Managerial Labor Market Competition and Incentive Contracts

Bo Hu*

Job Market Paper

November 12, 2018

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 job turnovers for executives from U.S. publicly listed firms. The structural estimates show that the model is capable of explaining and predicting the empirical puzzles that executives of larger firms experience higher compensation growth, and receive higher performance-based incentives.

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

The latest version can be downloaded here (pdf).

*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. 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, and 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, to react to the "battle for talent", IBM targets to 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) competitors."[1] 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".[2]

Despite its relevance for the industry, a characterization of how heterogeneous firms compete for executives is still missing in the literature, and its consequence on executive contracts has 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 in the paper.[3] The model considers two types of agents: executives and firms. Executives are heterogeneous in their managerial productivity, which evolves stochastically depending on their current and past effort.[4] Firms are heterogeneous in time-invariant asset size.[5] As in Gabaix

[1]See IBM proxy statement, https://www.ibm.com/annualreport/2017/assets/downloads/IBM_Proxy_2018.pdf (visited on Oct 27, 2018).

[2]See Johnson and Johnson proxy statement, http://www.investor.jnj.com/gov/annualmeetingmaterials.cfm (visited on Oct 27, 2018).

[3]The executive job ladder is relevant 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 increases from 15.3% in the 1970s to 30.0% at the beginning of 1990s. A similar pattern is reported by Murphy and Zabojnik (2007).

[4]I consider the *general* managerial productivity rather than the firm-specific one.

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

and Landier (2008), the marginal impact of an executive's productivity increases with the value of the firm. 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 frictional labor market. By making use of poaching offers, the executive can renegotiate with the current firm, or transits to a larger poaching firm, where the compensation contract follows a sequential auction (Postel-Vinay and Robin 2002, Cahuc et al. 2006). Essentially, the current and the poaching firms are engaged in a Bertrand competition for the executive.

The competition from poaching offers impacts managerial 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 it 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 needs 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 past effort. Together, they imply that taking 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 a new source of incentives, which I call *labor market incentives* in this paper. Consequently, firms can take advantage of the labor market incentives, and give less performance-based incentives to executives, but still resolve the moral hazard issue.

These two channels enable the model to shed light on two firm-size premiums: the *firm-size pay-growth premium* and the *firm-size incentive premium*. The firm-size pay-growth premium refers to the empirical fact that starting with the same total compensation, the executive of a larger firm experiences a higher compensation growth. For a 1% increase in firm size, the total compensation-growth rate increases by 15.4%. This fact is firstly documented in this paper, and it complements the stylized facts on executive pay levels in the literature (see e.g., Edmans et al. 2017). My explanation is intuitive: executive compensation grows because of the competition from poaching offers; larger firms are more capable of countering outside offers due to the firm size effect in production, thus their executives tend to experience higher pay growth.

The firm-size incentive premium refers to the empirical fact that performance-based incentives increase with firm size, a relationship that holds after controlling for executives total compensation, firm performance, etc. In the data, a 1% increase in firm size leads to a 0.35% increase in performance-based incentives.[6] The channel that poaching offers generate labor market incentives helps to explain this premium. An executive is motivated by two sources of incentives which substitute each other: performance-based incentives and labor market incentives. I show that labor market incentives decrease with firm size. Thus, performance-based incentives increase with firm size.

There are two reasons why labor market incentives decrease with firm size. The first reason is due to the job ladder. Executives from larger firms are located "higher" on the job ladder. Consequently, the chance to receive an outside offer that beats the current value is lower. Thus, labor market incentives for executives from larger firms are smaller. 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 marginal utility of a more favorable poaching offer is smaller. As a result, labor market incentives are smaller for executives of larger firms.

To provide empirical evidence and structurally estimate the model, I assembled a new dataset on executive job turnover by merging ExecuComp and BoardEX databases. ExecuComp is the standard data source for executive compensation studies. It contains annual records on top executives' compensation of firms comprising the S&P 500, MidCap, and SmallCap indices. BoardEX contains detailed executive employment history, and it helps to identify the employment status of each executive after the spells in ExecuComp. For executives that are not identified in BoardEX, I further search for information using Linkedin and Bloomberg. In the final data sample, there are 35,088 executives and a total number 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 around 5%, which is stable over the years and across industries. Moreover, there is a job ladder on the firm size dimension: about 60% of job-to-job transitions are towards larger firms, and for other transitions, 20% of them are due to a promotion from a non-CEO title to a CEO title. 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. 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 both firm-size premiums in 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 average of the general ability index (Custódio et al. 2013).

[6] Performance-based incentives are measured by the dollar change in firm-related wealth per percentage change in firm value. 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 2 for more details.

To numerically solve for the optimal contract is difficult 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 this approach, to solve for the optimal contract, I only need to solve for one Lagrangian multiplier. 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, optimal incentive pay and promised values can be solved.

Using the simulated method of moments (SMM), I estimate the model by targeting data moments of executive compensation, incentives, and job turnovers, as well as moments 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 reassures 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 counter-factual exercise to quantitatively account for the sharp increase in executive total pay, performance-based incentives, the inequality across executives, and the correlation between executive compensation and firm size since the mid-1970s, as is documented by Frydman and Saks (2010). I show that all the above increases in data moments can be explained by an exogenous increase in the arrival rate of poaching offers. The intuition is that the managerial labor market is much thinner before the 1970s, which is also supported by the evidence in the literature (Frydman 2005, 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 more important to lower the bids for executives from small and medium firms, via various reforms such as a more independent compensation committee, greater mandatory pay (or pay ratio) disclosure, say-on-pay legislation, etc. In this way, the competitive pressure in the overall managerial labor market decreases. In the model, there is a spillover effect that higher bids from a set of firms not only boosts the executive pay in those firms but also increases the pay in all firms that are higher on the job ladder. In a comparative static analysis based on the estimated model, I show that, compare to the increase in the bids from large firms, the same increment in that 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. Section 6 estimates the model. Section 7 explains the sharp increase in executive pay since the mid-1970s. Section 8 discusses the policy implications. 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 has a similar prediction on 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, the view on the pay differentials between small and large firms is different. In this paper, executives are paid much more in larger firms, not because they are talented (e.g., Gabaix and Landier 2008), but because they are lucky to be matched with a large firm that can 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 in 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 it highlights a hierarchical job ladder. Besides the explanation given in Edmans et al. (2009), the job ladder in my model gives rise to the 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, hence higher equity incentives are needed to satisfy the incentive compatibility condition. Gayle et al. (2015) embed the model of Margiotta and Miller (2000) into a generalized Roy model and they find that the quality of the signal is unambiguously poorer in larger firms, and that explains most of the 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 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 to larger firms, whereas the driving force of my explanation is a hierarchical job ladder that executives move from small to large firms. The different modeling approach of job-to-job transitions explains why labor market incentives contribute much less in their framework. Using executives' job-to-job transition data, I show that the hierarchical job ladder over firm size does exist.

This paper also contributes to the literature in explaining the rise of executive compensation in recent decades. My paper is closest in spirit to the explanation based on executive mobility. Literature show that the increases in compensation coincides with the increased occupational mobility of executives, while the mobility 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 the job hopping opportunities can help explaining 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 explain the increase in total pay, performance-based incentives and the correlation between firm size and pay levels.

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 the 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(or) 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 risk-neutral workers 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 of their framework. In real life, it happens very often that executives are contacted and "auctioned" by competing firms for promotion.[7]

3 Motivating Facts

This paper is motivated by two firm-size premiums: the pay-growth premium and the incentive premium. I first present these premiums, then I show that both premiums are related to the managerial labor market. Using proxies for the activeness of executive labor markets, I find that both premiums are higher in industries where the managerial labor market is more active. The primary data source for the analyses in this section is Standard & Poor's ExecuComp database. The proxies for executive labor markets include the job-to-job transition rate, general ability index, and fraction of inside CEOs. These variables come from my newly assembled dataset and the datasets provided by Custódio et al. (2013) and Martijn Cremers and Grinstein (2013).[8]

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 compensation growth is measured by the first-order difference of $\log(tdc1)$ where $tdc1$ is 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. Column (1) in table 1 presents the regression of $\Delta\log(tdc1)$ on firm size, controlling for total compensation, tenure, age, and year times industry dummies. The estimates indicate that start from the same total compensation, for a 1% increase in firm size, the compensation growth rate increase by 11.2%.[9]The premium is slightly larger with an estimate of 0.154 in column (2) after further controlling for *operating profitability*, *market-book ratio*, *annualized stock return*, title dummies such as *director*, *CEO*, *CFO*, and *interlock*, etc.

To link the premium with the managerial labor market, I explore the variation across industries. An industry is an appropriate sub labor market since more than 60% job-to-job transitions are within the industry (see details in section 4). As a direct test of whether size premium is related with a more active labor market, I use four proxies to measure how active the managerial labor market is, and test if the interactions between

[7]For example, one can find related description in the book of Khurana (2004)

[8]All nominal quantities are converted into constant 2016 dollars using a measure of the price level the DGP deflater from the Bureau of Economic Analysis. Please refer to section 4 for a statistical description of data.

[9]To control for the beginning of period compensation level, I include the lagged $\log(tdc1)$. The result is robust by controlling lagged total compensation using 200 group dummies that equally cut the sample according to $\log(tdc1)$.

Table 1: Compensation 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)$_{-1}$ × *EE90*			0.0711*			
			(0.0403)			
log(firm size)$_{-1}$ × *EE190*				0.0759**		
				(0.0353)		
log(firm size)$_{-1}$ × *gai*					0.0233***	
					(0.00546)	
log(firm size)$_{-1}$ × *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 contorls		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 × industry	X	X				
Observations	129068	106819	106820	106820	58188	106820
adj. R^2	0.157	0.216	0.260	0.260	0.233	0.262

Note: This table reports evidence on firm-size premium in compensation growth. 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 value of retirement and long-term compensation schemes. The key independent variable is *log(firm size)* where firm size is measured by the market capitalization defined by the common shares outstanding times the fiscal year close price. The key control variable is the lagged total compensation. Whenever possible, I control for dummies for age, tenure, and year times industry. Other controls include *operating profitability*, *market-book ratio*, *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. The standard error (clustered at the firm × fiscal year level) are shown in parentheses, and we denote symbols of significance by * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

these proxies and firm size are significant. The first two proxies are job-to-job transition rates in each industry-year (Fama-French 48 industries and fiscal years). *EE90* defines a job-to-job transition by that an executive leaves the current firm and starts to work in another firm within *90* days. Similarly, *EE190* defines a job-to-job transition with a gap of no more than 190 days. The third proxy *gai* is the mean of general ability index of CEOs at the industry-year level. The general ability index itself is the first principal component of five proxies that measure the generality of the CEO's human capital based on the CEO's lifetime work experience.[10] The last proxy *inside CEO* is the percentage of insider CEOs in the industry in which the firm operates. It counts for all new CEOs between 1993 and 2005 using Fama-French 48-industry groups.

The picture that emerges in the last four columns is not ambiguous: All four interaction terms are statistically and economically significant, and the signs confirm that the size growth premium is larger in industry/years where the executive labor market is more active. Specifically, the coefficient of 0.0759 on the interaction with *EE190* and a standard deviation 0.0224 of *EE190*, means that a one-standard-deviation increase in job-to-job transition measured by *EE190* implies the pay-growth rate increase from 7.3% to 7.47%. Similar, a one-standard-deviation increase in job-to-job transition measured by *EE90* implies the pay-growth rate increase from 7.3% to 7.43%. The standard deviation of *gai* is 0.253, together with the coefficient of 0.0233, means that with a one-standard deviation increase in *gai* the pay-growth rate increase from 7.3% to 7.89%. The standard deviation of *inside CEO* is 0.122, together with the coefficient of 0.0233, means that with a one-standard deviation increase in *inside CEO* the pay-growth rate increase from 7.3% to 7.58%.

Size incentive premium

I measure performance-based incentives in executive contract by *"delta"*. By definition, *delta* the dollar increase in executives' firm-related wealth for a percentage increase in firm value. It measures the incentives before the performance is realized.[11] 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) ar-

[10] *insider CEO* is provided by Martijn Cremers and Grinstein (2013). *gai* is provided by Custódio et al. (2013). The five proxies to measure general ability of CEO's are: the number of positions that 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.

[11] *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 JensenMurphy 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" here). The JensenMurphy statistic is the correct measure of incentives for activities whose dollar impact is the same regardless of the size of the firm, and the value of equity at stake is appropriate for actions whose value scales with firm size.

Table 2: Performance-based incentives (delta) increase with firm size

	log(*delta*)					
	(1)	(2)	(3)	(4)	(5)	(6)
log(firm size)	0.585***	0.360***	0.525***	0.529***	0.561***	0.571***
	(0.0141)	(0.0247)	(0.00512)	(0.00499)	(0.00310)	(0.0139)
log(firm size) $\times$ *EE90*			0.359*			
			(0.118)			
log(firm size) $\times$ *EE190*				0.415**		
				(0.101)		
log(firm size) $\times$ *gai*					0.0648***	
					(0.00156)	
log(firm size) $\times$ *inside CEO*						-0.0458*
						(0.0202)
log(tdc1)		0.609***	-0.251***	-0.251***	-0.304***	-0.253***
		(0.0350)	(0.00173)	(0.00173)	(0.00267)	(0.00173)
Other contorls	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	146747	128006	125858	125858	75747	125858
adj. R^2	0.442	0.514	0.521	0.521	0.531	0.521

Note: This table reports evidence on firm size premium in executives' performance-based incentives. 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. The total compensation *tdc1*, including the sum of salary and bonus, the value of restricted stocks and options granted, and value of retirement and long-term compensation schemes. Whenever possible, I control for dummies for age, executive tenure, and year $\times$ industry. Other controls including *operating profitability*, *market-book ratio*, *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. The standard error (clustered at the firm $\times$ fiscal year level) are shown in parentheses, and we denote symbols of significance by * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

gued that because executives in larger firms are paid higher, they require more incentives to induce effort. However, this does not explain the entire size 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 the firm scale, *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 *firm size incentive premium* that I aim to explain. As I will show in section 5, the estimates of size premium in both columns (1) and (2) can be quantitatively predicted by the model. I further test if the firm-size incentive premium is related to the managerial labor market using the same four proxies as before: *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 industry/years where the executive labor market is more active. Specifically, the coefficient of 0.415 on the interaction with *EE190* and a standard deviation 0.0224 of *EE190*, means that a one-standard-deviation increase in *EE190* implies the elasticity increases from 0.525 to 0.534. Similar, a one-standard-deviation increase in job-to-job transition measured by *EE90* implies a higher elasticity of .0.532. The standard deviation of *gai* is 0.253, together with the coefficient of 0.0648, means that with a one-standard deviation increase in *gai* the elasticity by 0.016. The standard deviation of *inside CEO* is 0.122, together with the coefficient of 0.0458, means that with a one-standard deviation increase in *inside CEO* the elasticity increase by 0.0056.

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

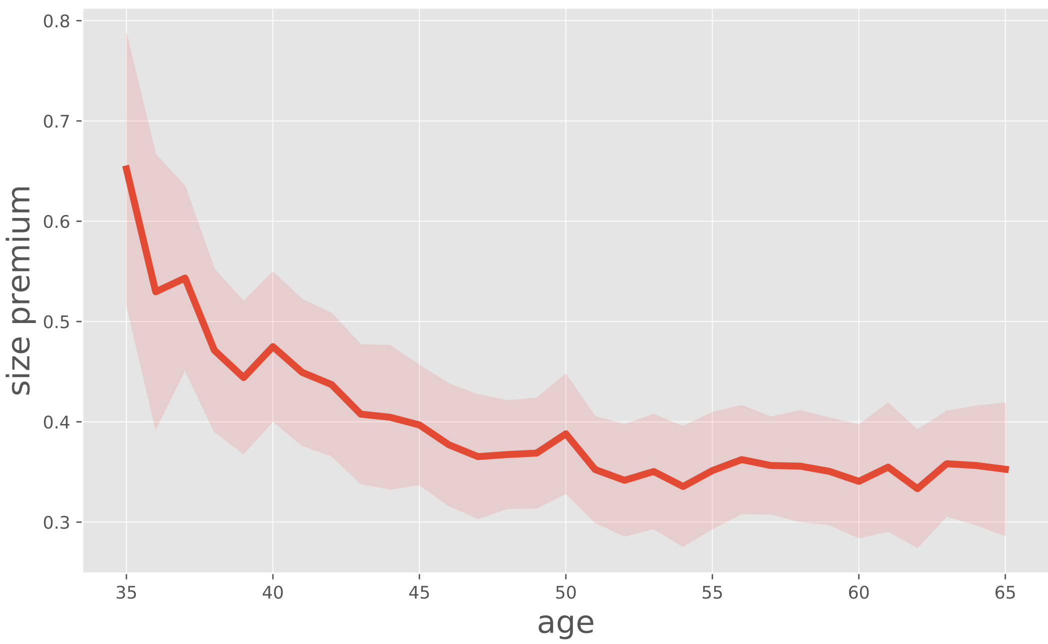

Figure 1: Size premium in performance-based incentives decreases in age

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

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

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

4 Model

In this section, I construct an equilibrium model of the managerial labor market. The model is featured with on-the-job search, poaching offers and 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 a compensation growth. Thus, the size growth premium is linked to firm's capability of overbidding poaching offers. Poaching offers also generate a new source of 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. Hence, the size premium in performance-based incentives is linked to labor market 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 are looking for management 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, $c(\cdot)$ 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, to firms that he or she meets, and can be carried with the individual through job-to-job transitions.[12]

Individual productivity z changes over time according to a Markov process. Denote z_t 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 shirk). To start the process, I assume all unmatched executive candidates have the same starting productivity $z = z_o$. In the following, whenever

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

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

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

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 taking effort is more likely to deliver a higher productivity, $g(z'|z)$ is non-increasing in z'.[13]

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.[14] A match between a worker of productivity z and a firm of size s produces a flow of output $y(s,z) = \alpha_0 s^{\alpha_1} z$, $\alpha_0 \in (0,1)\alpha_1 \in (0,1]$. This function form entails that executive's 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 $\alpha_1 < 1$.[15]

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 exogenous, so we are in partial equilibrium.

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 is triggered. Otherwise, the executive has an interest in transiting to the outside firm.

[13]This is the Monotone likelihood ratio property (MLRP).

[14]From the view of labor search 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. If one interprets firm size as the total number of employees, then it can be endogenized by modeling the labor market of normal workers.

[15]There has been a discussion on the appropriate production function for executives. Taking 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 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.

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 is how the labor market competition is characterized 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 the contract of a continuation value W^0. An on-the-job executive enters a period with his beginning of period productivity z and the current firm of size s. The timing line is shown in figure 2.

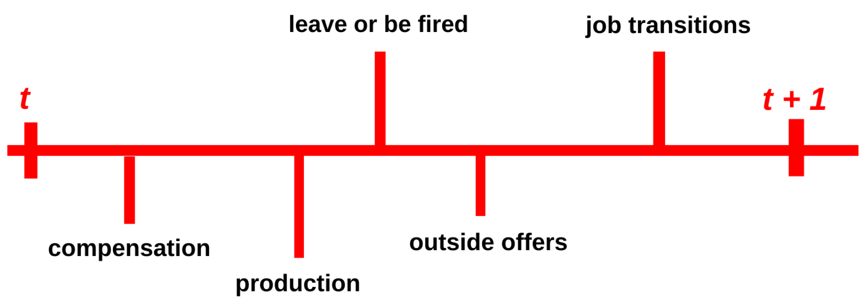

Figure 2: Timing

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 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 for the executive and the firm within a match 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. Denote $\tilde{s} = s_o$ if there is no poaching firm. The history of productivity and 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 history — productivity and poaching firm — correspond to the frictions we have in the contracting problem — moral hazard and search frictions. First, while productivity is included in history and is contractible, the 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 history, I allow a contract to stipulate whether and how to counter poaching offers. This is how I characterize competing for executives in the contracting problem.

Countering outside offers is optimal (or subgame perfect in terms of game terminology), it is, therefore, necessary to allow limited commitment for both sides — to terminate the contract when the surplus is negative. Firms can not 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 can not 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 leaves the firm voluntarily. In the latter case, the executive transits to the poaching firm.

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

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

that satisfied 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 some incentive package. Secondly, I assume a reasonable support of productivity z such that the value of a job is always positive. As a result, firing is excluded, $I_t(h^t) = 1$ exclusively means there is a job-to-job transition.[16]

[16]If we allow a large domain of z such that for some z the profit is negative even when the firm only offers a value W^0, then firing happens. This is an interesting extension that can be analyzed in the future.

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, w is the compensation, and 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).[17]

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 in terms of promised utilities.

Sequential auction

Here I illustrate the the sequential auction using value functions.[18] 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 the maximum bidding values $\overline{W}(z, s)$ are defined by

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

The firm would rather fire the executive (and the vacancy value is normalized to 0) if he or she demands a value higher than $\overline{W}$. I also define $\overline{W}(z, s^o) \equiv W^0$, meaning 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 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 a Bertrand competition won by the larger firm. 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

[17] 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 the PC described above, 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.

[18] What distinguishes this model from the original sequential auction framework is here the wage is not flat. Firms compete on a sequence of wages contingent on all possible future histories, as summarized by $\overline{W}(z, s)$. More importantly, it brings a new source of incentives into the contracting problem. Firms appreciate higher productivities, and are willing to bid more. The *bidding frontier* $\overline{W}(z, s)$ increases in z. The sequential auction therefore begets incentives for managers' 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 the worker become better aligned and the need for short-term compensation incentives decreases.

$\overline{W}(z', s)$. When $\tilde{s} > s$, the executive accepts to move 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 less 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 current firm, and be promoted to the continuation value $\overline{W}(z', \tilde{s})$ that makes him 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, a set of promised values $W(z', \tilde{s})$ depending on the state $(z', \tilde{s})$. For the ease of notations, I denote the effective discount factor $\tilde{\beta} = \beta(1 - \eta)$, and write the mixture distribution of outside offers by

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

The expected profit of the firm can be expressed recursively as

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

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 $f(s, z') - w$, 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 the value V to the executive, and 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 says the continuation value of taking effort is higher than not taking effort. This creates incentives for the executive to pursue the shareholders' interests rather than his 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 that of the current firm, $\overline{W}(z', s)$.

4.3 Equilibrium definition

Before turning to characterize 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.

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

Proposition 1. *The equilibrium exists.*

4.4 Contract characterization

In this section, I derive a characterization of the optimal contract. The characterization builds on and extend 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 the beginning of the 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 utility 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 $\overline{W}(z',\tilde{s}) \geq \overline{W}(z',s)$, 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 says, abstract 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) says 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 IC constraint is not binding ($\mu = 0$, μ is the multiplier of IC constraint), $W(z') = V$ keeps constant. Thus, the pay is also constant over the time. Generally speaking, a higher V induces a higher $W(z')$. That is, an optimal dynamic contract has some memory.

When the outside offers are realized such that the participation constraint is binding, the contract dispenses of the history dependence, and the continuation value depends 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 equals 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 Explain size growth premium

With the characterization of the optimal contract, we are ready to explain the size premium in pay-growth and incentives. I start with a definition on two sets of poaching firms $\tilde{s}$ depending on whether it is larger than the current firms.

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

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 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 $\mathcal{M}_1$ and $\mathcal{M}_2$ is not competitive in the sense that they can not be used to negotiate compensation with the incumbent firm.

Accordingly, the Bellman equation of executives can be written as

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

(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 as well decrease as an executive is willing to sacrifice the current pay for a 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

[19] If compensation information in both the original and target firms are available, 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.

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 keep the executive with a compensation growth while the firm s_1 can not overbid and consequently loses the executive. Therefore, the total pay increases faster in larger firms.

4.6 Explain size incentive premium

To explain the firm-size incentive premium, I start with the definitions of "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 each other given a constant effort cost. Finally, I explain that labor market incentives decrease in firm size. Thus, performance-based incentives increase in firm size.

I first define an "incentive operator" 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 are in three parts: i) the incentives brought by larger firms in $\mathcal{M}_1$; ii) the incentives brought by smaller firms in $\mathcal{M}_2$; and iii) the incentives in the performance-related pay when there are no poaching firms from $\mathcal{M}_1$ or $\mathcal{M}_2$.[20]

The incentives when there is no competitive poaching offers are the *performance-based incentives* in the model, denoted by Ξ_p,

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

[20] We can similarly rewrite the Bellman equations of firms using the optimal continuation value, and this equation is consistent with Postel-Vinay and Robin (2002).

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

and the incentives due to poaching offers are the *labor market incentives*, denoted by Ξ_m,

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

Ξ_m would be zero if there were no poaching offers. Mathematically, Ξ_m is an expectation of incentives in each poaching offer. When the poaching firm is larger, the incentives are from the bidding frontier of the current firm. When the poaching firm is smaller, 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 doesn't. On the other hand, $W(z')$ determines the lower bound of set $\mathcal{M}_2$. The larger the promised continuation value $W(z')$ is, the less likely a poaching firm can be used to renegotiate with the current firm, and the lower labor market incentives are.

Based on this, there is a simple "job ladder" explanation for the size premium when comparing executives of different pay levels (see column (1) in table 2). Since executives of larger firms tend to have higher total compensation, the corresponding continuation value is higher, they are thus higher on the job ladder. Accordingly, the chance to meet 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. 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. [21]

Labor market incentives decrease in firm size

Now I compare labor market incentives for executives with 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 in 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, $s_1 < s_2$, and they have the same total compensation. Figure 3 illustrates the possible poaching firms for the two executives and the associated incentives. The poaching firm size ranges from $\underline{s}$ to $\bar{s}$. I denote 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 life-time utilities rather than current period compensation. Although the two executives have the same total compensation, the one in s_2 has higher life-time utility. The left side of the axis depicts sets $\mathcal{M}_1$, $\mathcal{M}_2$ and corresponding

[21] 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).

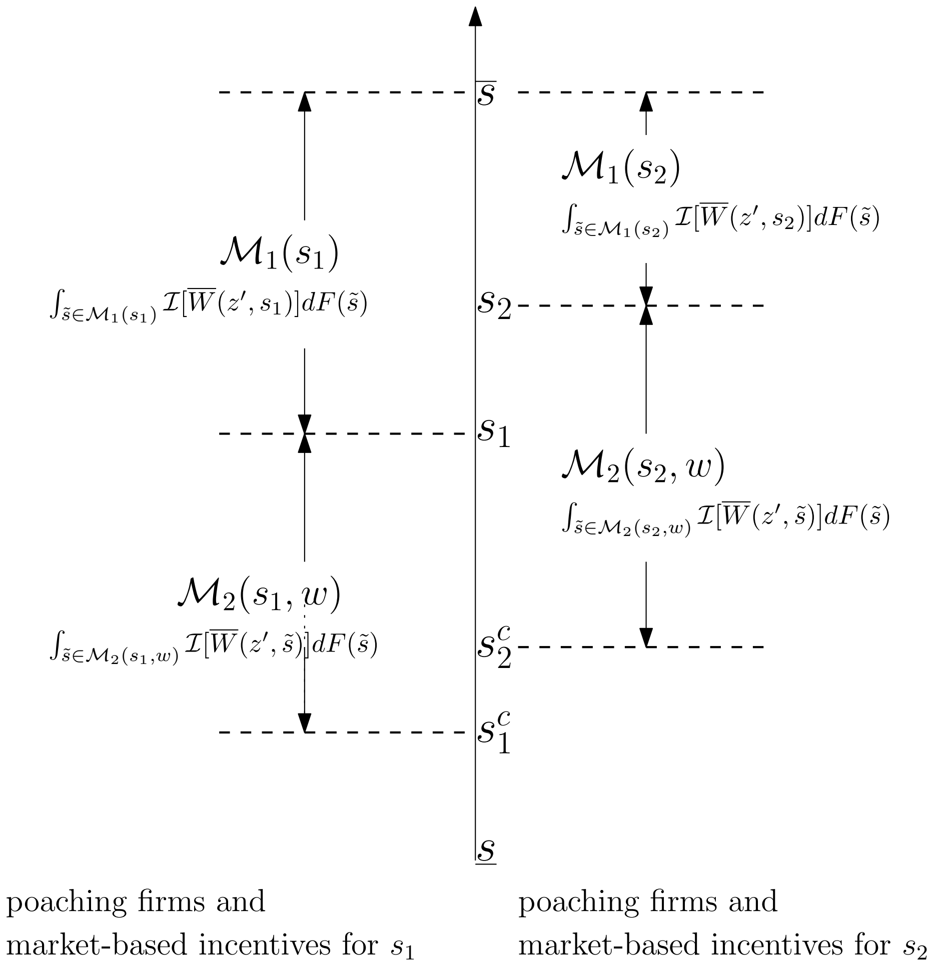

Figure 3: Compare labor market incentives

Note: The figure illustrates labor market incentives for executive 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 in the two sets for the executive in s_1. And 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

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

The difference in their labor market incentives lies in two parts. First, for 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 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 firm s_1 and s_2 are drawn on different bidding frontiers. This corresponds 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 $\overline{W}(z', s_1)$.[22] 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, $\mathcal{I}\Big(\overline{W}(z', s_1)\Big) > \mathcal{I}\Big(\overline{W}(z', \tilde{s})\Big)$ for $\tilde{s} > s_1$. This is a wealth effect of poaching offers — a wealthier executive is harder to incentivize.[23]This wealth effect holds as long as the utility function is concave enough. In the following, I give a sufficient condition under the restriction that the utility is of a CRRA form and effort cost c equals 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*

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

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

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

where $\overline{w}$ is the per-period compensation (wage) corresponds to $\overline{W}$. The first equality follows from implicit differentiation. In the second equality,

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

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

because keeping the promised value, all increasing output is accrued to the firm. $\tilde{\alpha} = \alpha \times$ adjust-factor adjusts for chance that the executive will leave the firm and the job is destructed.

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

follows directly from the optimal contract condition (1) in proposition 2.

There are two opposing effect of s. 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. On the other hand, the incentives in terms of utilities can actually be lower because the marginal utility for extra returns from the executiveial labor market is lower now ($\overline{w}$ increases in s making $u'(\overline{w})$ lower). The second force dominates when the utility function has enough concavity as stated in the proposition.

The requirement of (7) is consistent with 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. The series of 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) approximately equals to one in the parameter space that are explored in my estimation.

Back to the firm-size incentive premium, given (6) and (7), labor market incentives Ξ_m are lower for the executive in firm s_2, and since the effort cost is the same for both executives, the one in firm s_2 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. The close scrutiny of the managerial labor market allows me to put together a rich array of data from various sources. Specifically, I assemble a new dataset on job turnovers from BoardEX and Linkedin, 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.

Table 3: Summary statistics

Variable	N	mean	sd	p25	p50	p75
age	218168	51.04	6.96	46	51	56
male	218168	0.936	0.244	1	1	1
CEO	218168	0.184	0.387	0	0	0
CFO	218168	0.096	0.295	0	0	0
director	218168	0.339	0.473	0	0	1
interlock	218168	0.013	0.112	0	0	0
tenure	218168	4.71	3.793	2	4	6
tdc1	198673	2555.527	5454.153	632.164	1270.806	2690.385
delta	146790	322.518	4736.982	16.966	50.634	154.411
mkcap	212271	7997.377	25810.758	598.919	1622.236	5169.379
at	216384	15594.888	98653.077	542.863	1796.467	6570.342
sales	216276	5472.709	17387.175	428.2	1217.738	3917.269
profit	209639	0.119	0.359	0.069	0.121	0.176
annual return	211067	0.181	0.802	-0.127	0.106	0.356
mbr	183565	1.669	2.21	0.811	1.198	1.913

Note: The table reports summary sample statistics for the ExecuComp/Compustat dataset, which covers named executive officers reported in ExecuComp over the period 1992 to 2016. *age* is the executive's age by the end of the fiscal year. The sample episodes with age lower than 35 or higher than 70 are dropped. Dummy variables *CEO*, *CFO*, *director* and *interlock* indicate whether the executive serve as a director, CEO, CFO and is involved in the interlock relationship during the fiscal year, respectively. *tenure* (in years) counts the number of fiscal years that the executive works as a named officer. *tdc1* is the total compensation comprised of the following: Salary, Bonus, Other Annual, Total Value of Restricted Stock Granted, Total Value of Stock Options Granted (using BlackScholes), Long-Term Incentive Payouts, and All Other Total. *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 base 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 Market

$$Value\ of\ Assets\ (MVA) = prcc\_f * cshpri + dlc + dltt + pstkl - txditc.$$

Variable definitions are provided in the main text.

5.1 Data

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

My final sample comprises 35,088 executive episodes with age between 30 and 65.[24] 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.

Here I describe the variables that are used in my analysis. Using information from ExecuComp, I identify the *gender*, *age* of executive in each year, the *tenure* in the current executive episode, whether he or she is a *CEO*, *CFO*, or *director* of the board or involved in a *interlock* relationship during the fiscal year. Table 3 reports summary statistics for my sample. 93% of the executives are male, and the average age is 51. The average length of episodes is 6.21 years. Among all executive-year observations, 18.4% are CEO spells, 9.6% are CFO spells.

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

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 million 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 BoardEX database.[26]BoardEX contains details of each executive's employment history, including start and end dates, firm names and positions. It also has extra information on education background, social networks, etc. I merge the two databases using three sources of information: the executive's first, middle and last names, the date of birth, and working experiences — in which year the executive worked in which firms. If all three aspects are consistent, the executive is identified. And for executives that cannot be identified in BoardEX, I search for their

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

[25]In my dataset, all nominal quantities are converted into constant 2016 dollars using a measure of the price level the DGP deflater from the Bureau of Economic Analysis.

Linkedin pages and manually collect the available employment information. By this way, I am able to identify more than 91% of executives in ExecuComp, 32,864 executives in total.

5.2 Job-to-job transitions

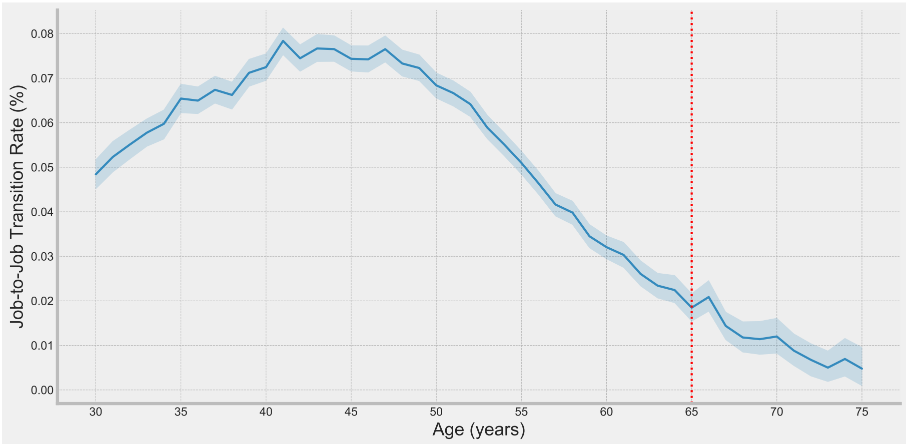

Figure 4: Job-to-job transition rate over age

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

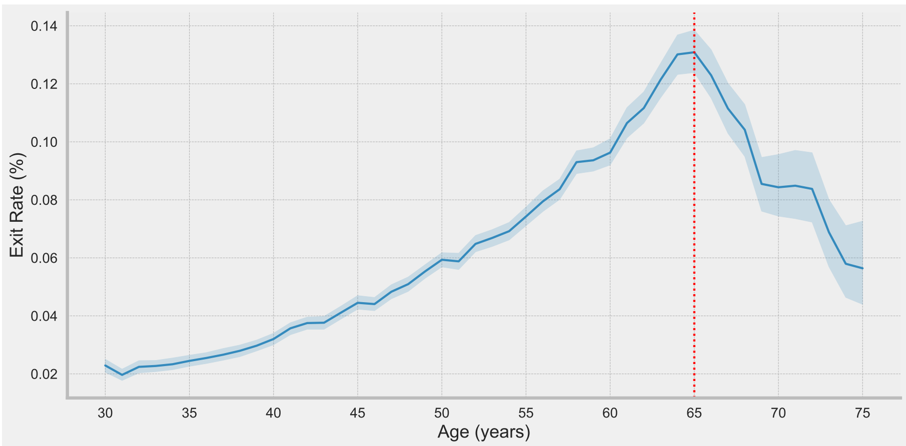

Figure 5: Exit rate over age

Note: The figure depicts the estimates of exit rates over age with the 95% confidence interval around the estimates. A job exit is defined as an executive leaves the current firm and does not work in another firm within 190 days.

[26]What is missing in 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 one *S&P* firms or simply retires. Moreover, the start and end dates of the current employment are also not known.

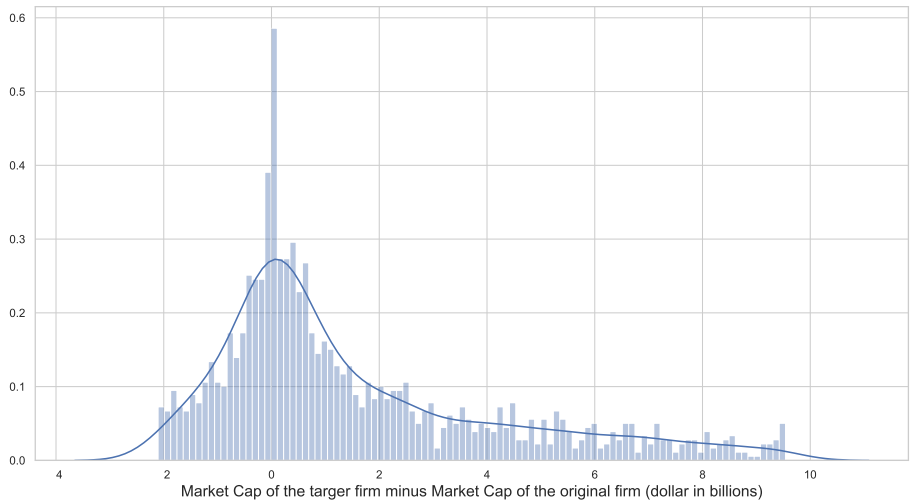

Figure 6: Distribution of change of firm size upon job-to-job transitions

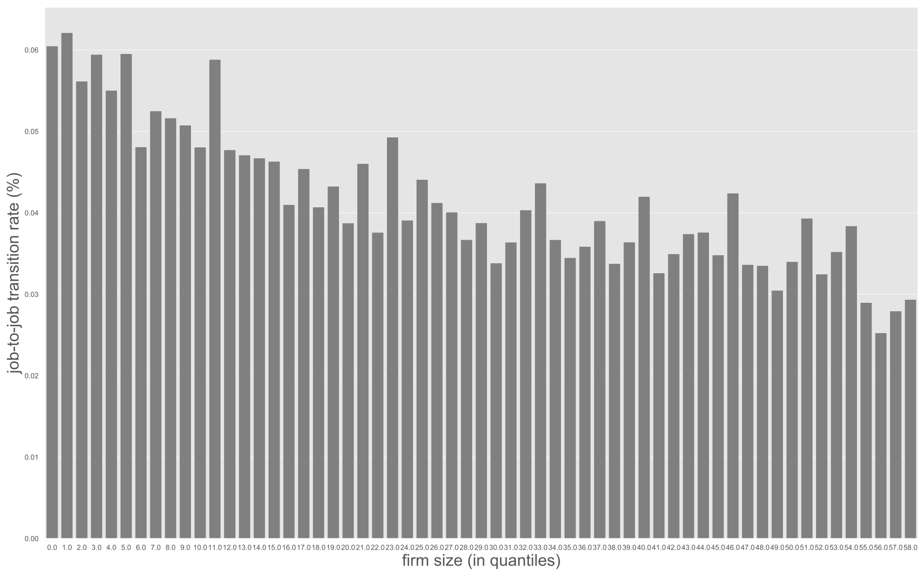

Figure 7: Job-to-job transition rate across firm size

I define a job-to-job transition by the executive leaving the current firm and starting to work in another firm within 190 days. Otherwise, it is defined as an exit from the managerial labor market. In the data, the job-to-job transition rate is 4.98% each year over 1992 to 2015, while the job exit rate is slightly higher 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 the age around 45, and goes down after 50. In contrast, job exit rate is meager before 55 and peaks sharply at the age 65 as expected.

Most job-to-job transitions are within the industry. Among transitions that industry information is observable, 1717 out of 2567 transitions are within the industry defined by the Fama-French 12 industry classification, and 1407 out of the 2567 cases by the Fama-French 48 industry.

Executives transit to larger firms

In my sample, there are 9138 job-to-job transitions from a CompuStat firm, and only 2567 of them have the size information on both original and target firms. The rest firms are private firms whose size information is not disclosed. Based on the selected sample where size information is observable, I find that approximately 60% 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 those cases are due to a title change from a non-CEO title to a CEO title, while this fraction is only 3.3% in transitions towards larger firms.

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

Job-to-job transitions decrease in firm size

Next, I check whether executives in larger firms have fewer transitions. 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 duration 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. Being different from the model's prediction that job-to-job transition is not related to compen-

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	2567	985 (39%)	1582 (61%)
Sales	2617	1051 (40%)	1566 (60%)
Book Assets	2616	1038 (40%)	1578 (60%)

Panel B: Across age groups

Age groups	Total obs.	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%)

sation level, in the data, when the total compensation rises by 1%, the hazard rate drops by 27% One possible explanation 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.

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 leaves the current firm (and does not return to the current firm within one year), and starts to work in another firm within 180 days. All variables have the same definition as in table 1. All dollar-related variables are adjusted by a GDP deflater. The standard error are shown in parentheses, and I denote symbols of significance by * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

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 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 solving and makes the estimation feasible. I leave more details to Appendix C.

6.2 Model Specification and Parameters

I estimated the model fully parametrically and make several parametric assumptions. Being consistent with the analysis before, I use the constant relative risk aversion utility function $u(w) = \frac{w^{1-\sigma}}{1-\sigma}$, and a production function of $f(z,s) = e^{\alpha_0} s^{\alpha_1} z$. 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 effort level $e = 0$ is normalized to zero. The process is transformed to a discrete Markov Chain using Tauchen (1986) on a grid of 6 points.[27] Furthermore, I set the sampling distribution of firm size $F(s)$ a truncated log-normal distribution with expectation of μ_s and standard deviation of σ_s.[28] 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.

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

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, 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 on the profitability of each firm-executive match,

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

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

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

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

[28] The upper and lower bound of the truncated normal distribution is calibrated to be the 0.99 and 0.01 quantile of market capitalization in the data.

Similarly, the higher σ_s is, the more heterogeneous the outside firms are, 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 following regression of log tdc1 on log firm size,

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

The final part of the identification concerns the parameters σ and c. These parameters govern the level of incentives and how the incentives change with 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 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. The higher σ is, the larger $\beta_{delta-tdc1}$ will be.

Firm-size premiums

I intentionally leave the firm-size pay-growth premium and incentive premium untargeted in the estimation. Instead, I estimate the premiums in the data and model-simulated dataset to examine if the model mechanism can match up with the data estimates. In both the real 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,

$$\Delta\log(tdc1_{it}) = \beta_3 + \beta_{\Delta tdc1-size}\log(size_{it}) + \beta_4\log(tdc1_{it}) + \epsilon_{it,3}, \tag{8}$$

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

$$\log(delta_{it}) = \beta_5 + \beta_{delta-size}\log(size_{it}) + \beta_6\log(tdc1_{it}) + \epsilon_{it,3}. \tag{9}$$

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

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		

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 arranged moments and parameters along the identification argument made in the previous subsection, all parameters are estimated jointly. Overall, the model provides a decent fit to the data.

Looking into the estimates, a job arrival rate $\lambda_1 = 31.64\%$ 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 a confirmed by a small mean of poaching firms. The magnitude of μ_s indicates that most offers are provided by relative small firms, though the magnitude of σ_s implies the variation of poaching firms is high. Compare the data and model-simulated mean and variance of $\log(size)$, it seems using a log-normal distribution is sufficient to match the firm size distribution in the data.

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

6.5 Predicting firm-size premiums

Table 8 reports the size-premium estimates in the data and the model generated data. There are three premiums. The first row is the size pay growth premium that is estimated in regression (8). The second row is the size incentive premium estimated in regression (9). The last row is also a size incentive premium using specification (9) except the total compensation is not controlled.

Table 8: Predictions on Size Premiums

	Benchmark		Variants		
Size premiums	Data	Benchmark	Ignore mkt inc	More offers	Less offers
pay-growth	0.1542	0.1450	0.1481	0.1624	0.0411
incentives	0.3473	0.3122	-0.0444	0.4299	0.1964
incentives (w/o tdc1)	0.6044	0.6507	0.4202	0.7093	0.4076

Column (1) are the premium estimates in the data, the same as reported in table 1 and table 2. Column (2) are the estimates in the benchmark model using the estimated parameters. Comparing columns (1) and (2), I find even without targeting on these premiums, the model can capture all three premiums quantitatively. 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 model mechanism to play an important role in explaining the firm size premium.

To further clarify that these premiums are due to poaching offers, in column (3) to (5), I report the premium estimates in several model variants. In columns (3), I simulate a counterfactual scenario where the firms ignore labor market incentives when designing the incentive contracts. In column (4), I simulate the model using a higher job arrival probability $\lambda_1 = 0.6$. And in column (5), I simulate the model with a smaller 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 in the second row essentially becomes zero. Therefore, the incentive premium is solely driven by labor market incentives. The incentive premium without controlling for *tdc1*, 0.4202, entirely reflect the compensation levels are higher in larger, which is the channel discussed by Edmans et al. (2009). Columns (4) and (5) show that when there are more job offers, both the pay-growth and incentive premiums are higher, whereas when there are fewer job offers, both premiums decrease. These exercises clarify that premiums indeed stem from poaching offers.

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 10 groups. The upper panel of figure 8 shows the box plot of $\log(delta)$ across firm size. Clearly, smaller firms are likely to suffer more by ignoring labor market incentives, in consistent with the job ladder mechanism. I further calculate the ratio of the delta's with versus without labor market incentives as in the lower panel of figure 8. The fraction of market incentives is surprisingly high for the smallest firm group: the delta will be 80% higher when job ladder is absent. The fraction quickly goes down to around 15% in the medium-size firms, and almost vanishes for top-size firms.

7 Understand the Long-run Trends in Executive Pay

Based on the structural estimation, I use a counter-factual exercise to quantitatively explain the sharp increase in executive total pay, performance-based incentives, the inequality of across executives, and the correlation between executive compensation and firm size since the mid-1970s. Using a new dataset, Frydman and Saks (2010) document that the level and inequality of executive pay were relatively low from the late 1930s to the mid-1970s and had soared since then. Similarly, the correlation between firm size and total compensation is much weaker before the mid-1970s.

In table 9, I select two representative periods 1970 - 1979 and 1990 - 1999, and replicate the data moments from Frydman and Saks (2010). The mean of total compensation rises from 1090 thousand dollars before 1979 to 4350 thousand dollars after 1990, and the mean of performance-based incentives increases by almost six folds from the 1970s to the 1990s. The interquartile range of third and first quartiles increases from 670 thousand dollars to 3080 thousand dollars. While firm size is closely related to executive pay in the data after 1992, it was weaker in previous decades. The coefficient increases from 0.199 to 0.264 from the 1970s to 1990s.

All these changes can be accounted by my model with an exogenous change of the outside managerial labor market, measured by the job arrival rate λ_1. While I do not have an endogenous mechanism for the increase in λ_1, there has been abundant evidence of a more active executive labor market since the mid-1970s. Murphy and Zabojnik (2007) document that an increasing number of CEO openings has been filled through external hires. Huson et al. (2001) document that the fraction of outsider CEOs increases

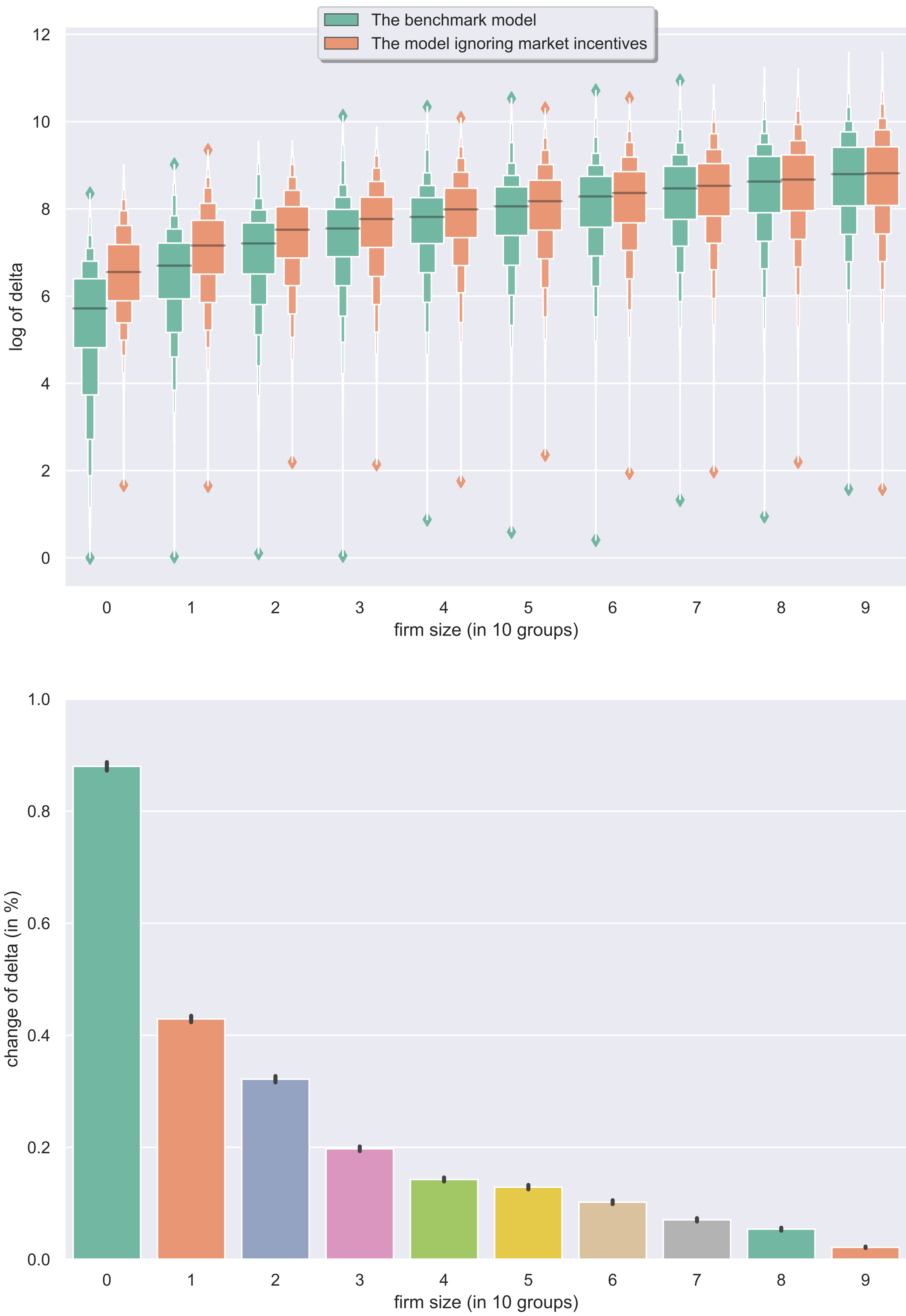

Figure 8: The Fraction of Market Incentives along Firm Size and Wage

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 can be transferred 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 equal to the estimated values in the last section. Since most firms in the sample of Frydman and Saks (2010) are within the rank 500. I only keep the largest 500 firms in the simulated data. The moments calculated by simulated datasets are reported in the last two columns of table 9.

The results are consistent with the intuition. As λ_1 increases, executives are more likely to use poaching offers to renegotiate contracts which leads to higher total compensation $tdc1$ and higher incentives $delta$. As firms bidding for executives, the correlation between pay and firm size becomes larger. Next, since the labor market is frictional, the inequality is amplified with more poaching offers: lucky executives receive many poaching offers, while unlucky ones get few job-hopping opportunities. 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 reflects that the poaching offer distributions of the two periods can be different.

My model also entails predictions on moments that are not disclosed in Frydman and Saks (2010). A more active labor market also induce a larger average firm size. The mean of firm size doubles as λ_1 increases (from 2425 millions to 5710 millions). 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 tests in the future.

Table 9: The 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)	21743	120342	24972	125310
$\beta_{tdc1-size}$	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

8 The Spillover Effect and Policy Implications

In this section, I discuss the spillover effect of the firm's 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 to the executive. The "spillover" refers to the effect that a higher willingness to bid from some firms not only raises the executive pay in those firms but also increases the pay in all firms that are higher on the job ladder. This is because executives that are higher on the job ladder can make use of these bids to negotiate with their present firm. Consequently, the renegotiation leads to higher pay and 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: By no means that α_0 should always be negatively correlated with corporate governance, and this assumption should be valid only in the range where α_0 is very high.

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. One is that α_0 doubles for firms that are smaller than the size median. I call these firms "small and medium firms". I denote this higher willingness to pay from small and medium firms by "worse governance in small firms". And it supposes to create a spillover effect on the pay of large firms. To compare to this spillover effect, I use the second counterfactual that α_0 doubles for firms that are larger than the median. I call these firm the large firms, and this case is denoted by "worse governance in large firms". Figure 9 plots the distribution of delta (in upper panel) and total compensation (in lower panel) across 10 equally divided firm size groups. There is one box plot for each group, and a median is marked as a horizontal line in the middle of the box.

Not surprisingly, the boosts in willingness to pay increase total compensation and incentives in each type of firms separately. A higher willingness of bidding from small firms (in green) raises the pay and incentives in relatively small firms (the first five size groups), while a higher α_0 from large firms (in blue) increases the pay and incentives of the largest five groups of firms. More importantly, the rise in α_0 in small and medium firms spillovers to large firms as well. Precisely, in terms of median, the spillover effect is as large as the effect of a higher willingness to bid in large firms themselves. There are 40% to 50% increases in the pay level and incentives in firms of the largest two groups, no matter the rise is originated from the higher bidding of small or large firms.

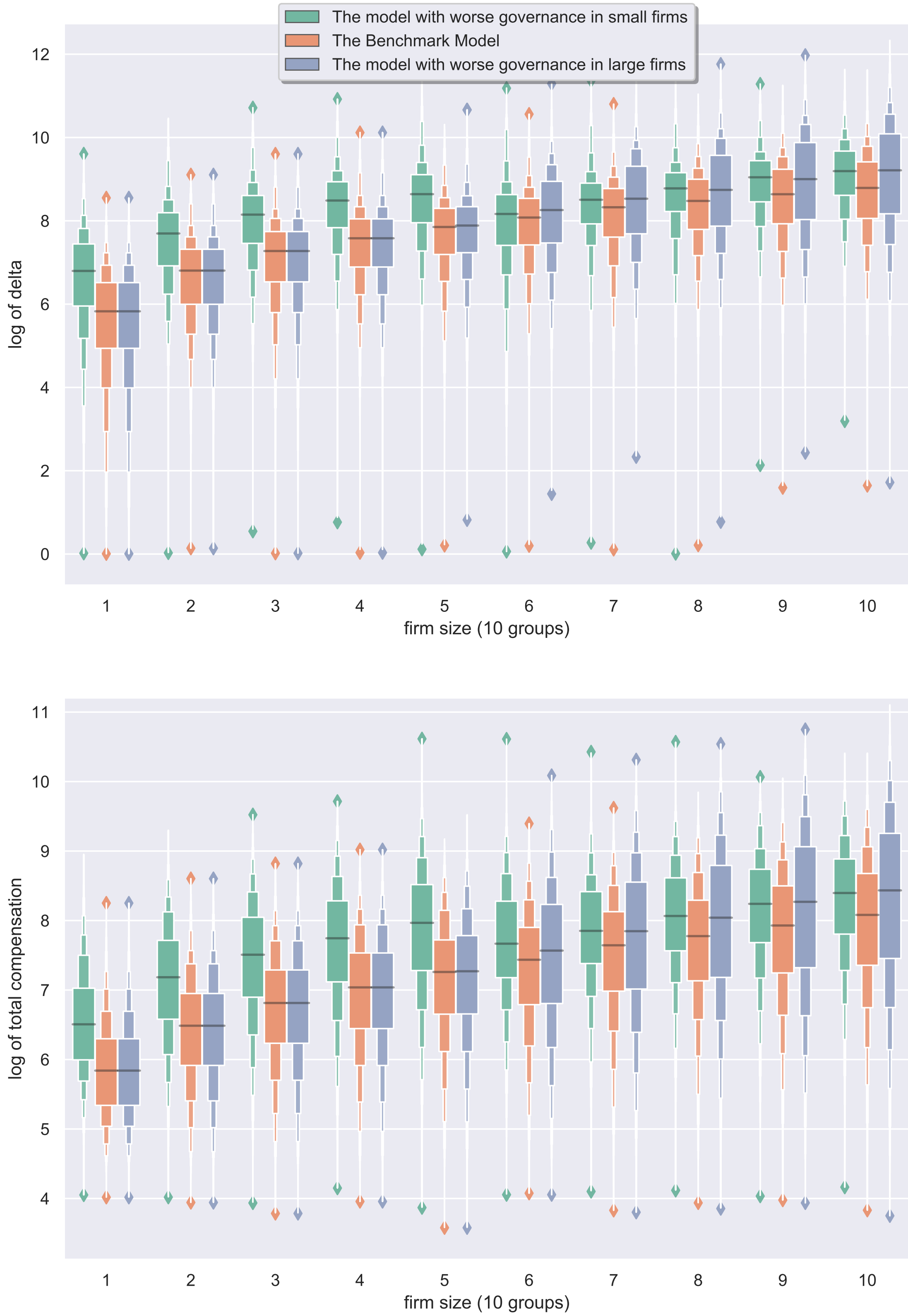

Figure 9: The Fraction of Market Incentives along Firm Size and Wage

The policy implication of this exercise is as follow. To regulate the compensation of highly paid executives, rather than only focusing on large firms, it is more important to lower the willingness to bid in small and medium firms. By doing so, large firms face less competitive pressure in the managerial labor market and the impact will reach large firms as well. As for particular regulation policies, the reforms that have been proposed or implemented should work in small and medium firms as well, including more independent compensation committee, greater mandatory pay (or pay ratio) disclosure, say-on-pay legislation, etc.

9 Conclusions

In this paper, I 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. I show 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 job turnover dataset supports assumptions and implications of the model.

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 close to the estimates in the data. A counterfactual analysis based on the structural estimation showing that with an exogenous increase of poaching offer arrival rate, my model can explain the sharp increase in total pay, performance-based incentives and the correlation between firm size and pay levels since the mid-1970's.

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

Appendix A. Model appendices

Proof for proposition 3

We 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 $\frac{\Delta\overline{W}(z',s)}{\Delta z'}$ decrease in s.

Step 1: show 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 space $\mathbb{Z} = \{z^{(1)}, z^{(2)}, ..., z^{(n_z)}\}$. Suppose there is a distribution of productivity when the executive takes the effort Γ, a distribution when the executive shirk 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. The fact that $\sum_{z\in\mathbb{Z}}(1 - g(z))\Gamma(z) = 0$ implies that

$$\gamma_1 + ... + \gamma_m - \gamma_{m+1} - ... - \gamma_{n_z-1} - \gamma_{n_z} = 0. \tag{10}$$

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 we have applied condition (10). By construction, ω_i is positive.

$\square$

Step 2: express $\frac{\Delta \overline{W}(z,s)}{\Delta z}$ 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

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

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 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

$$\begin{aligned}
\frac{\Delta \Pi(z,s,\overline{W})}{\Delta z} &= \sum_{z' \in \mathbb{Z}} \Big(\alpha_0 s^{\alpha_1} z' + \tilde{\beta} \int_{\tilde{s} \leq s} \Pi(z',s,\overline{W}(z',s) d\tilde{F}(\tilde{s})\Big)\frac{\Delta \Gamma(z'|z)}{\Delta z} \\
&\quad + \lambda \tilde{\beta} \sum_{z' \in \mathbb{Z}} \Big(\int_{\tilde{s}} \overline{W}(z',s) d\tilde{F}(\tilde{s})\Big)\frac{\Delta \Gamma(z'|z)}{\Delta z} \\
&\quad + \mu \tilde{\beta} \sum_{z' \in \mathbb{Z}} \Big(\int_{\tilde{s}} \overline{W}(z',s) d\tilde{F}(\tilde{s})\Big)\frac{\Delta\Big((1-g(z'|z)\Gamma(z'|z)\Big)}{\Delta z} \\
&= \alpha_0 s^{\alpha_1} \sum_{z' \in \mathbb{Z}} z' \frac{\Delta \Gamma(z'|z)}{\Delta z} + \tilde{\beta} \sum_{z' \in \mathbb{Z}} \int_{\tilde{s}} \overline{W}(z',s) d\tilde{F}(\tilde{s}) \Big(\lambda \frac{\Delta \Gamma(z'|z)}{\Delta z} + \mu \frac{\Delta\Big((1-g(z'|z)\Gamma(z'|z)\Big)}{\Delta z}\Big).
\end{aligned} \tag{12}$$

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

$$\frac{\frac{\Delta\Pi(z,s,\overline{W})}{\Delta z}}{\alpha_0 \sum_{z' \in \mathbb{Z}} z' \frac{\Delta\Gamma(z'|z)}{\Delta z}} = s^{\alpha_1} + \frac{\tilde{\beta} \sum_{z' \in \mathbb{Z}} \int_{\tilde{s}} \overline{W}(z', s) d\tilde{F}(\tilde{s}) \left(\lambda \frac{\Delta\Gamma(z'|z)}{\Delta z} + \mu \frac{\Delta\left((1-g(z'|z)\Gamma(z'|z) \right)}{\Delta z} \right)}{\Delta z} / \alpha_0 \sum_{z' \in \mathbb{Z}} z' \frac{\Delta\Gamma(z'|z)}{\Delta z}$$
$$= s^{\alpha_1} + \psi(s), \tag{13}$$

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

Insert (12) and (13) into (11), we have

$$\frac{\Delta\overline{W}(z,s)}{\Delta z} = u'(\overline{w}(s)) \frac{\Delta\Pi(z,s,\overline{W})}{\Delta z} = u'\left(\alpha_0 s^{\alpha_1} \sum_{z' \in \mathbb{Z}} z' \Gamma(z'|z) \right) \left(s^{\alpha_1} + \psi(s) \right) \alpha_0 \sum_{z' \in \mathbb{Z}} z' \frac{\Delta\Gamma(z'|z)}{\Delta z}. \tag{14}$$

Step 3: show 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 (14)

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

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.

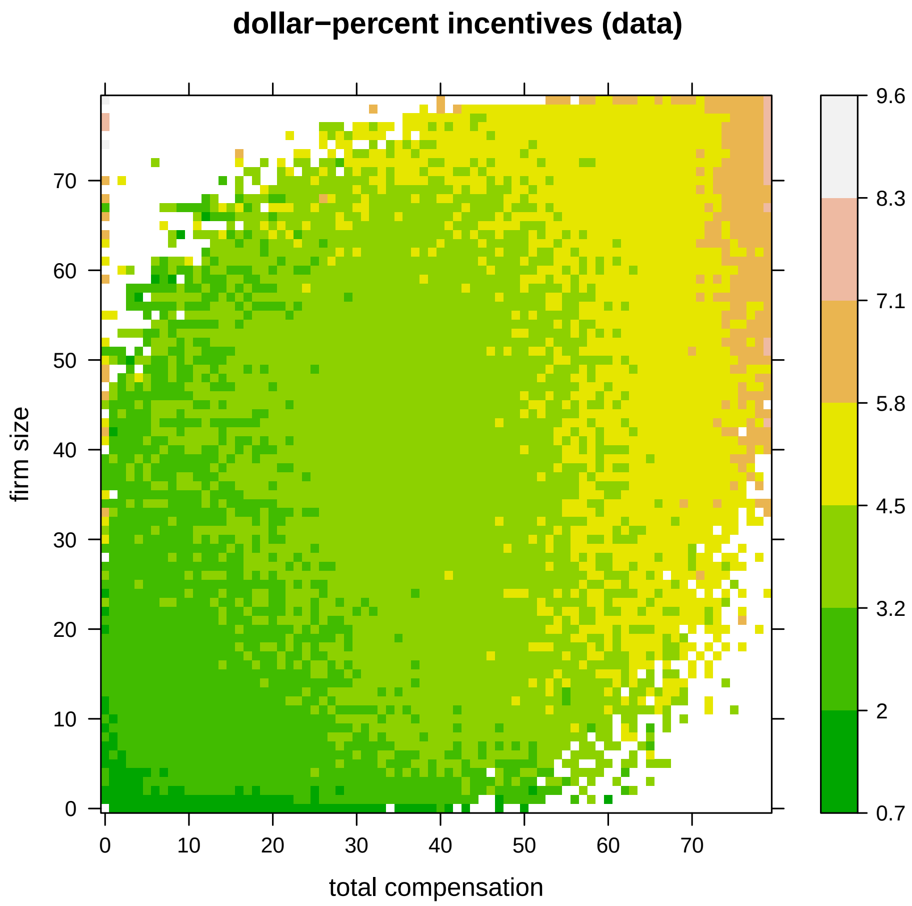

Figure 10: $\log(delta)$ over firm size and total compensation

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 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. The estimates in column (2) are used to plot figure 1.

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	146747	128006	128006	128006	109730
adj. R^2	0.442	0.514	0.523	0.524	0.595

Note: This table reports evidence on firm size premium in executives' 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. 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. The key control variable is the total compensation *tdc1*, 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 grouped into 50 and 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*, 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. The standard error (clustered at the firm × fiscal year level) are shown in parentheses, and we denote symbols of significance by $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$.

Table 11: Firm-size incentive premium increases with managerial labor market compeiti-ton

	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) $\times$ *EE190*	0.415***			
	(0.101)			
EE90		2.611**		
		(0.903)		
log(firm size) $\times$ *EE90*		0.359**		
		(0.118)		
gai			-1.211***	
			(0.0941)	
log(firm size) $\times$ *gai*			0.0648***	
			(0.0118)	
inside CEO				-0.00566***
				(0.00156)
log(firm size) $\times$ *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 premium in executives' performance-based incentives increases as the managerial labor market competition is more fierce. 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 when executive leaves the current firm and starts 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. Regression in column (3) measures labor market activity by the general ability index `gai` averaged by (Fama-French 48) industries $\times$ fiscal years. This index was composed by Custódio et al. (2013). Column (4) uses the percentage of new CEO's who are insiders at the industry level which 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. The standard error (clustered at the firm $\times$ fiscal year level) are shown in parentheses, and we denote symbols of significance by * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

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

Notes for table 12

This table reports evidence that firm size premium in executives' performance-based incentives decreaes 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 are 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. The standard error (clustered at the firm $\times$ fiscal year level) are shown in parentheses, and we denote symbols of significance by * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Appendix C. Estimation appendices

Recursive multiplier method

To further characterize the optimal solution, we resort to the tools developed by Marcet and Marimon (2017, hereafter MM).[29]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 planners 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

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

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

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

Pareto weight evolves according to

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

and

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

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

$$u'(w) = \frac{1}{\lambda}, \tag{15}$$

$$W(z') = W(z',s,\lambda'). \tag{16}$$

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

[29]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 $\mu^{(0)}$. Then depending on the realization of z', the weight of the executive will be updated to

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

The evolve of λ continues as such till the match breaks. When there is an outside offer such that the executive moves from his current firm to the outside firm, then the new match starts with a $\lambda^{(n)}$ such that $W(z, s', \lambda^{(n)}) = \overline{W}(z, s)$, where I have denoted the current productivity by z, current firm by s, and the outside firm by 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 above. In a nutshell, proposition 4 allows us to solve the optimal contract in the space of Pareto weight λ instead of in the space of the promised utility. At any moment, we can transfer from the metrics of λ back to the metrics of utilities using (15) and (16).

The advantage of this method is I do not need to find the promised utilities $W(z')$ in each state of the world for the next period. Instead, λ and μ are enough to trace all $W(z')$. 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.

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.

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."

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.

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. 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." *V Columbia University*, 5.

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.

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.

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.