Why Do Larger Firms Pay Executives More For Performance?

Performance-Based versus Labor Market Incentives

Bo Hu*

Job Market Paper

October 24, 2018

Abstract

This paper evaluates the impact of managerial labor market competition executive incentive contracts. I construct an equilibrium search model that features a hierarchical managerial job ladder towards larger firms. The competition for managerial skills between heterogeneous firms increases executive compensation, and generates a new source of incentives, called the *labor market incentives*, which substitutes for the performance-based incentives. The model is capable of explaining the stylized facts that executives of larger firms experience higher compensation growth and receive higher performance-based incentives. I estimate the model using the data of executives from U.S. publicly listed firms. The structural estimates suggest that for executives in small firms more than 40% of the total incentives come from the managerial labor market, and the fraction decreases to less than 5% for executives in large firms.

Key Words: executive compensation, firm-size premium, managerial labor market, 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 the performance-based incentives are designed to align the interests of executives with those of shareholders. In recent decades, however, we have seen that competition for executive talents has increasingly become influential in shaping the design of managerial incentive contracts. According to Apple's latest proxy statement, "experienced personnel in the technology industry are in high demand, and competition for executive talent is intense." Facing fierce competition, Apple designed its performance-based incentives to "attract and retain a talented executive team". Similarly, to react to the intensive competition for executives, Amazon declared that the core philosophy concerning executive incentive package is "to attract and retain the highest caliber employees". These examples highlight the role of competition for executive on incentive contracts.

Contrary to its relevance in the industry, the underlying mechanism linking the executive labor market to the design of incentive contracts has so far remained unclear. To make theoretical progress, I construct and estimate a tractable model that embeds the dynamic contracting problem into an equilibrium labor search framework. The model retains key intuitions of the previous static models while delivers insights to explain new empirical puzzles. As pointed out in the "directions for future research" by Edmans et al. (2017), "Most models of incentives in market equilibrium are static. It would be useful to add a dynamic moral hazard problem where incentives can be provided not only through contracts, but also by ... the promise of being hired by a larger firm. This would, among other things, analyze how contracting incentives interact with ... hiring incentives. These different incentive channels may conflict with as well as reinforce each other." This statement describes exactly what the current paper does and how it contributes to the literature.

I construct a framework that combines the dynamic moral hazard problem and the equilibrium labor search. There are two types of agents: firms and executives. Firms are heterogeneous in (time-invariant) asset size. Executives are heterogeneous in managerial productivity which evolves stochastically depending on the current and past effort. They search in the managerial labor market to find a match. What distinguishes this paper from previous studies is that the managerial labor market here is *search frictional* and allows *on-the-job search*.[1] These two features together give rise to a hierarchical job ladder from small to large firms, which lies in the core of the model. Once a firm and an executive are matched, they sign a long-term incentive contract and produce. In the production, firm size and executive productivity are complementary, analogous to Gabaix and Landier (2008). While the output is observable, the effort is not. Thus, a moral

[1]The managerial labor market in previous studies are mostly frictionless and perfectly competitive (Gabaix and Landier 2008, Edmans et al. 2009, Gayle et al. 2015).

hazard problem arises. Importantly, the moral hazard problem is not isolated from the outside managerial labor market. Both the firm and the executive have limited commitment to the relationship, and may encounter outside poaching offers. By making use of poaching offers, the executive can renegotiate with the current firm or transit to poaching firms, as in Postel-Vinay and Robin (2002).

The impacts from the competition of outside offers on managerial incentive contracts are shown in two mechanisms. The first is about the impact on the compensation level. When the poaching firm is smaller than the current firm, the executive may use it to negotiate a higher pay with the current firm. When the poaching firm is larger, the poaching firm can always outbid the present firm, because firm size contributes to the production. Thus, the executive transits to the poaching firm, and use the previous firm as a threat to negotiate with the poaching firm. In either case, the executive climbs up the job ladder towards a higher compensation level and a larger firm size.

The second mechanism is about the performance-based incentives in the contract. Poaching firms are willing to bid higher for more productive executives, thanks to the contribution of executive productivity in the production function. The productivity of an executive, on the other hand, is stochastically determined by his past effort. Together, it implies that taking effort today will lead to a more favorable offer from the poaching firm. This incentive effect is called the *labor market incentives* in this paper (or hiring incentives by Edmans et al. 2017). Consequently, firms can take advantage of the incentives from outside offers, and give less performance-based incentives to the executive, but still solve the moral hazard issue.

These two mechanisms enable the model to go beyond the static framework in the literature and get the explaining power on two important empirical puzzles. The first puzzle is the firm-size premium in compensation growth, namely, executives of larger firms experience a higher compensation growth compared to the counterparts in smaller firms with the same compensation at the beginning. This fact is firstly documented in this paper, and it extends the scaling of total compensation with firm-size, as documented in the literature, to the time dimension. My explanation is very intuitive: In line with the first mechanism in the previous paragraph, executive compensation grows because of the competition of poaching offers. Larger firms are more capable to counter outside offers, thus their executives tend to experience higher compensation growth.

The second empirical puzzle is the firm-size premium in performance-based incentives (hereafter, the *firm-size incentive premium*), namely, the performance-based incentives increase with firm size, which holds even confined to executives with the same total compensation.[2] In accordance with the second mechanism above, in the model, an executive is motivated by two sources of incentives: the performance-based incentives

[2]The performance-based incentives are measured by the dollar change in firm-related wealth per percentage change in firm value. The firm size premium in performance-based incentives holds after controlling for total compensation and firm performance. See section 2 for more details.

and the labor market incentives. The fact that performance-based incentives increase in firm size reflects that labor market incentives decrease in firm size, given that the required total incentives are fixed.

Then why do labor market incentives decrease in firm size? My model gives two reasons. First, executives from larger firms are on "higher" positions of the job ladder, and the chance to receive an outside offer that can improve upon the current value is lower. Thus, the labor market incentives are smaller. Second, because executives of larger firms are expected to be more compensated in the future as is stated and explained in the first puzzle, the certainty equivalents of their future expected utilities are higher. According to the diminishing marginal utility, at a higher certainty equivalent, the marginal utility of a more favorable outside offer is smaller. Therefore, the labor market incentives are smaller. Taking together, both reasons contribute to explain that the labor market incentives decrease in firm size.

I provide reduced-form empirical evidence to support the model. Firstly, I justify the hierarchical job ladder featured in the model. By merging different databases, I am able to identify job turnovers using detailed employment histories. I document a job-to-job transition rate around 5%, which is stable over the years and across industries. Moreover, there exists a job ladder on the firm size dimension: about 60% of job-to-job transitions are towards larger firms, and for the rest transitions 20% of them are due to a title change from a non-CEO title to a CEO title. Secondly, in the model the "position" on the job ladder matters for job-to-job transitions. In particular, the model predicts that controlling for firm size, more compensated executives are less likely to transit; and controlling for the total compensation, executives in larger firms are less likely to transit. Both predictions are verified by the data.

Numerically solving for the optimal contract becomes difficult in the presence of the incentive compatibility constraint, limited-commitment constraints, together with shocks of large support. The reason is one needs to solve for the promised value in each state of the world, which is computationally infeasible. Following Marcet and Marimon (2017), I address this issue using the recursive Lagrangian approach. In this approach, I only need to solve for one Lagrangian multiplier in order to solve the optimal contract. This multiplier represents the weight of the executive in a constructed Pareto problem, and it keeps track of the incentive compatibility constraint, limited commitment constraints and job-to-job transitions. It is then easy to recover optimal incentive pay and promised values from this multiplier.

Using the simulated method of moments (SMM), I estimate the model by targeting the data moments of executive compensation, incentives and job turnovers, as well as firm size and profitability. Importantly, I do not explicitly target the firm-size premium on compensation growth and performance-based incentives. The estimated model quantitatively captures both, which reassures that the model mechanism plays an essential role in explaining both premiums. A counter-factual decomposition shows that the labor market incentives proposed in my model 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.

Finally, I discuss the policy implication of the model. I interpret an improved corporate governance as a lower firm's willingness to bid for executives. I show that an improved governance has a spillover effect on executive compensation through the managerial labor market. Corresponding to the two mechanisms in the model, a lower willingness to bid of the firm decreases the total compensation of the executives who have received offers from this firm; it also increases the labor market incentives for executives who are expecting to receive an offer from this firm. Overall, the improved corporate governance leads to lower total compensation and lower performance-based incentives through spillover effect. A counterfactual analysis based on the estimated model shows that the effect is larger when the governance is improved in small and medium-size firms. The policy implication is, to regulate executive compensation, rather than only focus on large firms, it is equally important to improve the governance in small and medium firms in order to lower the competition pressure in the overall managerial labor market.

This paper contributes to the literature in explaining the scaling of incentive pay with firm size. Edmans et al. (2009) and Edmans and Gabaix (2011) provide an explanation based on that total compensation increases with firm size.[3] Their model, however, does not explain why after controlling for total compensation, size premium in pay-for-performance incentives still exists. My model is a search-frictional version of their assignment framework. While preserving their key intuition, the search frictions added in this paper give rise to the labor market incentives that explains the firm-size incentive premium. The driving force of my explanation is the hierarchical job ladder that goes from small to large firms. This job ladder contrasts with the managerial labor market in Gayle et al. (2015) where job-to-job transitions are based on Roy model and are in general not directed to larger firms.[4] This is 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.

[3]This line of research starts from Gabaix and Landier (2008) and Tervio (2008), where competitive assignment models of the managerial labor market, absent an agency problem, are presented to explain why total compensation increases with firm size. By assuming CEO's effort has a multiplicative effect on both CEO utility and firm value, Edmans et al. (2009) embed a moral hazard problem into the competitive assignment model. The model quantitatively generates predictions on wealth performance sensitivities that are consistent with the data. Edmans and Gabaix (2011) extend the model further to risk-averse executives.

[4]In this line of literature, 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. While Gayle et al. (2015) include the career concern through human capital into the model, the channel does not explain the size premium in their estimation.

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). Builds on this literature, I embed an optimal contracting problem with moral hazard and two-sided limited commitment into an equilibrium search model. In doing so, the exogenous environment is endogenized and 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; Menzio and Shi (2010), Tsuyuhara (2016) and Lamadon (2016) among others construct theoretical models, estimate or calibrate structural models using directed search. Compared to this literature, I add a dynamic moral hazard problem which allows me to understand how market frictions influence a long-term contract. The extreme case in my model where firms with size below a threshold only pay a flat wage corresponds to this type of models.

This paper is also closely related to 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. Other than the differences in topics, there is a critical difference that distinguishes my model from theirs: the productivity (or output) of agents is persistent over time in my model while it is not in their model. Therefore, in my model, working hard today rewards the agent in the future labor market. 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.

The rest of the paper is organized as follows. In Section 2, I present the motivating facts of the firm-size premiums in compensation growth and performance-based incentives. I further show both premiums increase when the managerial labor market is more active, indicating they are related to the labor market competition. I then set up the model in section 3, where I characterize the optimal contract and explain the puzzles. Section 4 presents the reduced-form evidence. Section 5 estimates the model. Section 6 discusses the policy implications. Section 7 concludes.

2 Motivating Facts

The analysis of this paper is motivated by three stylized facts. I first investigate the *firm size premium* in performance-based incentives controlling for total compensation, firm performance, etc. I then illustrate that the firm size premium decreases as executives approach retirement age. Since age is a proxy for career concerns, this seems to indicate that the size premium is related to career prospectives. Finally, I bring the size premium directly to the managerial labor market. I trace the firm size premium across different industries and find the size premium is higher in industries where the managerial labor market is more active.

Size premium in performance-based incentives

I measure the pay-for-performance incentives by "delta", which is defined by the dollar increase in executives' firm related wealth for a percentage increase in firm value.[5] As is documented in Edmans et al. (2009) and replicated in table 1 column (1), "delta" is positively correlated with firm size: For 1% increase in firm size, measured by market capitalization, performance-based incentives increase by 0.59%. Edmans et al. (2009) argued that because executives in larger firms are paid higher, they require more incentive pay to induce effort.

However, this does not explain the whole size premium. The positive correlation between performance-based incentives and firm size continues to hold after controlling for total compensation in various forms in table 1 column (2) to (4) and extra controls in column (5).[6] For 1% increase in the firm scale, delta increases by more than 0.3%, which accounts for more than half of the size premium estimated in column (1). Indeed, checking the heatmap of *log(delta)* on total compensation and firm size in figure 1, we find that among executives with similar total compensation, those in larger firms get higher performance-based incentives. This is the *firm size premium in performance-based incentives*. This fact serves as a puzzle that urges an explanation. The next two facts hint that the size premium is related to the executive labor market.

Size premium decreases as executives approach retirement age

Started from Gibbons and Murphy (1992), *age* had been used as an indicator for career concerns. The older the executive is, the less influential that managerial labor market is on executive contract design. If the size premium is, at least partly, caused by the managerial labor market, we would expect the premium to decrease as executives become older. This is indeed true as shown in figure 2. The size premium 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.

[5] Delta is also known as "dollar-percentage incentive" or "wealth-performance sensitivity".

[6] In table 1 column (2) to (4), I control total compensation using *log(delta)*, delta divided into 50 equal sized groups, and delta divided into 100 equal sized groups, respectively. Other control variables in column (5) includes *operating profitability*, *market-book ratio*, *annualized stock return*, *director* defined by whether the executive served as a director during the fiscal year, *CEO* and *CFO* defined by whether the executive served as a CEO (and CFO) during the fiscal year, *interlock* defined by whether the executive is involved in the interlock relationship.

Size premium increases as the managerial labor market is more active

The effect of labor market not only varies over age, but also varies across industry. An industry is an appropriate sub labor market. As I will discuss in section 5, more than 60% job-to-job transiations are within the industry. As the final set of facts, in table 2, I use four proxies to measure how active the managerial labor market is, and test if the interactions between these proxies and firm size is significant.

The first two variables are job-to-job transition rates in each industry-year (Fama-French 48 industries and fiscal years). *EE 190* defines a job-to-job transtion by an 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 set to *90* days. Variable *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 last 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 to measure the generality of the CEO's human capital based on the CEO's lifetime work experience. [7]

In all regressions, I have controlled for total compensation, age dummies, year times industry fixed effects. The results show that firm size premium is increasing in all four measures of labor market activity. In words, the size premium is larger in industries/years where managerial labor market is more active.

In the following, I will start to build a model to explain the firm size premium in performance-based incentives. The key notation is labor market incentves. As the labor market becomes more active, labor market incentives are larger, the size premium in performance-based incentives is higher. The model, therefore, is consistent with the facts above.

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

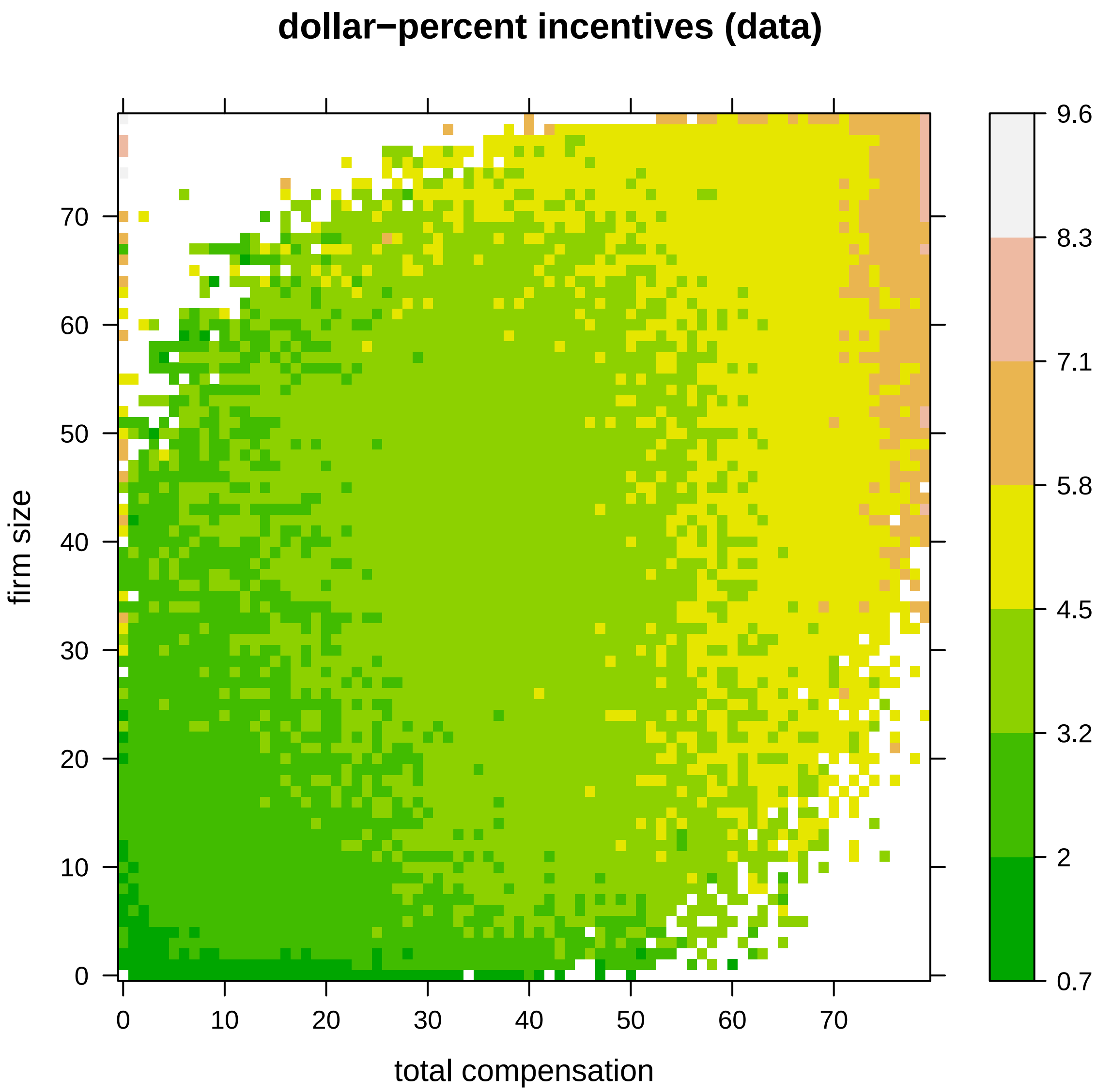


Figure 1: $\log(delta)$ over firm size and total compensation

Note: This is a heatmap of $\log(delta)$ on total compensation and firm size. *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. It is the variable `tdc1` in ExecuComp dataset. 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. These variables will be used throughout the paper. 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 1: Pay-for-performance 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 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

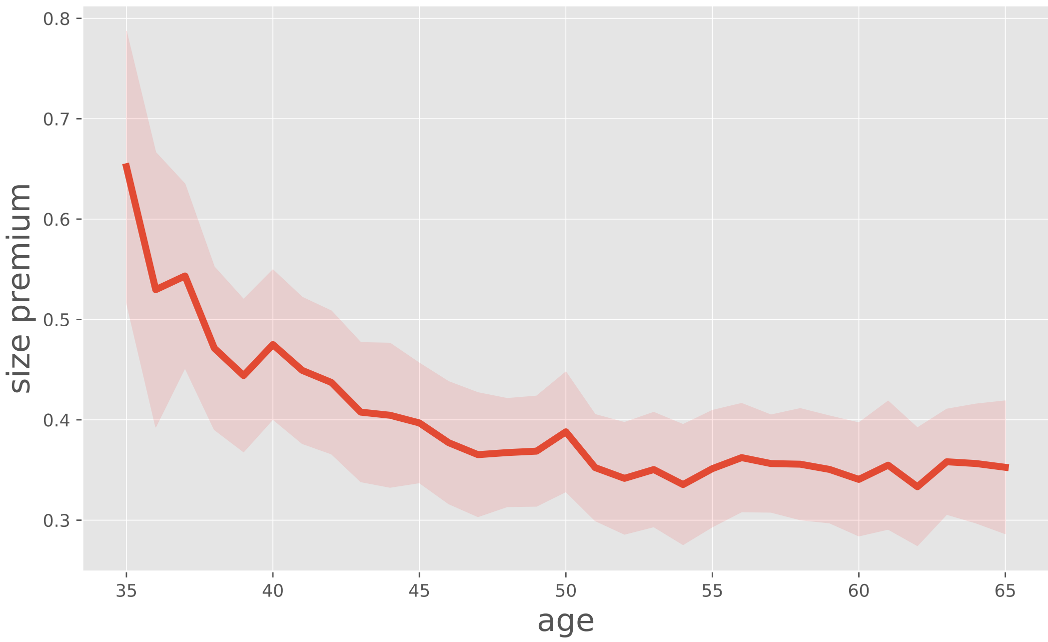


Figure 2: Size premium in performance-based incentives decreases in age

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

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

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.

Table 2: Size premium increase with managerial labor market compeititon

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

Note: This table reports evidence that the firm size 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 × 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 times industry fixed effects, *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 × fiscal year level) are shown in parentheses, and we denote symbols of significance by * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3 Model

I now construct an equilibrium model of the managerial labor market. The model is featured with strategic bargaining, on-the-job search and counteroffers. I embed the bilateral moral hazard problem into the labor market equilibrium. The model generates labor market incentives, a wedge between the total incentives required to motivate exeucutives and the pay-for-performance incentives offered by firms. Hence, the size premium in performance-based incentives is linked to labor market incentives. I now formally introduce the model.

3.1 Ingredients

Agents

There is a fixed measure of individuals. They are either employed as executives, or not employed 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 will enter the economy.

Individuals want to maximize expected lifetime utility,

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

where $\beta \in (0,1)$ is the discount factor, $\delta \in (0,1)$ 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, to firms that he meets, and can be carried with the individual through job-to-job transitions.[8]

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 it is not confusing, I will denote z_t by z, and z_{t+1} by z'.

While z and z' are observable to the firm, effort e is not. Hence, there is *moral hazard.*

[8] 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.

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:

$$E_\Gamma[z'g(z'|z)] < E_\Gamma[z'];$$

- taking effort is more likely to deliver a higher productivity (Monotone likelihood ratio property, MLRP):

$$g(z'|z) \text{ is non-increasing in } z'.$$

The other side of the managerial labor market are firms are characterized by the scale of assets, called firm size, denoted by $s \in \mathbb{S} = [\underline{s}, \bar{s}]$. Firm size is observable to all agents, and it is permanent and exogenous.[9] A match between a worker of productivity z and a firm of size s produces a flow of output $f(s,z) = \alpha sz$, where the constant $\alpha \in (0,1)$. This function form entails that executive's efforts rolled out across the entire firm up to a scale of α. Hence, a manager has a greater effect in a larger company.[10]

Managerial labor market

Labor market is where individuals are matched to executive jobs. Executives and jobs are imperfectly informed about executive types and the location of the job. The search friction precludes the optimal assignments assumed in Gabaix and Landier (2008). Yet, when two agents meet, both are informed about each other's types, i.e. firm size and executive productivity.

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 a partial equilibrium.

[9] 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 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 interpret firm size as the total number of employees, then it can be endogenized by modeling the labor market of normal workers.

[10] There has been a discussion on the appropriate production function form 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 $f(s,z) = s + z$ implies the effect of executives on firm value is independent of firm size. This specification is appropriate for a perk consumption. An multiplicative production function such as $f(s,z) = y(s)z$ where $y(s)$ is increasing in s, is appropriate for executives' actions that can roll out across the entire firm to some extent and thus have a greater effect in a larger company. The latter is the function form adopted here.

When a candidate meets a firm, they bargain on a contract. Suppose the continue 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 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 to transit to the outside firm. I allow incumbent firm to respond to outside offers: a sequential auction is played between the executive and both jobs as in Postel-Vinay and Robin (2002). If the poaching firm is larger, the executive moves to the alternative firm. This is because the poaching firm can always pay more than the current one can match. Alternatively, if the alternative firm is smaller, then the executive may use the outside offer to negotiate up his compensation. This process will be illustrated using value functions in the next section.

This sequential auction mechanism is the core of the labor market competition in this paper. It highlights that executives in larger firms usually experience higher compensation growth because their firms are able to match more and better poaching offers. Consequently, these executives are less sensitive to performance-based incentives.

Timing

I now state the timing with elements introduced. Time is discrete, indexed by t and continues forever. The period of an executive candidate is simple — he is matched to a firm with some probability and starts with the contract of a continuation value W^0. Now consider an on-the-job executive enters a period with his beginning of period productivity z and current firm of size s. The timing line is shown in figure 3.

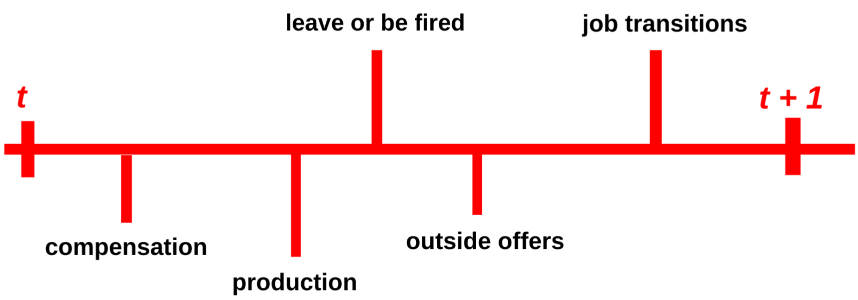


Figure 3: Timing

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

2. **Production:** Then the executive enters the production phase. He 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 $f(z, s)$ not the effort e. This is the moral hazard problem.

3. **Labor Market:** With probability δ the executive dies, otherwise with probability λ, 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 get a higher compensation, or transit to the poaching firm. The value of the contract to the executive therefore will be determined by a sequential auction between the current and poaching firms.

The transfers w, effort choice e and job-to-job transitions along the time ling are be 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 poaching firm size. Denote $\tilde{s} = s_o$ if there is no poaching firm. The history of productivity and poaching firms $h^t = (h_1, h_2, .., h_t)$ is common knowledge to the executive and the firm, and is fully contractible.

The two elements included in the history — productivity and poaching firms — correspond to the frictions we have in the contracting problem — moral hazard and countering outside offers. On the one hand, while the productivity is included in the history and is contractible, the manager's effort is not and needs to be induced by incentives. Hence, an incentive compatibility constraint is required. On the other hand, by having the information of bidding firms s' in the history, I allow the contract to stipulate whether and how to counter outside offers, conditional on the size of the competing firm. In this way, the competition for executives is introduced into the contract.

However, countering outside offers should be optimal (or subgame perfect in terms of game terminology), it is therefore necessary to allow limited commitment for both sides — terminate the contract when the surplus is negative. Firms can not commit to the relationship if the profits are negative. This happens when the outside offer comes from a larger firm. In this case, the firm's participation constraint binds and the match separate. Likewise, executives can not 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 of a poaching firm. In the former case, the manager leave the firm voluntarily. In the latter case, the manager transfer to the poaching firm.

Given the information structure, I define a feasible contract by a plan that stipulates the compensation $w_t(h^{t-1})$, the 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 the participation constraints of both sides and incentive compatibility constraint.

To further simplify, I impose two assumptions. First, I assume taking effort $e = 1$ is always optimal. This 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. Under the hook, I require a relative small effort cost. 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.[11]

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

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

where e is the effort level suggested by the contract, w is the compensation, and W is the promised value given for productivity z', and Φ is the set of feasible and incentive compatible expected utilities that can be derived following Abreu et al. (1990).[12]

3.2 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 (Postel-Vinay and Robin 2002) using value functions.[13] Let $\Pi(z, s, V)$ denote the discounted profit of a firm with size s, executive of

[11] 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 will be analyzed in the future.

[12] 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.

[13] 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.

beginning of period productivity z, and promised value to the executive V. The maximum the firm would like to give to the executive, the maximum bidding value $\overline{W}(z,s)$, is 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 demands a value higher than $\overline{W}$. To simplify notations later, I also define $\overline{W}(z,s^o) = W^0$. This means when there is no outside offer, manager's outside value is simply W^0.

The sequential auction works as follows. When a size s firm's executive receives an outside offer from a size $\tilde{s}$ firm, both firms enter a Bertrand competition won by the larger firm. Consider this sort of auction over an executive z' by a firm of size s and one of size $\tilde{s} > s$. Since it is willing to extract a positive marginal profit out of every match, the best the firm s can do is to provide a promised utility $\overline{W}(z',s)$. Accordingly, the executive accepts to move to a potentially better match with a firm of size $\tilde{s}$ if the latter offers at least the $\overline{W}(z',s)$. Any less generous offer on the part of the size $\tilde{s}$ firm is successfully countered by the size s firm.

Now, if $\tilde{s}$ is less than s, then $\overline{W}(z',s) > \overline{W}(z',\tilde{s})$, in which case the size $\tilde{s}$ firm 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 the size $\tilde{s}$ firm.

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 executives in the 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' and $\tilde{s}$. For the ease of notations, I denote the effective discount factor $\tilde{\beta} = \beta(1-\delta)$, and write the mixture distribution of outside offers as

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

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 manager 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 (happens with probability δ) or transits to another firm (happens with probability $\sum_{\tilde{s} > s} \tilde{F}(s)$.

The promise keeping constraint (PKC) makes sure that the choices of the firm honors 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 continue value of taking effort is higher than not taking effort. This create incentives for the manager to pursue the shareholders' interests rather than his own. As described above, I assume that $e = 1$ is the optimal choice of the firm, consistent with the data and literature (Gayle et al., 2015).

Finally, the participation constraints are stated in (PC-E) and (PC-F). The firm commits to the relationship as long as the promised value for the future is not more than $\overline{W}(z', s)$. The sequential auction pins down an outside value of the manager, $\min\{\overline{W}(z', \tilde{s}), \overline{W}(z', s)\}$. The term $\overline{W}(z', \tilde{s})$ show how the market competition gets involved into the optimal contract. As the executive is hit by outside offers over time, the promised utilities will be lifted up due to the term $\overline{W}(z', \tilde{s})$.

3.3 Equilibrium definition

Before turning to characterize the optimal contract, I define the equilibrium. An equilibrium is the executive unemployment value W^0, the value function of employed executives W satisfies (PKC), the profit function of the firms Π and an optimal contract policy $\sigma = \{w, e, W(z')\}$ for $z' \in \mathbb{Z}$ that solves the contracting problem (BE-F) with associated constraints (PKC), (IC), (PC-E) and (PC-F), the stochastic process of executive productivity Γ 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 in Menzio and Shi (2010).

Proposition 1. *The equilibrium exists.*

3.4 Contract characterization

In this section, I derive characterizations of the optimal contract. These results build on and extend the dynamic limited commitment literature, pioneered by Thomas and Worrall (1988) and Kocherlakota (1996), 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. 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 present and future compensations. Equation (3) says, abstract from participation constraints, the continuation utility $W(z')$ only changes to motivate manager's effort. In the extreme case that IC constraint is not binding ($\mu = 0$, μ is the multiplier of IC), $W(z') = V$ keeps constant. Thus, the wage is also a constant. Generally, higher V induces a higher $W(z')$. Thus, an optimal dynamic contract has some memory.

However, when the outside offers are realized such that the participation constraint binds, the contract disposes of all history dependence and makes the continuation value depend only on the current state $(z',s,\tilde{s})$. 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 $W(z')$. Notice these results are not by assumption, but are stipulated in the optimal contract.

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 manager's productivity z' which is stochastically determined by the effort in the past. Therefore, the boundaries of participation constraints carries memory of the past effort choice. This is where the labor market incentives come into effect.

3.5 Explain why do performance-based incentives increase in firm size

With the characterization of the optimal contract, we are ready to explain the size premium in performance-based incentives. We start

Define incentives

Let me start with an explicit definition of "performance-based incentives" and "labor market incentives" in terms of utilities in the model. I define two sets of poaching firm size $\tilde{s}$ depending on whether it is larger than the current firm.

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

For poaching firms belonging to set $\mathcal{M}_1$, the executive will transit to such firms and receives the full surplus of his previous job $\overline{W}(z, s)$. For poaching firms 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 that they can not be used to negotiate compensation with the incumbent firm.

Accordingly, the incentives are separated into three parts, incentives brought by larger firms in $\mathcal{M}_1$, by smaller firms in $\mathcal{M}_2$, and by the incentive pay when there is no poaching firms from either $\mathcal{M}_1$ or $\mathcal{M}_2$,[14]

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

where $\mathcal{I}\Big(W(z')\Big) \equiv \int_{z'} W(z')(1 - g(z, z'))\Gamma(z, z')$ is an "incentive operator" that defines on a continuation utility scheme and calculates the incentives one get from it.

IC$'$ has an incentive part offered by the current firm, the *performance-based incentives*, denoted by Ξ_p,

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

and an incentive part due to the labor market competition, the *labor market incentives*, denoted by $\Xi_m(s)$,

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

[14] We can similarly rewrite the Bellman equations of firms and executives using the optimal continuation value, and these equations are consistent with Postel-Vinay and Robin (2002).

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

$\Xi_m(s)$ is an expectation of the incentives an executive receives from all possible bidding frontiers as long as the bidding frontier is higher than $W(z')$. When the poaching firm is larger, the executive receive the incentive of his current firm' bidding frontier, and when the poaching firm is smaller, the incentive comes from the bidding frontiers of the poaching firm. Notice that $\Xi_m(s)$ depends on the firm size even though the moral hazard problem fundamentally doesn't.

The lower bound of $\mathcal{M}_2$ determines whether a poaching offer is competitive, which itself is determined by the current promised continuation values $W(z')$ — the larger $W(z')$, the less likely a poaching firm can be used to negotiate up the value, the smaller the set $\mathcal{M}_2$ will be, the lower the labor market incentives are. For instance, when $W(z')$ equals to the bidding frontier $\overline{W}(z', s)$, the set $\mathcal{M}_2$ becomes empty. Labor market incentives only exist if the poaching firm is larger than s.

Bases on this, here is a simple "job ladder" explanation for the size premium when total compensation is not controlled.[15] In this model, as in the frictionless matching model of Gabaix and Landier (2008), executives from larger firms tend to have higher total compensation, therefore the lower bound of $\mathcal{M}_2$ is higher, and the labor market incentives are lower. As a result, more performance-based incentives in compensation are required. Essentially, executives in larger firms are higher on the job ladder and there are smaller moving-up space to further advance careers, hence they get less labor market incentives. As we will see immediately in the next subsection, this "job ladder" argument also applies to explain the size premium among executives receiving the same total compensation.

How do labor market incentives change with firm size?

Consistent with the stylized facts in Section 2, here I compare incentives between two executives with the same total compensation. One is in a smaller firm s_1, the other is in a larger firm s_2, $s_1 < s_2, s_1, s_2 \in \mathcal{S}$. 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 income. Although the two executives have the same total compensation, the one in s_2 has slightly higher life-time utility. Figure 4 illustrates the possible poaching firms for the two executives and the associated incentives.

To see that why labor market incentives decrease in firm size, i.e. $\Xi_m(s_2) < \Xi_m(s_1)$, I

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

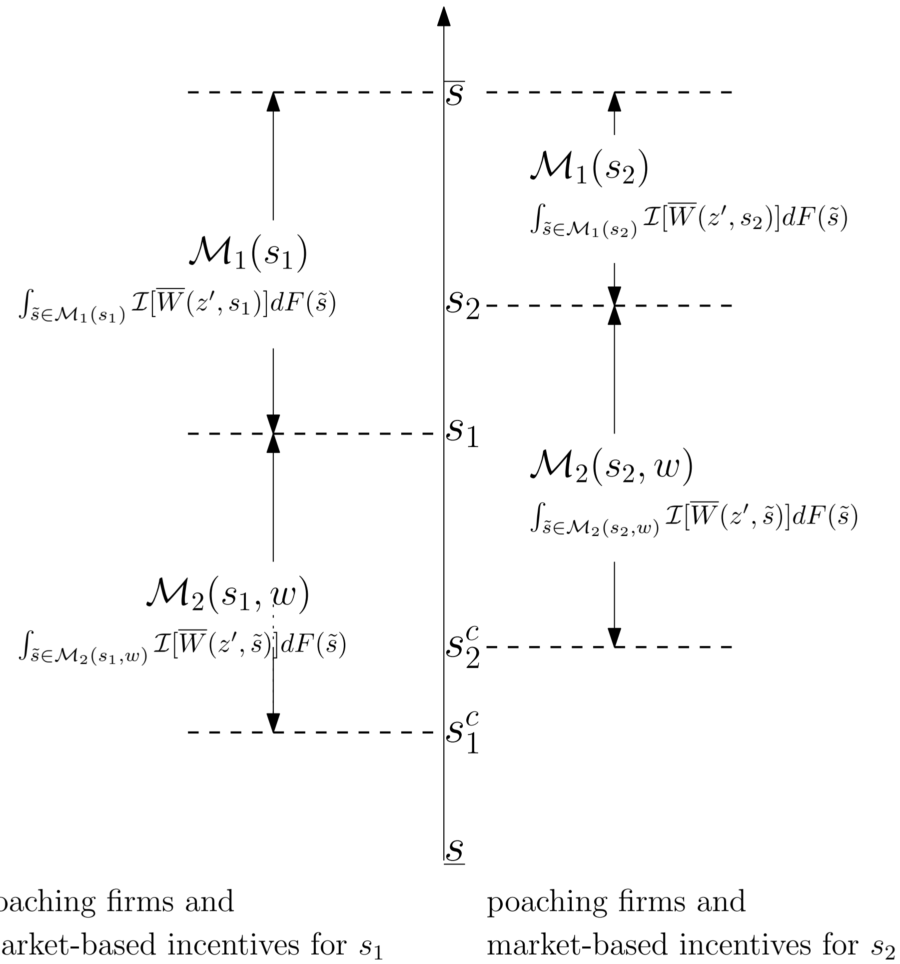


Figure 4: Compare labor market incentives

Note: The figure illustrates the 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 $[\underline{s}, \bar{s}]$. 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.

take the difference

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

The three items are due to two reasons. The first item is due to the difference in positions on the job ladder. This extends the argument of last subsection that executives in larger firms, though have the same total compensation, are higher on the job ladder, thus receive less incentives from the managerial labor market. The last two items are due to the difference in bidding frontiers. Apart from the different positions on the job ladder, s_2 is able to counter better outside offers. As a result, the labor market incentives for s_2 are derived from higher bidding frontiers.

Clearly, if higher bidding frontiers yield less incentives, then $\Xi_m(s_2) < \Xi_m(s_1)$. Numerically, this is true as long as the utility function is concave enough. In the following, I will analytically derive a sufficient condition when c equals to a particular value.

Proposition 3 (Labor market incentives and Firm size)**.** *Suppose the executives' utility is of the CRRA form,*

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

and the cost of effort $c = \bar{c}(s)$*,*

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

then $\mathcal{I}\left(\overline{W}(z', s)\right)$ *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*.*

Proof. See Appendix A. □

Immediately, (6) and (7) implies that labor market incentives $\Xi_m(s)$ decrease in firm size under the conditions stated in proposition 3. Numerically, I find that in almost everywhere of the parameter space that are explored in the estimation, the right hand side of (7) is approximately 1, so the condition is simplified as $\sigma > 1$.

To understand the intuition, first notice that $\mathcal{I}\left(\overline{W}(z', s)\right)$ is simply a weight sum of $\frac{\Delta \overline{W}(z',s)}{\Delta z'}$ over the domain of z' — the steeper $\overline{W}(z', s)$ is with respect to z', the larger 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,

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

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.

It is clear that 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 managerial 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.

Condition (7) is consistent with the literature in this context. If any, previous 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) use $\sigma > 1$. Using an employer-employee matched data from Sweden for the general labor market, Lamadon (2016) estimates that $\sigma = 1.68$.

4 Empirical Evidence

In this section, I present evidence that is consistent with the mechanism of the model. First, I examine executives' job-to-job transitions, and whether they climb the job ladder towards larger firms. Then I explore whether the job-to-job transition rate decreases with firm size and compensation as predicted by the model. Finally, in the model the larger the current firm is, the more that executives' compensation grows due to the between-firm competition. I test whether this can be rejected in the data.

4.1 Data

The empirical analysis and estimation mainly rely on information from ExecuComp 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. Finally, I get job-turnover information by matching BoardEX database with ExecuComp.

Our sample comprises $35,088$ executive episodes with age between 30 and 65.[16] Of

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

these, 26,972 episodes cover the full tenure of the executive from beginning to end. The total number of executive-fiscal year observations in our 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.

Using information from ExecuComp, we identify the gender, *age* of executive in each year, the *tenure* in the current executive episode, whether he 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, variable *salary* only contains the base salary, *totalcurr* also contains bonus, and *tdc1* is the total compensation including salary, bonus, values of stock and option granted, etc. All compensation variables are in thousand dollars. On average, the base salary is 424 thousand dollars per year, and with bonus added, the average is 706 thousand dollars per year. 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 the performance-based incentives are in the firm-related wealth. It is defined by the dollar change in wealth associated with a 1% change in the firms stock price (in \$000*s*). The distribution of $delta$ is right-skewed, with a mean of 323 thousand dollars, even larger than its 75th percentile of 154 thousand dollars.

For the firm side information, I use market capitalization *mkcap*, the market value of a company's outstanding shares, to measure the firm size. In some robustness checks, 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.[17]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 match 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. By this way, I am able to identify most of executives in ExecuComp, 32,864 executives in total.

[17]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

I define job-to-job transition by the executive leaves the current firm and starts 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 6.92% each year over 1992 to 2015, while the job exit rate is slightly lower 6.48%. Figure 5 illustrates how job-to-job transition changes with age and figure 6 shows how job exit changes with age. To illustrate the trend, the figures also include those who did not retire after age 65. As show 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 very low before 55, and peaks sharply at the age 65 as expected.

Finally, 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 Fama-French 12 industry classification, and 1407 out of the 2567 cases by Fama-French 48 industry.

position in one *S&P* firms or simply retires. Moreover, the start and end dates of the current employment are also not known.

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
salary	218168	424.278	290.884	241.903	351.346	525.208
totalcurr	176856	706.41	990.367	323.869	485.251	784.222
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. *salary* (in $000*s*) is the dollar value of the base salary earned by the named executive officer during the fiscal year. *totalcurr* (in $000*s*) is the dollar value of the base salary and bonus earned by the named executive officer during the fiscal year. *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.

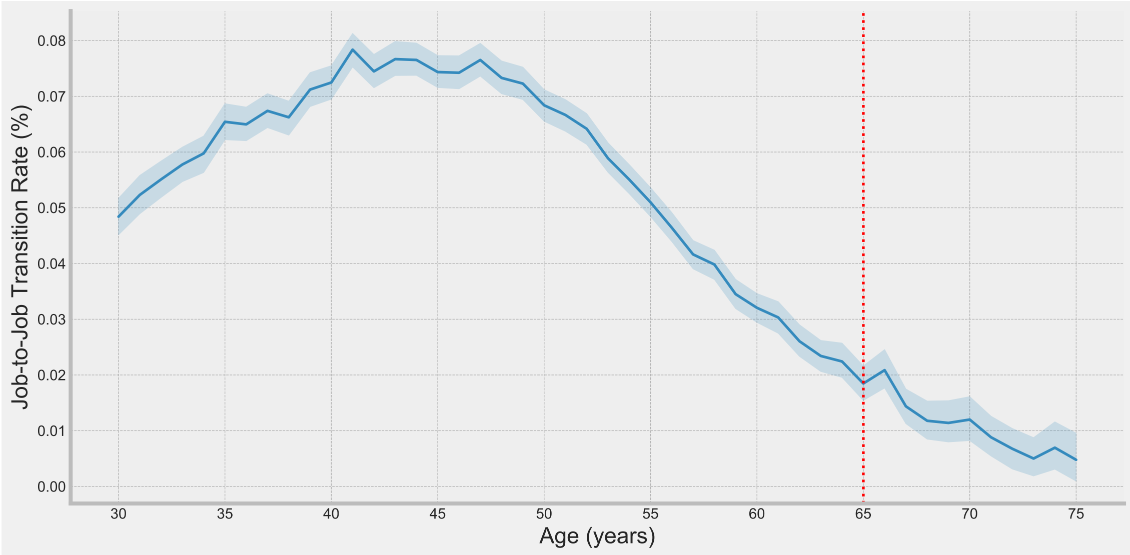


Figure 5: Job-to-job transition rate over age

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 the executive leave the current firm and starts to work in another firm within 190 days. The identification of job-to-job transition is based on BoardEX data.

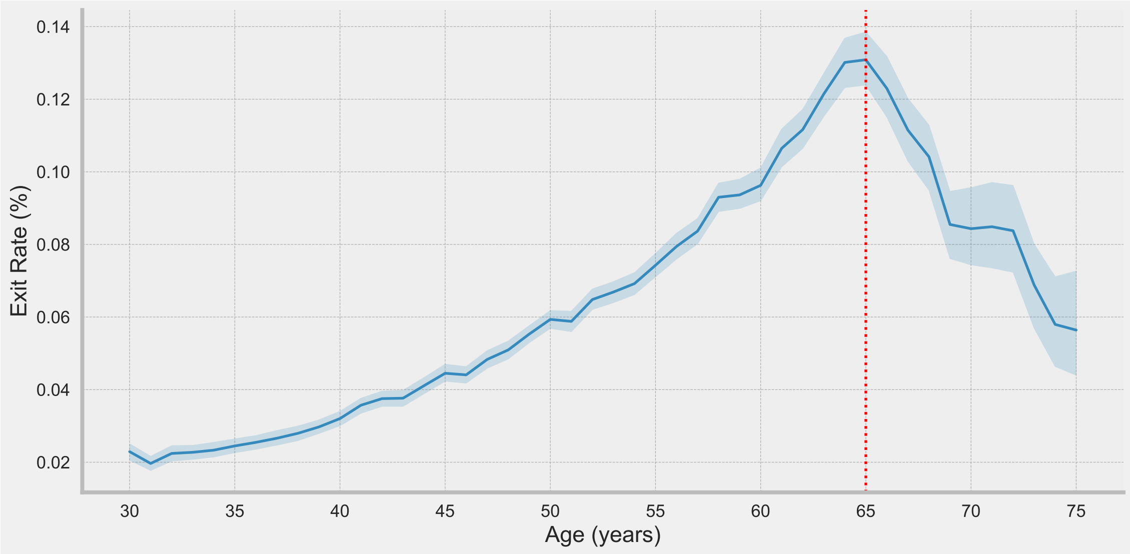


Figure 6: Exit rate over age

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 the executive leave the current firm and does not work in another firm within 190 days. The identification of exit is based on BoardEX data.

4.2 Empirical evidence of the key mechanisms

Do most executives transit towards larger firms? Are executives with higher compensation and/or in larger firms less likely to have job-to-job transitions? Do executives in larger firms experience more compensation growth? The model confirms all of these question, here I provide data evidence for each of them.

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 this data, approximately 60% job-to-job transitions are associated with 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, 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 7 depicts the distribution of the change of firm size upon a transition. While most of the transition are 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 managerial labor market where executives engage in random on-the-job search.

Job-to-job transitions decrease in firm size

As a first pass, figure 8 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 and executive compensation, I estimate a Cox model on how firm size and total compensation changes the duration to job-to-job transitions, controlling for executive age, firm performance indicators, year and industry dummies. For 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 consistent with the model, when the total compensation increase by 1%, the hazard rate decreases by 27%.

Executives of larger firms tend to experience higher compensation growth

I start with on a discrete version by dividing the whole sample into four groups: "large and high profit", "large and low profit", "small and high profit" and "small and low profit". I label a firm as a "large" firm if the market capitalization is higher than the industry median, and label a firm as "high profit" if the operating profitability is higher than the industry median. Taking the small and low profit group as the control group, and other groups as treatment groups. For each executive in one of the three treatment groups, I use the nearest neighborhood matching to find a matched executive in the control group, based on three variables: the fiscal year that the executive starts the episode, the tenure in the episode, and the total compensation of the first year of the episode. By doing so, I can compare growth of total compensation over tenures between treatment groups and the control groups. The results are depicted in the upper panel of figure 9. The lower panel adds to each line the wage growth trend of the control group. As figure 9 shows, executives in large firms experience significant a higher compensation growth compared to those in small firms, while the performance seems not so important.

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%)

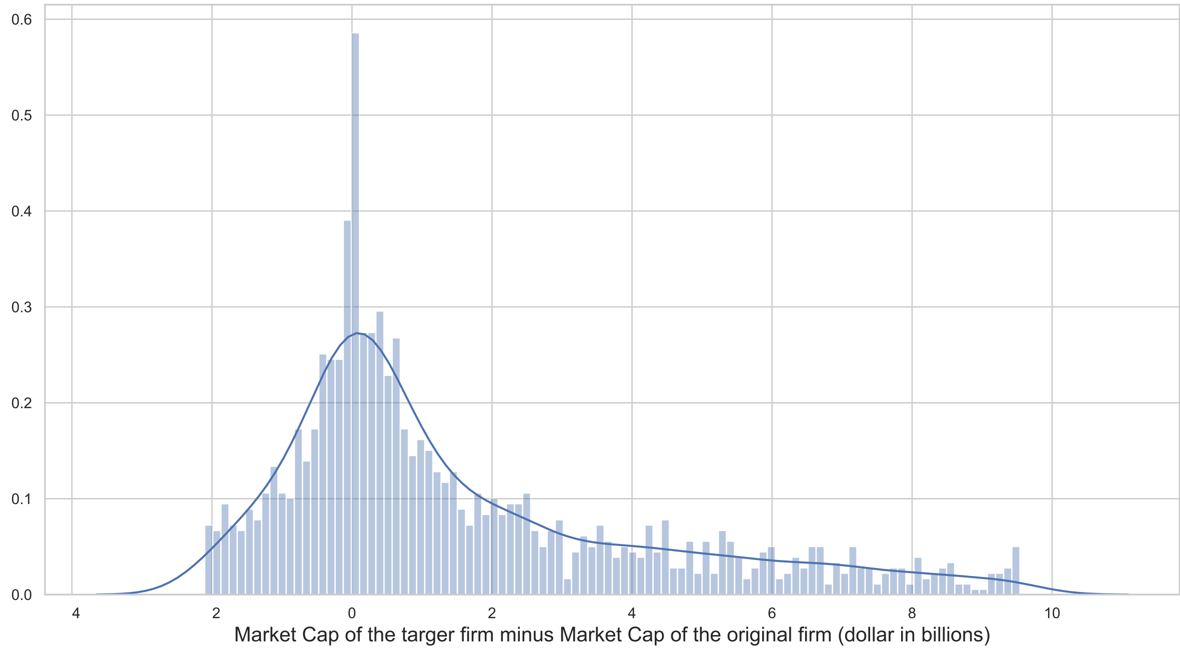


Figure 7: Distribution of change of firm size upon job-to-job transitions

Note: This bar plot depicts the distribution of change in firm size (measured by market capitalization in billions) and the kernel estimates. A job-to-job transition is defined as the executive leave the current firm and starts to work in another firm within 190 days. The identification of job-to-job transition is based on BoardEX data.

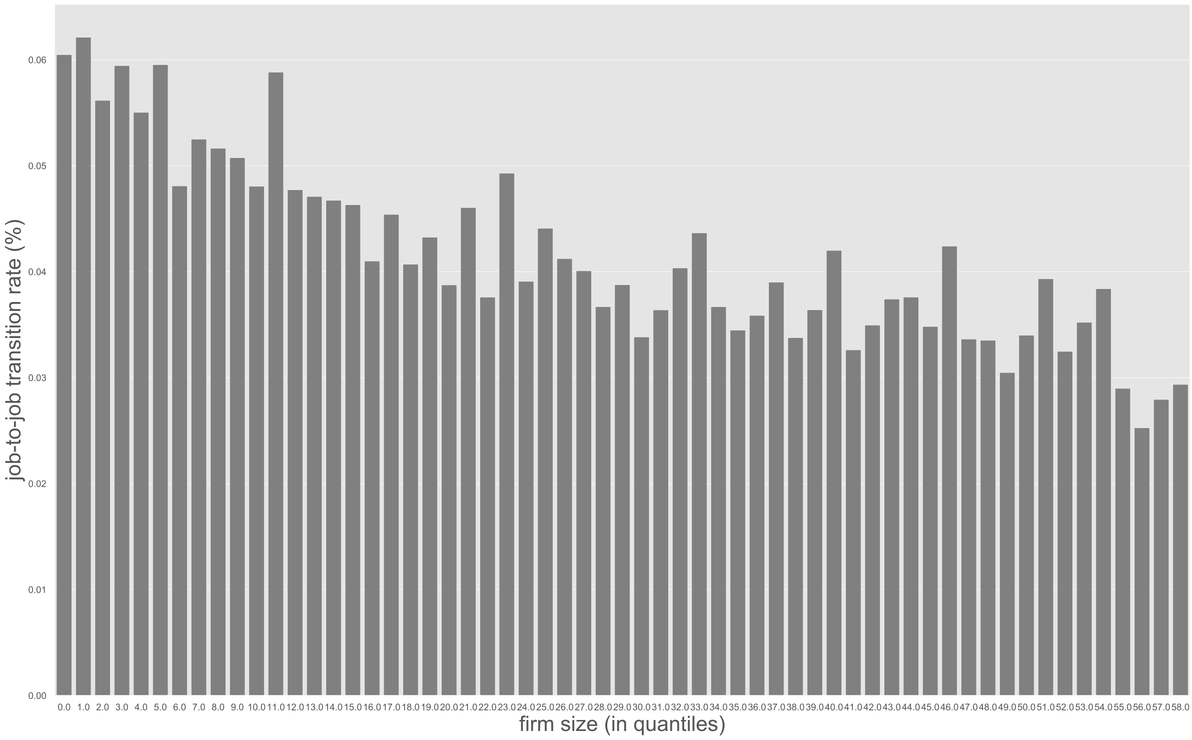


Figure 8: Job-to-job transition rate across firm size

Note: The figure depicts the estimates of job-to-job transition rates across firm size quantiles. A job-to-job transition is defined as the executive leave the current firm and starts to work in another firm within 190 days. The identification of job-to-job transition is based on BoardEX data.

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)
Market-Book Ratio	0.942****	0.939****
	(0.0150)	(0.0157)
Market Value Leverage	1.033**	1.035**
	(0.0139)	(0.0142)
Profitability	0.913****	0.905****
	(0.0197)	(0.0199)
Year FE	Yes	Yes
Industry FE	Yes	Yes
N	154635	118119
chi2	496.1	491.4

Note: The standard error are shown in parentheses, and I denote symbols of significance by * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. I estimate a Cox proportional hazards model with the event of job-to-job transitions. 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. Other variables are defined as follows. `Firm Size` is the market capitalization of the firm, $\texttt{prcc\_f} \times \texttt{csho}$. `Market-Book Ratio` is defined by `market value of assets/total book assets`, where `market value of assets` is calculated using the formulate `market value of assets` $= \texttt{prcc\_f} \times \texttt{cshpri} + \texttt{dlc} + \texttt{dltt} + \texttt{pstkl} - \texttt{txditc}$. `Profitability` is defined by `EBITDA/at`, where `EBITDA` stands for earnings before interest, taxes, depreciation and amortization, and `at` stands for total assets. All variables are adjusted by GDP deflater.

Table 6: Total compensation growth and firm size

	log(tdc1) (1)	(2)	(3)
profit	0.183***	0.339***	0.556***
	(0.0184)	(0.0555)	(0.0711)
l. log(mkcap)	0.124***		
	(0.0102)		
l. log(sale)		0.113***	
		(0.00716)	
l. log(at)			0.121***
			(0.00747)
profit × *l. log(mkcap)*	0.0336***		
	(0.00792)		
profit × *l. log(sale)*		0.0544***	
		(0.00837)	
profit × *l. log(at)*			0.0934***
			(0.0125)
l. log(tdc1)	0.681***	0.692***	0.680***
	(0.0192)	(0.0167)	(0.0171)
N	91054	89771	89793
adj. R^2	0.666	0.654	0.656

Note: The standard error are shown in parentheses, and I denote symbols of significance by * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. I estimate a Cox proportional hazards model with the event of job-to-job transitions. 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. Other variables are defined as follows. `Firm Size` is the market capitalization of the firm, `prcc_f` × `csho`. `Market-Book Ratio` is defined by `market value of assets/total book assets`, where `market value of assets` is calculated using the formulate `market value of assets` = `prcc_f` × `cshpri` + `dlc` + `dltt` + `pstkl` − `txditc`. `Profitability` is defined by `EBITDA/at`, where `EBITDA` stands for earnings before interest, taxes, depreciation and amortization, and `at` stands for total assets. All variables are adjusted by GDP deflater.

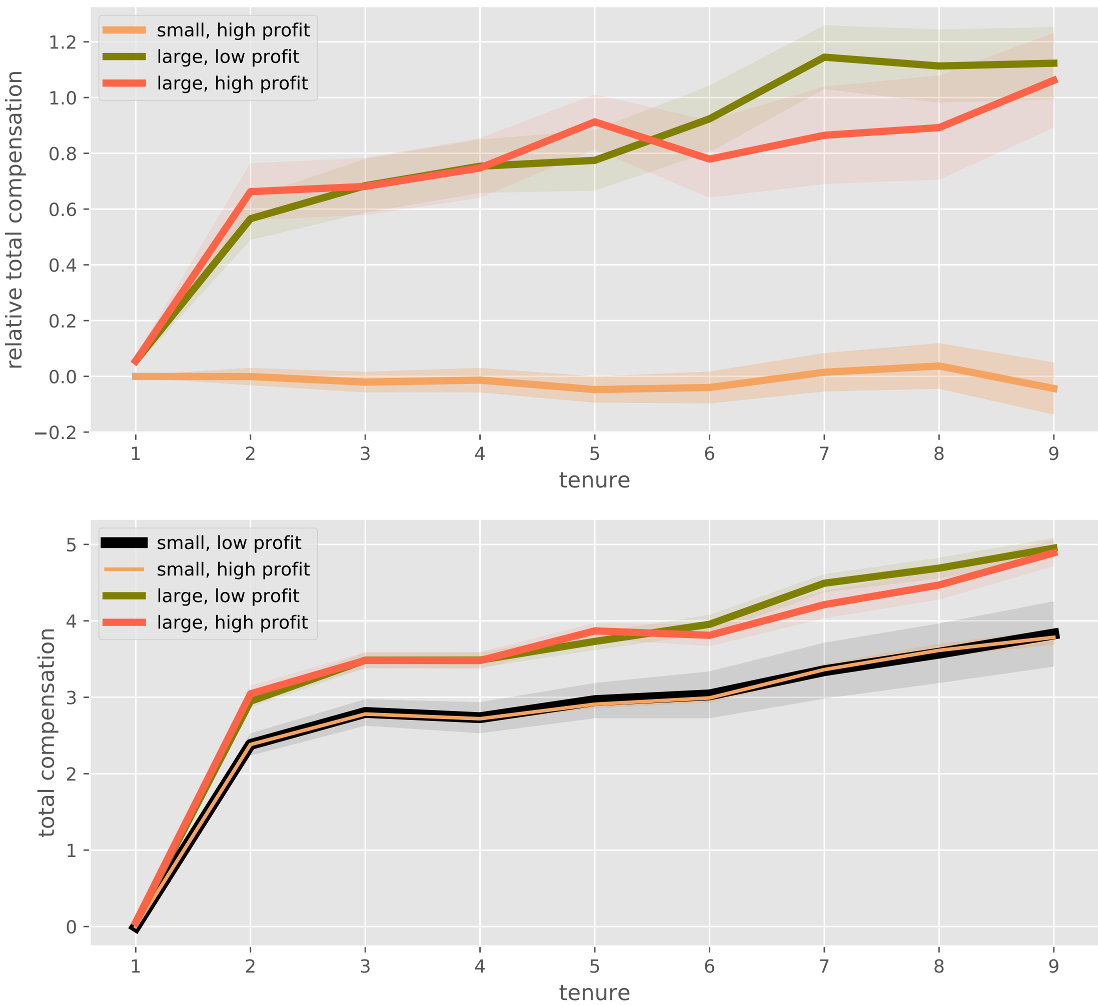


Figure 9: Executives of larger firms tend to experience higher compensation growth

Note: The figure depicts the estimates of total compensation growth. I divide the whole sample into four groups: ”large, low profit” firms, ”large, high profit” firms, ”small, low profit” firms and ”small, high profit” firms. In the upper pandel,

5 Estimation

I estimate the model's parameters using Simulated Methods of Moments. That is, I use a set of moments that are informative for the model's 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 numerical method that I employ to solve the dynamic contracting problem. Then I describe the model specifications and moments used for identification. After reporting the parameter estimates, I conclude the section a visual comparison of pay-for-performance sensitivity in the data and the model.

5.1 Numerical Method

To solve the contracting problem, one need to find 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 in order to enforce an incentive compatible allocation. This Pareto weight becomes the 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 the [Appendix].

5.2 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 $f(z,s) = \alpha_0 s^{\alpha_1} z$. 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 5 points.[18] Furthermore, I set the sampling distribution of firm size $F(s)$ a log-normal distribution with expectation of μ_s and standard deviation of σ_s. 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 7 lists the complete set of parameters that I estimate.

Table 7: Parameters

Parameters	Description
δ	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

5.3 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 δ. 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,

$$\texttt{profit}_{it} = \beta_0 + \rho_z \texttt{profit}_{it-1} + \epsilon_{it,0},$$

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, $\log(size)$ is larger. Similarly, the higher σ_s is, the more heterogeneous the outside firms are, both mean and variance of $\log(size)$ increases.

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

[18] This is for speed of estimation. However, the simulated moments are very robust to this choice.

compensation. Therefore, α_0 and α_1 are identified by the mean and variance of $\log(tdc1)$ and $\beta_{wage-size}$ in following regression of log wage on log firm size,

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

The final part of the identification concerns the parameters σ and c. These parameters governs the level of incentive pay and how the incentive pay changes with total compensation. To be consistent with the incentive measurement delta in the data, I construct in the simulated data a "delta" variable defined by the dollar change in wage 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,

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

and use $\beta_{delta-wage}$ to inform σ.

Predict The Firm Size Premium

The most important moment in this quantitative exercise is the the firm-size premium in the performance-based incentives. In particular, it is the coefficient $\beta_{delta-size}$ in the following regression,

$$\log(delta_{it}) = \beta_3 + \beta_{delta-size} \log(size_{it}) + \text{wage dummies} + \epsilon_{it,3}.$$

I intentionally leave it untargeted in the estimation, but use it examine if the model mechanism is able to match up with it.

5.4 Results

Panel A in table 8 reports the targeted values of moments in the data and the corresponding values in the estimated model. The last two columns lists the parameter estimates and the standard errors. While I arranged moments and parameters along the identification argument made in the previous subsection, all parameters are estimated jointly. Panel B is the prediction of $\beta_{delta-size}$ based on the estimates.

Overall, the model provides a decent fit to the data. In particular, even without targeting on $\beta_{delta-size}$, the model is able to quantitatively capture the negative correlation between delta and firm size. In the model, this relationship is entirely driven by the labor market incentives. There is nothing mechanical that forces these two measure to coincide. The fact that the predicted $\beta_{delta-size}$ matches up so closely with the counterpart in the data is reassuring for the model mechanism to play an important role in explaining the firm size premium.

Table 8: Moments and Estimates

A. Targeted Moments

Moments	Data	Model	Estimates	Standard Error
Exit Rate	0.0691	0.0691	$\delta = 0.0691$	0.0012
EE Rate	0.0523	0.055	$\lambda_1 = 0.2759$	0.0017
$\hat{\rho_z}$	0.8111	0.5499	$\rho_z = 0.7$	0.0036
Mean(z)	0.1284	0.1763	$\mu_z = 0.06$	0.0006
Var(z)	0.0141	0.0141	$\sigma_z = 0.12$	0.0014
Mean(log(size))	7.44379	8.7934	$\mu_s = 1.7847$	0.228385
Var(log(size))	7.44379	8.7934	$\sigma_s = 1.3982$	0.0314657
Mean(log(wage))	7.17714	6.5241	$\alpha_0 = 1.7847$	0.228385
Var(log(wage))	7.17714	6.5241		
$\beta_{wage-size}$	0.370295	0.3196	$\alpha_1 = 1.7847$	0.228385
$\beta_{delta-wage}$	0.717209	2.1228	$\sigma = 2.50748$	0.0046
Mean(log(delta))	4.01842	3.8080	$c = 1.91385$	0.0259
Var(log(delta))	4.01842	3.8080		

B. Untargeted Moments

Moments	Data	Model
$\beta_{delta-size}$	0.297673	0.2941

Looking into the estimates, a job arrival rate $\lambda_1 = 0.2759$ is required to match the job-to-job transition rate in the data. The magnitude of λ_1 indicates that, on average, the executive will receive an outside offer every three years. The levels of $\log(wage)$, $\log(size)$ and coefficient $\beta_{wage-size}$ are matched reasonably well, showing the on-the-job search and sequential auction in the model capture the main features of the executive labor market. The magnitude of μ_s indicates that most offers are provided by relative small firms, though σ_s implies the variation is high.

The optimal dynamic contracting provides a reasonable prediction on the mean of $\log(delta)$, and the slop between delta and total compensation $\beta_{delta-wage}$. A further inspection of the simulated data and the real world data shows that in the real data, there is a large number of observations which have a small `tdc1` and a high delta, while in the simulated data, a low wage is usually associated with a small delta. Hence, there are some heterogeneity of executive-firm match that are not captured in the model. Finally, the exogenous processes on productivity and outside offers are matched quite well.

5.5 Compare Data and Model-Simulated delta

I conclude this section by providing some visual comparisons of delta from the dataset and the simulated data based on the estimates. The comparison are along two dimensions: total compensation and firm size. In both the real data and simulated data, I create a variable wage group which divides the sample into 100 groups according to the value of wage. Similarly, I create a variable size group based on the firm size. In this way, data is segmented into 100×100 cells. I calculate the mean of $\log(delta)$ in each cell and exclude those cells with less than 10 observations. The left panel of figure 10 shows $\log(delta)$ from the real data, and the right panel of figure 10 shows $\log(delta)$ of the simulated data.

It is clear that in both figures, delta increases with wage in a similar manner. The upper-left corner represents executives in big firms with low total compensations. This part is missing in the real data because only top 5 executives are observable.

As in the real data, the simulated data with labor market incentives has enough variation of delta along firm size within each wage group. In contrast, in the simulated data ignoring labor market incentives, there is little variation along firm size within each wage group, though the variation along wage group exists. Moreover, it seems the model (with labor market incentives) can fit the data better for low and medium wage levels. There are still large variations along the firm size for high wages in the data, while the model does not have enough labor market incentives to generate that large variations.

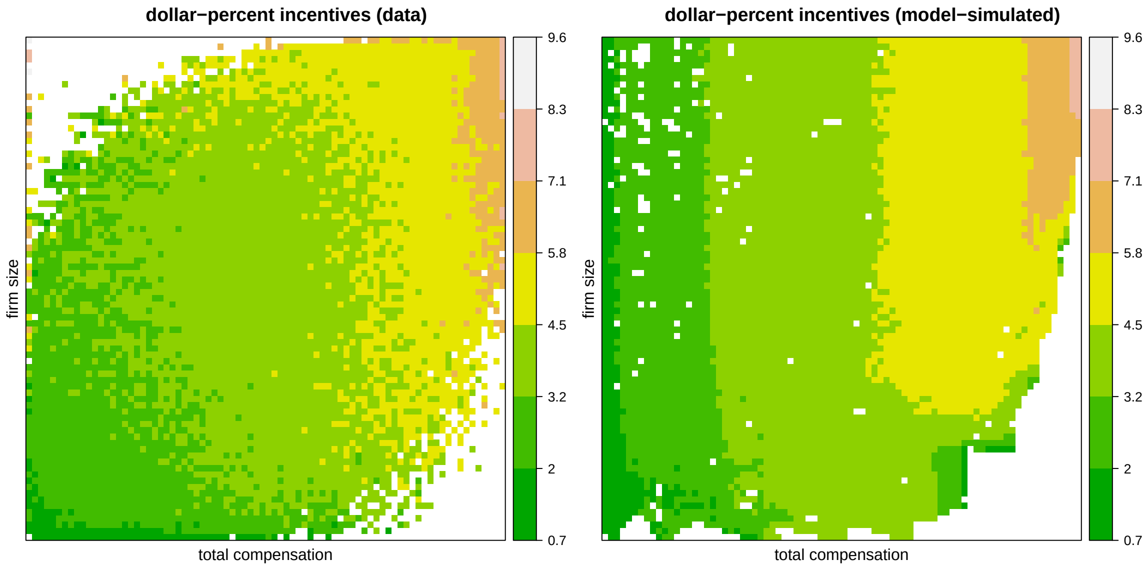


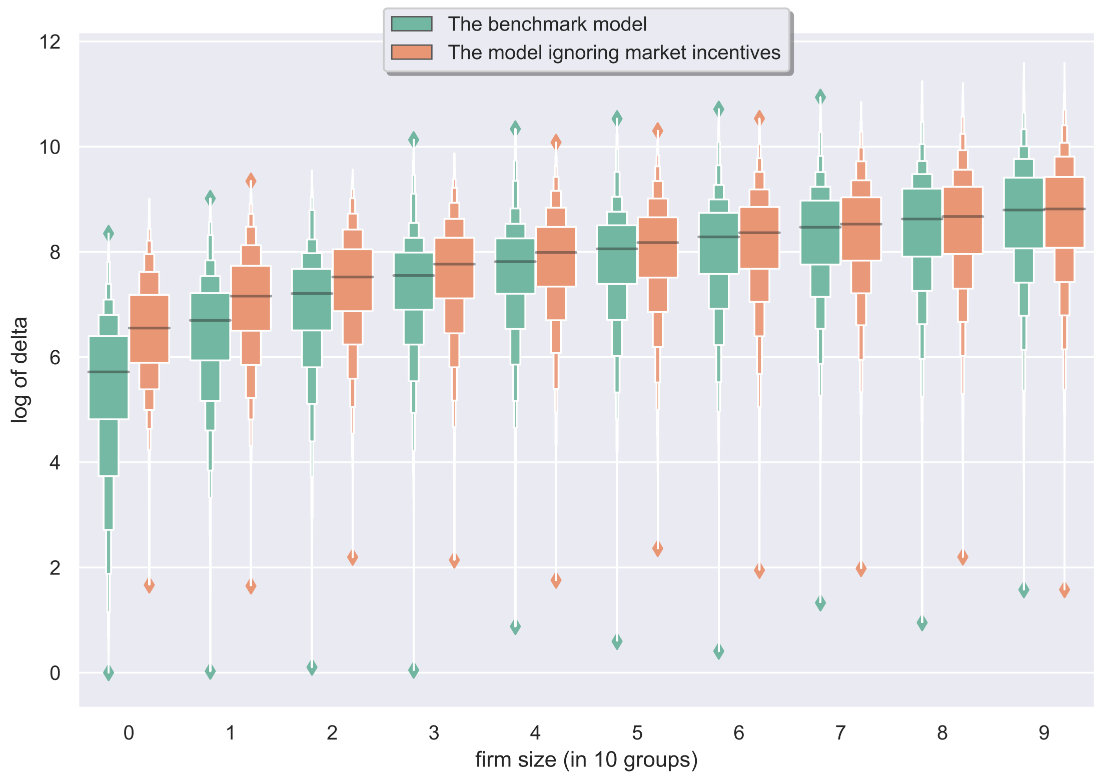
Figure 10: $\log(delta)$ in Data

6 Quantitative Analyses

The last section analyzes two questions. The first one is how much of the executive contract incentives is contributed by the managerial labor market. To answer this question, I simulate a counterfactual scenario where the firms ignore labor market incentives at all, and then compare the delta with and without labor market incentives. The second one is policy related: does an improvement of corporate governance have spillover effect on the overall executive labor market? If so, what is the policy implication?

6.1 Decomposition

To evaluate the contribution of labor market incentives, I solve a separate model where labor market incentives are ignored by both the firm and the executive when signing the contract. As expected, since labor market incentives are not counted, the performance-based incentives will increase, and delta will be higher. To see the influence across firms of different size, I cut the firm size into 10 groups. The upper panel of figure 11 shows the box plot of log delta across firm size. Clearly, smaller firms are likely to suffer more by ignoring labor market incentives. This is because executives of small firms are likely to be lower on the job ladder, and both direct and indirect effect of the ladder indicates they receive more labor market incentives. I further calculate the ratio of the delta's with versus without labor market incentives as in the lower panel of figure 11. 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.



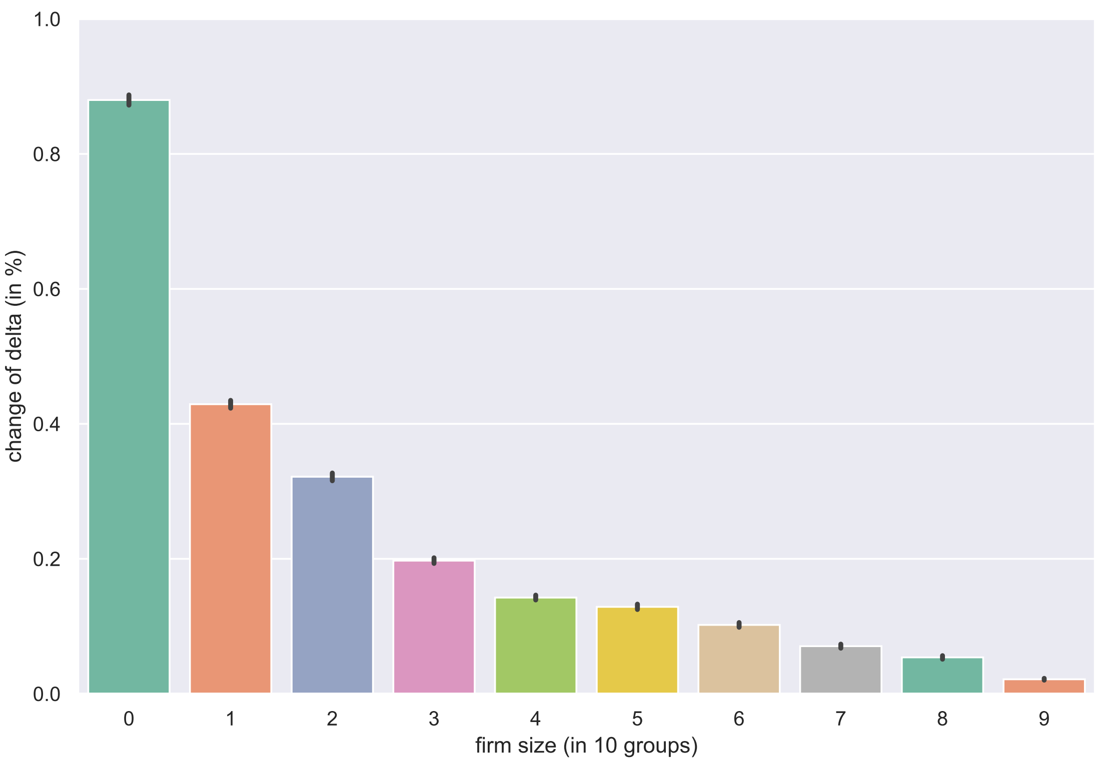


Figure 11: The Fraction of Market Incentives along Firm Size and Wage

6.2 Policy Implications: The Spillover Effect

A worse governance means α_0 is higher. It has two effects:

1. ***Directly,* firms are willing to bid higher, so total compensation and delta increases;**
2. ***Indirectly,* firms are willing to bid higher, so the labor market incentives become smaller, delta needs to be higher for any give total compensation.**

The influence of both the direct and indirect effects are stronger and wider when the governance is worse in small and medium-size firms.

Executive pay is an important part of corporate governance, and is often determined by a company's board of directors. When pay is inefficient, it is often a symptom of an underlying governance problem brought on by conflicted boards and dispersed shareholders. The parameter α_0 in the production function of the model can be interpreted as the firm's (or the board's) willingness to pay to the executive which is directly linked to the quality of corporate governance. For example, an entrenched executive tend to have a higher bargaining power and face a higher α_0, whereas a more independent board may impose a lower α_0 on executives. The role of managerial labor market is the effect of α_0 is not confined within the firm-executive pair. But rather, it spills over to the overall market through the bidding behavior for executives.

I analyze two counterfactual scenarios: one is the *worse corporate governance in small firms*, that is α_0 doubles for firms that are smaller than the median; another is the *worse corporate governance in large firms*, that is α_0 doubles for firms that are larger than the median. Figure 12 plots the distribution of delta (in upper panel) and total compensation (in lower panel) across firm size.

Not surprisingly, the deterioration in large firms' governance significantly increases both the total and the incentive pay for executives in those firms. However, there is This is because all firms

7 Conclusions

By the law of diminishing utility, wealthy agents are costly to motivate. In a dynamic relationship, the wealth is traced from a life-time utility. Executives in larger firms are expected to be wealthier, less sensitive to monetary rewards in the future than their equally paid counterparts in smaller firms. Thus, labor market incentives are lower in larger firms. This implies higher contract incentive pays are required from big firms. With this intuition, this paper provides a framework for empirically analyzing market-induced and performance-based incentives in a frictional labor market. I use this framework to investigate why larger firms pay more for performance. I show that in the counterfactual when labor market incentives are ignored, the incentive pays would be much higher in small and medium firms.

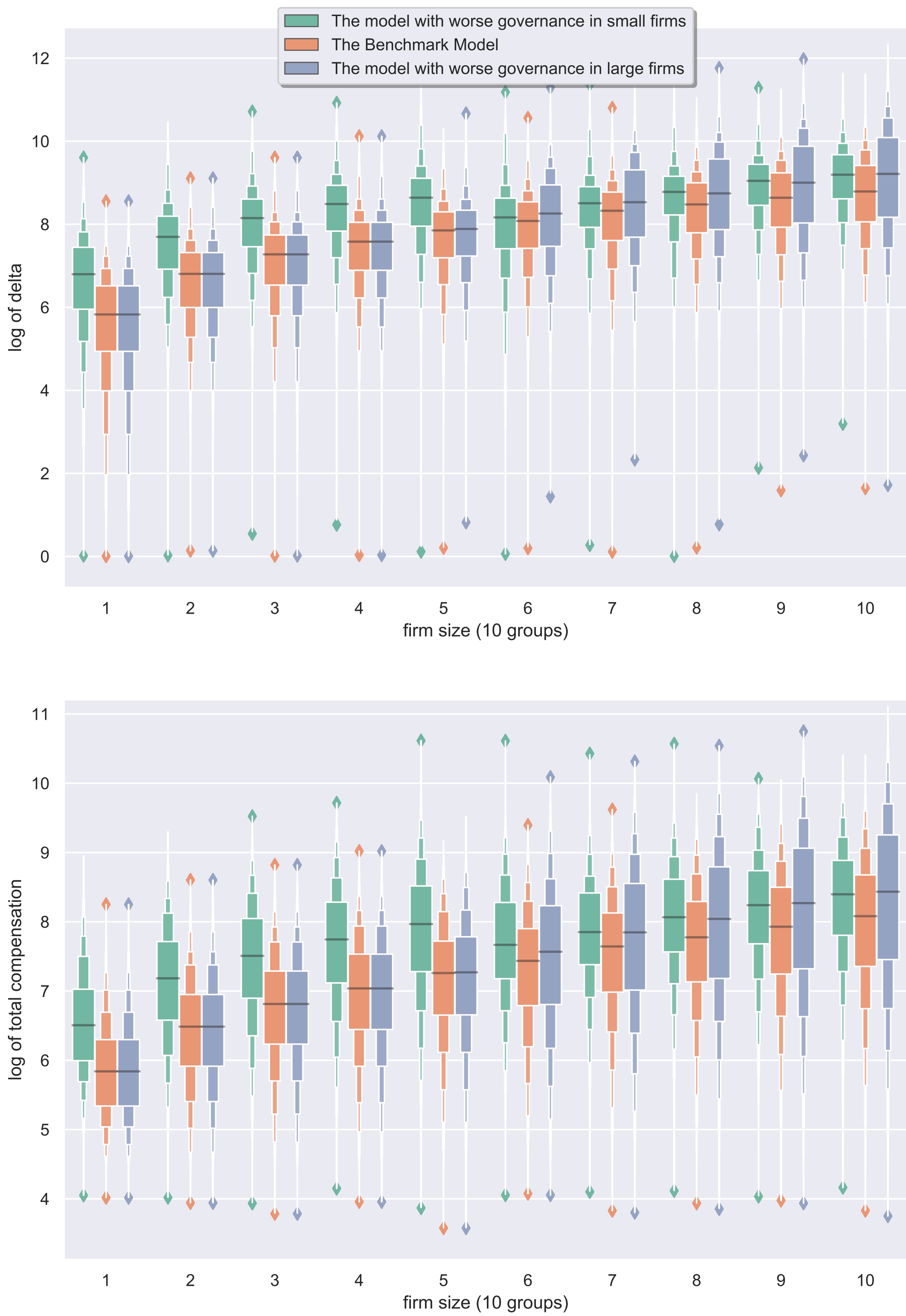


Figure 12: The Fraction of Market Incentives along Firm Size and Wage

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

$$\gamma_1 + ... + \gamma_m - \gamma_{m+1} - ... - \gamma_{n_z-1} - \gamma_{n_z} = 0. \tag{8}$$

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 (8). By construction, ω_i is positive.

□

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

$$\frac{\Delta\overline{W}(z,s)}{\Delta z} = -\frac{\Delta\Pi(z,s,\overline{W})/\Delta z}{\Delta\Pi(z,s,\overline{W})/\Delta\overline{W}} = u'(\overline{w}(s))\frac{\Delta\Pi(z,s,\overline{W})}{\Delta z}, \tag{9}$$

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

$$\begin{aligned}
\frac{\Delta\Pi(z,s,\overline{W})}{\Delta z} &= \sum_{z' \in \mathbb{Z}} \Big(\alpha_0 s^{\alpha_1} z' + \tilde{\beta}\int_{\tilde{s} \le s} \Pi(z',s,\overline{W}(z',s)d\tilde{F}(\tilde{s})\Big)\frac{\Delta\Gamma(z'|z)}{\Delta z} \\
&\quad + \lambda\tilde{\beta}\sum_{z' \in \mathbb{Z}} \Big(\int_{\tilde{s}} \overline{W}(z',s)d\tilde{F}(\tilde{s})\Big)\frac{\Delta\Gamma(z'|z)}{\Delta z} \\
&\quad + \mu\tilde{\beta}\sum_{z' \in \mathbb{Z}} \Big(\int_{\tilde{s}} \overline{W}(z',s)d\tilde{F}(\tilde{s})\Big)\frac{\Delta\Big((1-g(z'|z)\Gamma(z'|z)\Big)}{\Delta z} \\
&= \alpha_0 s^{\alpha_1}\sum_{z' \in \mathbb{Z}} z'\frac{\Delta\Gamma(z'|z)}{\Delta z} + \tilde{\beta}\sum_{z' \in \mathbb{Z}}\int_{\tilde{s}} \overline{W}(z',s)d\tilde{F}(\tilde{s})\Big(\lambda\frac{\Delta\Gamma(z'|z)}{\Delta z} + \mu\frac{\Delta\Big((1-g(z'|z)\Gamma(z'|z)\Big)}{\Delta z}\Big).
\end{aligned} \tag{10}$$

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

$$\begin{aligned}
\frac{\frac{\Delta\Pi(z,s,\overline{W})}{\Delta z}}{\alpha_0\sum_{z' \in \mathbb{Z}} z'\frac{\Delta\Gamma(z'|z)}{\Delta z}} &= s^{\alpha_1} + \frac{\tilde{\beta}\sum_{z' \in \mathbb{Z}}\int_{\tilde{s}} \overline{W}(z',s)d\tilde{F}(\tilde{s})\Big(\lambda\frac{\Delta\Gamma(z'|z)}{\Delta z} + \mu\frac{\Delta\Big((1-g(z'|z)\Gamma(z'|z)\Big)}{\Delta z}\Big)}{\Delta z} / \alpha_0\sum_{z' \in \mathbb{Z}} z'\frac{\Delta\Gamma(z'|z)}{\Delta z} \\
&= s^{\alpha_1} + \psi(s),
\end{aligned} \tag{11}$$

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

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 (10) and (11) into (9), we have

$$\frac{\Delta\overline{W}(z,s)}{\Delta z} = u'(\overline{w}(s))\frac{\Delta\Pi(z,s,\overline{W})}{\Delta z} = u'\left(\alpha_0 s^{\alpha_1}\sum_{z'\in\mathbb{Z}} z'\Gamma(z'|z)\right)\left(s^{\alpha_1}+\psi(s)\right)\alpha_0\sum_{z'\in\mathbb{Z}} z'\frac{\Delta\Gamma(z'|z)}{\Delta z}. \tag{12}$$

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 (12)

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

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

Table 9: Pay-for-performance incentives increase in 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	
profit					0.651***
					(0.122)
annual return					0.0984*
					(0.0466)
mbr					0.113***
					(0.0203)
director					0.754***
					(0.0330)
interlock					0.486***
					(0.0954)
CEO					0.608***
					(0.0381)
CFO					0.0870***
					(0.0131)
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	
Observations	146747	128006	128006	128006	109730
adj. R^2	0.442	0.514	0.523	0.524	0.595

Table 10: Size 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 10: Size 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

Table 11: Size premium increases with managerial labor market competition

	log(*delta*)			
	(1)	(2)	(3)	(4)
log(firm size)	0.525*** (0.00512)	0.529*** (0.00499)	0.561*** (0.00310)	0.571*** (0.0139)
EE190	1.919* (0.776)			
log(firm size) × *EE190*	0.415*** (0.101)			
EE90		2.611** (0.903)		
log(firm size) × *EE90*		0.359** (0.118)		
gai			-1.211*** (0.0941)	
log(firm size) × *gai*			0.0648*** (0.0118)	
inside CEO				-0.00566*** (0.00156)
log(firm size) × *inside CEO*				-0.000458* (0.000202)
profit	0.538*** (0.0309)	0.541*** (0.0309)	0.632*** (0.0366)	0.528*** (0.0309)
annual return	0.109*** (0.00447)	0.108*** (0.00447)	0.166*** (0.00736)	0.110*** (0.00447)
mbr	0.145*** (0.00242)	0.146*** (0.00242)	0.130*** (0.00272)	0.143*** (0.00244)
director	0.937*** (0.0104)	0.937*** (0.0104)	1.006*** (0.0125)	0.932*** (0.0104)
interlock	0.492*** (0.0292)	0.492*** (0.0292)	0.457*** (0.0298)	0.498*** (0.0292)
CEO	0.755*** (0.0118)	0.755*** (0.0118)	0.736*** (0.0138)	0.757*** (0.0118)
CFO	0.0765*** (0.0133)	0.0764*** (0.0133)	0.0162 (0.0275)	0.0757*** (0.0133)
tenure dummies	Yes	Yes	Yes	Yes
age dummies	Yes	Yes	Yes	Yes
year dummies	Yes	Yes	Yes	Yes
industry dummies	Yes	Yes	Yes	Yes
year × *industry dummies*	Yes	Yes	Yes	Yes
Observations	125858	125858	75747	125858
adj. R^2	0.521	0.521	0.531	0.521

Notes for table 9:

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 $\times$ 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 *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 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Notes for table 10

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 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Notes for table 11:

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 *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. 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 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Appendix C. Estimation appendices

Computing Algorithm

To further characterize the optimal solution, we resort to the tools developed by Marcet and Marimon (2017, hereafter MM).[19] 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

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*

$$u'(w) = \frac{1}{\lambda}, \tag{13}$$

$$W(z') = W(z',s,\lambda'). \tag{14}$$

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 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

[19] 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)).

outside firm by s'. It means the new match will assign a new weight to the executive so that he 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 (13) and (14).

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).

[Incomplete Reference]

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.

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."

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

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.

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.

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.

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

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.

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.

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.

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

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.