Managerial Labor Market Competition and Incentive Contracts

Bo Hu*

Job Market Paper

February 1, 2019

Check the latest version (pdf).

Abstract

This paper assesses the impact of managerial labor market competition on executive incentive contracts. I develop a dynamic contracting framework that embeds the moral hazard problem into an equilibrium search environment. Competition for executives increases total compensation, and generates a new source of incentives, called *labor market incentives*, which substitutes for performance-based incentives (e.g. bonus, stocks, options, etc.). The model is estimated using a newly assembled dataset on executive turnovers of U.S. publicly listed firms. The structural estimates show that the model is capable of explaining and predicting the puzzling facts that executives of larger firms experience higher compensation growth and receive higher performance-based incentives.

Key Words: executive compensation, dynamic moral hazard, firm-size premium, managerial labor market, search frictions

*Department of Economics, Vrije University Amsterdam, and Tinbergen Institute, De Boelelaan 1105, 1081 HV Amsterdam. Email: huboshufe@gmail.com. Website: bohuecon.github.io. I am greatly indebted to my PhD supervisors Makoto Watanabe and Pieter Gautier for their guidance and encouragement. I would like to thank Sarolta Laczó, Gregory Jolivet, Helene Turon, Sebastian Gryglewicz, Alvaro Remesal, Giulio Fella, Yanos Zylberberg, and Aniol Llorente-Saguer for helpful discussions and valuable comments. All errors are my responsibility.

1 Introduction

Executives are incentivized by having their compensation closely tied to firm performance in the form of bonuses, stocks, options, etc. Traditionally, it is believed that incentive contracts are designed to align the interests of executives with those of shareholders. In recent decades, however, we have seen that competition for executives is increasingly influential in shaping incentive contracts. For example, in the "battle for talent", IBM targets the 50th percentile of both cash and equity compensation among a large group of benchmark companies. The contract of individual executives is further adjusted according to "the skills and experience of senior executives that are highly sought after by other companies and, in particular, by our (IBM's, added) competitors." Similarly, Johnson & Johnson compare "salaries, annual performance bonuses, long-term incentives, and total direct compensation to the Executive Peer Group companies" who compete with Johnson & Johnson "for executive talent".[1]

Despite its relevance for the industry, a characterization of how heterogeneous firms compete for executives is still missing in the literature, and the consequences for executive contracts have remained unclear. For example, in the assignment models (e.g., Gabaix and Landier 2008, Edmans et al. 2009), equilibria are static and dynamic features such as career concerns or job ladder effects are absent. In the multiple-period models (e.g., Holmström 1999, Oyer 2004, Giannetti 2011), it is usually assumed that all companies compete with the same spot market wage, and executives cannot transit to potentially more productive companies. Other dynamic models concentrate more on the firm/rank choice of executives rather than competition between firms (e.g., Gayle et al. 2015).

This paper focuses on the competition between heterogeneous firms in a tractable framework that combines dynamic moral hazard and equilibrium labor search. In particular, I allow executives to *search on-the-job* along a *hierarchical job ladder towards larger firms* as in Postel-Vinay and Robin (2002). This feature, which is missing in the existing studies on managerial labor markets, drives the key results.[2] The model considers two types of agents: executives and firms. Executives are heterogeneous in the *general* managerial productivity, which evolves stochastically depending on their current and past effort. Firms are heterogeneous in time-invariant asset size.[3] As in Gabaix and Landier

[1]See detailed compensation policy in IBM and Johnson & Johnson proxy statements, which are accessible at `https://www.ibm.com/annualreport/2017/assets/downloads/IBM_Proxy_2018.pdf` and `http://www.investor.jnj.com/gov/annualmeetingmaterials.cfm` (visited on Oct 27, 2018).

[2]The executive job ladder exists in the real world. The career path of Richard C. Notebaert is a good example as is described by Giannetti (2011): "Notebaert led the regional phone company Ameritech Corporation before its 1999 acquisition by SBC Communication Inc.; after, he held the top job at Tellabs Inc., a telecom-equipment maker; finally, in 2002, he became CEO of Qwest Communications International Inc." This anecdotal description is consistent with the data evidence in the literature. Huson et al. (2001) report that the fraction of outsider CEOs increased from 15.3% in the 1970s to 30.0% at the beginning of the 1990s. A similar pattern is reported by Murphy and Zabojnik (2007).

[3]I measure firm size by market capitalization (value of debt plus equity).

(2008), the marginal impact of an executive's productivity increases with the value of the firm under his or her control. While output is observable, the effort is not. Thus, a moral hazard problem arises. To resolve the problem, the firm and the executive sign a long-term incentive contract. Moreover, the executive has limited commitment to the relationship and may encounter outside poaching offers from an external labor market. By making use of poaching offers, the executive can renegotiate with the current firm or transit to a larger poaching firm, where the compensation contract is determined in a sequential auction. Essentially, the current and the poaching firms are engaged in a Bertrand competition for the executive.

The competition from poaching offers impacts executive incentive contracts via two channels. *First*, as in Postel-Vinay and Robin (2002), competition from outside offers increases total compensation. When the poaching firm is smaller than the current firm, the executive may use the offer to negotiate with the current firm for a higher pay. When the poaching firm is larger, it can always outbid the current firm since firm size contributes to the production. Thus, the executive uses the current firm as a threat point to negotiate with the poaching firm and transits to the poaching firm. In either case, the executive climbs up the job ladder towards a higher compensation level and (or) a larger firm size. *Second*, poaching offers generate a new source of incentives and consequently reduce the need for performance-based incentives. Poaching firms are willing to bid more for more productive executives. Meanwhile, the productivity of an executive is stochastically determined by his or her past effort. Together, these factors imply that effort today will lead to a more favorable offer from the same poaching firm in the future. This potential gain from labor market competition becomes what I call *labor market incentives* in this paper. Firms can take advantage of these labor market incentives and give fewer performance-based incentives to executives, but still resolve the moral hazard issue.

These two channels enable the model to shed light on two puzzling facts: the *firm-size pay-growth premium* and the *firm-size incentive premium*, both of which are firstly documented in this paper, complementing other stylized facts in the literature (see, e.g., Edmans et al. 2017). The *firm-size pay-growth premium* refers to the empirical finding that starting with the same total compensation, the executive of a larger firm experiences a higher compensation growth. Based on the data for U.S. listed firms, I find that for a 1% increase in firm size, the annual pay-growth rate increases by 15.4%. This big gap in pay-growth rate significantly contributes to the pay differentials between small and large firms. My explanation for this premium is as follows. Executive compensation grows because firms desire to retain executives in response to poaching offers. Due to a firm-size effect in production, larger firms are more capable of countering poaching offers; hence, their executive compensation tends to grow faster.

The *firm-size incentive premium* refers to the empirical fact that performance-based incentives embedded in bonuses, stocks, options, etc. increase with firm size after controlling for total compensation. As will be elaborated in Section 3, a 1% increase in firm size leads to a 0.35% increase in performance-based incentives.[4] This incentive premium has excluded the size differentials in pay levels, and it reflects that larger firms tend to allocate a higher fraction of the compensation package to performance-related pay. My explanation for this premium is based on labor market incentives. In the model, an executive is motivated by two sources of incentives which substitute for each other: performance-based incentives and labor market incentives. I show that labor market incentives decrease with firm size. To motivate executives in larger firms, the performance-based incentives are required to be higher.

There are two reasons why executives in larger firms receive less labor market incentives. The first reason lies in the job ladder. Executives from larger firms are located "higher" on the job ladder. Consequently, the chance of receiving an outside offer that beats the current value is lower. Thus, labor market incentives for larger firms are smaller. Indeed, for individuals that are at the bottom of the job ladder, labor market incentives can be large enough that no performance-related pay is required, which goes back to the original model of Postel-Vinay and Robin (2002). The second reason is based on a wealth effect. Executives of larger firms are expected to receive higher compensation in the future, i.e., the size premium in pay-growth; thus, the certainty equivalents of their future expected utilities are higher. Given a diminishing marginal utility, at a higher certainty equivalent, the utility gain from a more favorable poaching offer is smaller. As a result, labor market incentives are smaller as firm size increases.

To provide empirical evidence and structurally estimate the model, I assembled a new dataset on executive job turnovers by merging the ExecuComp and BoardEX databases. ExecuComp is the standard data source for executive compensation studies. It contains annual records on top executives' compensation in firms comprising the S&P 500, MidCap, and SmallCap indices. BoardEX contains detailed executive employment history, and it helps to identify the employment status after executives leave the S&P firms. For executives that are not identified in BoardEX, I further searched for executive profiles and biographies using LinkedIn and Bloomberg. In the final data sample, there are 35,088 executives and a total of 218,168 executive-year observations spanning the period 1992 to 2016.

I first provide reduced-form evidence to support the model set-up and implications. Using the merged data, I document a job-to-job transition rate of around 5%, which is stable over the years and across industries. Moreover, there is a job ladder in the firm-size dimension: about 65% of job-to-job transitions are towards larger firms. This justifies the hierarchical job ladder featured in the model. Second, I test whether the job ladder "position" of an executive matters for his/her chance of job-to-job transitions.

[4]Performance-based incentives are measured by the dollar change in firm-related wealth per percentage change in firm value. It is an ex ante measure of incentives in compensation package, before the realization of firm performance. Consistent with the literature (Jensen and Murphy 1990, Hall and Liebman 1998), I use all firm-related wealth instead of only current compensation to calculate performance-based incentives. See Section 3 for more details on motivating facts.

Specifically, using a Cox model, I find that executives in larger firms are less likely to experience job-to-job transitions, which is in line with the model's prediction. Finally, using the variation across industries, I find that firm-size premiums in both pay-growth and incentives are higher in industries where the managerial labor market is more active. I proxy the activeness of an executive labor market by the job-to-job transition rate, the fraction of outside CEOs and the average of the general ability index (Custódio et al. 2013).

It is difficult to numerically solve for the optimal contract in the presence of an incentive compatibility constraint, limited-commitment constraints, and shocks of large support, as one needs to solve for the promised value in each state of the world. I address this issue by using the recursive Lagrangian approach (Marcet and Marimon 2017), under which I only need to solve for one Lagrangian multiplier to find the optimal contract. This multiplier represents the weight of the executive in a constructed Pareto problem, and it keeps track of various constraints and job-to-job transitions. Based on this multiplier, the optimal incentive pay and promised values can be solved.

Using the simulated method of moments (SMM), I estimate the model by targeting data moments on executive compensation, incentives, and turnovers, as well as on firm size and profitability. Importantly, I do not explicitly target the firm-size premiums in compensation growth and performance-based incentives. Yet, the estimated model quantitatively captures both. The predictions of the estimated model are very close to the premium estimates from the data, which corroborates that the model mechanism plays an essential role in explaining both premiums. A counter-factual decomposition shows that labor market incentives account for more than 40% of total incentives among small-firm executives, around 15% for medium-firm executives and less than 5% for large-firm executives.

Based on the structural estimation, I use a counterfactual exercise to quantitatively account for the sharp increases in executive compensation since the mid-1970s. I show that with an exogenous rise in the arrival rate of poaching offers, the model generates increases in total compensation and performance-based incentives, more inequality of compensation across executives, and a higher correlation between compensation and firm size. Quantitatively, the above changes in model-simulated data match well with the data facts documented by Frydman and Saks (2010). The intuition is that the managerial labor market was much thinner before the 1970s, which is supported by the evidence presented in Frydman (2005) and Murphy and Zabojnik (2007).

Finally, there is a clear policy implication of the model regarding how to regulate the compensation of highly paid executives, especially in large firms. Rather than only focusing on large firms, it is important to lower the bids for executives from small and medium firms. This could be achieved via various reforms such as more independent compensation committees, greater mandatory pay (or pay ratio) disclosure or say-on-pay legislation. In this way, the competitive pressure in the overall managerial labor market will decrease. In the model, there is a spillover effect whereby higher bids from a set of firms boost executive pay not only in those firms but also in all firms that are higher on the job ladder. In a comparative static analysis, I show that, compared to an increase in the bids from large firms, the same increment in those of small and medium firms has a similar effect on the compensation of large firms, and has a more substantial impact on the compensation of the whole managerial market.

The rest of the paper is organized as follows. In Section 2, I provide a detailed literature review. In Section 3, I present the motivating facts of the firm-size pay-growth premium and incentive premium. I further show that both premiums significantly increase when the executive labor market is more active. I then set up the model in Section 4, where I characterize the optimal contract and explain the premiums. Section 5 presents reduced-form evidence and Section 6 estimates the model. Section 7 explains the sharp increase in executive pay since the mid-1970s. Section 8 discusses the policy implications of the research. Finally, Section 9 concludes.

2 Literature Review

This paper contributes to two strands of literature in understanding pay differentials between small and large firms. The first strand explains the differentials using assignment models. Gabaix and Landier (2008), Tervio (2008) and Eisfeldt and Kuhnen (2013) present competitive assignment models to explain why total compensation increases with firm size. Consistent with these studies, I use a multiplicative production function to characterize the contribution of executives. My model provides a similar prediction of the relationship between total compensation and firm size. Since my model is dynamic, it also captures the growth of total compensation, which is absent in the existing literature. More importantly, it afford a different view on the pay differentials between small and large firms. In this paper, executives are paid much more in larger firms not because they are more talented (e.g., Gabaix and Landier 2008) but because they are lucky to be matched with a firm whose size makes it better able to counter outside offers. Further along this strand of research, Edmans et al. (2009) and Edmans and Gabaix (2011) add a moral hazard problem to the assignment framework and explain why performance-based incentives increase with firm size. Their explanation is based on the notation that total compensation increases with firm size. Yet, these models do not explain why, after controlling for total compensation, a firm-size incentive premium still exists. My model is a dynamic and search-frictional version of their framework and highlights a hierarchical job ladder. Besides the explanation given in Edmans et al. (2009), the job ladder in my model gives rise to labor market incentives, which contributes to the understanding of the firm-size incentive premium.

The second strand of literature explains the pay differentials using agency problems.

Margiotta and Miller (2000) derive and estimate a multi-period principal-agent model with moral hazard. Based on this model, Gayle and Miller (2009) show that large firms face a more severe moral hazard problem, so that higher equity incentives are needed to satisfy the incentive compatibility condition. Gayle et al. (2015) embed the model developed by Margiotta and Miller (2000) into a generalized Roy model. They find that the quality of the signal is unambiguously poorer in larger firms, and this explains most pay differentials between small and larger firms. In contrast to my focus on managerial labor market competition, Gayle et al. (2015) find that the career concern channel does not explain the size premium in their estimation. The critical difference between the model used by Gayle et al. (2015) and my model is that in their model job-to-job transitions are based on a Roy model and are in general not directed towards larger firms, whereas the driving force of my explanation is a hierarchical job ladder where executives move from small to large firms. The different approach to modeling job-to-job transitions explains why labor market incentives contribute much less in the framework of Gayle et al. (2015). Using executives' job-to-job transition data, I show that the hierarchical job ladder does exist.

This paper also contributes to the literature explaining the rise of executive compensation in recent decades. My paper is closest in spirit to the explanation based on executive mobility. The literature shows that the increases in compensation coincide with the increased occupational mobility of executives, which is brought about by an increased importance of executives general managerial skills in comparison to firm-specific knowledge (Frydman 2005, Frydman and Saks 2010, Murphy and Zabojnik 2007). Giannetti (2011) develops a model to show that job-hopping opportunities can help explain not only the increase in total pay, but also the structure of managerial contracts. In Section 7, I provide a counterfactual analysis showing that with an exogenous increase of poaching offer arrival rate from 5% to 40% per year, my model can account for the sharp increase in total compensation and performance-based incentives, as well as a much higher correlation between firm size and total compensation.

The fourth stream of the literature to which I contribute is the one on executive turnovers, see e.g., Taylor (2010), Jenter and Kanaan (2015), Kaplan and Minton (2012) and Peters and Wagner (2014). In particular, on the incentive effect of turnovers, Remesal et al. (2018) estimate a dynamic moral hazard model allowing for endogenous compensation and dismissals. Their estimation shows that dismissal threats play a small role in CEO incentives, whereas the bulk of CEO incentives comes from the flow and deferred compensation. These results justify my focus on performance-related compensation. Wang and Yang (2016) study the optimal termination in a dynamic contract with moral hazard and stochastic market value shocks. The model generates rich insights on voluntary and involuntary dismissals, and termination plays distinct roles depending on the level of market values. Their analysis, however, is abstract from firm size, and the market value shocks are homogeneous to all executives. In contrast, poaching offers in my model are heterogeneous across executives.

Finally, this paper is closely related to the work of Abrahám et al. (2016), who aim to explain wage inequality in the general labor market by combining repeated moral hazard and on-the-job search. Besides the differences in topics, there is a critical difference that distinguishes the two papers: The productivity of agents is independent over time in the model of Abrahám et al. (2016), while it is persistent in my model. Therefore, in my model, working hard today rewards the agent in the future. It is this feature that gives rise to labor market incentives and explains the firm-size incentive premium. This feature is absent in their model.

In terms of modeling, this paper links two strands of literature. One strand is the extensive literature on optimal long-term contracts with private information and commitment frictions, e.g., Townsend (1982), Rogerson (1985), Spear and Srivastava (1987), Phelan and Townsend (1991), Harris and Holmstrom (1982), Thomas and Worrall (1990) and Phelan (1995). I build on this literature by embedding an optimal contracting problem with moral hazard and two-sided limited commitment into an equilibrium search model. In doing so, the outside environment is endogenized, which significantly changes the optimal contract. Another important strand of literature uses structural search models to evaluate wage dispersions. Postel-Vinay and Robin (2002), Cahuc et al. (2006) and Lise et al. (2016) among others estimate models with job ladders and sequential auctions. Compared to this literature, I add a dynamic moral hazard problem, which allows me to understand how search frictions influence a long-term contract. The model of Postel-Vinay and Robin (2002) is a special case of my model when the moral hazard problem is absent. In addition, the managerial labor market is an appropriate environment for their framework. In real life, it happens very often that executives are contacted and ”auctioned” by competing firms for promotion, as is described by Khurana (2004).

3 Motivating Facts

This paper is motivated by two firm-size premiums: the pay-growth premium and the incentive premium. As far as I am aware, these facts are firstly documented. Moreover, I show that both premiums are larger in industries where the managerial labor market is more active, where labor market thickness is measured by job-to-job transition rates, the general ability index and the fraction of inside CEOs at the industry or industry-year level. The primary data source for the analyses is Standard & Poor’s ExecuComp database. Variables about executive labor markets come from a newly assembled dataset on executive turnovers and the other two datasets provided by Custódio et al. (2013) and Martijn Cremers and Grinstein (2013). All nominal quantities are converted into constant 2016 dollars using GDP deflater. Section 5 presents a statistical description of the data.

Size pay-growth premium

I measure firm size by market capitalization, defined by the common shares outstanding times the fiscal year close price. The executive annual compensation-growth rate is measured by the first-order difference of $\log(tdc1)$ where $tdc1$ is the total compensation including the sum of salary and bonus, the value of restricted stocks and options granted, and the value of retirement and long-term compensation schemes. Column (1) in table 1 presents a regression of $\Delta\log(tdc1)$ on firm size, controlling for lagged $\log(tdc1)$ and dummies on tenure, age and year times industry.[5] The estimation indicates that starting from the same level of total compensation, for a 1% increase in firm size, the annual compensation-growth rate increases by 11.2%. The premium is slightly larger with an estimate of 15.4% in column (2) after further controlling for *operating profitability*, *market-book ratio*, *annualized stock return*, title dummies such as *director*, *CEO*, *CFO*, etc. Definitions of these variables are provided in the table note.

To link the size pay-growth premium with managerial labor markets, I explore the variation across industries. An industry is an appropriate sub-labor market since more than 60% of executive job-to-job transitions are within the industry (see details in Section 5). As a direct test of whether size pay-growth premium is related to a more active managerial labor market, I use four proxies to measure the labor market thickness and test if the interactions between these proxies and firm size are significant. The first two proxies are job-to-job transition rates on the industry-year level (Fama-French 48 industries and fiscal years). *EE90* is the job-to-job transition rate where a job-to-job transition is defined by an executive leaves the current firm and starts to work in another firm within *90* days. Similarly, *EE190* is the job-to-job transition rate where a job-to-job transition is defined by a gap of no more than 190 days. The third proxy *gai* is the mean of the general ability index of CEOs at the industry-year level. The general ability index itself is the first principal component of five proxies that measure the generality of a CEO's human capital based on his or her lifetime work experience (Custódio et al. 2013).[6] The last proxy, *inside CEO*, is the industry-level percentage of the CEOs who are promoted inside the firm (Martijn Cremers and Grinstein 2013). It accounts for all new CEOs between 1993 and 2005 using Fama-French 48-industry categories.

The picture that emerges in the last four columns of table 1 is not ambiguous: All four interaction terms are statistically and economically significant, and the signs confirm that the size growth premium is larger in industries/years where the executive labor market is more active. Specifically, the coefficient of 0.0759 on the interaction with *EE190* and a standard deviation of 0.0224 for *EE190*, imply that a one-standard-deviation increase in

[5] The result is robust by controlling for lagged total compensation using 100 or 200 group dummies that equally divide the sample according to the value of $\log(tdc1)$.

[6] The five proxies to measure the general ability of CEO's are: the number of positions that CEO performed during his/her career, the number of firms where a CEO worked, the number of industries at the four-digit SIC level where a CEO worked, a dummy variable that equals one if a CEO held a CEO position at another firm, and a dummy variable that equals one if a CEO worked for a multi-division firm.

Table 1: Pay-growth increases with firm size

	$\Delta\log(tdc1)$					
	(1)	(2)	(3)	(4)	(5)	(6)
log(firm size)$_{-1}$	0.112***	0.154***	0.108***	0.107***	0.141***	0.127***
	(0.00903)	(0.0129)	(0.00183)	(0.00189)	(0.00177)	(0.00489)
log(firm size)$_{-1}$ $\times$ *EE90*			0.0711*			
			(0.0403)			
log(firm size)$_{-1}$ $\times$ *EE190*				0.0759**		
				(0.0353)		
log(firm size)$_{-1}$ $\times$ *gai*					0.0233***	
					(0.00546)	
log(firm size)$_{-1}$ $\times$ *inside CEO*						-0.0232***
						(0.00696)
log(tdc1)$_{-1}$	-0.290***	-0.390***	-0.251***	-0.251***	-0.304***	-0.253***
	(0.0200)	(0.0262)	(0.00173)	(0.00173)	(0.00267)	(0.00173)
other controls		X	X	X	X	X
tenure dummies	X	X	X	X	X	X
age dummies	X	X	X	X	X	X
year dummies	X	X	X	X	X	X
industry	X	X				
year $\times$ *industry*	X	X				
Observations	129,068	106,819	106,820	106,820	58,188	106,820
adj. R^2	0.157	0.216	0.260	0.260	0.233	0.262

Note: This table presents firm-size pay-growth premium and its correlation with the activeness of executive labor markets. The dependent variable is the first-order difference of $\log(tdc1)$ where *tdc1* is the total compensation including the sum of salary and bonus, the value of restricted stocks and options granted, and the value of retirement and long-term compensation schemes. Firm size is measured by the market capitalization, defined by the common shares outstanding times the fiscal year close price. I control for lagged *tdc1*. Whenever possible, I also control for age, tenure, and year times industry dummies. Other controls include *operating profitability*, *market-book ratio*, *annualized stock return*, *director*, *CEO*, *CFO* and *interlock*. *director* is a dummy which equals to 1 if the executive served as a director during the fiscal year. *CEO* and *CFO* are dummies defined by whether the executive served as a CEO (and CFO) during the fiscal year. *interlock* is a dummy which equals to 1 if the executive is involved in an interlock relationship. An interlock relationship generally involves one of the following situations: (1) The executive serves on the board committee that makes his or her compensation decisions; (2) the executive serves on the board of another company that has an executive executive serving on the compensation committee of the indicated executive's company; (3) the executive serves on the compensation committee of another company that has an executive executive serving on the board of the indicated executive's company. I use four variables to measure the activeness of the executive labor market at the industry or industry-year level. *EE90* is the industry-year level job-to-job transition rate where a job-to-job transition is defined by an executive leaves the current firm and starts to work in another firm within *90* days. *EE190* defines a job-to-job transition with a gap of no more than 190 days. *gai* is the mean of general ability index of CEOs at the industry-year level. The original data is provided by Custódio et al. (2013). *insider CEO* is the industry level percentage of internally promoted new CEOs between 1993 and 2005. The original data on this variable is provided by Martijn Cremers and Grinstein (2013). For all variables, an industry is based on Fama-French 48 categories, a year is based on fiscal years. Standard errors clustered on the firm $\times$ fiscal-year level are shown in parentheses, and I denote symbols of significance by * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$.

job-to-job transition measured by *EE190* gives an increase in pay-growth premium from 7.3% to 7.47%. Similarly, a one-standard-deviation increase in job-to-job transition measured by *EE90* implies that pay-growth premium increases from 7.3% to 7.43%. Given a standard deviation of 0.253 for *gai* and a coefficient of 0.0233, a one-standard deviation increase in *gai* leads to an increase in pay-growth premium from 7.3% to 7.89%. Given a standard deviation of 0.122 for *inside CEO* and a coefficient of 0.0233, a one-standard deviation increase in *inside CEO* leads to an increase of pay-growth premium from 7.3% to 7.58%.

Size incentive premium

I measure performance-based incentives in executive contracts by "*delta*". By definition, *delta* equals the dollar increase in executives' firm-related wealth for a percentage increase in firm value. It measures incentives before firm performance is realized. Thus, it is ex ante.[7] As has been documented in Edmans et al. (2009) and is replicated in table 2 column (1), *delta* is positively correlated with firm size: For a 1% increase in firm size, measured by market capitalization, performance-based incentives increase by 0.59%. Edmans et al. (2009) argued that because executives in larger firms are paid higher, they require more incentives to induce effort.

However, the level of total compensation does not explain the entire size incentive premium. The positive correlation between performance-based incentives and firm size remains after controlling for total compensation, $\log(tdc1)$, in table 2 column (2): For a 1% increase in firm size, *delta* increases by 0.36%, which accounts for more than half of the size premium estimated in column (1). The estimated elasticity 0.36 of incentives to size in column (2) is the *size incentive premium* that I aim to explain. It excludes the pay-level effects and only reflects the proportion of incentive-related pay. As I will show in Section 6, the estimates of size incentive premium in both columns (1) and (2) can be quantitatively captured by the model.

I further test if size incentive premium is related to managerial labor market using the same four proxies as in the last subsection: *EE90*, *EE190*, *gai* and *inside CEO*. As presented in columns (3) to (6) in table 2, all interaction terms are statistically and economically significant, and the signs indicate that the size incentive premium is larger in industries/years where the executive labor market is more active.[8]

[7] *delta* is also known as "the value of equity at stake" or "dollar-percentage incentives". Empirical studies of pay-to-performance have used a wide range of specifications to measure this relationship. Two common alternatives are the dollar change in executive wealth per dollar change in firm value (the Jensen-Murphy statistic) and the dollar amount of wealth that an executive has at risk for a 1% change in the firms value (the value of equity at stake or *delta*). The Jensen-Murphy statistic is the correct measure of incentives for activities whose dollar impact is the same regardless of firm size, and the value of equity at stake is appropriate for actions whose value scales with firm size. The latter is the modeling approach of this paper.

[8] Specifically, given a coefficient of 0.415 on the interaction with *EE190* and a standard deviation of 0.0224 for *EE190*, a one-standard-deviation increase in *EE190* leads to an increase in the elasticity from 0.525 to 0.534. Similarly, a one-standard-deviation increase in job-to-job transition measured by *EE90* implies a

Table 2: Performance-based incentives increase with firm size

	log(*delta*)					
	(1)	(2)	(3)	(4)	(5)	(6)
log(firm size)	0.585***	0.360***	0.315***	0.316***	0.325***	0.316***
	(0.0141)	(0.0247)	(0.0029)	(0.0029)	(0.0036)	(0.0029)
log(firm size) $\times$ *EE90*			0.772*			
			(0.1228)			
log(firm size) $\times$ *EE190*				0.716**		
				(0.1054)		
log(firm size) $\times$ *gai*					0.055***	
					(0.0112)	
log(firm size) $\times$ *inside CEO*						-0.087***
						(0.0196)
log(tdc1)		0.609***	0.693***	0.692***	0.687***	0.684***
		(0.035)	(0.046)	(0.046)	(0.056)	(0.046)
other controls	X	X	X	X	X	X
tenure dummies	X	X	X	X	X	X
age dummies	X	X	X	X	X	X
year dummies	X	X	X	X	X	X
industry	X	X				
year $\times$ *industry*	X	X				
Observations	146,747	128,006	128,006	128,006	79,476	128,006
adj. R^2	0.442	0.482	0.486	0.487	0.482	0.485

Note: This table presents firm-size incentive premium and its correlation with the activeness of executive labor markets. The dependent variable is $\log(delta)$ where *delta* is the dollar change in firm related wealth for a percentage change in firm value. Firm size is measured by the market capitalization, defined by the common shares outstanding times the fiscal year close price. *tdc1* is the total compensation including the sum of salary and bonus, the value of restricted stocks and options granted, and the value of retirement and long-term compensation schemes. Whenever possible, I control for age, tenure, and year times industry dummies. Other controls include *operating profitability*, *market-book ratio*, *annualized stock return*, *director*, *CEO*, *CFO* and *interlock*. *director* is a dummy which equals to 1 if the executive served as a director during the fiscal year. *CEO* and *CFO* are dummies defined by whether the executive served as a CEO (and CFO) during the fiscal year. *interlock* is a dummy which equals to 1 if the executive is involved in an interlock relationship. Please refer to the footnote of table 1 for a definition of *interlock*. I use four variables to measure the activeness of the executive labor market at the industry or industry-year level. *EE90* is the industry-year level job-to-job transition rate where a job-to-job transition is defined by an executive leaves the current firm and starts to work in another firm within *90* days. *EE190* defines a job-to-job transition with a gap of no more than 190 days. *gai* is the mean of general ability index of CEOs at the industry-year level. The original data is provided by Custódio et al. (2013). *insider CEO* is the industry level percentage of internally promoted new CEOs. The original data on this variable is provided by Martijn Cremers and Grinstein (2013). For all variables, an industry is based on Fama-French 48 categories, a year is based on fiscal years. Standard errors clustered on the firm $\times$ fiscal-year level are shown in parentheses, and I denote symbols of significance by * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$.

Finally, I show that size incentive premium decreases as executives approach retirement age. Starting from Gibbons and Murphy (1992), *age* has been used as an indicator for career concerns: The older the executive is, the less influential that managerial labor market is on incentive contract design. If size incentive premium is at least partly caused by managerial labor markets, we would expect the premium to decrease with age. This is indeed the case, as is shown in figure 1. The size incentive premium starts with 0.652 at age 35, and gradually decreases to around 0.35 after age 50. This pattern holds with or without controls.

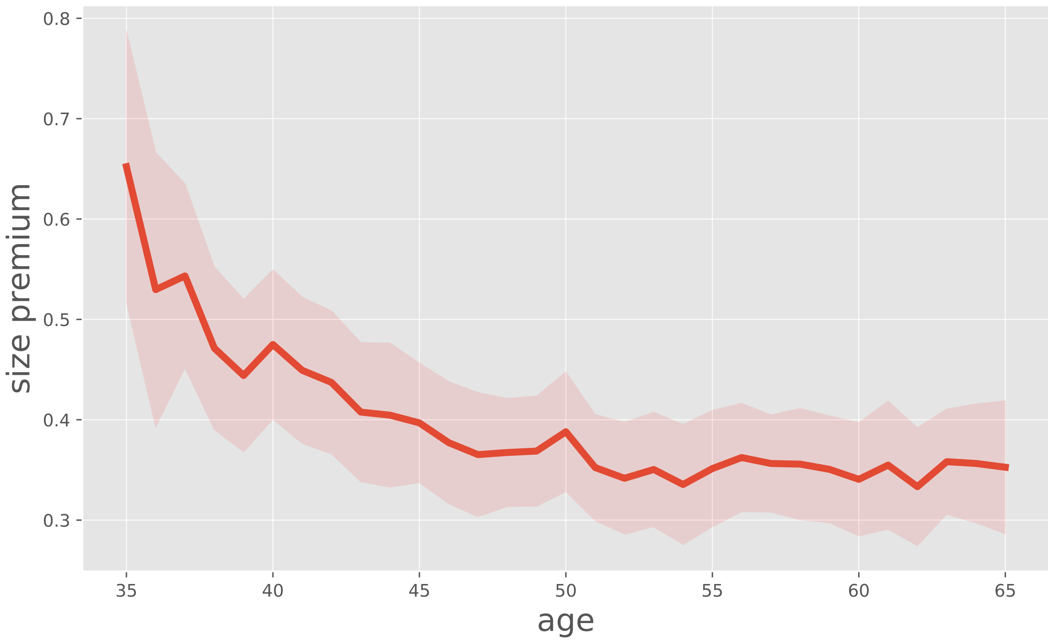

Figure 1: Size premium in performance-based incentives decreases with age

Note: The figure depicts the size premium in performance-based incentives at each age from 35 to 65. Each point is one estimated coefficient of the interaction term between one *age dummy* and *log(firm size)* in the following regression,

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

where i denotes an executive, t denotes the fiscal year, *age dummies* is a set of 31 dummies for each age from 35 to 65, *firm size* is measured by market capitalization by the end of the fiscal year, calculated by a firm's common shares outstanding times the close price by fiscal year, X denotes a vector of control variables including a constant term. I control for total compensation *log(tdc1)* and dummies of executive tenure, age, and fiscal year time industry. A 95% confidence interval is plotted using the standard error clustered on firm times fiscal year. The full regression result is provided in Appendix B.

higher elasticity of 0.532. A standard deviation of 0.253 for *gai*, together with the coefficient of 0.0648, means that with a one-standard deviation increase in *gai*, the elasticity increases by 0.016. A standard deviation of 0.122 for *inside CEO*, together with the coefficient of 0.0458, means that a one-standard deviation increase in *inside CEO* leads to an increase of 0.0056 of the elasticity.

4 Model

In this section, I construct an equilibrium model of the managerial labor market. The model features on-the-job search, poaching offers and contract renegotiation. I embed a bilateral moral hazard problem into the labor market equilibrium. Poaching offers are used to renegotiate with the current firm, leading to compensation growth. Thus, the size growth premium is linked to a firm's capability of overbidding poaching offers. Poaching offers also generate new incentives, called labor market incentives in the model, which constitute a wedge between the total incentives required to motivate executives and performance-based incentives provided by firms. That is, the size premium in performance-based incentives is linked to labor market incentives. These mechanisms are used to explain the size premiums in both annual pay-growth and performance-based incentives. I now formally introduce the model.

4.1 Ingredients

Agents

There is a fixed measure of individuals: They are either employed as executives or not hired as executives but looking for executive jobs. I call the latter "executive candidates". Individuals die with some probability. Once an individual dies, a new-born enters the economy.

Individuals want to maximize expected lifetime utility,

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

where $\beta \in (0,1)$ is the discount factor, $\eta \in (0,1)$ is the death probability, utility of consumption $u : \mathbb{R} \to \mathbb{R}$ is increasing and concave and $c(\cdot)$ is the dis-utility of effort. The effort e_t takes two values, $e_t \in \{0,1\}$, and cost of $e_t = 0$ is normalized to zero. I denote $c(1)$ by c.

Executives are heterogeneous in *general* managerial skills, or productivity, denoted by $z \in \mathbb{Z} = \{z^{(1)}, z^{(2)}, ..., z^{(n_z)}\}$. z is observable to the executive himself or herself and to firms that he or she meets, and can be carried with the individual through job-to-job transitions.[9]

Individual productivity z changes over time according to a Markov process. Denote z_t as the beginning of period t productivity. Given z_t and effort e_t, the end of period t productivity z_{t+1} follows $\Gamma_z(z_{t+1}|z_t, e_t)$. I denote the process by $\Gamma_z(z_{t+1}|z_t)$ for $e_t = 1$, and $\Gamma_z^s(z_{t+1}|z_t)$ for $e_t = 0$ (s is for shirking). To start the process, I assume all unmatched exec-

[9] Here I treat productivity as general management skills rather than firm-specific skills. However, firm-specific skills could be included using a productivity discount upon a job-to-job transition. This is left as a future extension.

utive candidates have the same starting productivity, $z = z_o$. In the following, whenever it is not confusing, I will denote z_t by z and z_{t+1} by z'.

While z and z' are observable to firms, effort e is not. Hence, there is *moral hazard.* To impose some structure on the moral hazard problem, I define the likelihood ratio as follows:

$$g(z'|z) \equiv \frac{\Gamma^s(z'|z)}{\Gamma(z'|z)}.$$

As a likelihood ratio, its expectation is one, $E[g(z'|z)] = 1$. I further assume that taking effort delivers a higher expected productivity, $E_\Gamma[z'g(z'|z)] < E_\Gamma[z']$, and that taking effort is more likely to deliver a higher productivity, i.e., $g(z'|z)$ is non-increasing in z'[10].

On the other side of the managerial labor market are firms characterized by the scale of assets, called firm size, denoted by $s \in \mathbb{S} = [\underline{s}, \bar{s}]$. Firm size is permanent and exogenous.[11] A match between a worker of productivity z and a firm of size s produces a flow of output (or a cash flow) $y(s,z) = \alpha_0 s^{\alpha_1} z$, $\alpha_0 \in (0,1), \alpha_1 \in (0,1]$. This function form entails that executive effort "roll out" across the entire firm up to a scale of α_0. It has constant return to scale if $\alpha_1 = 1$ and decreasing return to scales if $\alpha_1 < 1$.[12]

Managerial labor market

The managerial labor market is search-frictional. Executives and firms are imperfectly informed about executive types and location of firms. The search friction precludes the optimal assignments assumed in Gabaix and Landier (2008). Agents are only informed about each other's types when they meet. Search is random; executives and executive candidates all sample from the same, exogenous job offer distribution $F(s)$. Unmatched candidates meet firms with probability λ_0, while on-the-job executives meet firms with probability λ. I treat these parameters as exogenous.[13]

When a candidate meets a firm, they bargain on a contract. Suppose the continuation value of an unmatched executive candidate is W^0. Then, the firm ultimately offers a contract with a continuation value W^0, for there is no other credible threat. The individual then enters the next period as an employed executive.

When an on-the-job executive meets an outside firm, a compensation renegotiation

[10]This is the monotone likelihood ratio property (MLRP).

[11]From the point of view of the labor economics literature, one could interpret firm size here as "the productivity of the job" or "firm type". Instead of using the total number of employees, I use total asset value as a proxy for firm size since the performance of the firm is usually measured by return on assets.

[12]There has been some discussion in this literature on the appropriate production function of executives, see e.g., Edmans et al. (2017). Take s as firm size and z as the executive's per unit contribution to shareholder values. An additive production function such as $y(s,z) = s + z$ implies that the effect of executives on firm value is independent of firm size. This specification is appropriate for a perk consumption. A multiplicative production function such as $y(s,z) = sz$ is appropriate for executives' actions that can roll out across the entire firm and thus have a greater effect in a larger company. The latter is the function form adopted here.

[13]So we are in a "partial" equilibrium, in contrast to the "general" equilibrium where the labor market tightness is determined in the equilibrium.

is triggered. Otherwise, the executive has an interest in transiting to the outside firm. I allow the incumbent firm to respond to outside offers: A sequential auction is played between the executive and both firms as in Postel-Vinay and Robin (2002). If the poaching firm is larger, the executive moves to the alternative firm, for the poaching firm can always pay more than the current one can match. Alternatively, if the poaching firm is smaller, then the executive may use the outside offer to negotiate up his/her compensation. This sequential auction mechanism characterizes labor market competition in this paper.

Timing

Time is discrete, indexed by t, and continues forever. The period of an executive candidate is simple — he or she is matched to a firm with some probability and starts with a contract of continuation value W^0. An on-the-job executive enters a period with a beginning-of-period productivity z and current firm of size s. The timing is shown in figure 2.

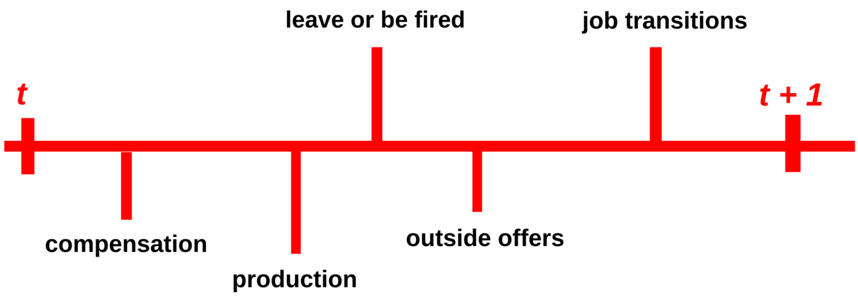

Figure 2: Timing

1. **Compensation:** The firm s firstly pays compensation w for this period, in accordance with the contract.

2. **Production:** Then, the executive enters the production phase. He or she chooses an effort level, $e \in \{0, 1\}$. His or her productivity z' is then realized according to $\Gamma(z'|z, e)$. The firm only observes the output $y(z, s)$, not the effort e. This is the moral hazard problem.

3. **Labor market:** With probability η the executive dies; otherwise, with probability λ_1, a job offer of firm size $\tilde{s} \sim F(s)$ arrives. The renegotiation game is triggered. The executive may stay in the current firm and receive higher compensation, or transit to the poaching firm. The value of the contract to the executive is determined by a sequential auction between the current and poaching firms.

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

Contractual environment

A contract defines the transfers and actions of the executive and the firm in a matched pair for all future histories, where a history summarizes the past states of the world. I define a history as follows. Call $h_t = (z'_t, \tilde{s}_t)$ the state of the world by the end of period t, where z'_t is the realized productivity by the end of t and $\tilde{s}_t$ is the size of a poaching firm. Let $\tilde{s} = s_o$ if there is no poaching firm. The history of productivity and the poaching firm up to period t, denoted by $h^t = (h_1, h_2, .., h_t)$, is common knowledge to the executive and the firm, and is fully contractible.

The two elements included in the history — productivity and poaching firm — correspond to the frictions I have in the contracting problem, namely moral hazard and search frictions. First, while productivity is included in the history and is contractible, executive's effort is not and needs to be induced by incentives. Hence, an incentive compatibility constraint is required. Second, by including poaching firms in the history, I allow a contract to stipulate whether and how to counter poaching offers. That is, competing for executives is included in the contracting problem.

Countering outside offers is optimal (or subgame perfect in game terminology); it is, therefore, necessary to allow limited commitment for both sides — to terminate the contract when the surplus is negative. Firms cannot commit to the relationship if the profits are negative. When the outside offer comes from a larger firm, the firm's participation constraint binds, and the match separates. Likewise, executives cannot commit to the match if the current firm cannot provide more than the outside value, be the unmatched value W^0 or the offer value of a poaching firm. In the former case, the executive is fired by the board. In the latter case, the executive transits to the poaching firm.

Given the information structure, I define a feasible contract as a plan that defines compensation $w_t(h^{t-1})$, a recommended effort choice $e_t(h^{t-1})$ and whether to terminate the contract $I_t(h^t)$ in every future history, represented by

$$\{e_t(h^{t-1}), w_t(h^{t-1}), I_t(h^t)\}_{t=0}^{\infty},$$

that satisfies the participation constraints of both sides and an incentive compatibility constraint.

To further simplify, I impose two assumptions. First, I assume taking effort, $e = 1$, is optimal. This assumption is consistent with Gayle et al. (2015) and in accordance with the fact that almost all executives in my data are provided with an incentive package. Secondly, I assume a reasonable support of productivity z such that the profits of a firm are always non-negative at the the unmatched value W^0. As a result, firing is excluded, and $I_t(h^t) = 1$ is equivalent to a job-to-job transition.[14]

A simplified contract state space

To recursively write up the contracting problem, I use the executive's beginning-of-period expected utility, denoted by V, as a co-state variable to summarize the history of productivities and outside offers. A dynamic contract, defined recursively, is

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

where e is the effort level suggested by the contract (optimal level is assumed to be 1), w is the compensation, W is the promised value given for a given state $(z', \tilde{s})$, and Φ is the set of feasible and incentive compatible expected utilities that can be derived following Abreu et al. (1990).[15]

4.2 Optimal contracting problem

In this section, I first characterize the participation constraints derived from the sequential auction, then I describe the contracting problem.

Sequential auction

Here I illustrate the sequential auction using value functions.[16] Let $\Pi(z, s, V)$ denote the discounted profit of a firm with size s, executive of beginning-of-period productivity z and a promised value to the executive V. The maximum bidding values $\overline{W}(z, s)$ are defined by

$$\Pi(z, s, \overline{W}) = 0.$$

The firm would rather fire the executive (normalizing the vacancy value to 0) if he or she demands a value higher than $\overline{W}$. I let $\overline{W}(z, s^o) \equiv W^0$, meaning that when there is no outside offer, the executive's outside value is simply W^0. I call $\overline{W}(z, s)$ the *bidding frontier* to highlight that it is a function (frontier) in terms of z and s.

The sequential auction works as follows. When the executive from a firm of size s (hereafter firm s) meets a poaching firm of size $\tilde{s}$ (hereafter firm $\tilde{s}$), both firms enter

[14] If I allow a large domain of z such that for some z the profit is negative at promised continuation value W^0, then firing happens. This is left as an extension in the future.

[15] Promised utilities as co-states have been used extensively in models with incentive or participation constraints. Among others, Phelan and Townsend (1991) studied a model of risk-sharing with incentive constraints, Kocherlakota (1996) analyzed the risk-sharing model with two-sided limited commitment, Hopenhayn and Nicolini (1997) studied a model of unemployment insurance, and Alvarez and Jermann (2000) studied a decentralized version of the above risk-sharing model with debt constraints.

[16] What distinguishes this model from the original sequential auction framework is that here the wage is not flat. Firms compete on a stream of wages contingent on all possible future histories. At each period, the contract defines a wage on each state (z, s).

a Bertrand competition won by the larger one. Since it is willing to extract a positive marginal profit out of every match, the best firm s can do is to provide a promised utility $\overline{W}(z', s)$. When $\tilde{s} > s$, the executive moves to a potentially better match with firm $\tilde{s}$, if the latter offers at least the $\overline{W}(z', s)$. Any less generous offer on the part of firm $\tilde{s}$ is successfully countered by firm s. If $\tilde{s}$ is smaller than s, then $\overline{W}(z', s) > \overline{W}(z', \tilde{s})$, in which case firm $\tilde{s}$ will never raise its offer up to this level. Rather, the executive will stay at his or her current firm, and be promoted to the continuation value $\overline{W}(z', \tilde{s})$ that makes him/her indifferent between staying and joining firm $\tilde{s}$.

The above argument defines outside values of the executive contingent on the state $(z', \tilde{s})$,

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

This is the participation constraint of the executive in a contracting problem.

The contracting problem

In designing the contract, the firm chooses a wage w and a set of promised values $W(z', \tilde{s})$ depending on the state $(z', \tilde{s})$. For ease of notation, I denote an effective discount factor, $\tilde{\beta} = \beta(1 - \eta)$, and write the mixture distribution of outside offers as follows:

$$\tilde{F}(s) = \mathbb{I}(s = s^o)(1 - \lambda_1) + \mathbb{I}(s \neq s^o)\lambda_1 F(s).$$

The expected profit of the firm can be expressed recursively as

$$\Pi(z, s, V) = \max_{w, W(z', \tilde{s})} \sum_{z' \in \mathbb{Z}} \left[y(s, z') - w + \tilde{\beta} \sum_{\tilde{s} \leq s} \Pi(z', s, W(z', \tilde{s}))\tilde{F}(\tilde{s}) \right] \Gamma(z'|z). \tag{BE-F}$$

subject to the promise-keeping constraint,

$$V = u(w) - c + \tilde{\beta} \sum_{z' \in \mathbb{Z}} \sum_{\tilde{s} \in \mathbb{S}} W(z', \tilde{s})\tilde{F}(\tilde{s})\Gamma(z'|z), \tag{PKC}$$

the incentive compatibility constraint,

$$\tilde{\beta} \sum_{z' \in \mathbb{Z}} \sum_{\tilde{s} \in \mathbb{S}} W(z', \tilde{s})\tilde{F}(\tilde{s})(1 - g(z, z'))\Gamma(z'|z) \geq c. \tag{IC}$$

and the participation constraints of the executive and the firm,

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

The objective function (Bellman Equation of the Firm, BE-F) includes a flow profit of $y(s, z') - w$, taking into account that the match may separate either because the executive dies, which happens with probability η, or transits to another firm, which happens with probability $\sum_{\tilde{s}>s} \tilde{F}(s)$.

The promise-keeping constraint (PKC) makes sure that the choices of the firm honor the promise made in previous periods to deliver a value V to the executive, and the promised value V contains all the relevant information in the history. The right-hand side of the constraint is the lifetime utility of the executive given the choices made by the firm. (PKC) is also the Bellman equation of an executive with state (z, s, V).

The incentive compatibility constraint (IC) differentiates itself from the promise-keeping constraint by the term $1 - g(z'|z)$. It asserts that the continuation value of effort is higher than no effort. This creates incentives for the executive to pursue the shareholders' interests rather than his or her own.

Finally, the participation constraints are stated in (PC-E) and (PC-F). The firm commits to the relationship as long as the promised value is no more than $\overline{W}(z', s)$. The sequential auction pins down the outside value of the executive, which is the minimum of bidding frontier of the poaching firm, $\overline{W}(z', \tilde{s})$, and of the current firm, $\overline{W}(z', s)$.

4.3 Equilibrium definition

Before turning to the characterization of the optimal contract, I define the equilibrium. An equilibrium is an executive unemployment value W^0, a value function of employed executives W that satisfies (PKC), a profit function of firms Π and an optimal contract policy $\sigma = \{w, e, W(z', \tilde{s})\}$ for $z' \in \mathbb{Z}$ and $\tilde{s} \in \mathbb{S}$ that solves the contracting problem (BE-F) with associated constraints (PKC), (IC), (PC-E) and (PC-F), a stochastic process of executive productivity Γ that follows the optimal effort choice, and a distribution of executives across employment states evolving according to flow equations.

Proposition 1. *The equilibrium exists.*[17]

4.4 Contract characterization

In this section, I derive a characterization of the optimal contract. The characterization builds on and extends the dynamic limited commitment literature, pioneered by Thomas and Worrall (1988) and Kocherlakota (1996), the dynamic moral hazard literature, pioneered by Spear and Srivastava (1987), and related literature in labor search such as Lentz (2014).

Proposition 2. $\Pi(z, s, V)$ *is continuous differentiable, decreasing and concave in* V*, and increasing in* z *and* s*. An optimal contract evolves according to the following updating rule. Given*

[17] The proof of the existence of the equilibrium is an exercise applying Schauder's fixed point theorem, as shown by Menzio and Shi (2010).

the beginning-of-period state (z, s, V)*, the current period compensation is given by* w^**,*

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

and the continuation value $W^*(z', \tilde{s})$ *follows*

$$W^*(z',\tilde{s}) = \begin{cases} \overline{W}(z',s) & \text{if } \ \overline{W}(z',\tilde{s}) \geq \overline{W}(z',s) \text{ or } W(z') > \overline{W}(z',s) \\ \overline{W}(z',\tilde{s}) & \text{if } \ \overline{W}(z',s) > \overline{W}(z',\tilde{s}) > W^i(z') \\ W(z') & \text{if } \ \overline{W}(z',s) \geq W^i(z') \geq \overline{W}(z',\tilde{s}) \end{cases} \tag{2}$$

where $W(z')$ *satisfies*

$$\frac{\partial \Pi(z', s, W(z'))}{\partial W(z')} - \frac{\partial \Pi(z,s,V)}{\partial V} = -\mu(1 - g(z,z')). \tag{3}$$

Proof. The properties of $\Pi(z,s,V)$ follow immediately from the proof of proposition 1. To characterize the optimal contract, I assign Lagrangian multipliers λ to (PKC), μ to (IC), $\tilde{\beta}\mu_0(z',\tilde{s})$ to (PC-E) and $\tilde{\beta}\mu_1(z',\tilde{s})$ to (PC-F). The first order condition w.r.t w gives

$$u'(w) = \lambda,$$

and the envelop theorem gives

$$-\frac{\partial \Pi(z,s,V)}{\partial V} = \lambda.$$

They together give (1). Participation constraints (PC-E) and (PC-F) can be simplified. If $\overline{W}(z',\tilde{s}) \geq \overline{W}(z',s)$, we have $W(z;,\tilde{s}) = \overline{W}(z',s)$. This is the first case in line 1 of (2). If $\overline{W}(z',\tilde{s}) \geq \overline{W}(z',s)$, the participation constraints become $\overline{W}(z',\tilde{s}) \leq W(z;,\tilde{s}) \leq \overline{W}(z',s)$. Use this to derive the first order condition w.r.t $W(z',\tilde{s})$:

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

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

Proposition 2 states that, if one abstracts from the participation constraints, an optimal contract inherits the essential properties of the classical infinite repeated moral hazard model (Spear and Srivastava 1987). Equation (1) states that the current period compensation w^* is directly linked to the promised continuation utility V, by equating the principal's and agent's marginal rates of substitution between the present and future compensation. Equation (3) says, abstract from participation constraints, the continuation utility $W(z')$ only changes to induce the executive effort. In the extreme case that the IC constraint is not binding ($\mu = 0$, μ is the multiplier of the IC constraint), $W(z') = V$ remains constant. Thus, the pay is also constant over time. Generally speaking, a higher V induces a higher $W(z')$. That is, an optimal dynamic contract has some memory.

When outside offers are realized such that the participation constraint is binding, the contract is no longer dependent on history, and the continuation value depends only on the current state. This is what Kocherlakota (1996) calls *amnesia*. More precisely, when the outside firm is larger $\tilde{s} \geq s$, the continuation value is equal to the bidding frontier of the current firm $W(z', \tilde{s}) = \overline{W}(z', s)$; when the outside firm is smaller, $\tilde{s} < s$, the continuation value depends on whether the bidding frontier of the outside firm $\overline{W}(z', \tilde{s})$ can improve upon $W(z')$.

Even when the participation constraint is binding, amnesia of the optimal contract is not "complete" — although $\overline{W}$ does not depend on the previously promised utility V, it does depend on the executive's productivity z', which is stochastically determined by past effort. Therefore, the boundaries of participation constraints carry the memory of the prior effort choice. This is where the incentives from the labor market come into effect.

4.5 Explaining the size pay-growth premium

With the characterization of the optimal contract, we are ready to explain the size premium in pay-growth and incentives. I start by defining two sets of poaching firms $\tilde{s}$: larger or smaller than the current firm.

$$\mathcal{M}_1(s) \equiv \{\tilde{s} \in \mathbb{S} | \tilde{s} > s\},$$
$$\mathcal{M}_2(z, s, W) \equiv \{\tilde{s} \in \mathbb{S} | \overline{W}(z, s) > \overline{W}(z, \tilde{s}), W < \overline{W}(z, \tilde{s})\}.$$

Given a poaching firm that belongs to the set $\mathcal{M}_1$, the executive will transit to such a firm and receive the full surplus of his or her previous job $\overline{W}(z, s)$. Given a poaching firm in $\mathcal{M}_2$, the executive will stay in the current firm but use the outside offer to renegotiate up to $\overline{W}(z, \tilde{s})$. Any poaching firm that is not in $\mathcal{M}_1$ or $\mathcal{M}_2$ is not competitive in the sense that it cannot be used to negotiate compensation with the incumbent firm.

Accordingly, the Bellman equation of executives can be written as:[18]

$$\begin{aligned} V = u(w) - c + \tilde{\beta} \sum_{z'} \Bigg[& \lambda_1 \sum_{s' \in \mathcal{M}_1} F(s') \overline{W}(z', s) + \lambda_1 \sum_{s' \in \mathcal{M}_2} F(s') \overline{W}(z', s') \\ & + \Big(1 - \lambda_1 \sum_{s' \in \mathcal{M}_1 \cup \mathcal{M}_2} F(s') \Big) W(z') \Bigg] \Gamma(z'|z), \end{aligned} \tag{PKC$'$}$$

(PKC′) shows that compensation grows mainly in two cases: i) There is a poaching firm from set $\mathcal{M}_2$ and total compensation increases without a job turnover; ii) There is a poaching firm from set $\mathcal{M}_1$ and total compensation grows upon a job-to-job transition.

The firm-size pay-growth premium observed in the data refers to the growth in the former case. In the latter case, compensation may also decrease if an executive is willing

[18] We can similarly rewrite the Bellman equations of firms using the optimal continuation value, and this

to make sacrifices on his or her current pay for the sake of higher pay in the future.[19]

To understand the firm-size pay-growth premium, consider two executives from a small firm s_1 and a large firm s_2, $s_2 > s_1$. For simplicity, suppose they have the same continuation value $W(z')$. Since the firm s_2 has a higher output and is more capable of overbidding outside offers, the corresponding set $\mathcal{M}_2$ is larger. That is, there exist poaching firms with a size between s_1 and s_2 such that the firm s_2 can overbid and retain the executive with compensation growth while the firm s_1 cannot overbid and consequently lose the executive. Therefore, the total pay increases faster in the larger firm s_2.

4.6 Explaining the size incentive premium

To explain the firm-size incentive premium, I define "performance-based incentives" and "labor market incentives" in the model. Using these definitions to rewrite the IC constraint, I then show that the two sources of incentives substitute for each other given a constant effort cost. Finally, I explain that labor market incentives decrease with firm size. Thus, performance-based incentives increase with firm size.

I first define an "incentive operator", $\mathcal{I}(\cdot)$, which calculates the incentives an executive receives from a continuation utility scheme:

$$\mathcal{I}\Big(W(z')\Big) \equiv \int_{z'} W(z')(1 - g(z,z'))\Gamma(z'|z).$$

I then rewrite the IC constraint using the incentive operator:

$$\begin{aligned} &\lambda_1 \int_{\tilde{s}\in\mathcal{M}_1} dF(\tilde{s})\mathcal{I}\Big(\overline{W}(z',s)\Big) + \lambda_1 \int_{\tilde{s}\in\mathcal{M}_2} \mathcal{I}\Big(\overline{W}(z',\tilde{s})\Big)F(\tilde{s}) \\ &+ \Big(1 - \lambda_1 \sum_{\tilde{s}\in\mathcal{M}_1\cup\mathcal{M}_2} F(\tilde{s})\Big)\mathcal{I}\Big(W(z')\Big) \geq c/\tilde{\beta}, \end{aligned} \tag{IC$'$}$$

The incentives comprise: i) incentives brought by larger firms in $\mathcal{M}_1$; ii) incentives brought by smaller firms in $\mathcal{M}_2$; iii) incentives in performance-related pay when there are no poaching firms from $\mathcal{M}_1$ or $\mathcal{M}_2$.

The incentives when there are no competitive poaching offers are the *performance-*

equation is consistent with Postel-Vinay and Robin (2002):

$$\begin{aligned} \Pi(z,s,V) = \max_{w,W(z')} \sum_{z'} \Bigg[y(s)z' - w + \tilde{\beta}\Bigg(\lambda_1 \sum_{s'\in\mathcal{M}_2} F(s')\Pi_1(z',s,\overline{W}(z',s')) \\ + \Big(1 - \lambda_1 \sum_{s'\in\mathcal{M}_1\cup\mathcal{M}_2} F(s')\Big)\Pi_1(z',s,W(z'))\Bigg)\Bigg]\Gamma(z'|z). \end{aligned} \tag{BE-F$'$}$$

[19] Where compensation information is available in both the original and target firms, it would be interesting to examine whether there is also a firm-size compensation-growth premium in job-to-job transitions. This is, however, not possible with the current dataset.

based incentives, denoted by Ξ_p:

$$\Xi_p\Big(W(z')\Big) \equiv \Big(1 - \lambda_1 \sum_{\mathcal{M}_1 \cup \mathcal{M}_2} F(s')\Big)\mathcal{I}\Big(W(z')\Big), \tag{4}$$

and the incentives from poaching offers are the *labor market incentives*, denoted by Ξ_m:

$$\Xi_m\Big(s, W(z')\Big) \equiv \lambda_1 \int_{\tilde{s}\in\mathcal{M}_1} dF(\tilde{s})\mathcal{I}\Big(\overline{W}(z', s)\Big) + \lambda_1 \int_{\tilde{s}\in\mathcal{M}_2} \mathcal{I}\Big(\overline{W}(z', \tilde{s})\Big)F(\tilde{s}). \tag{5}$$

Ξ_m would be zero if there were no poaching offers.

Because of labor market incentives, the need for performance-based incentives is less. Intuitively, firms appreciate higher productivities and are willing to bid higher for a more productive executive. The sequential auction in the model therefore begets labor market incentives for executive effort: If working hard today is not only an input into current production but also an investment in the (inalienable and transferable) human capital, then it is intuitive that the objectives of the firm and of the executive become better aligned and the need for short-term compensation incentives decreases.

Mathematically, Ξ_m is an expectation of incentives from all possible poaching offers. When the poaching firm is larger than the current firm, the incentives are from the bidding frontier of the current firm. When the poaching firm is smaller than the current firm, the incentives are from the bidding frontier of the poaching firm.

The magnitude of Ξ_m is determined by current firm size s and the promised continuation value $W(z')$. In particular, firm size s enters Ξ_m via bidding frontiers. That is, Ξ_m depends on s even though the moral hazard problem fundamentally does not. On the other hand, $W(z')$ determines the lower bound of set $\mathcal{M}_2$. The larger the promised continuation value $W(z')$, the less likely a poaching firm can be used to renegotiate with the current firm, and the lower the labor market incentives.

Based on this, there is a simple "job ladder" explanation for the size premium when comparing executives of different pay levels. Such an incentive premium is reported in column (1) in table 2. Since executives of larger firms tend to have higher total compensation, the corresponding continuation values are larger; they are thus higher on the job ladder. Accordingly, the chance of encountering a competitive poaching offer that beats the current value is smaller. Hence, labor market incentives are lower. As a result, executives in larger firms require more incentives in performance-related pay.[20] As we will see in the following section, this "job ladder" argument also applies to explaining the size incentive premium among executives with the same total compensation.

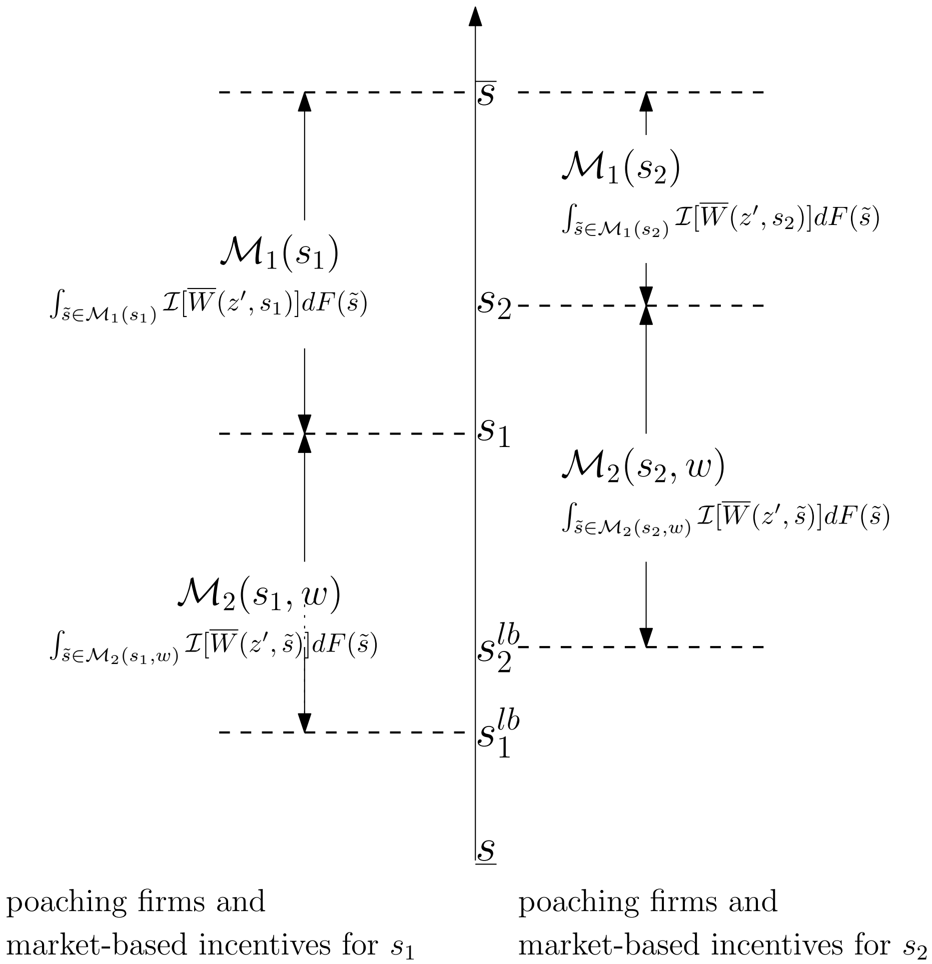

Figure 3: Compare labor market incentives

Note: The figure illustrates labor market incentives for executives with the same compensation w from firms of size s_1 and s_2. The vertical axis labels the size of poaching firms $[\underline{s}, \overline{s}]$. s_1^{lb} is the lower bound of set $\mathcal{M}_2(s_1, w)$ and s_2^{lb} is the lower bound of set $\mathcal{M}_2(s_2, w)$. The labor market incentives of s_1 and s_2 are on the left and right of the vertical axis, respectively. The notation for each interval is followed by the value of incentives from poaching firms of that interval.

Labor market incentives decrease with firm size

Now I compare labor market incentives for executives who have the same total compensation but come from firms of different size. I show that when there is enough concavity in the utility function, labor market incentives decrease with firm size. Therefore, larger firms need to provide more performance-based incentives. This explains the firm-size incentive premium.

Consider two executives from firms s_1 and s_2, $s_1 < s_2$. The executives have the same total compensation. Figure 3 illustrates the possible poaching firm sizes for the two executives and the associated incentives. The poaching firm size ranges from $\underline{s}$ to $\overline{s}$. I denote

[20]This is an alternative explanation in addition to the current explanations based on moral hazard (Gayle and Miller, 2009) and on multiplicative utility (Edmans et al., 2009).

the lower bound of $\mathcal{M}_2$ for firms s_1 and s_2 by s_1^{lb} and s_2^{lb}, respectively. Notice that $s_2^{lb} > s_1^{lb}$ because they are determined by lifetime utilities rather than current period compensation. Although the two executives have the same total compensation, the one in s_2 has higher lifetime utility. The left side of the axis depicts sets $\mathcal{M}_1$, $\mathcal{M}_2$ and corresponding labor market incentives in the two sets for the executive in s_1. The right side of the axis depicts the counterparts for s_2. Taking the difference between $\Xi_m(s_2)$ and $\Xi_m(s_1)$, we have

$$\begin{aligned}\Xi_m(s_2) - \Xi_m(s_1) = &- \int_{s_1^{lb}}^{s_2^{lb}} d\tilde{F}(\tilde{s})\mathcal{I}(\overline{W}(z',\tilde{s})) \\ &+ \int_{s_2}^{\bar{s}} d\tilde{F}(\tilde{s})\Big(\mathcal{I}(\overline{W}(z',s_2) - \mathcal{I}(\overline{W}(z',s_1)\Big) \\ &+ \int_{s_1}^{s_2} \Big(\mathcal{I}(\overline{W}(z',\tilde{s}) - \mathcal{I}(\overline{W}(z',s_1)\Big) d\tilde{F}(\tilde{s}). \end{aligned} \tag{6}$$

Their labor market incentives are different in two respects. First, with poaching firms in $[s_1^{lb}, s_2^{lb}]$, the executive in s_1 receives an incentive of $\int_{s_1^{lb}}^{s_2^{lb}} d\tilde{F}(\tilde{s})\mathcal{I}(\overline{W}(z',\tilde{s}))$, while the executive in s_2 has no incentive from the labor market. This is the first item in (6), and it corresponds to the job ladder argument previously mentioned — since $s_2^{lb} > s_1^{lb}$, the executive in s_2 is less likely to receive a competitive outside offer, and labor market incentives are lower.

Second, for poaching firms in the range of $[s_1, \bar{s}]$, labor market incentives for firms s_1 and s_2 are drawn on different bidding frontiers, which correspond to the second and third items in (6). With poaching firms in this range, the bidding frontier for the executive of firm s_1 is always $\overline{W}(z', s_1)$, since any poaching firm larger than s_1 can bid just $\overline{W}(z', s_1)$ to attract the executive. In contrast, the bidding frontiers for the executive in firm s_2 are either $\overline{W}(z', s_2)$ or $\overline{W}(z', \tilde{s})$ with $\tilde{s} > s_1$, both of which are larger than $\overline{W}(z', s_1)$.[21] Consequently, the certainty equivalent of the executive in s_2 is higher. By diminishing marginal utility, the incentives from these higher bidding frontiers are lower:

$$\mathcal{I}\Big(\overline{W}(z',s_1)\Big) > \mathcal{I}\Big(\overline{W}(z',\tilde{s})\Big) \text{ for } \tilde{s} > s_1.$$

This is a wealth effect of poaching offers — a wealthier executive is harder to incentivize. This wealth effect holds as long as the utility function is sufficiently concave. In the following, I give a sufficient condition under the restriction that the utility function is of the CRRA form and effort cost c is equal to a particular value.

Proposition 3 (Labor market incentives and firm size)**.** *Suppose the executives' utility is of the CRRA form, and the cost of effort $c = \bar{c}(s)$, then $\mathcal{I}\Big(\overline{W}(z',s)\Big)$ decreases in s if*

$$\sigma > 1 + \frac{s^{1-\alpha_1}}{\alpha_1}\psi'(s), \tag{7}$$

[21] $\overline{W}(z', s)$ is strictly increasing in s.

where $\psi(s)$ is a function of s that is positive and increasing in s and

$$\bar{c}(s) \equiv \tilde{\beta} \sum_{z' \in \mathbb{Z}} \overline{W}(z', s)(1 - g(z'|z))\Gamma(z'|z).$$

Proof. See Appendix A. □

To understand the proposition, first notice that $\mathcal{I}\left(\overline{W}(z', s)\right)$ is simply a weight sum of $\frac{\Delta \overline{W}(z',s)}{\Delta z'}$ over the domain of z' — the steeper $\overline{W}(z', s)$ with respect to z', the higher the incentives to induce effort. So it would be sufficient to show that $\frac{\Delta \overline{W}(z',s)}{\Delta z'}$ decreases in s under the stated condition. To proceed, it follows that

$$\frac{\Delta \overline{W}(z, s)}{\Delta z} = -\frac{\Delta \Pi(z, s, \overline{W})/\Delta z}{\Delta \Pi(z, s, \overline{W})/\overline{W}} = \frac{\tilde{\alpha} \times s}{1/u'(\overline{w})}, \tag{8}$$

where $\overline{w}$ is the optimal compensation for the current period, corresponding to a promising continuation value $\overline{W}$. The first equality follows from an implicit differentiation. In the second equality,

$$\Delta \Pi(z, s, \overline{W})/\Delta z = \tilde{\alpha} \times s$$

because keeping the promised value, all increasing output is accrued to the company. In particular, $\tilde{\alpha}$ is α multiplied by a factor that adjusts for the possibility that the executive will leave the firm and the job is destructed. On the denominator,

$$\Delta \Pi(z, s, \overline{W})/\overline{W} = -1/u'(\overline{w})$$

follows directly from condition (1) in Proposition 2.

There are two opposing effects of s involved in (8). On the one hand, the maximum value that larger firms are able to bid changes more with respect to z due to the multiplicative production function. This will generate more labor market incentives, and it is reflected in the numerator of (8). On the other hand, the incentives in terms of utilities can actually be lower because the marginal utility for extra returns from the executive labor market is lower now ($\overline{w}$ increases in s making $u'(\overline{w})$ lower). This is reflected in the denominator of (8). The second force dominates when the utility function has enough concavity, as stated in the proposition.

The requirement stated in (7) is consistent with the literature in this context. The existing studies usually estimate or calibrate a higher σ value. For example, a careful calibration study on CEO incentive pay by Hall and Murphy (2000) uses σ between 2 and 3. Calibration exercises on CEO incentive compensation convexity starting from Dittmann and Maug (2007) are based on $\sigma > 1$. Using an employer-employee matched data from Sweden for the general labor market, Lamadon (2016) estimates that $\sigma = 1.68$. Numerically, I find the right-hand side of (7) is approximately equal to one in the parameter space explored in my estimation.

Back to the firm-size incentive premium. If (7) is satisfied, and given (6), the labor market incentives Ξ_m are lower for the executive in firm s_2. Since the effort cost is the same for both executives, the executive in the larger firm s_2 demands more incentives from the performance-related pay. This explain the firm-size incentive premium.

5 Empirical Evidence

To quantitatively evaluate the model, I use data on executives employed in U.S. publicly listed firms. Close scrutiny of the managerial labor market allows me to put together a rich array of data from various sources. Specifically, I assemble a new dataset on job turnovers from BoardEX, and merge the job turnover data with two sets of standard data, the executive compensation from ExecuComp, and firm-level information from CompuStat. In the following, I provide a brief description of the relevant data features. In particular, I examine executives' job-to-job transitions, and whether they climb the job ladder towards larger firms. These are the key features of the managerial labor market in the model. Additionally, I examine whether the job-to-job transition rate decreases with firm size as predicted by the model.

5.1 Data

The empirical analysis and estimation mainly rely on the ExecuComp database, which provides rich information on executive compensation of the top five to eight executives in companies included in the $S\&P$ 500, MidCap and SmallCap indices for the period of 1992 to 2016. The accounting information from CompuStat and stock returns from CRSP are merged with ExecuComp. The dataset provided by Coles et al. (2006) and Coles et al. (2013) contains performance-based incentives *delta* calculated based on ExecuComp. To collect job turnover information, I extract the full employment histories of executives from the BoardEX database, and supplement them with the information from executives' LinkedIn and Bloomberg pages.

My final sample comprises $35,088$ executives with age between 30 and 65.[22] Of these, $26,972$ episodes cover the full tenure of the executive from beginning to end. The total number of executive-fiscal year observations in my sample is $218,168$. The minimum number of firms covered in a given year is $1,556$ in 1992, and the maximum is $2,235$ in 2007. All nominal quantities are converted into constant 2016 dollars using a GDP deflater from the Bureau of Economic Analysis.

Here I describe the variables that are used in my analysis. Using information from ExecuComp, I identify the *gender* and *age* for each executive, the *tenure* in the current executive episode, whether he or she is a *CEO*, or *CFO*, or *director* of the board or is involved

[22]I select this age range because the managerial labor market is more relevant than for those passing the retirement age.

Table 3: Summary statistics

Variable	N	mean	sd	p25	p50	p75
age	218,168	51.04	6.96	46	51	56
male	218,168	0.936	0.244	1	1	1
CEO	218,168	0.184	0.387	0	0	0
CFO	218,168	0.096	0.295	0	0	0
director	218,168	0.339	0.473	0	0	1
interlock	218,168	0.013	0.112	0	0	0
tenure	218,168	4.71	3.793	2	4	6
tdc1	198,673	2,555.527	5,454.153	632.164	1,270.806	2,690.385
delta	146,790	322.518	4,736.982	16.966	50.634	154.411
mkcap	212,271	7,997.377	25,810.758	598.919	1,622.236	5,169.379
at	216,384	15,594.888	98,653.077	542.863	1,796.467	6,570.342
sales	216,276	5,472.709	17,387.175	428.2	1,217.738	3,917.269
profit	209,639	0.119	0.359	0.069	0.121	0.176
annual return	211,067	0.181	0.802	-0.127	0.106	0.356
mbr	183,565	1.669	2.21	0.811	1.198	1.913

Note: The table reports summary sample statistics for my dataset, which covers named executive officers reported in ExecuComp over the period of 1992 to 2016. *age* is the executive's age by the end of the fiscal year. Sample episodes with age lower than 35 or higher than 70 are dropped. Dummy variables *CEO*, *CFO*, *director* and *interlock* indicate whether the executive served as a CEO, or a CFO, or a director, or is involved in the interlock relationship during the fiscal year, respectively. An interlock relationship is described in the note of table 1. *tenure* (in years) counts the number of fiscal years that the executive works as a named officer. *tdc1* is the total compensation, composed of the following: Salary, Bonus, Other Annual, Total Value of Restricted Stock Granted, Total Value of Stock Options Granted (using BlackScholes), Long-Term Incentive Payouts, and All Other Total. *delta* is the dollar change in wealth associated with a 1% change in the firms stock price (in $000*s*). *mkcap* (in millions) is the market capitalization of the company, calculated by *csho* (Common Shares Outstanding, in millions of shares) multiplied by *prcc_f* (fiscal year end price). *prcc_f* and *csho* are reported in CompuStat Fundamentals Annual file. *at* (in millions) is the Total Book Assets as reported by the company. *sales* (in millions) is the Net Annual Sales as reported by the company. *profit* is the profitability, calculated by EBITDA/Assets. *annual return* is the annualized stock return which is compounded based on CRSP MSF (Monthly) returns. MSF returns have been adjusted for splits, etc. *mbr* is the Market-to-Book Ratio, calculated by Market Value of Assets divided by Total Book Assets. Market Value of Assets is calculated according to $\textit{Market Value of Assets (MVA)} = prcc\_f * cshpri + dlc + dltt + pstkl - txditc$. Variable definitions are provided in the main text.

in an *interlock* relationship during the fiscal year. Table 3 reports summary statistics for my sample. Ninety-three percent of the executives are males and the average age is 51. The average length of an episode is 6.21 years. Among all executive-year observations, 18.4% are CEO spells and 9.6% are CFO spells.

In terms of the compensation information, *tdc1* is the total compensation including salary, bonus, values of stock and options granted, etc. The total compensation has an average of $2,555,000$ dollars, with a 25th percentile of $632,000$ dollars and a 75th percentile of $2,690,000$ dollars. In terms of means, only 16.5% of the total compensation is fixed base salary; the rest is all incentive-related. Performance-based incentives come not only from the total compensation each year, but also from the stocks and options that are granted in previous years. The variable *delta* measures how strong performance-based incentives are in firm-related wealth. It is defined by the dollar change in wealth (in $000*s* in table 3) associated with a 1% change in the firms stock price. The distribution of $delta$ is right-skewed, with a mean of $323,000$ dollars, even larger than its 75th percentile of $154,000$ dollars.

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

The job turnover information comes from the BoardEX database.[23]BoardEX contains details of each executive's employment history, including start and end dates, firm names and positions. It also has extra information on educational background, social networks, etc. I merge the two databases using three sources of information: the executive's first, middle and last names, date of birth, and working experiences, i.e. in which years the executive worked in which firms. If all three aspects are consistent, the executive is identified. For executives that cannot be identified in BoardEX, I search for the respective LinkedIn and Bloomberg pages and manually collect the available employment information. In this way, I am able to identify more than 93% of executives in ExecuComp, $32,864$ in total.

5.2 Job-to-job transitions

I define a job-to-job transition as the executive leaving their current firm and starting to work in another within 190 days. Otherwise, the event is defined as an exit from

[23]What is missing in the ExecuComp database is the information on executives' employment history. For example, there is no information to identify whether the executive transits to another firm after the current position in an *S&P* firm or whether they simply retire. Moreover, the start and end dates of current employment are not known.

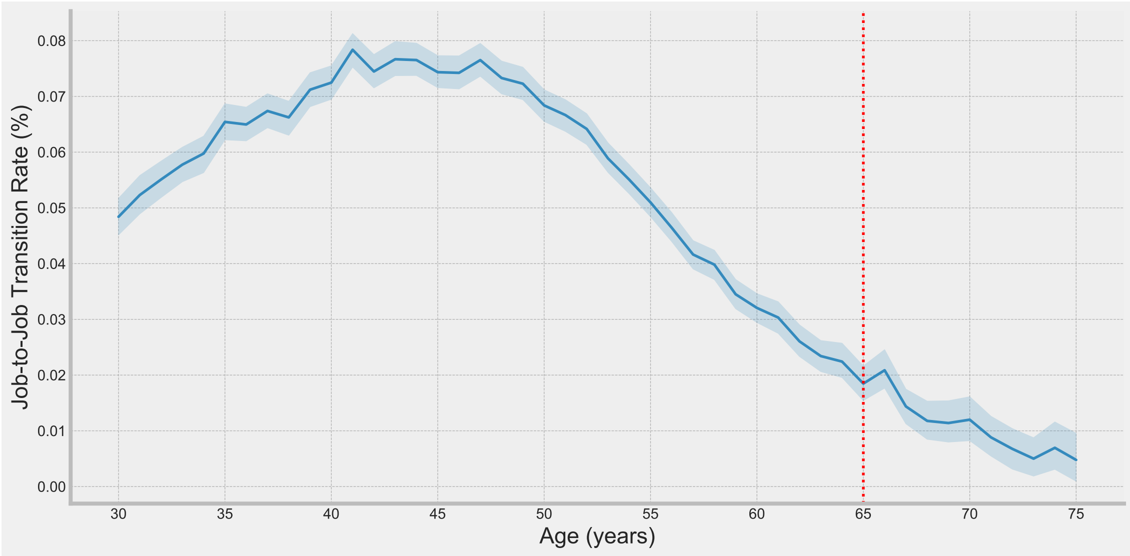

Figure 4: Job-to-job transition rate over age

Note: The figure depicts estimated job-to-job transition rates over age with the 95% confidence interval. A job-to-job transition is defined as an executive leaving the current firm and starting to work in another firm within 190 days.

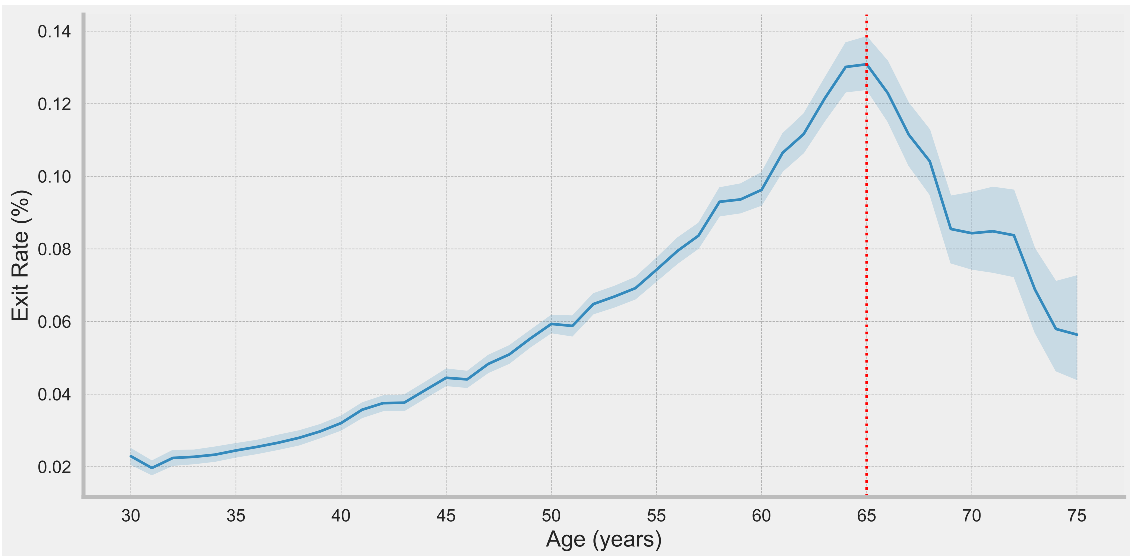

Figure 5: Exit rate over age

Note: The figure depicts the estimated exit rates over age with the 95% confidence interval. A job exit is defined as an executive leaving the current firm and not working in another firm within 190 days.

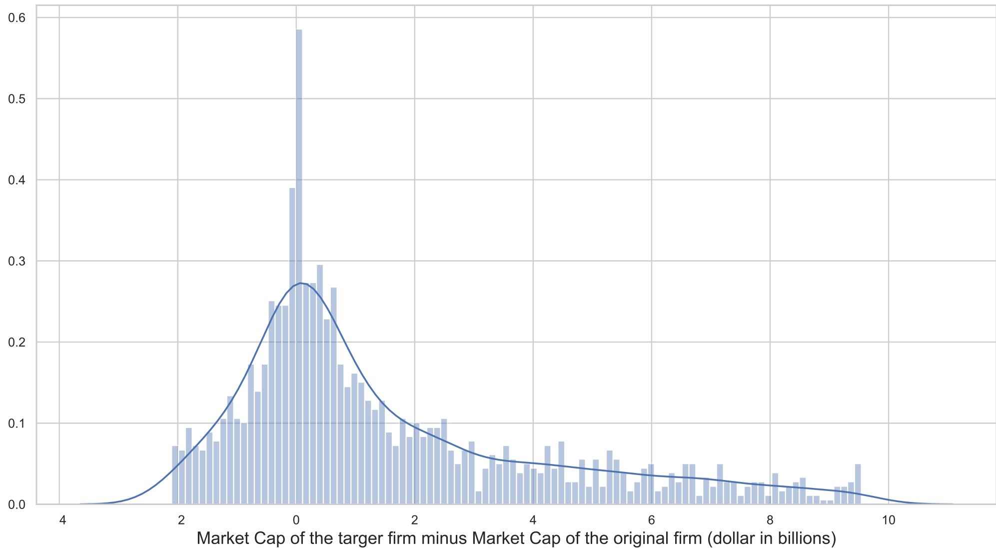

Figure 6: Distribution of the change of firm size upon job-to-job transitions

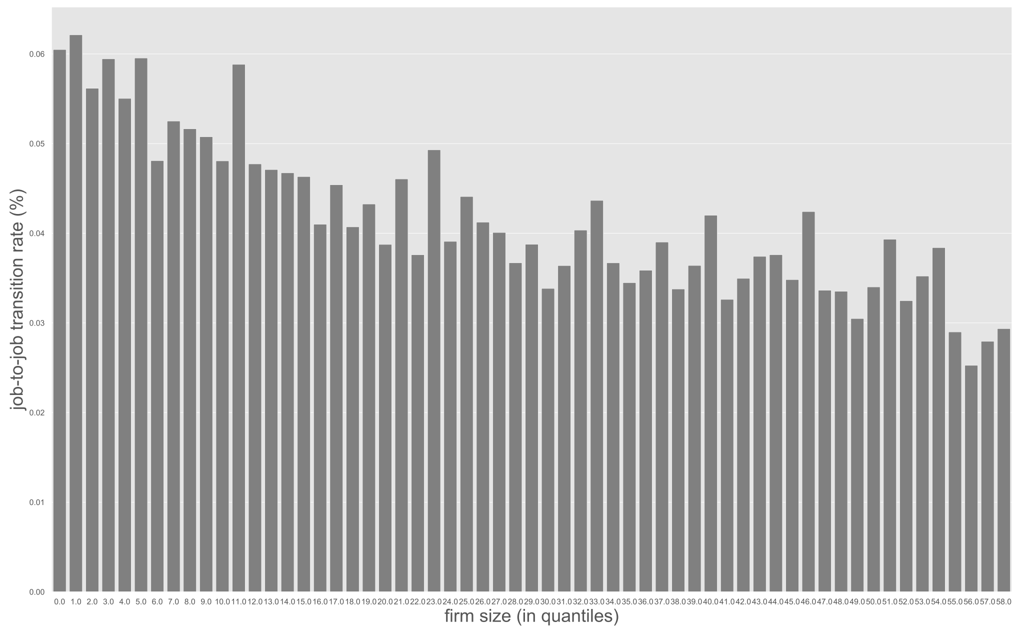

Figure 7: Job-to-job transition rates across firm size

the managerial labor market. In the data, the job-to-job transition rate is 4.98% each year over the period of 1992 to 2015, while the job exit rate is slightly higher, at 6.91%. Figure 4 illustrates how job-to-job transition changes with age, and figure 5 shows how job exit changes with age. To illustrate the trend, the figures also include those who did not retire after age 65. As shown in the figure, the job-to-job transition rate increases gradually before 40 and peaks at around age 45 before decreasing after 50. In contrast, the job exit rate is lower before 55 and peaks sharply at age 65 as expected.

Most job-to-job transitions are within the industry. Among transitions for which industry information is available, 1,717 out of 2,567 transitions are within the industry as defined by the Fama-French 12 industry classification, and 1,407 out of the 2,567 cases as defined by the Fama-French 48 industry classification.

Executives transit to larger firms

In my sample, there are $9,138$ job-to-job transitions from a CompuStat firm; only $2,567$ have firm size information on both the original and target firms. The rest are private firms whose size information is not disclosed. Based on the selected sample where size information is available, I find that approximately 60% of job-to-job transitions are associated with a firm size increase. The pattern is stable across age-groups and industries, as shown in table 4. I further check the transitions towards smaller firms. It turns out that 20% of these cases are due to a title change from a non-CEO title to a CEO title, while this fraction is only 3.3% in transitions towards larger firms.

Figure 6 portrays the distribution of the change of firm size upon a transition. While many transitions are between firms with similar size, there are a lot of "leap" transitions where the target firm is much larger. This lends support to my modeling of the managerial labor market, where executives engage in *random* on-the-job search.

Job-to-job transitions decrease with firm size

Next, I check whether executives in larger firms have fewer transitions, which is predicted by the model. As a first pass, figure 7 depicts the transition rates across firm size quantiles. The transition rate decreases from more than 6% at the 5th percentile of firm size to around 3% at the 95th percentile of firm size. To further investigate how job-to-job transitions vary with firm size, I estimate a Cox model on how firm size affects the time to job-to-job transitions, controlling for executive age, firm performance indicators, year and industry dummies. For a 1% increase in the firm scale, the hazard rate decreases by 8.3% without controlling for total compensation, and by 2.8% after controlling for total compensation. That is, larger firms have significantly lower job-to-job transition rate.[24]

[24]Contradicting the model's prediction that job-to-job transition is not related to compensation level, in the data, when the total compensation rises by 1%, the hazard rate drops by 27% One possible explanation

Table 4: Change of firm size upon job-to-job transitions

Panel A: All executives

Firm size proxy	Total obs.	Firm size decrease obs. (%)	Firm size increase obs. (%)
Market Cap	2,567	985 (39%)	1,582 (61%)
Sales	2,617	1,051 (40%)	1,566 (60%)
Book Assets	2,616	1,038 (40%)	1,578 (60%)

Panel B: Across age groups

Age groups	Total obs.	Firm size decrease obs. (%)	Firm size increase obs. (%)
≤ 40	100	34 (34%)	66 (66%)
[40, 45)	381	135 (35%)	246 (65%)
[45, 50)	701	262 (37%)	439 (63%)
[50, 55)	766	304 (40%)	462 (60%)
[55, 60)	261	179 (43%)	82 (67%)
[60, 65)	73	52 (39%)	21 (61%)
[65, 70)	30	7 (25%)	23 (75%)
≥ 70	6	1 (16%)	5 (84%)

Panel C: Across industries

Fama-French industries (12)	Total obs.	Firm size decrease obs. (%)	Firm size increase obs. (%)
1	119	39 (33%)	80 (67%)
2	88	33 (38%)	55 (61%)
3	281	98 (35%)	183 (65%)
4	120	58 (48%)	62 (52%)
5	71	30 (42%)	41 (58%)
6	609	229 (38%)	380 (62%)
7	60	20 (33%)	40 (67%)
8	96	48 (50%)	48 (50%)
9	381	142 (37%)	239 (63%)
10	197	89 (45%)	108 (65%)
11	314	115 (37%)	199 (63%)
12	231	84 (36%)	147 (64%)

Table 5: Job-to-job transitions and firm size

	Job-to-Job transition	
	(1)	(2)
log(firm size)	0.917****	0.972*
	(0.0109)	(0.0139)
age	0.985****	0.967***
	(0.00273)	(0.0112)
log(tdc1)		0.830****
		(0.0150)
other controls	X	X
year x *industry*	X	X
N	154635	118119
chi2	496.1	491.4

Note: I estimate a Cox proportional hazards model with the event of a job-to-job transition. A job-to-job transition is defined as the executive leaving the current firm (and not returning to the current firm within one year), and starting to work in another firm within 190 days. All variables have the same definition as in table 1. All dollar-related variables are adjusted by a GDP deflater. The standard errors are shown in parentheses, and I denote symbols of significance by * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$.

6 Estimation

I estimate the model parameters using Simulated Methods of Moments. That is, I use a set of moments that are informative for the parameters and minimize the distance between data moments and model-generated moments. My moments are partly coefficients from auxiliary regressions, so the approach could alternatively be presented as Indirect Inference. I first introduce the numerical method that I employ to solve the dynamic contracting problem. Then I describe the model specifications and moments used for identification. Specifically, I do not explicitly target the firm-size pay-growth and incentive premiums. After reporting the parameter estimates, I compare the estimates of the premiums in the data and in the model simulated data. I show that the model quantitatively captures both premiums.

6.1 Numerical method

To solve the contracting problem, one needs to find the optimal promised values in each state of the world for the next period. This becomes infeasible as soon as reasonable supports are considered for $\mathbb{Z}$ and $\mathbb{S}$. Instead of solving for promised values directly, I

is that the compensation level contains information on ranks which are related to the production function parameters α_0 and α_1. Perhaps a more accurate way to measure production is by "effective firm size", which combines both firm asset scales and executive rank information.

use the recursive Lagrangian techniques developed in Marcet and Marimon (2017) and extended by Mele (2014). Under this framework, the optimal contract can be characterized by maximizing a weighted sum of the lifetime utilities of the firm and the executive, where in each period the social planner optimally updates the Pareto weight of the executive to enforce an incentive compatible allocation. This Pareto weight becomes a new state variable that recursifies the dynamic agency problem. In particular, this endogenously evolving weight summarizes the contract's promises according to which the executive is rewarded or punished based on the performance and outside offers. Ultimately, solving an optimal contract is to find the sequence of Pareto weights that implements an incentive-compatible allocation. Once these weights are solved, the corresponding utilities can be recovered. This technique improves the speed of computation and makes the estimation feasible. I leave more details to Appendix C.

6.2 Model specification and parameters

I estimate the model fully parametrically and make several parametric assumptions. Being consistent with the analysis above, I use the constant relative risk aversion utility function:

$$u(w) = \frac{w^{1-\sigma}}{1-\sigma},$$

and a production function:

$$y(z,s) = e^{\alpha_0} s^{\alpha_1} z.$$

I model the process of productivity by an $AR(1)$ process:

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

where ϵ follows a normal distribution $N(0, \sigma_\epsilon)$, and the mean for no effort, $\rho_0(0)$, is normalized to zero. The process is transformed into a discrete Markov Chain using Tauchen (1986) on a grid of 6 points.[25] Furthermore, I set the sampling distribution of firm size $F(s)$ as a truncated log-normal distribution with expectation of μ_s and standard deviation of σ_s.[26] Finally, the discount rate β is set to be 0.9 for the model is solved annually. I set the number of grid points for the Pareto weight to be 50 and for firm size s to be 20. Table 6 lists the complete set of parameters that I estimate.

6.3 Moments and identifications

I next make a heuristic identification argument that justifies the choice of moments used in the estimation. Firstly, for the identification of the productivity process, the exit rate,

[25]The choice of grid points is for speed of estimation. The simulated moments are very robust in this choice.

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

Table 6: Parameters

Parameters	Description
η	the death probability
λ_1	the offer arrival probability
ρ_z	the $AR(1)$ coefficient of productivity shocks
μ_z	the mean of productivity shocks for $e = 1$
σ_z	the standard deviation of productivity shocks
μ_s	the mean of $F(s)$
σ_s	the standard deviation of $F(s)$
c	cost of efforts
σ	relative risk aversion
α_0, α_1	production function parameter

and offer arrival rate, there are direct links between the model and the data. The exit rate directly informs η. Likewise, the incidence of job-to-job transitions is monotonically related to λ_1. The parameters of the productivity process, namely ρ_z, μ_z and σ_ϵ, are informed directly by the estimates of an $AR(1)$ process relating to the profitability of each firm-executive match:

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

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

Secondly, the two parameters governing the job offer distribution, μ_s and σ_s, are disciplined by the mean and variance of firm size. Given $\lambda_1 > 0$, the higher μ_s, the more likely executives can transit to larger firms and the larger the mean of $\log(size)$. Similarly, the higher σ_s, the more heterogeneous the outside firms, and both mean and variance of $\log(size)$ increase.

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

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

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

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

and use $\beta_{delta-tdc1}$ to inform σ. Numerical exercises show that $\beta_{delta-tdc1}$ is closely related to σ. The higher σ, the larger $\beta_{delta-tdc1}$.

Firm-size premiums

I intentionally leave the firm-size pay-growth premium and incentive premium untargeted in the estimation. Instead, in the real data and the simulated data by the estimated model, I separately estimate these premiums using the same regression specification in order to examine whether the model mechanism can match up with the real world. In both the data and model-generated data, the premiums are estimated as follows. The firm-size pay-growth premium is the coefficient $\beta_{\Delta tdc1-size}$ in the following regression:

$$\Delta \log(tdc1_{it}) = \beta_3 + \beta_{\Delta tdc1-size} \log(size_{it}) + \beta_4 \log(tdc1_{it}) + \epsilon_{it,3}; \tag{9}$$

and the firm-size incentive premium is the coefficient $\beta_{delta-size}$ in the following regression:

$$\log(delta_{it}) = \beta_5 + \beta_{delta-size} \log(size_{it}) + \beta_6 \log(tdc1_{it}) + \epsilon_{it,4}. \tag{10}$$

The estimates of both premiums in the data are shown in column (2) of table 1 and table 2 in the section of motivating facts, respectively.

6.4 Estimates

Table 7 reports the targeted values of moments in the data and the corresponding values in the estimated model. The last two columns list the parameter estimates and standard errors. While I arrange moments and parameters along the identification argument made in the previous subsection, all parameters are estimated jointly. Overall, the model provides a decent fit to the data.

Looking into the estimates, a job arrival rate $\lambda_1 = 31.64\%$ is required to match the job-to-job transition rate 4.98% in the data. The magnitude of λ_1 indicates that, on average, the executive will receive an outside offer every three years. Most job offers (about 84%) are from poaching firms that are smaller than the current firm and are used to negotiate compensation with the current firm. This is confirmed by a small mean of poaching firm size. The magnitude of μ_s indicates that most offers are provided by relatively small firms, though the magnitude of σ_s implies the variation of poaching firm size is high. Comparing the data and the model-simulated mean and variance of $\log(size)$, it seems using a log-normal distribution is sufficient to match the firm size distribution in

Table 7: Moments and estimates

Moments	Data	Model	Estimates	Standard Error
Exit rate	0.0691	0.0691	$\eta = 0.0695$	0.0127
J-J transition rate	0.0498	0.0473	$\lambda_1 = 0.3164$	0.0325
$\hat{\rho}_{profit}$	0.7683	0.6299	$\rho_z = 0.8004$	0.0366
$Mean(profit)$	0.1260	0.1144	$\mu_z = 0.0279$	0.0014
$Var(profit)$	0.0144	0.0160	$\sigma_z^2 = 0.1198$	0.0044
$Mean(\log(size))$	7.4515	7.4806	$\mu_s = 1.2356$	0.0365
$Var(\log(size))$	2.3060	2.1610	$\sigma_s = 2.5795$	0.1211
$Mean(\log(tdc1))$	7.2408	7.2665	$\alpha_0 = -1.5534$	0.0147
$Var(log(tdc1))$	1.1846	0.8960	$\alpha_1 = 0.5270$	0.0217
$\beta_{tdc1-size}$	0.3830	0.2822		
$\beta_{delta-tdc1}$	1.1063	1.1997	$\sigma = 1.1038$	0.0030
$Mean(\log(delta))$	8.4994	8.478	$c = 0.0814$	0.0259
$Var(\log(delta))$	3.4438	3.35872		

the data.

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

6.5 Predicting firm-size premiums

Table 8 reports the size-premium estimates in the data and the model simulated data. There are three premiums. The first row is the size pay-growth premium estimated in regression (9). The second row and the third row are both the size incentive premiums estimated in regression (10) except that the total compensation is not controlled in estimating *incentive premium* $_{\text{(w/o tdc1)}}$ in the last row. Therefore, it includes premiums for both level and compositional reasons, while the second row is the incentive premium that cannot be attributed to pay levels of total pay, which is the focus of my explanation. Nevertheless, I show all the premiums can be replicated by my model.

Table 8: Predictions on size premiums

	Benchmark		Model Variants		
	Data	Model	w/o mkt inc	More offers	Less offers
Size premiums	(1)	(2)	(3)	(4)	(5)
pay-growth premium	0.1542	0.1450	0.1481	0.1624	0.0411
incentive premium	0.3473	0.3122	-0.0444	0.4299	0.1964
incentive premium **(w/o tdc1)**	0.6044	0.6507	0.4202	0.7093	0.4076

Column (1) shows the premium estimates in the data, as reported in table 1 and table 2. Column (2) shows the estimates in the benchmark model using the estimated parameters. Comparing columns (1) and (2), I find that even without targeting these premiums, the model can quantitatively capture all three premiums. In the model, the size pay-growth premium is driven by the renegotiation, and the size incentive premium is driven by labor market incentives. There is nothing mechanical that forces these estimates to coincide between the data and the model. The fact that the predicted premiums match up so closely with the estimates in the data is reassuring for the ability of the model mechanism to play an important role in explaining the firm size premium. In particular, since my model carries the insights of Edmans et al. (2009), I am able to predict the size incentive premiums with or without controlling for total compensation.

To further clarify the mechanisms behind the premium predictions, in columns (3) to (5), I report the premium estimates in several model variants. In column (3), I simulate a counterfactual scenario where firms ignore labor market incentives when designing incentive contracts. In column (4), I simulate the model using a higher job arrival probability $\lambda_1 = 0.6$. In column (5), I simulate the model with a lower job arrival probability $\lambda_1 = 0.1$.

Column (3) shows that once labor market incentives are ignored, while the pay-growth premium remains almost the same as in column (2), the incentive premium (after controlling for total compensation) in the second row essentially becomes zero. Therefore, the incentive premium of columns(2) is solely driven by labor market incentives. The incentive premium estimated at 0.4202 without controlling for total compensation reflects the notion that total compensation is higher in larger firms, which is the channel proposed by Edmans et al. (2009). Columns (4) and (5) show that when there are more (less) job offers, both the pay-growth and incentive premiums are higher (lower). These exercises illuminate that it is indeed the poaching offers that drive the predicted premiums.

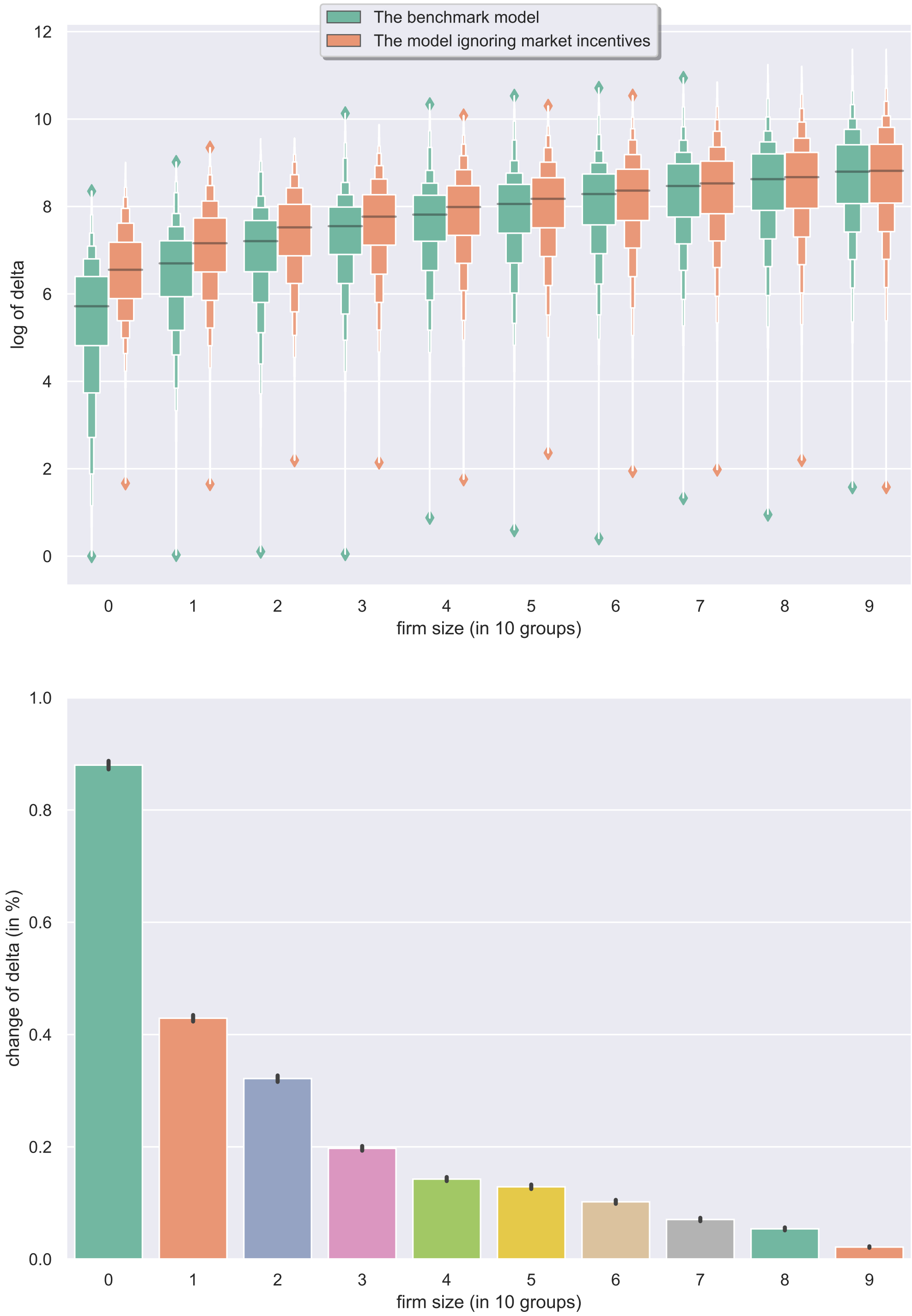

Figure 8: Fraction of market incentives is higher in smaller firms

6.6 Decomposition

To further evaluate the contribution of labor market incentives, in the data generated by a model where labor market incentives are ignored (column (3) in table 8), I cut the firm size into ten groups. The upper panel of figure 8 shows the box plots of $\log(delta)$ across ten firm size groups. Clearly, smaller firms are likely to suffer more by ignoring labor market incentives, in consistent with the job ladder mechanism. Indeed, firms that are lower on the job ladder benefit more from executives' concerns of climbing the ladder. I further calculate the ratio of *delta* with and without labor market incentives in the lower panel of figure 8. The fraction of market incentives is very high for the smallest firm group: The *delta* will be 80% higher when the job ladder is absent. The fraction quickly decreases to around 15% in the medium-size firms, and almost vanishes for top-size firms.

7 Understanding the Long-run Trends in Executive Compensation

Based on the structural estimation, I use a counterfactual exercise to quantitatively explain the sharp increases in executive total pay and performance-based incentives, more inequality across executives, and a higher correlation between executive compensation and firm size since the mid-1970s, as documented by Frydman and Saks (2010). In table 9, I select two representative periods 1970 - 1979 and 1990 - 1999 and replicate the data moments from Frydman and Saks (2010). The average total compensation rises from $1,090,000$ dollars before 1979 to $4,350,000$ dollars after 1990, and the average performance-based incentives increase almost six-fold from the 1970s to the 1990s. The interquartile range of third and first quartiles increases from $670,000$ dollars to $3,080,000$ dollars. While firm size is closely related to executive pay in the data after 1992, it was weaker in previous decades. The coefficient increases from 0.199 to 0.264 from the 1970s to the 1990s.

All these changes since the 1970s can be accounted for in my model by an exogenous change of the external executive labor market, measured by the job arrival rate λ_1. While the model does not provide an endogenous mechanism for the increase in λ_1, there is abundant evidence of more active executive labor markets since the mid-1970s. Murphy and Zabojnik (2007) document that an increasing number of CEO openings have been filled through external hires. Huson et al. (2001) document that the fraction of outsider CEOs increased from 15.3% in the 1970s to 30.0% at the beginning of the 1990s. One explanation for the trend is that executive jobs have increasingly placed greater emphasis on general rather than firm-specific skills (Frydman 2005). This is also the view taken by this paper. The executive productivity in the model is "general" and "transferable" between firms.

For the exercise, I calibrate λ_1 to be 5% for 1970 - 1979 and 40% for 1990 - 1999. These values are chosen to match the data moments under the constraint that all other parameters are equal to the estimated values in the structural estimation. Since most firms in the sample of Frydman and Saks (2010) are within the top 500, I keep the largest 500 firms in the simulated data as well. The moments calculated by model-simulated datasets are reported in the last two columns of table 9.

Table 9: Long-run trend in executive compensation

Moments	Data		Model	
(dollar value in year 2000)	1970 - 1979	1990 - 1999	$\lambda_1 = 0.05$	$\lambda_1 = 0.4$
Mean tdc1 (thousand)	1090	4350	985	4296
Mean size (million)	-	-	2426	5710
Mean delta (thousand)	21.743	120.342	24.972	125.310
$\beta_{tdc1-size}$	0.199	0.264	0.175	0.240
Percentiles of tdc1 (thousand)				
25th percentile	640	1350	109	1217
50th percentile	930	2360	478	2957
75th percentile	1310	4430	1596	5860

The results are consistent with the model intuition and are quantitatively matched with the counterparts in data. As λ_1 increases, executives are more likely to use poaching offers to renegotiate contracts, which leads to higher total compensation $tdc1$ (from $985,000$ dollars to $4,296,000$ dollars) and higher incentives $delta$ (from $24,972$ dollars to $125,310$ dollars for a 1%increase in fir's rate of return). Moreover, as firms bid for executives, the correlation between pay and firm size becomes larger (from 0.175 to 0.240). Finally, since the executive labor market in the model is search-frictional, inequality is amplified with more poaching offers: Lucky executives receive many poaching offers, while unlucky ones get few job-hopping opportunities.[27]

My model also entails predictions for moments that are not disclosed in Frydman and Saks (2010). A more active labor market also induces a larger average firm size. The mean of firm size doubles as λ_1 increases (from 2,425 million to 5,710 million). The predictions for firm-size pay-growth and incentive premiums (not shown in table 9) follow a similar pattern as in the last two columns of table 8. These predictions require further examination in the data.

[27] While the simulated moments are mostly very close to the data, there are some exceptions. In particular, the model generates much lower *tdc1* in the first two percentiles when $\lambda_1 = 0.05$. This may indicate that the poaching offer distributions of the 1970s and 1990s are different. Thus, separate estimations are required for different periods.

8 The Spillover Effect and Policy Implications

In this section, I discuss the spillover effect of firms' willingness to bid for executives using comparative statics. The parameter α_0 in the production function of the model represents the firm's (or the board's) willingness to pay for executives. The "spillover" refers to the effect where higher bids from some firms not only raises executive pay in those firms but also increases pay in all firms that are higher on the job ladder. This is because executives who are higher on the job ladder can make use of these bids to negotiate with their present firms. Consequently, the renegotiation leads to higher pay and higher performance-based incentives.

From the perspective of a regulator, executive pay is an essential part of corporate governance and is often determined by a company's board of directors. When compensation is inefficient, it is usually a symptom of an underlying governance problem brought on by conflicted boards and dispersed shareholders. For this reason, I assume that α_0 is negatively correlated with the quality of corporate governance. For example, an entrenched executive tends to have higher bargaining power and face a higher α_0, while a more independent board may impose a lower α_0 on executives. A caveat of this assumption must be emphasized: It is by no means that α_0 should always be negatively correlated with the quality of governance. This assumption should be valid only in the range where α_0 is too big.

Quantitatively, I use counterfactuals of higher α_0 values in firms of different size to evaluate how sizeable such spillover effect can be. I consider two counterfactual scenarios. In one scenario, α_0 doubles for firms that are smaller than the size median, the "small/medium firms". I denote this higher bid of small/medium firms as "worse governance in small firms". And it supposes to create a spillover effect on the pay of large firms. To compare this spillover effect, I use the second counterfactual that α_0 doubles for firms that are larger than the median. And this case is denoted as "worse governance in large firms". Figure 9 plots the distribution of delta (in the upper panel) and total compensation (in the lower panel) across ten equally divided firm size groups. There are three box plots for each size group, i.e., worse governance in small firms, the benchmark model and worse governance in large firm,s, and the medians are marked as a horizontal line in the middle of each box.

Not surprisingly, the boosts in bids increase total compensation and incentives in each separate type of firms. A higher bidding willingness in small/medium firms (in green) raises pay and incentives in firms of the first five groups, while a higher α_0 in large firms (in blue) increases pay and incentives in the largest five groups. Importantly, the rise in α_0 in small/medium firms spillovers to large firms as well. In terms of median, this spillover effect is as large as the effect of higher bids from large firms themselves. As shown in figure 9, there are 40% to 50% increases in pay and incentives of the largest two groups of firms, in both cases of higher bids from small/medium firms and higher

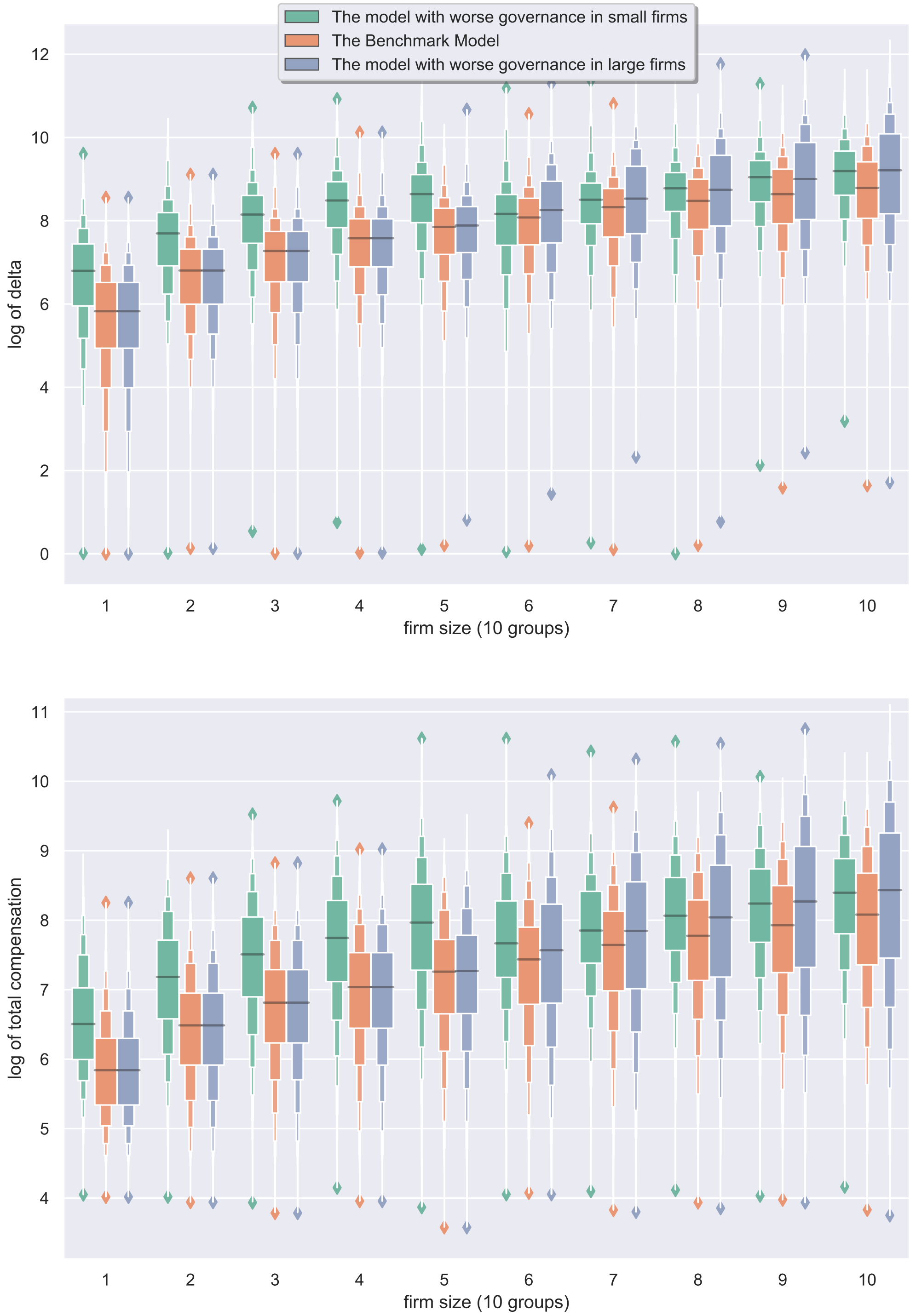

Figure 9: Compare higher bids from small/medium firms and from large firms

bids from large firms.

The policy implications of this exercise are as follow. To regulate the compensation of highly paid executives, rather than only focusing on large firms, it is important to lower the bids in small/ medium firms. Thus, large firms will face less competitive pressure. As for particular regulation policies, reforms that have been proposed or implemented including more independent compensation committee, greater mandatory pay (or pay ratio) disclosure, say-on-pay legislation should work in small and medium firms as well.

9 Conclusions

This paper has studied the impact of labor market competition on managerial incentive contracts. I developed a dynamic contracting model where executives use poaching offers to renegotiate with the current firm, and showed that poaching offers have both a level and an incentive effect on compensation. The model explains the firm-size pay-growth premium and incentive premium. Empirical evidence from a new dataset on job turnovers supports the job ladder mechanism.

I structurally estimated the model without explicitly targeting firm-size pay-growth and incentive premiums, yet the predicted premiums of the model match up very closely with the estimates in the data. A counterfactual analysis based on the structural estimation showed that with an exogenous increase of poaching offer arrival rate, my model can account for the sharp increase in total pay, performance-based incentives, and the correlation between firm size and pay levels since the mid-1970s.

Quantitative analysis showed that there is a spillover effect from the deterioration of corporate governance in small and medium firms to the compensation growth of the overall executive labor market. The policy implication is that to regulate the compensation of highly paid executives especially in large firms, it is important to improve the corporate governance of small and medium firms and reduce their bids. This will lower the competitive pressure faced by board members of large firms.

Appendix A. Model appendices

Proof for proposition 3

I start with a lemma showing that $\mathcal{I}\Big(\overline{W}(z',s)\Big)$ is a weighted sum of $\frac{\Delta\overline{W}(z',s)}{\Delta z'}$ over the domain of z'. And then show $\frac{\Delta\overline{W}(z',s)}{\Delta z'}$ decreases in s.

Step 1: Showing that $\mathcal{I}\Big(\overline{W}(z',s)\Big)$ is a weighted sum of $\frac{\Delta\overline{W}(z',s)}{\Delta z'}$

lemma 1. *Consider a productivity set $\mathbb{Z} = \{z^{(1)}, z^{(2)}, ..., z^{(n_z)}\}$. Suppose there is a distribution of productivity when the executive takes the effort Γ, a distribution when the executive shirks Γ^s, a likelihood ratio $g = \Gamma/\Gamma^s$ and a value function W. All functions are defined on $\mathbb{Z}$, then the incentive the executive receives from W is*

$$\mathcal{I}\Big(W(z)\Big) = \sum_{i=1}^{n_z-1} \omega_i \frac{\Delta W(z^{(i)})}{\Delta z^{(i)}},$$

where $\Delta z^{(i)} = z^{(i+1)} - z^{(i)}$ and $\omega_i \geq 0$.

Proof. Without lose of generality, I assume $g(z) \geq 1$ for $z \in \{z^{(1)}, z^{(2)}, ..., z^{(m)}\}$ and $g(z) < 1$ for $z \in \{z^{(m+1)}, ..., z^{(n_z)}\}$ where $m < n_z$, and define $\gamma(z) \equiv |1 - g(z)| \times \Gamma(z)$. I further denote $W(z^{(i)})$ by W_i and $\gamma(z^{(i)})$ by γ_i. Moreover, $\sum_{z\in\mathbb{Z}}(1-g(z))\Gamma(z) = 0$ implies that

$$\gamma_1 + ... + \gamma_m - \gamma_{m+1} - ... - \gamma_{n_z-1} - \gamma_{n_z} = 0. \tag{11}$$

It follows that

$$\begin{aligned}
\mathcal{I}\Big(W\Big) =& \sum_{z\in\mathbb{Z}} \Big(W(z)(1-g(z))\Gamma(z)\Big) \\
=& -\gamma_1 W_1 - \gamma_2 W_2 - ... - \gamma W_m + \gamma_{m+1} W_{m+1} + \gamma_{n_z} W_{n_z} \\
=& \gamma_1(W_2 - W_1) + (\gamma_1 + \gamma_2)(W_3 - W_2) + ... \\
&+ (\gamma_1 + ... + \gamma_m)(W_{m+1} - W_m) + (\gamma_1 + ... + \gamma_m - \gamma_{m+1})(W_{m+2} - W_{m+1}) + ... \\
&+ (\gamma_1 + ... + \gamma_m - \gamma_{m+1} - ... - \gamma_{n_z-1})(W_{n_z} - W_{n_z-1}) \\
&+ (\gamma_1 + ... + \gamma_m - \gamma_{m+1} - ... - \gamma_{n_z-1} - \gamma_{n_z})W_{n_z} \\
=& \gamma_1 \Delta z_1 \frac{W_2 - W_1}{\Delta z_1} + (\gamma_1 + \gamma_2)\Delta z_2 \frac{(W_3 - W_2)}{\Delta z_2} + ... \\
&+ (\gamma_1 + ... + \gamma_m)\Delta z_m \frac{(W_{m+1} - W_m)}{\Delta z_m} \\
&+ (\gamma_1 + ... + \gamma_m - \gamma_{m+1})\Delta z_{m+1} \frac{W_{m+2} - W_{m+1}}{\Delta z_{m+1}} + ... \\
&+ (\gamma_1 + ... + \gamma_m - \gamma_{m+1} - ... - \gamma_{n_z-1} - \gamma_{n_z-1})\Delta z_{n_z-1} \frac{W_{n_z} - W_{n_z-1}}{\Delta z_{n_z-1}} \\
=& \omega_1 \frac{W_2 - W_1}{\Delta z_1} + \omega_2 \frac{(W_3 - W_2)}{\Delta z_2} + ... \\
&+ \omega_m \frac{(W_{m+1} - W_m)}{\Delta z_m} + \omega_{m+1} \frac{W_{m+2} - W_{m+1}}{\Delta z_{m+1}} + ... + \omega_{n_z-1} \frac{W_{n_z} - W_{n_z-1}}{\Delta z_{n_z-1}} \\
=& \sum_{i=1}^{n_z-1} \omega_i \frac{\Delta W(z^{(i)})}{\Delta z^{(i)}}.
\end{aligned}$$

The first equality follows from the definition of the incentive operator $\mathcal{I}$, the rest steps are simple algebraic transformations, where I have applied condition (11). By construction, ω_i is positive.

□

Step 2: Expressing $\frac{\Delta \overline{W}(z,s)}{\Delta z}$ in terms of s.

Given lemma 1, it is sufficient to show that $\frac{\Delta \overline{W}(z,s)}{\Delta z}$ decreases in s for all $z \in \mathbb{Z}$. Notice that

$$\frac{\Delta \overline{W}(z,s)}{\Delta z} = -\frac{\Delta \Pi(z,s,\overline{W})/\Delta z}{\Delta \Pi(z,s,\overline{W})/\Delta \overline{W}} = u'(\overline{w}(s))\frac{\Delta \Pi(z,s,\overline{W})}{\Delta z}, \tag{12}$$

where $\overline{w}(z,s)$ is the compensation corresponding to $\overline{W}(z,s)$ and satisfies (1).

To derive $\overline{w}$, suppose the effort cost is

$$c = \bar{c}(s) \equiv \tilde{\beta} \sum_{z' \in \mathbb{Z}} \overline{W}(z',s)(1 - g(z'|z))\Gamma(z'|z),$$

such that the optimal contract indicates that the promised value equals to the bidding frontier

$$W(z',\tilde{s}) = \overline{W}(z',s).$$

Under the optimal contract, the continuation value (profit) of the firm is zero.

According to the Bellman equation of the firm,

$$\begin{aligned}
\Pi(z,s,\overline{W}(z,s)) &= \sum_{z' \in \mathbb{Z}} \Big(\alpha_0 s^{\alpha_1} z' - \overline{w} + \tilde{\beta} \int_{\tilde{s}} \Pi(z',s,W(z',\tilde{s}) d\tilde{F}(\tilde{s})\Big)\Gamma(z'|z) \\
&= \sum_{z' \in \mathbb{Z}} \Big(\alpha_0 s^{\alpha_1} - \overline{w} + \tilde{\beta} \int_{\tilde{s}} \Pi(z',s,\overline{W}(z',s) d\tilde{F}(\tilde{s})\Big)\Gamma(z'|z) \\
&= \sum_{z' \in \mathbb{Z}} \Big(\alpha_0 s^{\alpha_1} - \overline{w}\Big)\Gamma(z'|z) = 0.
\end{aligned}$$

Therefore,

$$\overline{w}(z,s) = \alpha_0 s^{\alpha_1} \sum_{z' \in \mathbb{Z}} z' \Gamma(z'|z)$$

To derive $\frac{\Delta \Pi(z,s,\overline{W})}{\Delta z}$, I use envelop theorem. It follows that

$$\begin{aligned}
\frac{\Delta \Pi(z,s,\overline{W})}{\Delta z} = &\sum_{z' \in \mathbb{Z}} \Big(\alpha_0 s^{\alpha_1} z' + \tilde{\beta} \int_{\tilde{s} \leq s} \Pi(z',s,\overline{W}(z',s) d\tilde{F}(\tilde{s})\Big)\frac{\Delta \Gamma(z'|z)}{\Delta z} \\
&+ \lambda \tilde{\beta} \sum_{z' \in \mathbb{Z}} \Big(\int_{\tilde{s}} \overline{W}(z',s) d\tilde{F}(\tilde{s})\Big)\frac{\Delta \Gamma(z'|z)}{\Delta z} \\
&+ \mu \tilde{\beta} \sum_{z' \in \mathbb{Z}} \Big(\int_{\tilde{s}} \overline{W}(z',s) d\tilde{F}(\tilde{s})\Big)\frac{\Delta\Big((1 - g(z'|z)\Gamma(z'|z)\Big)}{\Delta z} \\
= &\alpha_0 s^{\alpha_1} \sum_{z' \in \mathbb{Z}} z' \frac{\Delta \Gamma(z'|z)}{\Delta z} + \tilde{\beta} \sum_{z' \in \mathbb{Z}} \int_{\tilde{s}} \overline{W}(z',s) d\tilde{F}(\tilde{s}) \Big(\lambda \frac{\Delta \Gamma(z'|z)}{\Delta z} + \mu \frac{\Delta\Big((1 - g(z'|z)\Gamma(z'|z)\Big)}{\Delta z}\Big).
\end{aligned} \tag{13}$$

Divide both sides by $\alpha_0 \sum_{z'\in\mathbb{Z}} z' \frac{\Delta\Gamma(z'|z)}{\Delta z}$,

$$\frac{\frac{\Delta\Pi(z,s,\overline{W})}{\Delta z}}{\alpha_0 \sum_{z'\in\mathbb{Z}} z' \frac{\Delta\Gamma(z'|z)}{\Delta z}} = s^{\alpha_1} + \frac{\tilde{\beta} \sum_{z'\in\mathbb{Z}} \int_{\tilde{s}} \overline{W}(z',s) d\tilde{F}(\tilde{s}) \left(\lambda \frac{\Delta\Gamma(z'|z)}{\Delta z} + \mu \frac{\Delta\left((1-g(z'|z)\Gamma(z'|z) \right)}{\Delta z} \right)}{\Delta z} / \alpha_0 \sum_{z'\in\mathbb{Z}} z' \frac{\Delta\Gamma(z'|z)}{\Delta z}$$
$$= s^{\alpha_1} + \psi(s), \tag{14}$$

where $\psi(s) \equiv \frac{\tilde{\beta} \sum_{z'\in\mathbb{Z}} \int_{\tilde{s}} \overline{W}(z',s) d\tilde{F}(\tilde{s}) \left(\lambda \frac{\Delta\Gamma(z'|z)}{\Delta z} + \mu \frac{\Delta\left((1-g(z'|z)\Gamma(z'|z) \right)}{\Delta z} \right)}{\Delta z} / \alpha \sum_{z'\in\mathbb{Z}} z' \frac{\Delta\Gamma(z'|z)}{\Delta z}$.

Since all items of $\psi(s)$ are positive, $\psi(s) > 0$. Because $\psi(s)$ only depends on s via $\overline{W}$ which is increasing in s, $\psi(s)$ is also increasing in s.

Insert (13) and (14) into (12), we have

$$\frac{\Delta\overline{W}(z,s)}{\Delta z} = u'(\overline{w}(s)) \frac{\Delta\Pi(z,s,\overline{W})}{\Delta z} = u'\left(\alpha_0 s^{\alpha_1} \sum_{z'\in\mathbb{Z}} z'\Gamma(z'|z) \right) \left(s^{\alpha_1} + \psi(s) \right) \alpha_0 \sum_{z'\in\mathbb{Z}} z' \frac{\Delta\Gamma(z'|z)}{\Delta z}. \tag{15}$$

Step 3: Showing that $\frac{\Delta\overline{W}(z,s)}{\Delta z}$ decreases in s under the stated condition.

To have

$$\lim_{\Delta s\to 0} \frac{\Delta\overline{W}(z,s+\Delta s)}{\Delta z} - \frac{\Delta\overline{W}(z,s)}{\Delta z} > 0,$$

using (15)

$$\frac{u'\left((s+\Delta s)^{\alpha_1} \alpha_0 \sum_{z'\in\mathbb{Z}} z'\Gamma(z'|z) \right)}{u'\left(s^{\alpha_1} \alpha_0 \sum_{z'\in\mathbb{Z}} z'\Gamma(z'|z) \right)} < \frac{s^{\alpha_1} + \psi(s)}{(s+\Delta s)^{\alpha_1} + \psi(s+\Delta s)}.$$

Applying $u'(w) = w^{-\sigma}$, we have

$$\left(\frac{s}{s+\Delta s} \right)^{-\alpha_1\sigma} < \frac{s^{\alpha_1} + \psi(s)}{(s+\Delta s)^{\alpha_1} + \psi(s+\Delta s)},$$

or

$$\sigma > \frac{\log \frac{s^{\alpha_1} + \psi(s)}{(s+\Delta s)^{\alpha_1} + \psi(s+\Delta s)}}{\frac{s}{s+\Delta s}}.$$

Take $\Delta s \to 0$ using L'Hopital's rule,

$$\sigma > 1 + \frac{s^{1-\alpha_1}}{\alpha_1} \psi'(s).$$

Appendix B. Empirical appendices

This appendix contains some extra regression results on firm-size incentive premium. Figure 10 is a heatmap of performance-based incentives *log(delta)* on total compensation and firm size. It shows that among executives with similar total compensation, those in larger firms get higher performance-based incentives.

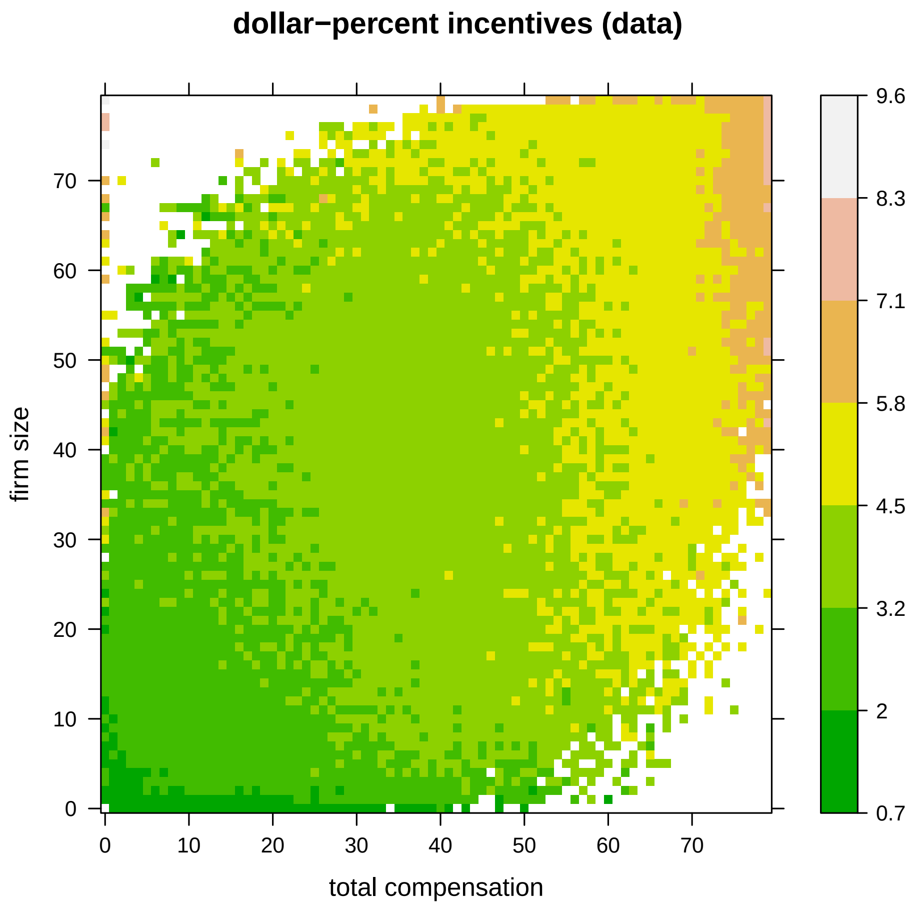

Figure 10: $\log(delta)$ over firm size and total compensation

Note: $delta$ is the wealth-performance sensitivity defined as the dollar change in firm-related wealth for a percentage change in firm value. The total compensation is the sum of salary and bonus, the value of restricted stocks and options granted, and the value of retirement and long-term compensation schemes. The firm size is the market capitalization by the end of the fisical year, calculated by `csho` $\times$ `prcc_f` where `csho` is the common shares outstanding and `prcc_f` is the close price by fiscal year. I divide the whole sample into 80×80 cells according to the total compensation and firm size, and compute the mean of $\log(delta)$ within each cell.

Table 10 shows more robustness check on firm-size incentive premium. Table 11 contains the full results on the interaction of firm size and proxies of labor market competition. Table 12 contains the full result of the size incentive premium for each age.

Table 10: Performance-based incentives increase with firm size

	log($delta$)				
	(1)	(2)	(3)	(4)	(5)
log(firm size)	0.585***	0.360***	0.331***	0.330***	0.440***
	(0.0141)	(0.0247)	(0.0237)	(0.0236)	(0.0236)
log(tdc1)		0.609***			0.334***
		(0.0350)			(0.0323)
tdc1 Dummies (50)			Yes		
tdc1 Dummies (100)				Yes	
Other contorls					Yes
tenure dummies	Yes	Yes	Yes	Yes	Yes
age dummies	Yes	Yes	Yes	Yes	Yes
year dummies	Yes	Yes	Yes	Yes	Yes
industry dummies	Yes	Yes	Yes	Yes	Yes
year × industry dummies	Yes	Yes	Yes	Yes	Yes
Observations	146,747	128,006	128,006	128,006	109,730
adj. R^2	0.442	0.514	0.523	0.524	0.595

Note: This table reports evidence on firm size premium in performance-based incentives. The dependent variable is the log of *delta* where *delta* is the dollar change in firm related wealth for a percentage change in firm value. Firm size is measured by the market capitalization defined by the common shares outstanding times the fiscal year close price. *tdc1* is the total compensation, including the sum of salary and bonus, the value of restricted stocks and options granted, and value of retirement and long-term compensation schemes. It is the variable *tdc1* in ExecuComp dataset. In all regressions, I have controlled for age dummies, executive tenure dummies, year × industry dummies. Column (1) is a regression of *log(delta)* on *log(firm size)*, which replicates the cross-sectional regression in the literature. From column (2) to column (4), I add *log(tdc1)*, *tdc1 dummies 50* and *tdc1 dummies 100* (tdc1 values are evenly divided into 50 or 100 groups and then transformed into dummies), respectively. In column (5), I add other controls including *operating profitability*, *market-book ratio*, *annualized stock return*, *director*, *CEO* and *CFO*, *interlock*. Standard errors clustered at the firm × fiscal year level are shown in parentheses, and I denote symbols of significance by * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$.

Table 11: Size incentive premium increases with managerial labor market competition

	log($delta$)			
	(1)	(2)	(3)	(4)
log(firm size)	0.525^{***}	0.529^{***}	0.561^{***}	0.571^{***}
	(0.00512)	(0.00499)	(0.00310)	(0.0139)
EE190	1.919^{*}			
	(0.776)			
log(firm size) × *EE190*	0.415^{***}			
	(0.101)			
EE90		2.611^{**}		
		(0.903)		
log(firm size) × *EE90*		0.359^{**}		
		(0.118)		
gai			-1.211^{***}	
			(0.0941)	
log(firm size) × *gai*			0.0648^{***}	
			(0.0118)	
inside CEO				-0.00566^{***}
				(0.00156)
log(firm size) × *inside CEO*				-0.000458^{*}
				(0.000202)
Controls	Yes	Yes	Yes	Yes
Observations	125858	125858	75747	125858
adj. R-sq	0.521	0.521	0.531	0.521

Note: This table reports evidence that the firm size incentive premium increases as the managerial labor market competition becomes thicker. The dependent variable is the log of *delta* where *delta* is the dollar change in firm related wealth for a percentage change in firm value. The independent variables include the log of firm size, several variables that measure the how active the competition in managerial labor markets, and the interaction terms between firm size and labor market competition. In column (1), labor market competition is measured by job-to-job transition rate in each (Fama-French 48) industries and fiscal years. A job-to-job transition is defined as an executive leaving the current firm and starting to work in another firm within *190* days. The same measure is used in column (2) except the gap between jobs is changed to *90* days. Column (3) measures labor market activeness by the average of the general ability index at the industry-year level. The original index is provided by Custódio et al. (2013). Column (4) uses the industry level percentage of new CEOs who are promoted inside the company. The data is provided by Martijn Cremers and Grinstein (2013). The control variables include executive tenure dummies, age dummies, fiscal year dummies, *operating profitability*, *market-book ratio*, *annualized stock return*, whether the executive served as a director, CEO or CFO during the fiscal year, whether the executive is involved in the interlock relationship. For regression including *inside CEO*, I use data from year 1992 to year 2006. For the rest, I use data from year 1992 to year 2015. Standard errors clustered at the firm times fiscal year level are shown in parentheses, and I denote symbols of significance by * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$.

Table 12: Size incentive premium decreases with executive age

	log($delta$)				
	(1)	(2)	(3)	(4)	(5)
age35 × log(firm size)	0.849***	0.652***	0.580***	0.541***	0.539***
	(0.0534)	(0.0649)	(0.0620)	(0.0614)	(0.0611)
age36 × log(firm size)	0.753***	0.530***	0.516***	0.484***	0.481***
	(0.0487)	(0.0658)	(0.0519)	(0.0538)	(0.0529)
age37 × log(firm size)	0.746***	0.543***	0.540***	0.508***	0.506***
	(0.0366)	(0.0440)	(0.0359)	(0.0365)	(0.0365)
age38 × log(firm size)	0.689***	0.471***	0.471***	0.438***	0.436***
	(0.0328)	(0.0390)	(0.0339)	(0.0341)	(0.0332)
age39 × log(firm size)	0.667***	0.444***	0.443***	0.410***	0.410***
	(0.0276)	(0.0365)	(0.0297)	(0.0297)	(0.0296)
age40 × log(firm size)	0.664***	0.475***	0.493***	0.462***	0.461***
	(0.0296)	(0.0358)	(0.0319)	(0.0341)	(0.0338)
age41 × log(firm size)	0.653***	0.449***	0.489***	0.459***	0.457***
	(0.0264)	(0.0350)	(0.0320)	(0.0325)	(0.0324)
age42 × log(firm size)	0.640***	0.437***	0.478***	0.453***	0.450***
	(0.0285)	(0.0342)	(0.0330)	(0.0338)	(0.0335)
age43 × log(firm size)	0.630***	0.408***	0.469***	0.445***	0.444***
	(0.0248)	(0.0333)	(0.0303)	(0.0311)	(0.0309)
age44 × log(firm size)	0.622***	0.405***	0.473***	0.449***	0.447***
	(0.0230)	(0.0345)	(0.0313)	(0.0314)	(0.0314)
age45 × log(firm size)	0.608***	0.397***	0.468***	0.447***	0.446***
	(0.0220)	(0.0287)	(0.0280)	(0.0280)	(0.0278)
age46 × log(firm size)	0.592***	0.377***	0.443***	0.424***	0.422***
	(0.0210)	(0.0293)	(0.0283)	(0.0286)	(0.0284)
age47 × log(firm size)	0.594***	0.365***	0.445***	0.428***	0.426***
	(0.0207)	(0.0297)	(0.0289)	(0.0296)	(0.0295)
age48 × log(firm size)	0.598***	0.367***	0.454***	0.435***	0.434***
	(0.0163)	(0.0259)	(0.0252)	(0.0256)	(0.0257)
age49 × log(firm size)	0.594***	0.369***	0.437***	0.419***	0.417***
	(0.0180)	(0.0264)	(0.0284)	(0.0276)	(0.0277)
age50 × log(firm size)	0.589***	0.388***	0.457***	0.439***	0.438***
	(0.0210)	(0.0287)	(0.0301)	(0.0316)	(0.0317)
age51 × log(firm size)	0.563***	0.352***	0.426***	0.410***	0.409***
	(0.0173)	(0.0254)	(0.0270)	(0.0273)	(0.0275)
age52 × log(firm size)	0.560***	0.342***	0.414***	0.399***	0.398***
	(0.0191)	(0.0268)	(0.0280)	(0.0282)	(0.0281)
age53 × log(firm size)	0.577***	0.350***	0.425***	0.409***	0.408***
	(0.0192)	(0.0274)	(0.0278)	(0.0286)	(0.0287)
age54 × log(firm size)	0.570***	0.335***	0.423***	0.409***	0.409***
	(0.0209)	(0.0288)	(0.0286)	(0.0290)	(0.0293)
age55 × log(firm size)	0.569***	0.351***	0.435***	0.423***	0.423***
	(0.0184)	(0.0279)	(0.0271)	(0.0273)	(0.0273)

Table 12: Size incentive premium decreases with executive age (continue)

	(1)	(2)	(3)	(4)	(5)
age56 × log(firm size)	0.592***	0.362***	0.454***	0.442***	0.441***
	(0.0157)	(0.0260)	(0.0271)	(0.0272)	(0.0270)
age57 × log(firm size)	0.593***	0.356***	0.440***	0.429***	0.428***
	(0.0141)	(0.0233)	(0.0237)	(0.0232)	(0.0230)
age58 × log(firm size)	0.592***	0.356***	0.442***	0.430***	0.429***
	(0.0175)	(0.0266)	(0.0261)	(0.0260)	(0.0261)
age59 × log(firm size)	0.593***	0.351***	0.435***	0.423***	0.422***
	(0.0172)	(0.0256)	(0.0258)	(0.0253)	(0.0254)
age60 × log(firm size)	0.579***	0.341***	0.424***	0.412***	0.412***
	(0.0175)	(0.0271)	(0.0259)	(0.0258)	(0.0259)
age61 × log(firm size)	0.600***	0.355***	0.438***	0.428***	0.427***
	(0.0216)	(0.0307)	(0.0302)	(0.0311)	(0.0309)
age62 × log(firm size)	0.587***	0.333***	0.420***	0.409***	0.408***
	(0.0192)	(0.0282)	(0.0268)	(0.0272)	(0.0272)
age63 × log(firm size)	0.605***	0.358***	0.448***	0.436***	0.435***
	(0.0196)	(0.0252)	(0.0256)	(0.0253)	(0.0255)
age64 × log(firm size)	0.593***	0.356***	0.440***	0.429***	0.429***
	(0.0242)	(0.0285)	(0.0296)	(0.0292)	(0.0289)
age65 × log(firm size)	0.596***	0.353***	0.435***	0.423***	0.423***
	(0.0246)	(0.0318)	(0.0339)	(0.0334)	(0.0332)
logtdc1		0.611***	0.345***		
		(0.0352)	(0.0339)		
tdc1 Dummies (50)				Yes	
tdc1 Dummies (100)					Yes
profit			0.619***	0.598***	0.602***
			(0.117)	(0.116)	(0.116)
annual return			0.102*	0.0999	0.0998
			(0.0488)	(0.0485)	(0.0485)
mbr			0.116***	0.120***	0.120***
			(0.0209)	(0.0213)	(0.0213)
director			0.754***	0.739***	0.737***
			(0.0326)	(0.0307)	(0.0306)
interlock			0.517***	0.529***	0.527***
			(0.0953)	(0.0948)	(0.0947)
CEO			0.593***	0.576***	0.574***
			(0.0387)	(0.0395)	(0.0397)
CFO			0.0837***	0.0711***	0.0711***
			(0.0130)	(0.0131)	(0.0130)
N	146750	128008	109732	109732	109732
adj. R^2	0.432	0.506	0.586	0.590	0.590

Note: This table reports the evidence that firm size incentive premium decreases in executive age. The dependent variable is the log of *delta* where *delta* is the dollar change in firm related wealth for a percentage change in firm value. The key independent variable is the log of firm size where firm size is measured by the market capitalization defined by the common shares outstanding times the fiscal year close price. I allow a different coefficients of firm size across ages from 35 to 65. Control variables include total compensation (*tdc1*), age dummies, executive tenure dummies, year times industry dummies, *profit*, the operating profitability, *mbr*, the market-book ratio, *annual return*, the annualized stock return, *director*, whether the executive served as a director during the fiscal year, *CEO* and *CFO*, whether the executive served as a CEO (and CFO) during the fiscal year, *interlock*, whether the executive is involved in the interlock relationship. Standard error (clustered at the firm $\times$ fiscal year level) are shown in parentheses, and I denote symbols of significance by * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$.

Appendix C. Estimation appendices

Recursive multiplier method

To further characterize the optimal solution, I resort to the tools developed by Marcet and Marimon (2017, hereafter MM).[28] In dynamic contracting problems with forward looking constraints such as the IC constraint here, the solution does not satisfy the Bellman equation. MM suggest to study a recursive Lagrangian. Under standard general conditions there is a recursive saddle-point functional equation (analogous to a Bellman equation) that characterizes a recursive solution to the planner's problem. The recursive formulation is obtained after adding a co-state variable λ_t summarizing previous commitments reflected in past Lagrange multipliers. The time-consistent continuation solution is obtained by using the endogenous λ_t as the vector of weights in the objective function. I summarize this method in the following proposition.

Proposition 4 (Marcet and Marimon)**.** *Define Pareto Frontier by*

$$P(z, s, \lambda) = \sup_{W} \Pi(z, s, W) + \lambda W,$$

where Π and W are defined as in (BE-F) and (PKC), and $\lambda > 0$ is a Pareto weight assigned to the executive. Then there exist positive multipliers of $\{\mu, \mu_0(z'), \mu_1(z')\}$ that solve the following problem

$$P(z, s, \lambda) = \inf_{\mu, \mu_0(z', \tilde{s}), \mu_1(z', \tilde{s})} \sup_{w} h(z, s, \lambda, w) + \hat{\beta} \sum_{z'} P(z', s, \lambda')\Gamma(z'|z),$$

where multiplier μ corresponds to the incentive compatibility constraint, multipliers $\mu_0(z', \tilde{s}), \mu_1(z', \tilde{s})$ correspond to participation constraints,

$$h(z, s, \lambda, w) = y(s)z' - w + \lambda u(w) - (\lambda + \mu)c,$$

Pareto weight evolves according to

$$\lambda' = \lambda + \mu(1 - g(z, z')) + \mu_0(z', \tilde{s}) + \mu_1(z', \tilde{s}),$$

and

$$\hat{\beta} = \tilde{\beta}(1 - \lambda_1 \sum_{\mathcal{M}_1 \cup \mathcal{M}_2} F(s')).$$

The optimal contract $\{w, W(z', \tilde{s})\}$ follows that

$$u'(w) = \frac{1}{\lambda}, \tag{16}$$

$$W(z', \tilde{s}) = W(z', \tilde{s}, \lambda'). \tag{17}$$

Proposition 4 can be illustrated intuitively using the Pareto weight of the executive λ and the multiplier μ of the incentive constraint. Suppose the match starts with a $\lambda^{(0)}$, and assume the

[28] This approach has been used in many applications. A few examples are: growth and business cycles with possible default (Marcet and Marimon (1992), Kehoe and Perri (2002), Cooley, et al. (2004)); social insurance (Attanasio and Rios-Rull (2000)); optimal fiscal and monetary policy design with incomplete markets (Aiyagari, Marcet, Sargent and Seppala (2002), Svensson and Williams (2008)); and political-economy models (Acemoglu, Golosov and Tsyvinskii (2011)).

participation constraints are not binding so that $\mu_0 = \mu_1 = 0$. $\lambda^{(0)}$ has to satisfy $W(z_O, s, \lambda^{(0)}) = W^0$. To deal with the moral hazard, the optimal contract indicates a $\mu^{(0)} > 0$. Then depending on the realization of z', the weight of the executive will be updated to

$$\lambda^{(i)} = \lambda^{(i-1)} + \mu^{i-1}(1 - g(z, z')) \text{ for } i \text{ in } 1, 2... \tag{18}$$

The evolve of λ continues as such till the match breaks.

When there is an outside offer such that the executive moves from his or her current firm to the outside firm, the new match starts with a weight denoted by $\lambda^{(n)}$ such that

$$W(z, \tilde{s}, \lambda^{(n)}) = \overline{W}(z, s),$$

were I have denoted the current productivity by z, current firm by s, and the outside firm by $\tilde{s}$. It means the new match will assign a new weight to the executive so that he or she gets the continuation value $\overline{W}(z, s)$. Then the new Pareto weight will evolve again as illustrated in (18). In a nutshell, proposition 4 allows me to solve the optimal contract in the space of Pareto weight λ instead of in the space of the promised utility. At any moment, I can transit from the metrics of λ back to the metrics of utilities using (16) and (17).

The advantage of this method is I do not need to find the promised utilities $W(z', \tilde{s})$ in each state of the world for the next period. Instead, λ and μ are enough to trace all $W(z', \tilde{s})$. Moreover, λ corresponds to the total compensation level (wage level), while μ corresponds to how much contract incentive is provided in the optimal contract. The two multipliers are enough to understand both theoretically and numerically why keeping the same wage level (the same λ), incentive pays increase with firm size (μ increases with firm size).

References

Abrahám, Arpád, Fernando Alvarez-Parra, and Susanne Forstner (2016), "The effects of moral hazard on wage inequality in a frictional labor market."

Abreu, Dilip, David Pearce, and Ennio Stacchetti (1990), "Toward a theory of discounted repeated games with imperfect monitoring." *Econometrica: Journal of the Econometric Society*, 1041–1063.

Agrawal, Ashwini and Alexander Ljungqvist (2014), "Managerial labor market frictions and corporate investment."

Alvarez, Fernando and Urban J Jermann (2000), "Efficiency, equilibrium, and asset pricing with risk of default." *Econometrica*, 68, 775–797.

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

Böhm, Michael J, Daniel Metzger, and Per Strömberg (2015), "'since you're so rich, you must be really smart': Talent and the finance wage premium."

Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin (2006), "Wage bargaining with on-the-job search: Theory and evidence." *Econometrica*, 74, 323–364.

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

Coles, Jeffrey L, Naveen D Daniel, and Lalitha Naveen (2013), "Calculation of compensation incentives and firm-related wealth using execucomp: Data, program, and explanation."

Cornelli, Francesca, Zbigniew Kominek, and Alexander Ljungqvist (2013), "Monitoring managers: Does it matter?" *The Journal of Finance*, 68, 431–481.

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

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

Dittmann, Ingolf, Ernst Maug, and Oliver Spalt (2010), "Sticks or carrots? optimal ceo compensation when managers are loss averse." *The Journal of Finance*, 65, 2015–2050.

Dittmann, Ingolf, Ko-Chia Yu, and Dan Zhang (2017), "How important are risk-taking incentives in executive compensation?" *Review of Finance*, 21, 1805–1846.

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

Edmans, Alex, Xavier Gabaix, and Dirk Jenter (2017), "Executive compensation: A survey of theory and evidence." Technical report, National Bureau of Economic Research.

Edmans, Alex, Xavier Gabaix, and Augustin Landier (2009), "A multiplicative model of optimal ceo incentives in market equilibrium." *The Review of Financial Studies*.

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

Fama, Eugene F (1980), "Agency problems and the theory of the firm." *Journal of political economy*, 88, 288–307.

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

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

Frydman, Carola and Raven E Saks (2010), "Executive compensation: A new view from a long-term perspective, 1936–2005." *The Review of Financial Studies*, 23, 2099–2138.

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

Gautier, Pieter A and Coen N Teulings (2015), "Sorting and the output loss due to search frictions." *Journal of the European Economic Association*, 13, 1136–1166.

Gautier, Pieter A, Coen N Teulings, and Aico Van Vuuren (2010), "On-the-job search, mismatch and efficiency." *The Review of Economic Studies*, 77, 245–272.

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

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

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

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

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

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

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

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

Hermalin, Benjamin E (2005), "Trends in corporate governance." *The Journal of Finance*, 60, 2351–2384.

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

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

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

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

Jenter, Dirk and Fadi Kanaan (2015), "Ceo turnover and relative performance evaluation." *The Journal of Finance*, 70, 2155–2184.

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

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

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

Lamadon, Thibaut (2016), "Productivity shocks, long-term contracts and earnings dynamics." *manuscript, University of Chicago*.

Lentz, Rasmus (2014), "Optimal employment contracts with hidden search." Technical report, National Bureau of Economic Research.

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

Marcet, Albert and Ramon Marimon (2017), "Recursive contracts."

Margiotta, Mary M and Robert A Miller (2000), "Managerial compensation and the cost of moral hazard." *International Economic Review*, 41, 669–719.

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

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

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

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

Murphy, Kevin J and Jan Zabojnik (2007), "Managerial capital and the market for ceos."

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

Peters, Florian S and Alexander F Wagner (2014), "The executive turnover risk premium." *The Journal of Finance*, 69, 1529–1563.

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

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

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

Remesal, Alvaro et al. (2018), "How important are dismissals in ceo incentives? evidence from a dynamic agency model."

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

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

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

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

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

Teulings, Coen N and Pieter A Gautier (2004), "The right man for the job." *The Review of Economic Studies*, 71, 553–580.

Thomas, Jonathan and Tim Worrall (1988), "Self-enforcing wage contracts." *The Review of Economic Studies*, 55, 541–554.

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

Townsend, Robert M (1982), "Optimal multiperiod contracts and the gain from enduring relationships under private information." *Journal of political Economy*, 90, 1166–1186.

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

Wang, Cheng and Youzhi Yang (2011), "Termination of dynamic contracts in an equilibrium labor market model." *Journal of Economic Theory*, 146, 74.

Wang, Cheng and Youzhi Yang (2015a), "Equilibrium matching and termination." *Journal of Monetary Economics*, 76, 208–229.

Wang, Cheng and Youzhi Yang (2015b), "Outside opportunities and termination." *Games and Economic Behavior*, 91, 207–228.

Wang, Cheng and Youzhi Yang (2016), "Optimal ceo contracting with moral hazard and external shocks." *Working paper*.