of-year status for each executive-year observation as follows. If an executive remains at the same firm in the subsequent year, their status is marked as continuing. If they leave their current firm and take up an executive position at another firm within six months after their current ExecuComp spell ends, their status is recorded as a job transition. If no subsequent executive role is recorded in BoardEx within six months, the executive's end-of-year status is marked as exiting the labor market.[6] Based on this definition, I document three stylized facts on executive job transitions.

Fact 1. *The executive labor market features active job-to-job transitions.*

I calculate an average yearly job-to-job transition rate of 4.43%. In other words, each year, 4.43% of the named executives employed by S&P-listed firms transition to another firm the following year.[7] Figure 1 demonstrates job-transition rates across the Fama-French 48-industry clas-

[4]ExecuComp provides another total compensation metric called $TDC2$, which includes the value of stock and options that have vested or been exercised during the fiscal year. Unlike $TDC1$, which reflects the awarded value, $TDC2$ captures the realized value of executive compensation. I use $TDC2$ for supplementary analysis (see the Online Appendix). I thank the referee for highlighting this important distinction.

[5]Detailed variable descriptions are provided in the note of Table 2.

[6]The facts documented below continue to hold if the time window is adjusted to 60 days or one year. The summary statistics of job-spell classifications are provided in the Online Appendix.

[7]ExecuComp back-fills information for executives who continue to work in S&P listed firms. In my sample, all obser-

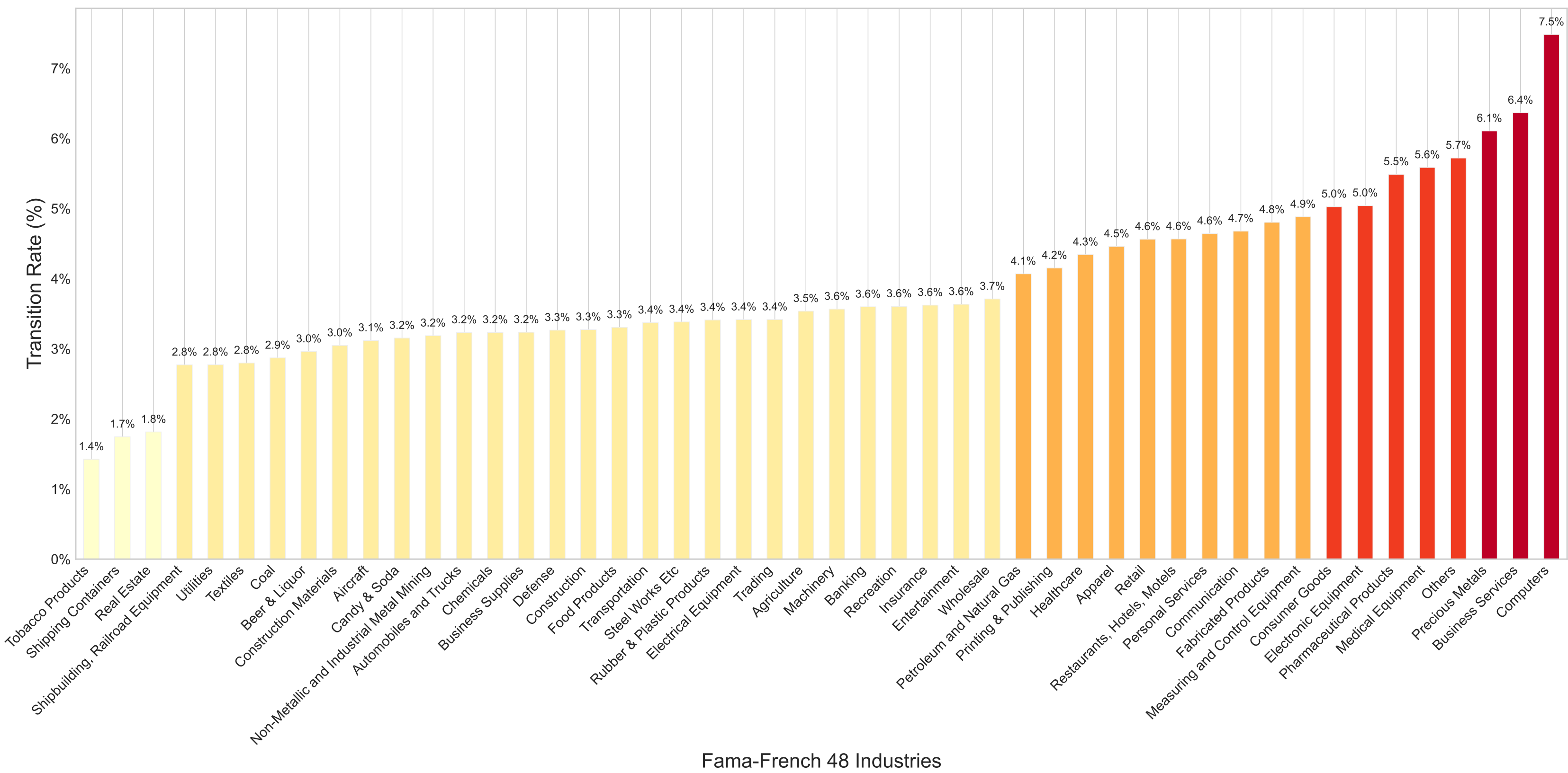


Figure 1: Executive job-transition rates across the Fama-French 48-industry classification

sification. This analysis reveals that transition rates vary significantly across industries, ranging from approximately 1.4% to 7.5%. Tobacco Products, Shipping Containers, and Real Estate have the lowest job-transition rates (the first three bars in light yellow in Figure 1), while Computers exhibits the highest transition rate at 7.5%. Additionally, Pharmaceutical Products, Medical Equipment, Business Services, and Electronic Equipment all have relatively high transition rates. There are 27 industries with transition rates between 2% and 4%. Later, I will use this variation to evaluate the impact of job transitions on executive incentive compensation.

Fact 2. *Executives tend to transition to larger firms.*

In my sample, there are 9,094 job-to-job transitions from Compustat firms, with 2,568 of these transitions including size information for both the original and target firms. This subset essentially captures job-to-job transitions between publicly listed firms. Notably, 60.51% of these transitions involve a move to a larger firm. This finding is robust across alternative firm size metrics, such as book asset value and net sales, and is consistent across various industries.[8]

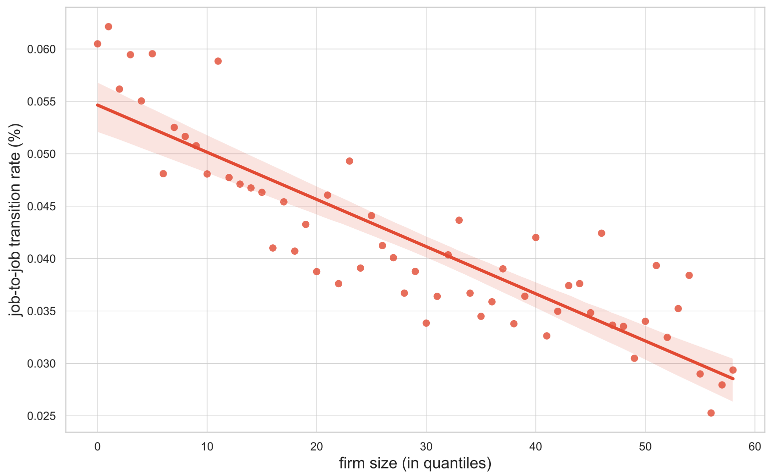


Figure 2: Job-to-job transition rate across firm size

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

Fact 3. *Executives in larger firms are less likely to move.*

---

vations from 1992 have an end-of-year status of ”continuing”. To avoid potential back-filling bias, I exclude observations from the 1992 fiscal year. Excluding 3,898 observations from 1992, I am left with 175,599 observations marked as continuing. The average job-transition rate is calculated by $9094/(175599 + 20621 + 9094) = 0.0443$.

[8] It is also important to consider why some executives transition to smaller firms. Intuitively, if were possible to construct an effective asset scale that combines firm size and title rank, it might be possible to observe that executives are moving into roles where they manage larger effective capital, even if the target firms themselves are smaller. In the theoretical model, executive titles are abstracted, with firm size as the single parameter representing the impact of title changes. However, in reality, the rank associated with different executive titles varies across firms. Nonetheless, it is widely recognized that moving from a non-CEO to a CEO position represents a significant elevation in rank. To illustrate this, I conduct a simple calculation: among transitions to smaller firms, 20% involve a change from a non-CEO to a CEO role; in comparison, only 3% of transitions to larger firms involve such a title change. This difference suggests that executives who move to smaller firms often attain higher rank and greater managerial authority.

As a first pass, Figure 2 illustrates transition rates across firm size quantiles, with a fitted line indicating the trend. The transition rate decreases from over 6% at the 5th percentile to under 3% at the 95th percentile. To further investigate the impact of firm size on the hazard of job-to-job transitions, I estimate a Cox proportional hazards model. Table 1 presents the results, showing a negative association between firm size and the hazard rate of job-to-job transitions across two regression specifications. Column (1) is a baseline model without control variables, while Column (2) includes controls for total compensation, executive roles (CEO, CFO, director, or interlocking relationship), and firm performance metrics such as the marking-to-ing-to-book ratio (MBR) and operating profitability. Overall, it is a robust result that executives in larger firms are *less* likely to have job-to-job transitions.

Table 1: Job-to-job transitions and firm size

	(1)	(2)
$\log(mktcap)$	-0.0680***	-0.0513***
	(0.00645)	(0.0100)
$\log(total\ pay)$		-0.0218
		(0.0161)
CEO		0.0314
		(0.0498)
CFO		0.00227
		(0.0313)
director		-0.905***
		(0.0458)
Obs.	212506	138478
χ^2	110.6	1300.6

Note: Columns (1) and (2) estimate a Cox proportional hazards model wherein the event of interest is a job-to-job transition. To control for industry-specific baseline hazards, the hazard function is stratified by the Fama-French 48-industry classification. Column (2) includes additional controls: $\log(total\,pay)$, *operating profitability*, *market-to-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 CFO during the fiscal year), *interlock* (whether the executive is involved in an interlocking relationship), and fiscal-year dummies. The table reports estimated coefficients instead of hazard ratios. *** refers to a significance level of $p < 0.001$.

2.3 Firm-size Incentive Premium

The firm-size incentive premium refers to the empirical observation that executives of larger firms receive a higher *fraction* of incentive pay in their compensation packages. This premium underscores the influence of firm size on the structure of incentive contracts. I document this phenomenon using various measures of incentives. Since job transitions are not needed for this analysis, I use the whole sample from ExecuComp (rather than just the sample that can be matched to BoardEx).

A straightforward approach is to correlate firm size with the proportion of incentive-related pay (e.g., newly granted shares) in the total compensation; this is one of the measures I adopt below (below, I shall denote this proportion as inc^f). However, since a large amount of the value of executive incentives is derived from previously granted stocks and options, I also follow the literature by using incentives that are embedded in executives' firm-related wealth.

The primary measure of contractual incentives that I use is wealth-performance sensitivity (denoted by *inc*), which is defined as the dollar change in the executive's firm-related wealth for a 100 percentage increase in firm value. In other words, *inc* represents the dollar value of the executive's stake in the firm, with options converted into stock equivalents according to their delta.[9]

An informal yet intuitive way to consider previously granted equities is as follows: the expected value of these equities has already been incorporated into the utility of previous periods. Therefore, in the current period, we may normalize the expected utility of those previous grants to zero. However, these holdings still generate fluctuations in executives' realized income. Consequently, the ratio *inc/totalpay* can be seen as the "effective" proportion of the incentive pay in the total compensation. In the following analysis, I regress *inc* on firm size while controlling for $totalpay$, allowing for a more flexible regression model.

In Table 2, Column (2) shows that *inc* is positively correlated to firm size—that is, if firm size doubles, *inc* increases by 28.17%. To put this number into context, take the median 4.05 million dollars of *inc* in 2012. If the firm size doubles, then *inc* will increase to 5.19 million dollars. This change is substantial (1.14 million) given a median annual total pay of 1.32 million dollars.

The size incentive premium also emerges when contractual incentives are calculated using alternative measures. In Column (4), I measure contractual incentives by the scaled wealth-performance sensitivity proposed by Edmans et al. (2009), which equals *inc* divided by $totalpay$, denoted as inc^s. Similar to *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 $totalpay$ (only the incentives in the flow pay), denoted as inc^f. Incentive pay includes bonuses, restricted stock grants, option grants, and other long-term incentive payouts. The results in both columns identify a positive firm-size incentive premium. For example, Column (5) indicates that if the firm size doubles, the percentage of incentive pay in the total compensation increases by 4.59 percentage points (the median fraction is 65%). I want to emphasize that the results above remain robust when controlling for *totalpay* in more flexible ways, such as by incorporating higher-order terms or categorizing *totalpay* into multiple groups. The findings are also consistent when firm size is measured using alternative metrics, such as the total book assets.[10]

[9] The data on *inc* is provided by Coles et al. (2006), who computed it based on the method of Core and Guay (2002). The variable *inc* is available from 1992 to 2014.

[10] The incentive premia estimated above are primarily identified by the between-firm variations. In the Online Appendix, I provide the estimated premia for each year separately.

Table 2: Firm-size incentive premium

	$\log(inc)$		$\log(inc^s)$		inc^f	$\log(inc)$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\log(mktcap)$	0.6186***	0.3581***	0.1523***	0.2610***	6.6240***	0.3251***	0.3555***	0.3264***
	(0.0153)	(0.0251)	(0.0101)	(0.0202)	(0.1485)	(0.0026)	(0.0031)	(0.0026)
$\log(total pay)$		0.6640 ***		-0.2793***		0.7139***	0.6821***	0.7092 ***
		(0.0324)		(0.0301)		(0.0041)	(0.0047)	(0.0041)
$\log(mktcap)$ $\times$ *JJ rate*						1.0803***		
						(0.0919)		
$\log(mktcap)$ $\times$ *gai*							0.0263**	
							(0.0095)	
$\log(mktcap)$ $\times$ *inside CEO*								-0.0266
								(0.0177)
Obs.	182,308	181,919	181,830	181,829	240,098	181,920	120,036	181,920
adj. R^2	0.415	0.490	0.219	0.258	0.283	0.458	0.480	0.457

Note: This table reports firm-size incentive premia (Columns 1 to 5) and their relationship with the activeness of the managerial labor market (Columns 6 to 8). Firm size is measured by market capitalization (*mktcap*), which is defined as the product of common shares outstanding and the fiscal-year-end closing price. Executives' compensation (*total pay*) is measured by annual awarded flow compensation ($TDC1$ in ExecuComp), which includes the sum of the salary and bonus, the value of the restricted stocks and options granted, and the value of any retirement and long-term compensation schemes. The dependent variable in Columns 1 to 2 and 6 to 8 is $\log(inc)$, where *inc* is the dollar change in firm-related wealth for a 100 percentage point change in firm value. The dependent variable in Columns 3 and 4 is the log of scaled *inc*, calculated as $inc^s = \frac{inc}{total pay}$. The dependent variable in Column 5 is the fraction of incentive pay in the total flow compensation: $inc^f = incentive\ pay / total\ pay \times 100$, where *incentive pay* is the sum of bonus, other annual compensation, restricted stock grants, option grants, and long-term incentive plan payouts. Columns 1 to 5 control for year-by-industry dummies and age dummies. Columns 6 to 8 control for year dummies and age dummies. Standard errors (clustered at the firm $\times$ fiscal-year level) are shown in parentheses. I denote symbols of significance as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Fact 4 (**Firm-Size Incentive Premium**). *Executives of larger firms receive proportionally higher incentive pay in their compensation packages.*

The Unconditional Premium. To align with the existing literature, I present the estimates of the firm-size incentive premium in Columns (1) and (3), where total compensation is not controlled for. These results have been documented by, e.g., Edmans et al. (2009). To differentiate from the premium that I document above (which is obtained after controlling for *totalpay*), I refer to the premium when *totalpay* is not controlled for as the *unconditional firm-size incentive premium.*

The "unconditional" premium captures the impact of firm size on the absolute amount of incentive pay, while my "conditional" premium reflects how firm size influences the proportion of incentive pay. In the following, I shall refer to the conditional premium as the firm-size incentive premium whenever the context is clear, and explicitly spell out "conditional" if necessary.

The wealth effect provides a straightforward explanation for the unconditional premium: larger firms typically offer higher total compensation, and because the utility function is concave, at higher pay levels, the same variation in income translates to a lower perceived incentive (in utility terms). As a result, larger firms must offer a greater absolute amount of incentive pay to achieve the same level of utility incentive.

However, this wealth effect does not explain the conditional premium. In fact, based on the explanation above, we would expect the proportion of incentive pay in the compensation package to remain constant across firms of different sizes, as noted by Edmans et al. (2009) and Edmans and Gabaix (2011).[11] The model I develop in the next section incorporates job ladder dynamics and poaching-offer incentives, providing a framework that can explain both the conditional and unconditional firm-size incentive premia.

2.4 The Firm-size Incentive Premium and the Managerial Labor Market

The following evidence highlights that the firm-size incentive premium is more pronounced in industries with higher executive job-transition rates. Each industry has intrinsically different job-transition rates for executives (see Figure 1), which can be attributed to factors such as industry-specific stability, risk, market structures, regulation, corporate governance practices, etc. In this analysis, I treat the industry-level variations as exogenous and use three proxies to measure the activeness of executive job transitions in an industry. The first proxy is the job-to-job transition rate for each industry-year (using the Fama-French 48-industry classification), calculated using the data and the definition of job transitions introduced above. The second proxy, *gai*, is the mean of the CEO general ability index at the industry-year level. The general ability index is the first principal component of five proxies that measure the generality of a CEO's human capital based

[11] Edmans et al. (2009) show that inc^s is not correlated with firm size. Their analysis focuses on CEOs of the top 200 or 500 firms, while my sample includes all named officers of S&P-listed firms.

on their lifetime work experience. The third proxy, *inside CEO*, is the percentage of insider CEOs in the industry. It counts all new CEOs hired between 1993 and 2005 using the Fama-French 48-industry.[12] In Table 2, Columns (6) to (8), I examine how the interaction terms are associated with my key measurement of contractual incentives *inc*. The results are unambiguous. All interaction coefficients support that the firm-size incentive premium is higher when executives have more job transition opportunities.

Fact 5. *The firm-size incentive premium is more pronounced when job transition opportunities are higher.*

Table 3: Job transitions and changes in the proportion of incentive pay

	Δinc^f (1)	$\Delta \log(inc^s)$ (2)
$\log(mktcap)_{dest}$	3.7848***	0.0628**
	(0.7380)	(0.0214)
$\log(mktcap)_{original}$	-4.9399***	-0.1786***
	(0.7006)	(0.0211)
intercept	17.591**	0.5074**
	(6.1952)	(0.188)
Obs.	958	973
adj. R^2	0.053	0.067

Note: The dependent variables are Δinc^f and $\Delta \log(inc^s)$, respectively, and the independent variables are the size of the original firm and the size of the destination firm. I denote symbols of significance as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Finally, I examine how the proportion of incentive pay in the compensation package changes when an executive transitions to a new job. Specifically, I measure the changes in two incentive proportion metrics, inc^f and $\log(inc^s)$, denoting these changes as Δinc^f and $\Delta \log(inc^s)$. For this analysis, I am limited to a relatively small sample of approximately 960 transitions, as I require available data on incentive shares both before and after the job change. Using this sample, I regress Δinc^f and $\Delta \log(inc^s)$ on the firm size of both the original firm and the destination firm and report the results in Table 3. Given that this is not a random sample, the results should be interpreted with caution. The estimates in both columns of Table 3 exhibit similar patterns, so I focus on interpreting Column (1). The results suggest that, after transitioning to a new job, the share of incentive pay increases by approximately 11.42 percentage points.[13] The change in inc^f is positively associated with the size of the destination firm and negatively associated with the size of the original firm.

[12] *Insider CEO* data is provided by Martijn Cremers and Grinstein (2013). *Gai* data is provided by Custódio et al. (2013). The five proxies used to measure a CEO's general ability are the number of positions held throughout their career, the number of firms they have worked for, the number of industries at the four-digit SIC level they have worked in, a dummy variable indicating whether the CEO has held a CEO position at another firm, and a dummy variable indicating whether the CEO has worked for a multi-division firm.

[13] The median of $\log(mktcap)_{original}$ is 7.34, and the median of $\log(mktcap)_{dest}$ is 7.95. Thus, the median effect of a transition on inc^f is $-4.9399 \times 7.34 + 3.7848 \times 7.95 + 17.591 = 11.42$.

Fact 6. *Following a transition, the share of incentive pay in the executive compensation package increases. Specifically, executives who transition to larger firms tend to experience a more substantial increase in the proportion of incentive pay.*

To conclude, I have presented six stylized facts on executive transitions and compensation. For the job transitions (Facts 1-3), my dataset goes beyond the focus on CEO transitions in prior research (see, e.g., Graham et al. (2021)) by including non-CEO executives such as CFOs and VPs. This broader scope provides a more complete view of mobility patterns in the executive labor market. The facts on the conditional firm-size incentive premium (Facts 4-6) are novel, and I additionally connect this premium to job transitions, suggesting a potential explanation from the perspective of executive mobility.

3 The Model

The empirical facts documented in the previous section motivate the development of a theoretical model that uses managerial job transitions to explain key features of executive incentive contracts. In this section, I introduce a job ladder model for the executive labor market and derive the optimal incentive contract. The model introduced here not only aligns with the empirical observations on executive job transitions but also provides a framework for understanding the firm-size incentive premium through the incentive generated by outside job offers. Specifically, it predicts that the increased executive mobility accentuates the firm-size incentive premium.

3.1 Setups

Executives. Time is discrete, is indexed by t, and continues forever. There is a continuum of individuals, each either employed as an executive or seeking an executive position as a *candidate*. Individuals face a probability of death, after which a *newborn* enters the labor market as a candidate searching for an executive job.[14] Each individual aims to maximize her expected lifetime utility,

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

where $\tilde{\beta} \in (0,1)$ is the discount factor and $\eta \in (0,1)$ is the death probability; the utility of consumption $u : \mathbb{R}_+ \to \mathbb{R}$ is twice differentiable, strictly increasing, and concave with $\lim_{w \to 0} u'(w) = \infty$; and $c(\cdot)$ is the disutility of effort. Effort e_t takes two values, $e_t \in \{0,1\}$, with 1 representing high effort and 0 representing a lack of effort. Normalize the cost of not making an effort $c(0)$ to 0, and denote the cost of making effort as $c \equiv c(1) > 0$. Let $\beta = \tilde{\beta}(1-\eta)$ be the effective discount factor.

[14] This model focuses on executives' on-the-job searches and their influence on compensation contracts. The inclusion of death and newborn candidates ensures that, in steady state, not all executives reach the top of the job ladder.

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 $\underline{z} = z^{(1)} < z^{(2)} < ... < z^{(n_z)} = \bar{z}$. z can be thought of as general managerial human capital that can be carried through job-to-job transitions between firms. z evolves according to a Markov process that demonstrates that a manager's productivity is affected by the current and previous effort. Formally, given a beginning-of-period productivity z, the executive decides whether to exert effort or not. At the end of the period, the next-period productivity z' is realized and becomes the beginning-of-the-next-period productivity. If the executive exerts effort ($e = 1$), z' is drawn from the distribution $\gamma(z'|z)$. If no effort is exerted ($e = 0$), z' is drawn from the distribution $\gamma^s(z')$, which, for simplicity, is assumed to be independent of z. We assume that a newly hired executive starts with an initial productivity $z^{(1)}$. The process of z has two common properties in the literature:

a. For each executive, z is positively correlated across time; therefore, a productive executive will likely remain productive in the next period. This requires that γ be monotonic 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. Exerting effort increases the likelihood of achieving high productivity. This requires the likelihood ratio defined by $g(z'|z) \equiv \frac{\gamma^s(z'|z)}{\gamma(z'|z)}$ to satisfy the monotone likelihood ratio property (MLRP), i.e., $g(z'|z)$ is non-increasing in z'.

Firms. Firms are different in scale of asset. I assume the asset scale is time-invariant that takes on values in a set: $s \in \{s^{(1)}, s^{(2)}, ..., s^{(n_s)}\}$, with $\underline{s} = s^{(1)} < s^{(2)} < ... < s^{(n)} = \bar{s}$. Since each firm has a single executive position, firm size is the scale of assets over which the executive can exert influence. A match between an executive of productivity z and a firm of size s generates a flow output $f(z, s)$ each period; this output is strictly increasing and concave in both arguments. To necessitate the use of incentive contracts, I impose a moral hazard setup whereby managerial productivity z and output $f(z, s)$ are observable but effort e is not.

Managerial Labor Market. The managerial labor market is characterized by random search. All individuals, whether employed or new-born, have a probability λ of encountering a poaching firm of size s' drawn from an exogenous distribution with probability $\tilde{p}(s')$.

An unemployed candidate has a continuation value of U, which is taken as given and is common for all candidates. With probability λ, the candidate is matched with a firm that offers a contract valued at U, and the candidate enters the next period as an employed executive.

With probability λ, an employed candidate is contacted by an outside firm. The incumbent and outside firms then engage in a Bertrand competition à la Postel-Vinay and Robin (2002) (details are provided in Section 3.2). Anticipating this subgame, a renegotiation-proof contract would specify counteroffers based on the executive's productivity and the size of both the incumbent and the poaching firms.

Finally, given that the primary focus of the model is on the dynamics of the executive side, rather than those of the firm or vacancy side, I abstract from the search decisions made by firms and concentrate solely on the one-sided search behaviors of executives. The continuation value of a vacancy is normalized to zero, a simplification that can be justified by the assumption of free entry by firms.

Timing. Consider an executive with a beginning-of-period productivity z who is currently matched with a firm of 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. New productivity z' is then drawn from $\gamma(z'|z)$ if the executive makes an 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'. The contract is then updated based on (z', s', s).

The compensation w, the 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. I will now turn to analyzing the optimal contract.

3.2 The Optimal Contract

The Contractual Environment. To recursively formulate the contracting problem, I adopt the executive's beginning-of-period expected utility, denoted by $V \in \mathbb{V}$, as a co-state variable used to encapsulate the history of productivity and outside offers. $\mathbb{V}$ represents the set of possible values that V can take. A dynamic contract, defined recursively, is comprised of a set of functions:

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

where $e(s,V)$ specifies the effort level required by the contract, and $w(s,V)$ determines the flow compensation, both of which are contingent on the initial promised value V and the size of the incumbent firm s. $W(z',s',s,V)$ denotes the promised continuation value, which is contingent on the realized shocks (z',s'), the initial promised value V, and the size of the incumbent firm s. For clarity, I will omit the explicit dependence of $e(\cdot)$, $w(\cdot)$, and $W(\cdot)$ on s and V in the subsequent analysis.

Given that publicly listed firms are generally large, I impose that $\underline{s}$ is sufficiently high to ensure the following assumptions hold.

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

b. Conditional on providing the executive with utility U, the sum of the future profits from the match remains positive. Consequently, dismissals are captured entirely by the exogenous probability η.

Point *a* justifies the inclusion of incentive pay in the executive's compensation package, while point *b* simplifies the analysis by focusing solely on exogenous dismissals. Extending the model to include endogenous dismissals can be a straightforward extension, following Lise et al. (2016). However, it is worth noting that dismissing an executive differs in nature from layoffs in the broader labor market (see, e.g., Taylor 2010). Given this distinction, I defer the exploration of executive dismissals to future research.

3.2.1 The Contracting Problem

In this subsection, I first characterize the maximum value a firm can bid and the competition for talent between firms (i.e., the sequential auction), and then set up the contracting problem.

Bidding Frontier and Productivity Shocks. Let $\Pi(W, z, s)$ represent the discounted sum of profits for a firm of size s that is matched with an executive whose beginning-of-period productivity is z and who has been promised a continuation value of W. The firm's maximum bidding value, denoted by $\overline{W}(z, s)$, is defined as

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

If the executive's demanded continuation value exceeds $\overline{W}$, the firm incurs a loss (noting that the value of a vacancy is normalized to zero). I refer to $\overline{W}(z, s)$ as the firm's *bidding frontier*. This frontier depends on (in fact, increases with) both z and s. Given the properties of $\Pi(\cdot)$ as established in Proposition 1 (see below), $\overline{W}(z, s)$ is well defined. The firm's participation constraint can thus be expressed as $W(z', s') \leq \overline{W}(z', s)$.

Sequential Auction. The competition between the incumbent and poaching firms proceeds as follows. When an executive from a firm of size s (hereafter firm s) encounters a poaching firm of size s' (hereafter firm s'), the two firms engage in a Bertrand competition. The maximum offer that firm s can make is a promised utility of $\overline{W}(z', s)$. If $s' > s$, the executive will move to firm s', which offers $\overline{W}(z', s')$. Any less competitive offer by firm s' will be successfully countered by firm s. If $s' \leq s$, the executive will remain at firm s, and her continuation value will be "promoted" to $\overline{W}(z', s')$, making her indifferent between staying and joining firm s'. The above argument defines the outside values of the executive contingent on the pair (z', s') as

$$W(z', s') \geq \min\{\overline{W}(z', s'), \overline{W}(z', s)\}.$$

A contract is considered renegotiation-proof only if it satisfies these participation constraints. For a rigorous proof, see Lentz (2014).

Before delving into the contracting problem, it is useful to briefly outline how the model's predictions correspond with the stylized facts presented in Section 2. First, the model inherently generates job-to-job transitions only when $s' > s$, meaning that executives tend to move to larger firms (Fact 2). Second, because the search process is random, those employed by larger firms are less likely to encounter an even larger poaching firm. Consequently, the likelihood of job-transition is lower (Fact 3).

The Contracting Problem. An executive always holds the outside value U independent of whether she receives a poaching offer or not. To write expressions compactly, I regard U the value as if the executive receives an offer from a "virtual" firm, whose size is denoted by $s^{(0)}$; the virtual firm has a bidding frontier $\overline{W}(z, s^{(0)}) \equiv U$ for all $z \in \mathbb{Z}$. I impose that $s^{(0)} < \underline{s}$ is sufficiently small such that $U < \overline{W}(\underline{z}, \underline{s})$. Recall that $\tilde{p}(s)$ is the probability of drawing s from $\mathbb{S} = \{s^{(1)}, s^{(2)}, ..., s^{(n_s)}\}$. Using this virtual firm, the poaching offer can be regarded as a draw from the mixture distribution, where the probability to take $s = s^0$ is $p(s^0) = 1 - \lambda$ and the probability to take $s \in \mathbb{S}$ is $p(s) = \lambda \tilde{p}(s)$, as follows:

$$p(s) = \mathbb{I}(s = s^{(0)})(1 - \lambda) + \mathbb{I}(s \neq s^{(0)})\lambda \tilde{p}(s).$$

Given a promised value to the executive, the firm chooses a current period compensation w and a set of promised values $W(z', s')$ for each realization of (z', s'). The expected discounted sum of the future profits of the firm can be expressed recursively as

$$\Pi(V, z, s) = \max_{w, W(z', s')} \Big\{ f(z, s) - w + \beta \sum_{s' \leq s} \sum_{z'} \Pi(W(z', s'), z', s)\gamma(z'|z)p(s') \Big\}, \tag{BE-F}$$

subject to the promise-keeping constraint

$$V = u(w) - c + \beta \sum_{s'} \sum_{z'} W(z', s')\gamma(z'|z)p(s'), \tag{PKC}$$

the incentive compatibility constraint

$$\beta \sum_{s'} \sum_{z'} \Big[W(z', s')(1 - g(z'|z)) \Big] \gamma(z'|z)p(s') \geq c, \tag{IC}$$

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:

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

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. The continuation value of a firm is normalized to zero if the match separates; this occurs either because the executive dies, which happens with probability η, or because the executive moves to another firm, which happens if the poaching firm is larger.

The promise-keeping constraint (PKC) ensures that the firm's choices honor the promises made in previous periods to deliver V to the executive. The incentive compatibility constraint (IC) says that the continuation value of making effort is higher than that of shirking. This creates an incentive 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 the continuation values between taking effort and shirking. I shall impose that c is sufficiently low to ensure that all firms, regardless of size, can provide the necessary incentives to induce executive effort.[15]

The participation constraints are stated in (PC-E) and (PC-F), and they shall hold state-by-state. The firm stays in the relationship if 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, $\overline{W}(z',s') = U$.

3.2.2 Characterization of the Optimal Contract

I now turn to the characterization of the optimal contract and its evolution over time and across different states. To build a clearer understanding, I begin by analyzing two simplified cases: dynamic contracting with only poaching offers and with only moral hazard. Afterward, I describe the optimal contract for the full model.

Poaching Offers Only. Without moral hazard, z' is drawn from $\gamma(z'|z)$ (the executive always exerts effort). Poaching offers are modeled as state-contingent participation constraints. Thus, the contracting problem is formulated with the objective given in (BE-F), subject to the promise-keeping constraint (PKC), and the participation constraints (PC-F) and (PC-E) for each realization of z' and s'.

Let ξ denote the Lagrangian multiplier associated with the promise-keeping constraint. Ap-

[15] A sufficient condition for this would be to guarantee that even the smallest firm, $\underline{s}$, can meet the incentive requirements. Consider the scenario where no poaching offers are present—a case that occurs with probability $1 - \lambda$. In this situation, firm $\underline{s}$ can freely choose any $W(z') \in [U, \overline{W}(z', \underline{s})]$. Suppose the following condition holds:

$$c < \beta(1-\lambda) \sup_{\{W(z')\}_{z' \in \mathcal{Z}}} \sum_{z'} \Big[W(z')(1 - g(z'|z)) \Big] \gamma(z'|z) \quad \text{for all } z \in \mathcal{Z}.$$

Under this condition, firm $\underline{s}$ can provide sufficient incentive to induce effort. Since the upper bound $\overline{W}(z', \underline{s})$ increases with s, it implies that condition (IC) holds for all firms with size $s > \underline{s}$.

plying the first-order conditions with respect to w and using the envelope theorem, we obtain

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

Two observations follow from this condition. First, the current period's flow compensation w is directly linked to the promised continuation utility V through the principal's and the agent's marginal rates of substitution between present and future compensation. As a result, a higher V is associated with a higher flow compensation w.

Second, since $u(\cdot)$ is strictly concave, Eq. (1) establishes that an increase in ξ corresponds to an increase in w. I denote this mapping by

$$w(\xi) = u'^{-1}\left(\frac{1}{\xi}\right), \tag{2}$$

where $w'(\xi) > 0$. Likewise, since $\Pi(\cdot)$ is strictly concave in V (see Proposition 1 below), Eq. (1) establishes that a higher ξ is also associated with a higher V. In the subsequent analysis, I use the dynamic changes in ξ to track the evolution of the contract.

Let $\mu_0(z',s')$ and $\mu_1(z',s')$ be the multipliers for (PC-F) and (PC-E), respectively. The first-order condition of $W(z',s')$ gives the evolution of ξ as follows:

$$\xi_{+1}(z',s') = \xi - \mu_0(z',s') + \mu_1(z',s'), \tag{3}$$

where $\xi_{+1}(z',s')$ represents the updated value of ξ when the new productivity is z' and the poaching firm is s'. If $s' \leq s$, the executive stays in the current firm and $\mu_0(z',s') = 0$. In addition, if s' is sufficiently competitive, then $\mu_1(z',s') > 0$, and ξ is updated to $\xi_{+1}(z',s') = \xi + \mu_1(z',s')$, which indicates that the next period pay is higher: $w_{+1}(z',s') > w$. If $s' > s$, the executive transitions to the outside firm, where the compensation is determined by the sequential auction.

It is important to emphasize that this contracting problem also captures the scenario described by Postel-Vinay and Turon (2010), where a sufficiently adverse productivity shock causes the maximum value the incumbent firm can offer to fall below the promised value $\overline{W}(z',s) < V$, leading to a downward revision of ξ: $\xi_{+1}(z',s') = \xi - \mu_0(z',s') < \xi$. In this case, compensation in the next period will be lower: $w_{+1}(z',s') < w$.

In summary, with poaching offers only, the optimal contract maintains a constant continuation value and compensation until either a competitive poaching offer arrives, resulting in an upward revision, or a negative productivity shock occurs, binding the participation constraint and prompting a downward revision in both the continuation value and pay.

Dynamic Moral Hazard. Assume there is moral hazard but no poaching offers. The problem essentially reduces to the repeated moral hazard problem described by Spear and Srivastava (1987), with one subtle difference. In Spear and Srivastava's model, exerting effort leads to a higher output today but has no impact on future output. Thus, the productivity shock is *i.i.d.* across time. Here, exerting effort is more likely to lead to higher productivity in all future periods. In other words, effort can be understood as an investment in human capital that has an inherently persistent effect. This dynamic nature of effort plays an important role in the full model with job-to-job transitions.

However, in the absence of poaching firms, the optimal contract is not substantially different from the one characterized by Spear and Srivastava, as I demonstrate now. Following Spear and Srivastava, I suppress the participation constraints. The contracting problem is then formulated with the objective given in (BE-F), subject to the promise-keeping constraint (PKC) and the incentive compatibility constraint (IC). I also drop s' from the equations. Let μ denote the multiplier of the incentive compatibility constraint. The first-order conditions and envelope theorem give the evolution of ξ,

$$\xi_{+1}(z') = \xi + \mu(1 - g(z'|z)). \tag{4}$$

The term $\mu(1 - g(z'|z))$ reflects the changes in ξ that are needed to illicit effort. Inserting (1) yields the wage dynamics:

$$\frac{1}{u'(w_{+1}(z'))} - \frac{1}{u'(w)} = \mu(1 - g(z'|z)), \tag{5}$$

where $w_{+1}(z')$ is the next period compensation when z' is realized. Since $\mu > 0$ and $g(z'|z)$ is decreasing in z', (5) implies that there exists $\hat{z}(z) \in (\underline{z}, \overline{z})$ that depends on z, such that $w(z') > w$ for $z' > \hat{z}(z)$ and $w(z') < w$ for $z' < \hat{z}(z)$. In other words, to motivate the executive to exert effort, the next period's compensation rewards high output and punishes low output. Moreover, both the next period flow pay $w_{+1}(z')$ and the continuation value $W(z')$ are strictly increasing in z'.[16] As such, the main properties established by Spear and Srivastava (1987) continue to hold.[17]

Dynamic Moral Hazard and Poaching Offers. In the full model, which incorporates both moral hazard and poaching offers, the evolution of the multiplier ξ is given by:

$$\xi_{+1}(z', s') = \xi + \mu(1 - g(z'|z)) - \mu_0(z', s') + \mu_1(z', s'). \tag{6}$$

[16] This is because $W(z)$ is the fixed point of the Bellman equation: $W(z) = u(w(z)) - c + \beta \sum_{z'} W(z')\gamma(z'|z)$. Take the right-hand side as an operator on $W(\cdot)$. Assume that $W(z')$ weakly increases in z'. With the strict monotonicity of $w(z)$ proved above, and the assumed monotonicity of γ, it is immediate that $W(z)$ is strictly increasing in z.

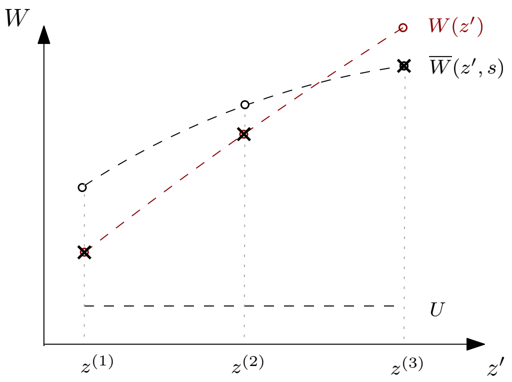
[17] However, since z has a persistent effect, meaning that z influences the distribution from which z' is drawn, some properties regarding the stationarity of the optimal dynamic contract, as outlined by Spear and Srivastava (1987), no longer hold. For instance, based on the first-order conditions, we have:

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

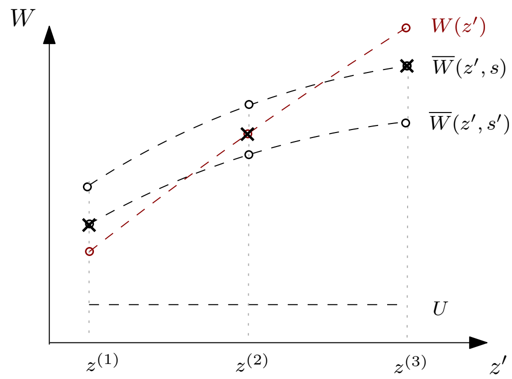
If $z' = \hat{z}(z)$, then in the model of Spear and Srivastava (1987), the contract at $t+1$ would be the same as the current contract, implying $W(\hat{z}(z)) = V$. However, in the present context, this does not generally hold because when $\hat{z} \neq z$, the distribution of future productivity changes.

This equation synthesizes (3) and (4), showing that $\xi_{+1}(z',s')$ adjusts in two distinct ways relative to ξ. The term $\mu(1-g(z'|z))$ reflects the change needed to induce effort, while $-\mu_0(z',s') + \mu_1(z',s')$ accounts for the influence of the binding participation constraints.

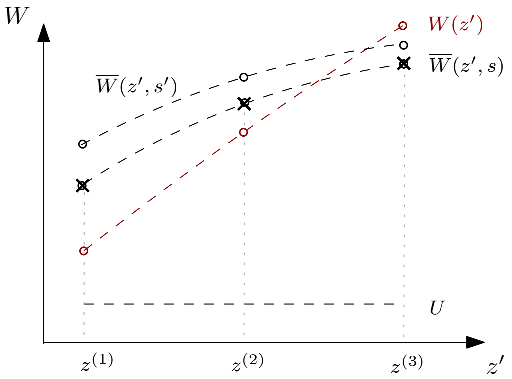
In the absence of binding participation constraints—when $\mu_0(z',s') = \mu_1(z',s') = 0$ for all (z',s')—Eq. (6) reduces to (4), and ξ evolves solely to induce effort. A higher ξ typically leads to a higher $\xi_{+1}(z',s')$. Consequently, an executive who is rewarded today is likely to experience a sequence of high flow payments in subsequent periods.



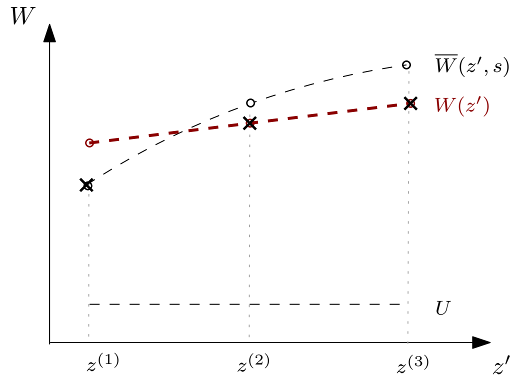
(a) No poaching offers $s' = s^{(0)}$



(b) Poaching firm $s' < s$ and competitive



(c) Poaching firm $s' > s$



(d) μ is small and no poaching offers

Figure 3: Illustration of $W(z',s')$ in four possible states

In other scenarios, participation constraints become binding. To illustrate, Figure 3 depicts the continuation values $W(z',s')$ prescribed by the optimal contract across four realizations of s'. For simplicity, I consider three possible realizations for z', denoted by $\{z^{(1)}, z^{(2)}, z^{(3)}\}$. The dashed black curves connect the circular markers, each of which represents the maximum value a firm is prepared to bid. These curves delineate the bidding frontiers for both the incumbent firm, $\overline{W}(z',s)$, and any potential poaching firm, $\overline{W}(z',s')$, should a poaching offer arise. The dashed red curve links the three circular markers corresponding to the continuation values, $W(z')$, which are derived directly from the multiplier update rule $\xi_{+1}(z',s') = \xi + \mu(1-g(z'|z))$. This curve reflects how the contract adapts solely in response to incentive considerations. The horizontal dashed line at the bottom represents the outside option value, U. The contract values offered to the executive are represented by the crossed points, showing the optimal contract under the binding constraints.

Panel (a) illustrates a scenario where no poaching offer is present, yet the participation constraint (PC-F) binds at $z' = z^{(3)}$. Absent this constraint, the continuation value would have been set higher, $W(z^{(3)}) > \overline{W}(z^{(3)}, s)$, but the incumbent firm s lacks the capacity to bid at that level.

Panel (b) depicts a situation where the poaching offer becomes competitive at $z' = z^{(1)}$. Here, $W(z')$ is constrained at both high and low productivity levels. The incumbent's participation constraint binds at $z' = z^{(3)}$, reflecting the incumbent firm's limited ability to bid, while the executive's participation constraint binds at $z' = z^{(1)}$ due to the poaching firm.

Panel (c) examines a case where the poaching firm is larger than the incumbent firm. In this scenario, for any realization of z', the executive transitions to the external firm, which offers a continuation value of $\overline{W}(z', s)$.

Panel (d) illustrates a scenario where the red dashed curve represents the moral hazard component of the continuation values and is flatter than the bidding frontier. This implies that the cost of effort is low, so a small μ is sufficient to motivate the executive's effort. In this case, the continuation value would be adjusted more sharply downward following adverse realizations of z'. The downward revision resulting from adverse productivity shocks, as proposed by Postel-Vinay and Turon (2010), can be considered a special case where $\mu = 0$.

I summarize the characterization of the optimal contract in the following proposition.

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

(i) The current period compensation w is determined by

$$\frac{\partial \Pi(V, z, s)}{\partial V} = -\frac{1}{u'(w)}. \tag{7}$$

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

$$\frac{\partial \Pi(W(z'), z', s,)}{\partial W(z')} - \frac{\partial \Pi(V, z, s)}{\partial V} = -\mu(1 - g(z'|z)), \tag{8}$$

where μ is the multiplier of the (IC) constraint, then

$$W(z', s') = \begin{cases} \overline{W}(z', s) & \text{if} \quad s' \geq s \ \text{or} \ W(z') > \overline{W}(z', s), \\ \overline{W}(z', s') & \text{if} \quad s' < s \ \text{and} \ W(z') < \overline{W}(z', s'), \\ W(z') & \text{otherwise.} \end{cases} \tag{9}$$

Proof. See the Appendix. □

Condition (7) merely restates condition (1). Condition (8) defines continuation value $W(z')$, which applies when the poaching offer does not trigger binding participation constraints. Condition (9) covers various situations where one of the participation constraints binds.

Notably, throughout these cases, whenever participation constraints are binding, the evolution of the contract value becomes independent of the previously promised value V (or, equivalently, the multiplier $\xi_{+1}(z', s')$ evolves independently of ξ). This feature of dynamic contracts, where future terms decouple from past values, is referred to by Kocherlakota (1996) as *amnesia*. However, in my model, the impact of prior effort is not entirely nullified. Specifically, the bidding frontiers of both the incumbent and poaching firms are functions of the executive's productivity z', which is itself influenced by effort exerted in earlier periods. As a result, the history of effort continues to affect the competitive landscape, as poaching firms base their bids on the productivity that has been shaped by the previous effort. This scenario introduces an additional layer of incentives for the executive— executives are motivated not only by the immediate compensation from their current firm but also by the potential for more lucrative offers from outside firms. This new incentive will play a central role in explaining the incentive premium.

3.2.3 The Wage Dynamics

A novel prediction of my model is that the within-job wage follows a constrained stochastic process with random boundaries. For the purpose of exposition, I impose a log-utility function. Inserting $\xi = 1/u'(w) = w$, I first define

$$\tilde{w}_{+1}(z') \equiv w + \mu(1 - g(z'|z))$$

as the component of the wage driven solely by the moral hazard problem. Since $\mathbb{E}[g(z'|z)] = 1$, this component follows a random walk process.

Furthermore, $\tilde{w}_{+1}(z')$ is bounded from above by the maximum pay that the incumbent firm is willing to give and from below by the maximum pay that the poaching firm is willing to give. The maximum wage that the incumbent firm s is willing to pay for an executive with productivity z' is

$$\overline{w}(z', s) \equiv -\Pi_W(\overline{W}(z', s), z', s).$$

And the maximum wage that the outside firm s' is willing to pay is

$$\overline{w}(z', s') \equiv -\Pi_W(\overline{W}(z', s), z', s').$$

Given that $s' \leq s$ (namely, that the executive stays with the current firm), $\overline{w}(z', s) > \overline{w}(z', s')$, and the wage in the next period, $w_{+1}(z', s')$, is determined by[18]

$$w_{+1}(z', s'|s' \leq s) = \min\Big\{ \max\{\tilde{w}_{+1}(z'), \overline{w}(z', s')\}, \overline{w}(z', s) \Big\}.$$

[18]If $s' > s$, then $w_{+1}(z', s') = -\Pi_W(\overline{W}(z', s), z', s')$, which can be lower than the previous wage.

There are three observations to make here. First, as in other models with sequential auctions, the upward pushes provided by bidding from outside firms, $\overline{w}(z', s')$, result in a series of positive steps for the wages, which gives rise to an upward drift over time. In other words, pay is back-loaded in expectation.

Second, despite the boundaries being imposed, since both $\tilde{w}_{+1}(z')$ and the boundaries are strictly increasing in z', $w_{+1}(z', s'|s' \leq s)$ is also increasing in z'. This contrasts with the wage dynamics in Postel-Vinay and Robin (2002), where wages either stay constant or increase in response to an outside offer.

Third, and most importantly, there is no downward wage rigidity within a job spell. Wages can increase either as a reward designed to induce effort or as a response to an outside offer. Wages can also decrease, either as a punishment intended to induce effort or as a result of adverse productivity shock that forces the incumbent firm to reduce wages. The downward renegotiation mechanism is triggered when the contract value is relatively close to the incumbent firm's bidding frontier.[19]

3.3 Explaining the Firm-size Incentive Premium

Thus far, I have developed and solved a dynamic contract model. It is important to emphasize that, under the assumptions of binary effort and homogeneous effort cost, the required incentive in utility terms is constant across executives. This implies that, for a given level of compensation, the amount of incentive pay in dollar terms should be uniform across executives, irrespective of firm size. As a result, within this framework, the moral hazard problem alone does not generate the *conditional* firm-size incentive premium.[20]

Nevertheless, because larger firms are better equipped to counter more generous poaching offers, their executives' total compensation tends to be higher. Given a concave utility function, this necessitates a higher incentive pay in dollar terms for executives in larger firms. Consequently, a job ladder model, when combined with moral hazard, naturally generates the *unconditional* firm-size incentive premium. This unconditional premium emerges in any model where total compensation scales with firm size. For example, Edmans, Gabaix and Landier (2009) extend the competitive talent assignment model developed by Gabaix and Landier (2008) to show that just as total pay increases with firm size, so does incentive pay under optimal contracting.

However, the fact that total pay increases in firm size is insufficient to generate a conditional

[19] As a side remark, the agency issue modeled here could plausibly explain wage decreases in the general labor market, where real wages might decline due to an employee's poor performance. There is growing evidence that, in many industries, firms are becoming increasingly reliant on performance-based pay structures. When an employee underperforms, this is often reflected in reduced bonuses, lower commissions, or even adjustments to base salary, all of which are intended to incentivize improved productivity. See Prendergast (1999) for further discussion.

[20] If the effort variable were continuous, the heterogeneity in returns to effort across executives and firms would naturally give rise to a size-related pattern. This suggests that the model could be easily extended to incorporate other firm-size-related characteristics.

firm-size incentive premium. Once total pay is controlled for, the size premium for incentives tends to disappear. For the conditional premium to exist, there must be a reason why larger firms need to provide more incentives in utility terms. My model presents one such reason, which I will now examine.

The idea is that poaching offers create additional incentives that can act as substitutes for direct contractual incentives. The strength of the poaching-offer incentives diminishes with the size of the incumbent firm. To compensate, larger firms are required to offer higher incentives in utilities through a compensation package containing a higher proportion of performance-based pay. This explains the conditional firm-size incentive premium.

In the following analysis, I assume that z' and s' follow continuous distributions on supports $[\underline{z}, \overline{z}]$ and $[\underline{s}, \overline{s}]$, respectively, with differentiable density functions. This leads to that $\Pi(V, z, s)$ is continuously differentiable in z and s. I denote the cumulative density functions by $\Gamma(z, z')$ for z' if $e = 1$, $\Gamma^s(z')$ if $e = 0$, and $P(s')$ for s'.[21]

Define an *incentive operator* $\mathcal{I}$ that quantifies the incentives an executive derives from any scheme $\{W(z')\}_{z' \in \mathbb{Z}}$, namely, the differential in utilities between exerting 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 $\mathcal{I}(\cdot)$, the incentive compatibility constraint can be reformulated as follows:

$$\underbrace{\lambda(1 - P(s))\mathcal{I}\Big(\overline{W}(z', s)\Big)}_{\text{poaching-offer incentives}} + \underbrace{(1 - \lambda)\mathcal{I}\Big(W(z', s^{(0)})\Big) + \lambda \int_{s' \leq s} \mathcal{I}\Big(W(z', s')\Big) P(ds')}_{\text{compensation incentives}} \geq \frac{c}{\beta}. \tag{IC$'$}$$

On the left-hand side, I decompose and label the incentives according to whether they are observable within the compensation package. If $s' > s$, which happens with probability $\lambda(1 - P(s))$, the executive transitions to the outside firm, and the incentives from this new employer will not be reflected in the compensation package offered by the current firm. I refer to these incentives as the poaching-offer incentives.

Conversely, if there is no poaching offer (equivalent to $s' = s^{(0)}$) or if $s' < s$, the executive remains with the incumbent firm, and the incentives are provided through performance-based compensation. It is crucial to highlight that, even in this scenario, poaching offers may play a role by setting the lower bound for the continuation value in each state. Despite their influence, the incentives generated by these poaching offers will still be captured in the compensation package provided by the current firm. Hence, I categorize them as compensation incentives.

To examine how poaching-offer incentives vary with s, consider a marginal increase in the incumbent firm's size, from s to $\tilde{s} > s$. This change introduces two effects on poaching-offer in-

[21] It is convenient to view these distributions as the limits of the discrete distributions in previous sections as $n_z, n_s \to \infty$, with $\underline{z}, \overline{z}, \underline{s}, \overline{s}$ fixed. I use continuous distributions because it allows for the use of differentiation to establish results. The predictions hold for discrete values of z and s, as demonstrated in the following numerical results.

centives. First, the job ladder effect: as s increases, the likelihood of encountering a larger outside firm, $1 - P(s)$, decreases. Second, the wealth effect: the poaching-offer incentives, $\mathcal{I}\Big(\overline{W}(z', s)\Big)$, become $\mathcal{I}\Big(\overline{W}(z', \tilde{s})\Big)$. Given that $\tilde{s} > s$, the bidding frontier shifts upward, such that $\overline{W}(z', \tilde{s}) > \overline{W}(z', s)$. With diminishing marginal utility, the incentives derived from this higher bidding frontier diminish, i.e., $\mathcal{I}\Big(\overline{W}(z', s)\Big) > \mathcal{I}\Big(\overline{W}(z', \tilde{s})\Big)$, assuming the utility function exhibits sufficient concavity. Proposition 2 provides a sufficient condition for the degree of concavity required in u.

These effects combine to imply that, to ensure the total incentives on the left-hand side of (IC′) are adequate to cover the adjusted effort cost c/β, firm $\tilde{s}$ must offer more incentives through performance-based pay relative to those offered by firm s. This mechanism underpins the observed firm-size incentive premium conditional on total pay, whereby executives in larger firms receive a higher fraction of incentive pay in their compensation packages.

Proposition 2. *Poaching-offer incentives decrease with the current firm size s if u is sufficiently concave, i.e., for all $z \in \mathbb{Z}$ and $s \in \mathbb{S}$, it holds that*

$$-\frac{u''(\overline{w})}{u'(\overline{w})} > \frac{\Pi_{zs}(\overline{W}, z, s)}{\Pi_z(\overline{W}, z, s)\Pi_s(\overline{W}, z, s)}, \tag{10}$$

where $\overline{W}$ satisfies $\Pi(\overline{W}, z, s) = 0$, and $\overline{w}$ is the flow compensation corresponding to $\overline{W}$.

Proof. See the Appendix. □

The intuition is as follows. The incentives $\mathcal{I}\Big(\overline{W}(z', s)\Big)$ are effectively a weighted sum of $\frac{\partial \overline{W}(z', s)}{\partial z'}$ over the range of possible z' outcomes. The steeper the slope of $\overline{W}(z', s)$ with respect to z', the greater the incentives generated to induce effort. To show that $\mathcal{I}\Big(\overline{W}(z', s)\Big)$ decreases in s, it is sufficient to show that $\frac{\partial \overline{W}(z', s)}{\partial z'}$ diminishes as s increases. Implicit differentiation yields

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

where $\overline{w}$ represents the flow compensation corresponding to $\overline{W}$. The numerator reflects the contribution of z to the firm's value, including both the current period's flow profit and future profits. The denominator directly follows from the optimal contract condition (7).

As s increases, two opposing forces are at play. On the one hand, if complementarity in firm value exists between z and s, such that $\Pi_{zs}(\overline{W}, z, s) > 0$, the numerator increases with s, potentially leading to higher poaching-offer incentives as the incumbent firm's size grows. On the other hand, larger firms are able to bid higher, resulting in an increase in $\overline{w}$ with s, which in turn reduces $u'(\overline{w})$, causing the denominator to rise with s. When the utility function is sufficiently concave, the second force dominates, effectively counteracting the supermodularity between executive productivity and firm size, as specified in the proposition. If the present firm value exhibits no complementarity between z and s, i.e., if $\Pi_{zs}(\cdot) \leq 0$, then condition (10) is trivially satisfied.

I shall clarify the distinction between complementarity in firm value, $\Pi_{zs}(\cdot) \geq 0$, and complementarity in the output function, $f_{zs}(\cdot) \geq 0$. The latter, and its stronger variants, are widely recognized in the literature as crucial conditions needed to generate sorting in equilibrium. To illustrate the difference, consider restricting the impact of z and s on Π to the current period only. Under this simplification, condition (10) reduces to

$$-\frac{u''(\overline{w})}{u'(\overline{w})} > \frac{f_{zs}}{f_z f_s},$$

where the right-hand side represents a normalized measure of supermodularity. In this case, risk aversion must outweigh the supermodularity present in the production function. However, it is crucial to recognize that a poaching firm's offer is not limited to a one-period flow of compensation; rather, it extends to a sequence of payments across all future contingencies. As a result, the relevant measure of complementarity in this context is that of the firm value, rather than merely of the flow output function.

Since the utility concavity is pivotal for the results, it is essential to determine the degree of concavity required. In the quantitative exercises (see Section 4), I employ a constant relative risk aversion (CRRA) utility function $u(w) = \frac{w^{1-\sigma}}{1-\sigma}$ with $\sigma > 0$, coupled with a multiplicative production function $f(z,s) = \alpha_0 s^{\alpha_1} z$, where $\alpha_0 \in (0,1)$ and $\alpha_1 \in (0,1]$. The multiplicative form of the production function is standard in this literature (see Gabaix and Landier 2008), and it captures the idea that the executive's effort scales across the entire firm, up to a factor of α_0. This production function exhibits constant returns to scale when $\alpha_1 = 1$ and decreasing returns to scale when $\alpha_1 < 1$.

Given the strong firm-size effect embedded in this production function, one might anticipate that $u(w)$ would need to exhibit substantial concavity to account for the firm-size incentive premium. However, I find that in numerical exercises, $\sigma > 1$ is sufficient to ensure that poaching-offer incentives decrease with firm size. This degree of concavity is relatively mild. Indeed, it aligns with the σ values commonly calibrated or estimated in related research.[22]

In summary, Proposition 2 implies that larger firms must offer their executives greater incentives in utility terms. Empirically, this suggests that the (conditional) firm-size incentive premium (Fact 4) must hold. Moreover, an increase in the poaching offer arrival rate, λ, amplifies the impact of poaching-offer incentives. Consequently, the conditional premium should be more pronounced when executives have more transition opportunities (Fact 5). Finally, Proposition 2 implies that, at the individual level, when an executive transitions to a larger firm, the proportion of incentive pay in their total compensation increases, with the increase being positively correlated with the size of the destination firm (Fact 6).

[22]For example, Hall and Murphy (2000) calibrate σ between 2 and 3 in their study of CEO incentive pay; Dittmann and Maug (2007) and subsequent works on the convexity of CEO incentive compensation assume $\sigma > 1$; Balke and Lamadon (2022) estimate $\sigma = 1.68$ using matched employer-employee data from Sweden for the general labor market.

4 Quantitative Analysis

To evaluate the model's ability to capture the firm-size incentive premium, I calibrate it using data on executive compensation and job transitions.

4.1 Calibration

I make the following parametric assumptions. I assume a CRRA utility function $u(w) = \frac{w^{1-\sigma}}{1-\sigma}$, and a production function of the form $f(z, s) = e^{\alpha_0} s^{\alpha_1} z$, where $\alpha_0 < 0$. I model the process of productivity as an $AR(1)$ process, namely,

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

where ϵ_t follows a normal distribution $N(0, \sigma_\epsilon)$, and the mean for effort level $e = 0$ is normalized to zero. The process is transformed into a discrete Markov Chain using the method of Tauchen (1986) on a grid of 6 points. I set the sampling distribution of firm size as a truncated log-normal distribution with mean μ_s and standard deviation σ_s. The upper and lower bounds of the truncated distribution are calibrated to the 0.99 and 0.01 quantiles of market capitalization in the data.

Table 4: Calibration

Parameters	Description	Value	Target
η	Death probability	0.0695	Exit rate
λ	Offer arrival probability	0.3162	Job-to-job transition rate
ρ_z	$AR(1)$ coefficient of z	0.8004	Moments of operating profitability
μ_z	Mean of z for $e = 1$	0.0279	
σ_z	Std. of z	0.3461	
μ_s	Mean of poaching firm size	1.2356	Moments of firm size
σ_s	Std. of poaching firm size	2.5795	
c	Cost of effort	0.0814	Moments of $\log(inc)$
σ	Relative risk aversion	1.1038	Reduced-form coefficient $\beta_{inc-pay}$
α_0	Production function parameter	-1.5534	Moments of $\log(totalpay)$ and
α_1	—	0.527	Reduced-form coefficient $\beta_{pay-size}$

I have 12 parameters to calibrate: the death probability η; the offer arrival probability λ; the mean, standard deviation, and autoregressive coefficient of executive productivity μ_z, σ_z, and ρ_z; the mean and standard deviation of firm size distribution μ_s and σ_s; the cost of effort c; the relative risk aversion σ; and the production function parameters α_0 and α_1. I set the discount rate $\tilde{\beta} = 0.9$ for the model is solved annually. I set $\eta = 0.0695$ to match the executive exit rate, which suggests that almost 7% of the executives exit the executive labor market each year.

Table 5: Data and model-simulated moments

Moments	Data	Model
J-J transition rate	0.0443	0.0473
β_{z1}	0.7683	0.6299
$\text{Mean}(profit)$	0.1260	0.1144
$\text{Var}(profit)$	0.0144	0.0160
$\text{Mean}(\log(mktcap))$	7.4515	7.4806
$\text{Var}(\log(mktcap))$	2.3060	2.1610
$\text{Mean}(\log(totalpay))$	7.2408	7.2665
$\text{Var}(\log(totalpay))$	1.1846	0.8960
$\beta_{pay-size}$	0.3830	0.2822
$\beta_{inc-pay}$	1.1063	1.1997
$\text{Mean}(\log(inc))$	8.4994	8.478
$\text{Var}(\log(inc))$	3.4438	3.3587

I calibrate the other 10 parameters simultaneously by simulating the model to match the targets in the data. Table 4 lists the model parameters, calibrated values, and the targeted moments. Table 5 lists the model simulated moments and the data moments. I calibrate $\lambda = 0.3164$ to match the job-to-job transition rate of 4.43%. The magnitude of λ indicates that, on average, an executive will receive an outside offer every three years.

Measuring the productivity of top executives within a firm is generally challenging. To approximate executive productivity, I use the firm-year level operating profitability when the executive is working at the firm. Operating profitability, also known as operating return on assets (OROA), is calculated by dividing earnings before interest, taxes, depreciation, and amortization (EBITDA) by total assets. I calibrate $\rho_z = 0.8004$, $\mu_z = 0.0279$, and $\sigma_z = 0.3461$ to match the mean and variance of operating profitability, and the autoregressive coefficient β_{z1} in the following regression:

$$profitability_{it} = \beta_{z0} + \beta_{z1} profitability_{it-1} + \epsilon_{it,0},$$

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

μ_s and σ_s govern the job offer distribution. The higher μ_s is, the more likely that executives can transition to larger firms, and the larger the mean of firm size. Similarly, the higher σ_s is, the more heterogeneous the poaching firms are, and both the mean and variance of firm size increase. I calibrate $\mu_s = 1.2356$ and $\sigma_s = 2.5795$ to match the mean and variance of $\log(mktcap)$. Comparing the data and model-simulated mean and variance of $\log(mktcap)$, it seems that using a log-normal distribution is sufficient to match the firm size distribution in the data.

In the production function, α_0 determines the maximum level of compensation that a firm is willing to give, and α_1 governs how the total compensation varies with firm size. I calibrate $\alpha_0 = -1.5534$ and $\alpha_1 = 0.527$ to match the mean and variance of $\log(totalpay)$ and the reduced-form coefficient $\beta_{pay-size}$ in the following regression:

$$\log(totalpay_{it}) = \beta_1 + \beta_{pay-size} \log(mkcap_{it}) + \epsilon_{it,1}.$$

The final part of the calibration concerns σ and c. These parameters govern the level of incentives and how the incentives change with total compensation. To align with the variable *inc* in the data, I construct an *inc* variable in the simulated data. I calibrate $c = 0.0814$ and $\sigma = 1.1038$ to match the mean and variance of $\log(inc)$, and the coefficient $\beta_{inc-pay}$ from regression

$$\log(inc_{it}) = \beta_2 + \beta_{inc-pay} \log(totalpay_{it}) + \epsilon_{it,2}.$$

4.2 Predicting Firm-size Incentive Premium

I intentionally leave the conditional and unconditional incentive premia untargeted. Instead, I estimate them using the model-simulated data, and compare the estimates with the real-world data estimates to examine whether the model mechanism matches up with the data. Specifically, I estimate the following regression in both the real-world and model-simulated data:

$$\log(inc_{it}) = \beta_5 + \beta_{inc-size} \log(size_{it}) + \beta_6 \log(totalpay_{it}) + \epsilon_{it,3}. \tag{12}$$

Here, $\beta_{inc-size}$ captures the conditional firm-size incentive premium. When $\log(totalpay_{it})$ is not controlled, $\beta_{inc-size}$ captures the unconditional firm-size premium.

Table 6: Predictions of firm-size incentive premia

	Data	Benchmark	Poaching-offer incentives ignored	Higher offer arrival rate ($\lambda = 0.6$)	Lower offer arrival rate ($\lambda = 0.1$)
	(1)	(2)	(3)	(4)	(5)
Conditional premium	0.3581	0.3122	-0.0444	0.4299	0.1964
Unconditional premium	0.6186	0.6507	0.4202	0.7093	0.4076

Table 6 reports the results. The first row shows the incentive premia estimated from regression (12) with $\log(totalpay)$ controlled, while the second row presents the premia without controlling for $\log(totalpay)$. Focusing on the first two columns: Column (1) displays the premia estimated using real data (replicating Columns 1 and 2 of Table 2), and Column (2) presents the model-simulated premia. Comparing these two columns, I find that the model quantitatively captures both the conditional and unconditional firm-size premia well. This result provides additional reassurance that the model's mechanism is important in explaining the firm-size premium.

To isolate the impact of poaching offers on incentive contracts, I simulate a counterfactual scenario in which firms, when designing contracts, (mistakenly) ignore the incentives generated by poaching offers. By "ignoring" poaching-offer incentives, I mean that firms assume that all incentives must be provided exclusively through their own compensation packages. This counterfactual scenario modifies the incentive compatibility constraint (IC$'$) by omitting the first term on the left-hand side, thereby relying solely on $W(z', s')$ to deliver the incentive. I shall emphasize that in this scenario, the job ladder structure is preserved: executives are still poached by outside firms, can still negotiate with incumbent employers using outside offers, and can still transition to outside firms. Consequently, while executive total compensation is expected to increase with firm size, the incentives embedded in potential poaching offers are eliminated.

Column (3) shows that in this case, the incentive premium with $totalpay$ controlled essentially becomes zero. Meanwhile, the incentive premium when $totalpay$ is not controlled decreases to 0.4202. The existence of the unconditional premium is driven by the fact that total compensation is expected to be higher for executives working at larger firms (because of the job ladder). Nevertheless, the unconditional premium is lower than the counterpart observed in the data.

To further evaluate the contribution of poaching-offer incentives, in Figure 4, I compare the model-generated inc in the benchmark, in which poaching-offer incentives are present (Column 2 of Table 6), and the model variant, in which poaching-offer incentives are ignored (Column 3). The higher inc in the model variant reflects the contribution made by poaching offers. I divide firms into ten groups based on firm size (Group 0 contains the smallest firms). The upper panel of Figure 4 shows the box plot of $\log(inc)$ across firm size. There are two observations to make here. 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. Specifically, smaller firms need to provide a higher $\log(inc)$ when poaching-offer incentives are ignored (the orange box) compared to that required by the benchmark model (the green box). In the lower panel of Figure 4, I calculate the contribution 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 proportion rapidly decreases to around 15% in medium-sized firms and almost vanishes for the largest firms.

Finally, in Columns (4) and (5) respectively, I simulate a version of the model with a high job-arrival probability $\lambda = 0.6$ and a low job-arrival probability $\lambda = 0.1$. These model variants show that when there are more (fewer) job offers, both the conditional and the unconditional firm-size premia are higher (lower).

5 Alternative Explanations for the Firm-size Incentive Premium

Thus far, I have demonstrated that poaching-offer incentives can explain the firm-size incentive premium both theoretically and quantitatively. There may be other reasons for the premium; I

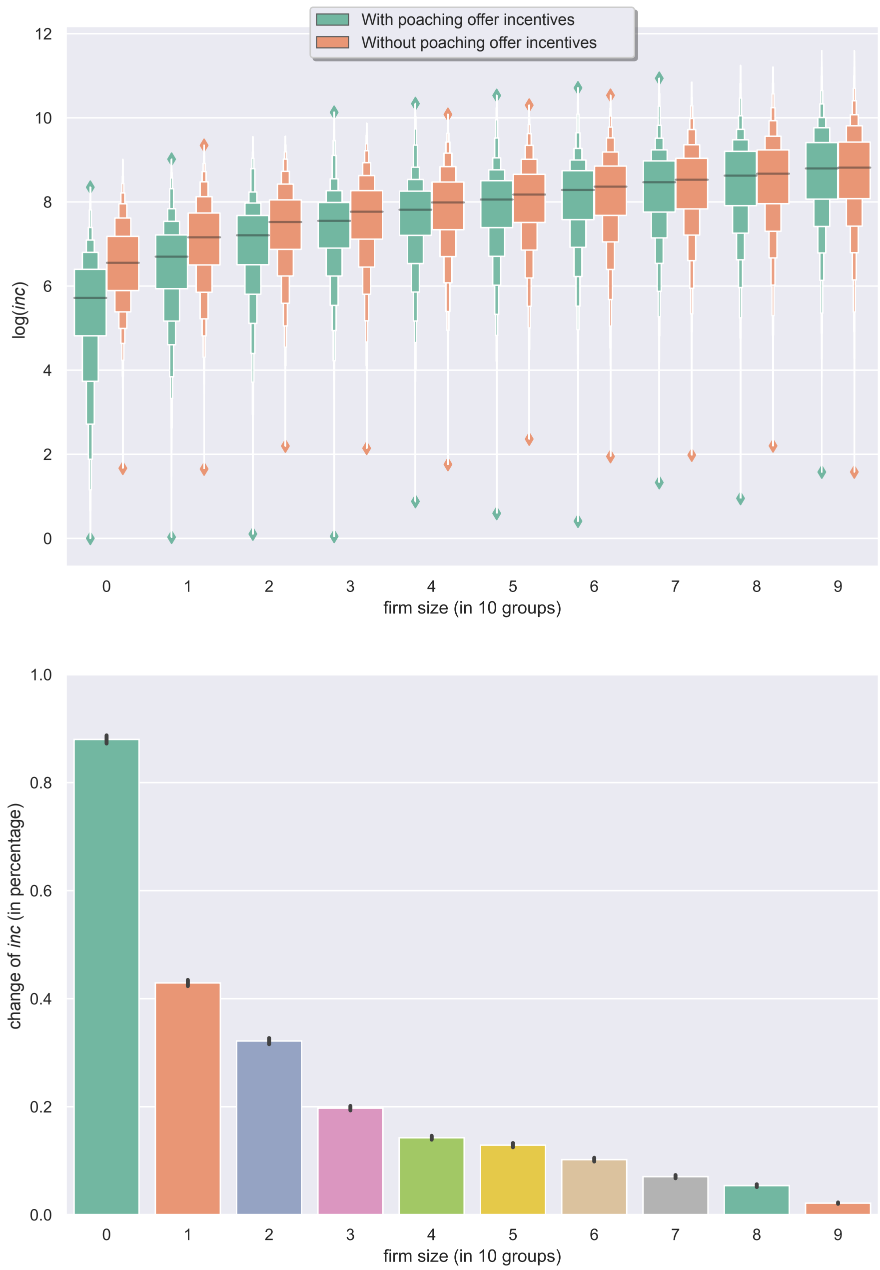


Figure 4: The distribution of $\log(inc)$ and the contribution of poaching-offer incentives

Note: The upper panel displays the distribution of $\log(inc)$ across ten firm-size groups. The green boxes represent the benchmark model, in which poaching-offer incentives are taken into account when designing the contract. The orange boxes represent a model variant in which poaching-offer incentives have been mechanically removed. The increase in the median for each group indicates that when there are no poaching-offer incentives, firms must provide higher contractual incentives measured by inc. The lower panel calculates the fraction of poaching-offer incentives using $\frac{inc^v - inc^b}{inc^b}$, where inc^b represents the wealth–performance sensitivity in the benchmark model, and inc^v represents the sensitivity in the model variant in which poaching-offer incentives are ignored.

now explore alternative explanations and compare them to my theory.[23]

Differences in Agency Problems. One potential explanation for the firm-size incentive premium lies in the differences in moral hazard problems across firms. Larger firms may have more precise metrics for evaluating executive performance, which could require higher incentive pay to better align the interests of executives and shareholders. Additionally, the returns to effort may be higher in larger firms due to the complementarity between the asset scale and executive productivity. This higher return to effort might make it optimal for larger firms to offer higher incentive pay to motivate executives to exert more effort. Lastly, if the selection process for executives is more competitive at larger firms, those who are hired may need differentiated incentive packages that account for their higher effort costs or distinct risk profiles. These factors could naturally lead to a firm-size incentive difference.

Talent Matching Between Executives and Firms. Another explanation for the incentive premium could be the use of incentive pay as a tool for selecting executive talent. High-ability executives are more likely to achieve superior performance, making them naturally inclined toward contracts with higher incentive pay. Consequently, the firm-size incentive premium could be seen as the equilibrium outcome of a search-frictionless matching process, where larger firms offer higher incentives to attract top talent. This explanation can be further enriched by a signaling story. That is, executives might begin their careers at smaller firms to demonstrate their managerial abilities. As they successfully transition to larger firms, it becomes optimal for these firms to offer higher incentives, recognizing that these executives have already proven their capabilities and potential for success.

Discussion. None of the aforementioned mechanisms are present in my model. The principal-agent problem is homogeneous across firms because I have assumed binary effort levels, a common effort cost for all executives, and a common distribution from which new productivity is drawn. Under these assumptions, the incentives required to motivate an executive are independent of other firm or executive characteristics. Additionally, talent matching does not exist in my model, as I have assumed random search, and signaling is unnecessary because I have imposed observable managerial skills. Consequently, my focus is solely on the role of poaching offers. However, I do not dismiss the possibility that these alternative explanations could complement my theory of the firm-size incentive premium, noting that they may indeed be important factors that have policy implications.

Nevertheless, reconciling these alternative explanations with the evidence presented in Section 2 is challenging. While most of the alternative explanations are not directly related to the

[23] I am grateful to the anonymous referee for suggesting several of the alternative explanations discussed in this section.

Table 7: Firm-size incentive premium (non-incentive pay $< 0.9M$)

	$\log(inc)$		$\log(inc^s)$		inc^f
	(1)	(2)	(3)	(4)	(5)
$\log(mktcap)$	0.5857*** (0.0172)	0.3663*** (0.0248)	0.1526*** (0.0114)	0.2666*** (0.0202)	6.6037*** (0.1504)
$\log(totalpay)$		0.6137*** (0.0298)		-0.3215*** (0.0285)	
Obs.	169,199	168,819	168,730	168,729	222,659
adj. R^2	0.378	0.442	0.213	0.260	0.272

Note: This table presents regression results using the sample where the sum of the salary and other compensation is less than 0.9 million dollars. Firm size is measured by market capitalization ($mktcap$). Executive total compensation ($totalpay$) is calculated using the annual awarded compensation ($TDC1$ in ExecuComp). The dependent variable in Columns (1) and (2) is $\log(inc)$. The dependent variable in Columns (3) and (4) is $\log(inc^s)$. The dependent variable in Column (5) is the fraction of incentive pay in the total compensation (inc^f), where incentive pay is defined as the sum of any bonus, other annual compensation, restricted stock grants, option grants, and long-term incentive plan payouts. All regressions control for industry-by-year fixed effects. Standard errors, clustered at the firm-year level, are reported in parentheses. Statistical significance $p < 0.001$ is denoted by ***.

executive labor market or job transitions, the incentive premium is. Specifically, in Fact 5, I demonstrate that the firm-size incentive premium is more pronounced in industries that offer more transition opportunities for executives. Therefore, to explain the incentive premium using differentiated moral hazard, a robust microfoundation would be required to show that variations in the moral hazard in large and small firms are positively associated with executive transition rates. Similarly, for the signaling explanation to be convincing, one would need to explain the variation of signaling across industries.[24]

The One-Million-Dollar Rule. The conditional firm-size incentive premium may also be due to tax considerations. Specifically, executive compensation is often structured to minimize tax liabilities. A notable example of this concerns the ”one-million-dollar rule”; this rule was introduced in the U.S. in 1993 and capped the tax deductibility of executive compensation at $1 million per year for publicly traded companies. However, performance-based compensation—including bonuses, stock options, and other forms of incentive pay—was exempt from this limit and could still be deducted for tax purposes. As a result, larger firms, which typically offer higher compensation, might structure pay packages to maximize tax deductibility by increasing the proportion of incentive pay. In contrast, smaller firms, where executive compensation often remains below the $1 million threshold, may face less pressure to do so.[25]

[24]To further support my explanation, I present two additional empirical findings in the online appendix that challenge alternative explanations but align well with my model. First, I demonstrate that the firm-size incentive premium is more pronounced among younger executives. This can be understood by extending my framework to include age as a factor influencing the arrival rate of poaching offers. Second, I show that wage back-loading is more prevalent in larger firms, which is consistent with their greater ability to counter outside offers.

[25]The Tax Cuts and Jobs Act of 2017 removed the performance-based pay exception to the one million dollar rule. Unfortunately, I do not have access to ExecuComp data after 2016.

Using ExecuComp data, I calculate the annual non-incentive (non-deductible) compensation for each executive-year by summing *salary* and *other compensation*. To ensure my estimated firm-size incentive premium is not driven by the one-million dollar rule, I conduct regressions of Table 2, but using a sample where annual non-incentive pay is lower than 0.9 million dollars. The estimates, reported in Table 7, show no substantial changes in the coefficients. I also run the regressions by restricting the non-incentive pay to be lower than 0.8 or 0.7 million. The estimates are consistent. This suggests that the one-million-dollar rule is unlikely to be a primary driver of the firm-size incentive premium.

6 Conclusions

This paper develops a dynamic contracting model in which executives can use poaching offers to negotiate with their current firm. The model introduces a job ladder and is supported by empirical findings from a newly assembled job-to-job transition dataset. The model shows that poaching offers impact executive compensation in two ways: they influence both the overall level of compensation and the structure of the incentive contract. Using poaching-offer incentives, the model explains the firm-size incentive premium, and the model-simulated premium closely aligns with the counterparts estimated in the data.

The model also generates rich wage dynamics. Specifically, with moral hazard embedded in the on-the-job search framework, the model can account for both large wage decreases following significant adverse productivity shocks and smaller wage decreases tied to performance-based pay. Further exploration of these mechanisms and their quantitative significance for the broader labor market presents a promising direction for future research.

References

Abrahám, Arpád, Fernando Alvarez-Parra, and Susanne Forstner, "The effects of moral hazard on wage inequality in a frictional labor market," 2017. Working paper.

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*, 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.

Chemla, Gilles, Liyan Shi, and Alejandro) Rivera, "Too much, too soon, for too long: The dynamics of competitive executive compensation," 2023. Working Paper.

Clementi, Gian Luca and Thomas Cooley, "CEO compensation: Facts," *Review of Economic Dynamics*, 2023, *50*, 6–27.

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.

Dey, Matthew and Christopher Flinn, "An equilibrium model of health insurance provision and wage determination," *Econometrica*, 2005, *73* (2), 571–627.

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, *22* (12), 4881–4917.

—, —, 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," 2005. Working Paper.

— 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," 2021. Nanyang Business School Research Paper No. 21-04.

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

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

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.

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. NBER Working Paper No. 19988.

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

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, "Managerial Capital and the Market for CEOs," 2006. Working Paper 1110, Economics Department, Queen's University.

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 Hélène Turon, "On-the-job search, productivity shock, and the individual earning process," *International Economic Review*, 2010, *51* (3), 599–629.

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

Prendergast, Canice, "The provision of incentives in firms," *Journal of Economic Literature*, 1999, *37* (1), 7–63.

Rogerson, William P., "Repeated moral hazard," *Econometrica*, 1985, *53* (1), 69–76.

Shi, Liyan, "Optimal regulation of noncompete contracts," *Econometrica*, 2023, *91* (2), 425–463.

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.

Taylor, Lucian A., "Why are CEOs rarely fired? Evidence from structural estimation," *The Journal of Finance*, 2010, *65* (6), 2051–2087.

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. Proofs

Proof of Proposition 1

Proof. Recall that $\Pi(W,z,s)$ represents the expected discounted sum of profits from a match between a firm of size s and an executive with productivity z, subject to a promised utility W. The function $\overline{W}(z,s)$ denotes the firm's maximum willingness to pay for the executive, while for the base firm $s^{(0)}$, we have $\overline{W}(z,s^{(0)}) = U$. Further, let $\overline{W}(s) = \sup\{\overline{W}(z,s) \mid z \in \mathbb{Z}\}$. The set of possible promised utility values is thus given by $\mathbb{X} = [U, \sup_{s\in\tilde{\mathbb{S}}} \overline{W}(s)]$. For notational convenience, the dependence of Π on s is suppressed throughout this proof, so that we write $\Pi(W,z)$ instead.

Rewrite the problem. The firm selects a contingent plan of promised continuation utilities for each possible realization of $(z',s') \in \mathbb{Z}\times\mathbb{S}$, denoted by $Y \equiv \{W(z',s')\}_{z'\in\mathbb{Z},s'\in\mathbb{S}}$. Y is the control variable. Assuming a monotone likelihood ratio, I initially conjecture that $W(z',s')$ is nondecreasing in z' and will verify this conjecture later. Let $\mathbb{Y} \equiv \mathbb{X}^{n_z\times n_s}$ denote the set of feasible continuation values. The update to the next period's state, conditional on the realized (z',s'), is given by the mapping $\phi : \mathbb{Y}\times\mathbb{Z}\times\mathbb{S} \to \mathbb{X}$:

$$\phi(Y,z',s') = W(z',s').$$

This mapping selects the realized promised value $W(z',s')$ from the menu $\{W(z',s')\}$.

Given that the promise-keeping constraint is always binding (noting that $u(\cdot)$ is strictly increasing), the current-period flow compensation w can be expressed as a function of Y:

$$w = u^{-1}\left(V + c - \beta\sum_{z'}\sum_{s'} W(z',s')p(s')\gamma(z'|z)\right),$$

where $u^{-1}(\cdot)$ is a convex function. The firm's flow profit is then defined by the function π : $\mathbb{X}\times\mathbb{Y}\times\mathbb{Z}\to\mathbb{R}$:

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

It follows that $\pi(V,Y,z)$ is continuous in V and Y. Given that both $\mathbb{X}$ and $\mathbb{Y}$ are compact-valued, $\pi(V,Y,z)$ is also bounded. Furthermore, since $u(\cdot)$ 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 both V and Y.

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

$$\Pi(V,z) = \max_{Y\in\Omega(z)}\left\{\pi(V,Y,z) + \beta\sum_{z'}\sum_{s'\le s}\Pi\left[\phi\left(Y,z',s'\right),z'\right]p(s')\gamma(z'|z)\right\}, \tag{13}$$

where the firm chooses $Y = \{W(z', s')\}$ from the set

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

To ensure that $\Omega(z)$ is non-empty, I impose the condition that for all $z \in \mathbb{Z}$, the effort cost c satisfies

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

where for all z', $W(z') \in [U, \overline{W}(z', \underline{s})]$, and I account for the probability that the executive does not receive a poaching offer $(1 - \lambda)$. The right-hand side represents the incentive that any firm is capable of providing. Note that lowering U raises the right-hand side, thereby the restriction on c is mild.

The set $\Omega(z)$ possesses the following properties: First, $\Omega(z)$ does not depend on V. Second, $\Omega(z)$ is convex because for any two elements $\{W_0(z', s')\}$ and $\{W_1(z', s')\}$ in $\Omega(z)$, their linear combination $W_\theta(z', s') = \theta W_0(z', s') + (1 - \theta) W_1(z', s')$ for $\theta \in (0, 1)$ also satisfies the constraints of $\Omega(z)$. Finally, $\Omega(z)$ is increasing in the sense that $z_1 \leq z_2$ implies $\Omega(z_1) \subset \Omega(z_2)$. This is evident from the fact that the left-hand side of the incentive compatibility constraint increases with 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 the monotonicity of γ.

Existence. Let $\mathbb{O} = \mathbb{X} \times \mathbb{Z}$ denote the product space. We define $C(\mathbb{O})$ as the space of bounded continuous functions $h : \mathbb{O} \to \mathbb{R}$, equipped with the sup norm: $||h|| = \sup_{o \in \mathbb{O}} |h(o)|$. Given that $\mathbb{Z}$ is a finite set, continuity here refers to the property that for each $z \in \mathbb{Z}$, the z-section of the function is continuous. Now, consider the right-hand side of Eq. (13) as a functional operator T. For a fixed $h \in C(\mathbb{O})$, since ϕ is continuous in Y, the term $\beta \sum_{z'} \sum_{s' \leq s} h \Big[\phi \Big(Y, z', s' \Big), z' \Big] p(s') \gamma(z'|z)$ is also continuous in Y. The problem then reduces to maximizing a continuous function over the compact set $\Omega(z)$. By Berge's Maximum Theorem, the maximum is attained, and thus Th remains bounded. Moreover, Th is continuous, implying that $T : C(\mathbb{O}) \to C(\mathbb{O})$. It is straightforward to verify that T satisfies the conditions of Blackwell's sufficient criteria for a contraction. Thus, T has a unique fixed point $\Pi \in C(\mathbb{O})$.

Monotonicity. To prove that $\Pi(V, z)$ is strictly decreasing in V, fix z and consider a function $h(V, z)$ belonging 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 we have:

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

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

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

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

where the second line uses the strict monotonicity of $\pi(V, Y, z)$ in z. Hence, the fixed point $\Pi(V, z)$ is strictly increasing in z. Similarly, it follows that $\Pi(V, z, s)$ is strictly increasing in s (reintroducing s into the argument list). These monotonicity properties, combined with the continuity of Π, ensure that $\overline{W}(z, s)$ is well defined.

Concavity. To establish that Π is strictly concave in V, fix z and consider a function $h(V, z)$ belonging to the set of bounded continuous functions that are weakly concave in V. The goal is to show that $Th(V, z)$ is strictly concave in V. Take any distinct $V_1 \neq V_2 \in \mathbb{X}$ and define $V_\theta = \theta V_1 + (1 - \theta) V_2$, where $\theta \in (0, 1)$. Let Y_i be the optimal control that attains $(Th)(V_i, z)$ for $i = 1, 2$. Since $\Omega(z)$ does not depend on V, it follows that $Y_\theta = \theta Y_1 + (1 - \theta) Y_2 \in \Omega(z)$. We then have:

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

where the strict inequality follows from the fact that $\pi(V, Y, z)$ is strictly concave in V and Y jointly, and that ϕ is concave in Y. Therefore, the fixed point $\Pi(V, z)$ is indeed strictly concave in V.

Differentiability. The strict concavity of $\Pi(V, z)$ implies that it is almost everywhere differentiable with respect to V. I now demonstrate that $\Pi(V, z)$ is, in fact, differentiable in V everywhere. Suppose, for a fixed z, that $\Pi(\cdot, z)$ is not differentiable at some point $\tilde{V}$, and let $\{\tilde{W}(z', s')\}$ denote the firm's optimal choice at $\tilde{V}$, with a corresponding flow pay $\tilde{w}$. For notational convenience, I will suppress the functional dependence of Π on z in the following analysis.

Now, consider a strategy whereby the firm delivers any V in the neighborhood of $\tilde{V}$ by adjusting the flow pay to $w^*(V) \equiv u^{-1}(V - \tilde{V} + u(\tilde{w}))$, while maintaining the contingent plan $\{\tilde{W}(z', s')\}$. By construction, this strategy satisfies all constraints, and I denote the firm's value function under this strategy by $\tilde{\Pi}(V)$. It follows that $\tilde{\Pi}(V) \leq \Pi(V)$ and $\tilde{\Pi}(\tilde{V}) = \Pi(\tilde{V})$.

Moreover, since V enters $\tilde{\Pi}(V)$ solely through $-w^*(V) = -u^{-1}(V - \tilde{V} + u(\tilde{w}))$, and given that $-u^{-1}(\cdot)$ is concave and twice differentiable, $\tilde{\Pi}(V)$ inherits these properties and is thus concave and differentiable at every point in the vicinity of $\tilde{V}$, including $\tilde{V}$. To summarize, $\tilde{\Pi}(V)$ is a function that is concave, twice differentiable, and coincides with $\Pi(V)$ at $\tilde{V}$. By the results of Benveniste and Scheinkman (1982), it follows that $\Pi(V)$ 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 ξ to (PKC), μ to (IC), $\beta\mu_0(z', s')$ to (PC-E), and $\beta\mu_1(z', s')$ to (PC-F). The first-order condition with respect to w gives $u'(w) = 1/\xi$, and the envelope theorem gives $-\frac{\partial \Pi(V,z)}{\partial V} = \xi$. Together, they yield (7). The 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 (9). If $\overline{W}(z', s') < \overline{W}(z', s)$, the participation constraints become $\overline{W}(z', s') \leq W(z', s') \leq \overline{W}(z', s)$. Using this, we derive the first-order condition with respect to $W(z', s')$:

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

If $\mu_0(z', s') = \mu_1(z', s') = 0$, $W(z', s') = W(z')$ as defined by (8). This is the case in line 3 of (9). If $\mu_0(z', s') > \mu_1(z', s') = 0$, $W(z', s') = \overline{W}(z', s')$. This corresponds to the case in line 2 of (9). Finally, if $\mu_1(z', s') > \mu_0(z', s') = 0$, $W(z', s') = \overline{W}(z', s)$. This is the second condition in line 1 of (9).

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 $W(z', s') = \overline{W}(z', s')$. Next, consider a non-binding

$W(z', s')$. Fix s' and z, and suppose $z'_2 > z'_1$. By the first-order conditions, we have

$$-\frac{\partial \Pi(W(z'_2, s'), z'_2)}{\partial W} = \xi + \mu(1 - g(z'_2|z)) > \xi + \mu(1 - g(z'_1|z)) = -\frac{\partial \Pi(W(z'_1, s'), z'_1)}{\partial W},$$

where I have used $g(z'_2|z) < g(z'_1|z)$. Since (PKC) is less binding as z becomes higher, $-\frac{\partial \Pi(W,z)}{\partial W}$ 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 the concavity of Π in W, we have $W(z'_2, s') > W(z'_1, s')$. □

Proof of Proposition 2

Proof. Consider a general incentive scheme $W(z)$ defined on $\mathbb{Z}$. The incentive $\mathcal{I}\Big(W(z)\Big)$ can be expressed as a weighted sum of $W_z(z)$:

$$\mathcal{I}(W(z)) \equiv \int_{\underline{z}}^{\overline{z}} W(z)(1 - g(z))\Gamma(dz) = \int_{\underline{z}}^{\overline{z}} \omega(z) W_z(z) dz, \tag{14}$$

where $\omega(z) = -\int_{\underline{z}}^{z}(1 - g(t))\Gamma(t)dt$. The identity in (14) can be easily verified using integration by parts, noting that $\int_{\underline{z}}^{\overline{z}}(1 - g(z))\Gamma(z)dz = 0$.

To establish that $\mathcal{I}(W(z, s))$ decreases in s, it suffices to show that the partial derivative with respect to z, $\overline{W}_z(z, s)$, decreases in s for all $z \in \mathbb{Z}$.

Applying the implicit function theorem to $\Pi(\overline{W}, z, s) = 0$ yields

$$\overline{W}_z(z, s) = -\frac{\partial \Pi(\overline{W}, z, s)/\partial z}{\partial \Pi(\overline{W}, z, s)/\partial W} = u'(\overline{w}(z, s))\Pi_z(\overline{W}, z, s).$$

Differentiating with respect to s gives

$$\frac{d\overline{W}_z(z, s)}{ds} = \Pi_{zs}(\overline{W}, z, s)u'(\overline{w}(z, s)) + \Pi_z(\overline{W}, z, s)u''(\overline{w}(z, s))\overline{w}_s(z, s),$$

where $\Pi_{zs}(\cdot)$ denotes $\frac{\partial^2 \Pi(\overline{W},z,s)}{\partial z \partial s}$.

Using $\Pi(\overline{w}, z, s) = 0$ and $\partial \Pi(\overline{w}, z, s)/\partial \overline{w} = -1$, we find that $\frac{\partial \overline{w}(z,s)}{\partial s} = \frac{\partial \Pi(\overline{W},z,s)}{\partial s}$. Therefore, the condition $\frac{d\overline{W}_z(z,s)}{ds} < 0$ is equivalent to

$$-\frac{u''(\overline{w})}{u'(\overline{w})} > \frac{\Pi_{zs}(\overline{W}, z, s)}{\Pi_z(\overline{W}, z, s)\Pi_s(\overline{W}, z, s)}.$$

□