Optimal Long Term Executive Contracts

Robert A. Miller and Anh Nguyen

Carnegie Mellon University (both of us)

August 2024

- Critics of capitalism sometimes claim that CEOs are driven by *short termism* . . .
    - the relentless pursuit of current profits,
    - at the expense of long term goals requiring investment.
- One response is that:
    - CEO compensation is positively correlated with firm returns.
    - A firm's return depends on both *dividends* (distributions to shareholders) and *investments* (impounded into capital gains).
    - Indeed 7 percent of CEOs earn negative income (primarily because the stocks and options they hold lose value relative to the market portfolio).

- Another tool for disciplining executive performance is the threat (and use) of *dismissal.*
- Dismissals are typically unobserved (to the econometrician).
- It is hard to *empirically* distinguish between separations that are:
  - voluntary (*quitting*)
  - involuntary (*firing*)
- How can these concepts be distinguished *analytically*?
- We develop and identify a model in which, roughly speaking, the firm uses:
  - *compensation to incentivize executives.*
  - *hiring and firing to find good job matches.*
- We use the Compustat Execucomp database from 1992 to 2022 to:
  - estimate the importance of dismissal as a tool for disciplining performance.
  - quantify the value of commitment by the firm.

- Job search and matching:
  - *Jovanovic, 1979; Miller, 1984; Antonovics and Golan, 2012.*
  - Here learning on the job is endogenous, a hidden action.
- Involuntary termination:
  - *Clementi and Hopenhayn, 2006; DeMarzo and Sannikov,2006; Spear and Wang, 2005.*
  - Here termination is also driven by limited liability.
- Role of risk in compensation contracts:
  - *Holmstrom, 1979; Demsetz and Lehn, 1985; Peters and Wagner, 2014.*
  - We focus on the gains from long term contracting.
- Structural estimation of moral hazard models:
  - *Margiotta and Miller, 2000; Gayle and Miller, 2009, 2015; Gayle, Golan and Miller, 2015.*
  - Previous work lacks learning about match quality and long term contracting.

Data

The setting

- Our analysis restricts attention to the chief executive officer (CEO).
    - Year-to-year CEO compensation is the most variable in the C-suite.
- We focus on four key variables of interest:
    1. *abnormal returns*
        - measuring current CEO performance relative to peers.
        - used to construct match value with firm.
    2. *annual pay, that is total compensation*
        - change in inside CEO wealth before consumption.
    3. *separations*
        - aggregate voluntary with involuntary separations.
    4. *external versus internal hires*
        - there might more to learn about the former.
- Compiled from Compustat Execucomp database:
    - SP500 in 1992 - 1993 and SP1500 from 1994 to 2022.
    - 3259 internal executives from 1785 firms
    - 444 external executives from 394 firms.

Data

Defining abnormal returns

- Denote
  - V_{tj} as market value of firm j at the beginning of period t
  - w_{tj} as CEO compensation in firm j for employment in period $t-1$
  - $Dividends_{tj}$ as dividends paid by firm j in period t
  - $\widetilde{x}_{tj}$ gross return to firm before CEO compensation defined as

$$\widetilde{x}_{tj} V_{t-1,j} \equiv V_{t,j} + w_{tj} + Dividends_{tj}$$

- Note that $\widetilde{x}_{tj}$ does not account for:
  - aggregate fluctuations in the economy that affect all returns.
  - industry effects arising from its covariance with aggregate returns.
- It is convenient to work in logs. Let:

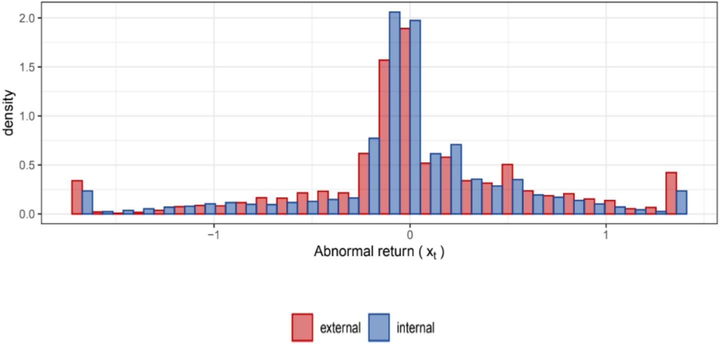
$$x_{tj} \equiv \log \widetilde{x}_{tj} - \log Marketreturn_t - \beta X_{t,j}$$

where:

  - $X_{t,j}$ are variables for industry fixed effects and firm size relating to (t, j)
  - $Marketreturn_t$ is the return on the stock index in t
  - x_{tj} is our measure of abnormal returns for firm j in period t.

Data

Abnormal returns by external and internal hiring



- By definition abnormal returns are centered on zero.
- Abnormal returns from external hires:
  - have fatter tails.
  - are more dispersed in the mid-section of the distribution.

Data

Defining executive performance

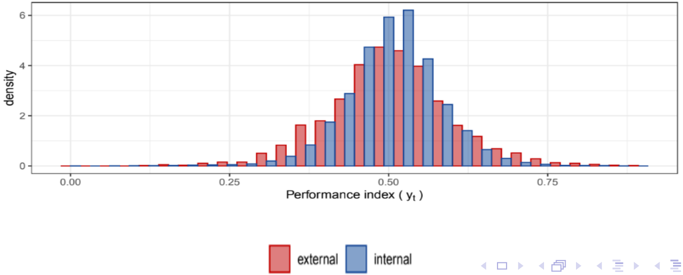
- Denote by:
  - $\widetilde{V}_{t,j} \equiv V_{t,j} + Dividends_{tj}$ the present value of current and future expected discounted profits from j in t.
  - $\widetilde{V}^{t,j} \equiv \left\{ \widetilde{V}_{\tau,j} \right\}_{\pi=1}^{t}$ the history of $\widetilde{V}_{t,j}$ from j up to t.
- Define the performance index of executive (t, j) as:

$$y_{tj} \equiv \Pr \left[\widetilde{V}_{t+1,,i} < \widetilde{V}_{t+1,,j} \middle| \widetilde{V}^{t,j}, X_{t,j} \right]$$

  where the probability is over all firms i.
- Let $M_{h,t,j}$ denote the h^{th} central moment of the distribution $\widetilde{V}_{t+1,,i}$ conditional on $\left(\widetilde{V}^{t,j}, X_{t,j} \right)$.
- We refer to $(M_{1,t,j}, M_{2,t,j}, M_{3,t,j})$ as the prior mean, variance, and skewness of shareholder beliefs.

Measuring executive performance

	All		Across-CEOs	
	External	Internal	External	Internal
Abnormal return	0.05	-0.01		
	(0.85)	(0.72)	(0.81)	(0.69)
Performance index	0.51	0.5		
	(0.1)	(0.07)	(0.08)	(0.05)
Prior mean	0	0.09		
	(1.51)	(1.52)	(0.57)	(0.45)
Prior variance	0.82	0.86		
	(0.56)	(0.59)	(0.49)	(0.52)
Prior skewness	0.04	-0.1		
	(0.86)	(0.94)	(0.72)	(0.83)



Data

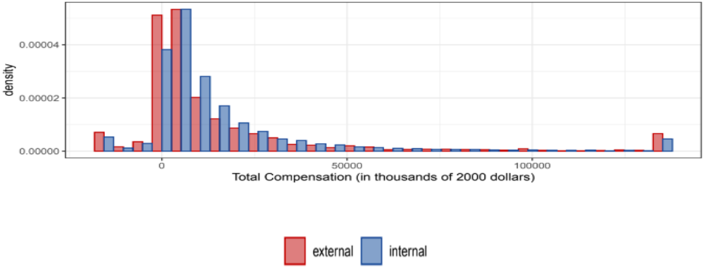
Components of compensation (thousands of $US 2000)

- An executive's total compensation consists of two components.
  1. The first component is the direct compensation, which includes:
     - salary
     - bonus
     - other annual and restricted stock and option grants
     - long-term incentive plan payouts.
  2. The second component is a measure of idiosyncratic change in wealth
     - from previously held options and stocks.
     - equals value of the options and stocks at the beginning of the period multiplied by the firm's abnormal return
- The second component arises from our assumption that:
  - *absent private information, the CEO would diversify out of specialized financial assets.*

Data

Components and distribution of total compensation (thousands of $US 2000)

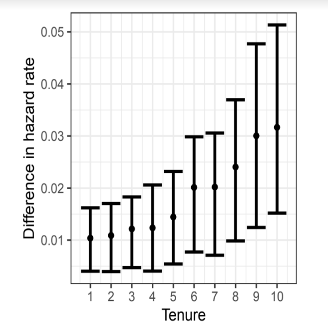
	All		Across-CEOs	
	External	Internal	External	Internal
Salary and bonus	3171	2950		
	(6116)	(6574)	(5414)	(5662)
Other payment	13829	14097		
	(40663)	(41318)	(36560)	(36873)
Change in wealth due to stocks and options	3178	3380		
	(62840)	(138296)	(61900)	(123892)
Total compensation	20189	20434		
	(86904)	(153137)	(83597)	(138266)



Data

Separations

- We focus on:
  - differences between external and internal hires
  - in the conditional probability of a separation at each point in tenure.



- At every tenure point, external hires are significantly more likely to leave the firm.

Model

Motivation

- Summarizing the patterns in the data:
    - Abnormal returns from external hires are more dispersed (than for internal hires)
    - Performance index for external hires more dispersed.
    - Internal hires have a greater dispersion in pay.
    - However compensation for externals have a fatter tail.
    - Match with externals more likely to break than with internals.
- In a model based on *matching* and *incentivizing*:
    - Match value with internals (externals) more or less (un)known.
    - Internals require motivation if effort is hidden.
    - Externals are also selected on a more provisional basis.

Model

Timeline

- Previous executive departs (either fired or quit).
- Firm writes long term contract for new executive:
  - differentiating between external and internal hires.
  - They are indifferent about what type signs on.
- The long term contract specifies (for the employment spell):
  - executive compensation
  - dismissal probability
- Each period shareholders:
  - paid executive for previous period's employment
  - randomly dismisses executive with prescribed probability
- If the executive is not dismissed he decides:
  - consumption for the period.
  - whether to quit or remain employed by the firm.
  - if employed, whether to work (in shareholder interests) or not (shirk).

Model

Worker choices and preferences

- c_t *consumption choices* in $t \in \{0, 1, \ldots\}$ tenure with firm
- (d_{0t}, d_{1t}, d_{2t}) *employment and effort choices* with $\sum_{j=1}^{2} d_{jt} = 1$ and:

$$d_{0t} = \begin{cases} 1 \text{ if worker stays and shirks} \\ 0 \text{ if not} \end{cases}$$

$$d_{1t} = \begin{cases} 1 \text{ if worker stays and works} \\ 0 \text{ if not} \end{cases}$$

$$d_{2t} = \begin{cases} 1 \text{ if worker quits} \\ 0 \text{ if not} \end{cases}$$

$$D_t = \begin{cases} 1 \text{ if worker is dismissed} \\ 0 \text{ if not} \end{cases}$$

- *Preferences*, where $\alpha_0 < \alpha_1$, and $\epsilon_{0t} \equiv (\epsilon_{0t}, \epsilon_{1t})$ is *iid*:

$$-\sum_{t=1}^{\infty} \delta^t \exp\left[-\gamma c_t\right] \left\{ \begin{array}{c} (1 - D_t)\left(d_{0t}\alpha_0 + d_{1t}\alpha_1\right)\epsilon_{1t} \\ + \left(d_{0t}D_t + d_{1t}D_t + d_{2t}\right)\epsilon_{2t} \end{array} \right\}$$

- Output:
  - x_t is a realization of random variable X_t in t
  - $x^t \equiv (x_1, \ldots, x_t)$ is history up until t
  - depends on match value $\theta \in \left(\underline{\theta}, \overline{\theta}\right)$ with *cdf* $H(\cdot)$ and *pdf* $h(\cdot)$
  - has *cdf* $F_j(x\,|\theta)$ with *pdf* $f_j(x\,|\theta)$ for effort $j \in \{0, 1\}$

Assumption 1. *For any* $\theta \in [\underline{\theta}, \bar{\theta})$ *and* $x, x' \in \mathbb{X}$ *where* $x' > x$

1. $\frac{F_1(x|\theta, x \leq x')}{F_1(x'|\theta)}$ *is decreasing in* θ.
2. $\frac{\partial}{\partial x}\frac{f_0(x|\theta)}{f_1(x|\theta)} \leq 0$, $\frac{\partial}{\partial \theta}\frac{f_0(x|\theta)}{f_1(x|\theta)} \geq 0$, $\frac{\partial^2}{\partial\theta\partial x}\frac{f_0(x|\theta)}{f_1(x|\theta)} \leq 0$, *and* $\lim_{x\to\infty}\frac{f_0(x|\theta)}{f_1(x|\theta)} = 0$.
3. *For* $1 \leq \tau < t$ *and* $\theta^* < \bar{\theta}$: $\lim_{x_\tau\to\infty}\Pr(\theta \leq \theta^*|x^t) = 0$.

- The expected value of the firm is:

$$E_0 \left\{ \sum_{t=0}^{T} \lambda_t Dividends_t + \lambda_{T+1} V_T \right\} \tag{1}$$

where:

$$x_t V_{t-1} = V_t + w_t + Dividends_t$$

and:

  - λ_t is the price of contingent claim to consumption in period t.
  - V_t is the value of the firm in period t.
  - $Dividends_t$ is the dividend payout to shareholders in period t.
  - at the (nonanticipating) time T the executive is dismissed or quits.
  - w_t is executive compensation at end of period t.
- Subject to *participation* and *incentive compatibility* constraints the firm chooses a long term contract $\{w_t, s_t\}_{t=1}^{T}$ to maximize (1) where:
  - $w_t \equiv w_t \left(x^t \right)$ is compensation at end of period t.
  - $s_t \equiv s_t \left(x^t \right)$ is probability that executive is dismissed at end of period t.

Optimality

The executive's career problem (when there is no firing)

- To characterize the two constraints we first analyze the executive's career choices for any given contract.
- Recursively define a measure of the executive's *human capital*, an exante continuation value, as:

$$
\begin{aligned}
& A_t\left(d_1^{t-1}, x^{t-1}\right) \equiv \\
& \quad p_{2t} E\left[\epsilon_{2t}^*\right] \alpha_1^{1/b_t} \\
& \quad + p_{1t} E\left\{\epsilon_{1t}^* \left[d_{0t}\alpha_0^{1/b_t} + d_{1t}\alpha_1^{1/b_t}\right] E_t\left[v_{t+1}A_{t+1}\left(d_1^t, x^t\right) \middle| d_{0t}\right]^{1-1/b_t}
\end{aligned}
$$

where:

- p_{jt} is the conditional choice probability, CCP, for his j^{th} choice.
- ϵ_{jt}^* is the truncated random variable obtained from ϵ_{jt} when $d_{jt} = 1$.
- $d_1^{t-1} \equiv \left(d_{10}, \ldots, d_{1,t-1}\right)$ is the executive's work history with the firm.
- b_t is the *bond price* in t.
- $v_t \equiv \exp\left(-\gamma w_t / b_t\right)$ is an *annuitized utility value from compensation.*

- Both constraints only apply when $\psi \neq 1$ (and the executive is not dismissed):
  - The *participation constraint* is:

$$\left(\alpha_1 \epsilon_{1t}\right)^{\frac{1}{b_t - 1}} \int_{\mathbb{X}} \int_{\Theta} v_{t+1} A_{t+1} f_1 \left(x \,|\theta \right) h\left(\theta, x^t, d^t\right) d\theta dx_t \leq \left(\alpha_2 \epsilon_{2t}\right)^{\frac{1}{b_t - 1}}$$

  - The *incentive compatibility constraint* is:

$$\begin{aligned} & \alpha_1^{1/(b_t - 1)} \int_{\mathbb{X}} \int_{\Theta} v_{t+1} A_{t+1} f_1 \left(x \,|\theta \right) h\left(\theta, x^t, d^t\right) d\theta dx_t \\ \leq \; & \alpha_0^{1/(b_t - 1)} \int_{\mathbb{X}} \int_{\Theta} v_{t+1} A_{t+1} f_0 \left(x \,|\theta \right) h\left(\theta, x^t, d^t\right) d\theta dx_t \end{aligned}$$

Model

A specialization and relaxation

- In the model the executive receives fixed compensation for shirking:
  - but our data shows it is highly volatile.
  - from which we infer the executive always works.
- Hence we focus on scenarios where it is optimal to induce working.
- The IC constraint must be satisfied for every history, including those not on the equilibrium path of working.
- However under our monotonicity assumptions, all future IC constraints are satisfied if the *executive has ever shirked*.
- This implies the firm's chooses $\{s_t, v_t, p_{2t}\}$ to maximize:

$$E_0 \left\{ \begin{array}{r} \sum_{t=1}^{T} \prod_{\tau=1}^{t-1} (1 - s_\tau) \, p_{1\tau} \left(x_{t+1} Val_t - Val_{t+1} + r^{-1} \ln v_{t+1} \right) \\ + \left[1 - (1 - s_t) \, p_{1t} \right] V_T \end{array} \right\}$$

subject to the constraint that the executive shirks at most once.

Model

The firm's relaxed optimization problem

- In the relaxed problem we do not impose IC constraints for (off equilibrium) histories in which the executive has shirked.
- Intuitively *he is better than he looks*, and rewarded relatively well for comparatively modest output demands given his true match:

Figure 4.2: Illustration of the difference between the full problem and the relaxed problem for the firm

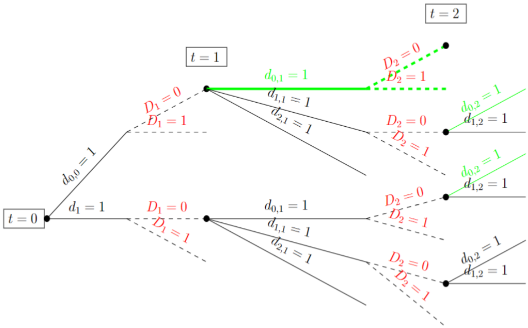
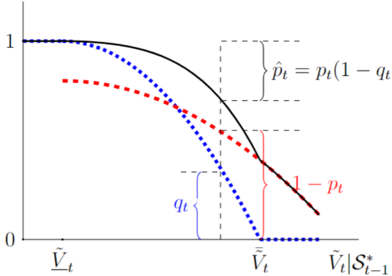


Figure 4.3: Illustration of the optimal separation probability



Note: The dashed red curve represents the quitting probability $1 - p_t$. This probability is not defined when an executive is fired with certainty (below $\underline{\tilde{V}}_t$). The dashed blue curve represents the firing probability q_t, which is equal to 1 for low outcomes $\tilde{V}_t \leq \underline{\tilde{V}}_t$ and equal to 0 when the executive's performance is sufficiently good ($\tilde{V}_t \geq \bar{\tilde{V}}_t$). the solid black line represents the separation probability in the data, which is equal to $1 - p_t(1 - q_t)$.

Identification and Estimation

Primitives and data generating process

- The primitives:
  - $h(\theta)$ job match *pdf*
  - $f_j(x|\theta)$ output *pdf* for $j \in \{0, 1\}$
  - α_j effort distaste parameters for $j \in \{0, 1\}$
  - γ and δ risk aversion and discount factors
  - $g(\epsilon_{1t}, \epsilon_{2t})$ *pdf* for taste shocks
- We assume $\ln \epsilon_{jt}$ is *iid* T1EV.
- The discount factor δ is not identified.
- The data generating process yields:
  - $f_1(x^t)$ the reduced form for firm returns given work
  - $w_t(x^t)$ compensation for any history on the work equilibrium
  - $\rho_t(x^t) \equiv s_t(x^t) + [1 - s_t(x^t)] p_{2t}(x^t)$ separation probability
  - Since $p_{0t}(x^t) = 0$ on optimal contract, $p_{1t}(x^t) + p_{2t}(x^t) = 1$

Identification and Estimation

Stage 1: When learning is completed

- There is no more (or negligible) learning if:
  - the executive survives *long enough*
  - he plans to retire next period (so $T < \infty$).
- At that point:
  - the kink disappears.
  - there are no dismissals.
  - the long term contract decomposes to a sequence of short term contracts.
- From these periods we can identify and estimate:
  - α_j effort distaste parameters for $j \in \{0, 1\}$.
  - γ the risk aversion and discount factors.
  - $f_j(x|\theta)$ output *pdf* for $j \in \{0, 1\}$ in an infinite horizon problem (or more generally in a problem where learning stops).
- Estimation is a standard (dynamic) RUM.
- The risk aversion parameter γ is identified because executives facing different lotteries that depend on past firm returns.

Identification and Estimation

Stage 2: The kink (or change point)

- $\overline{x}$ (the kink point) is identified from:

$$\lim_{x \downarrow x^*} \frac{\partial \rho_t \left(x^{t-1}, x^*\right)}{\partial x} - \lim_{x \uparrow x^*} \frac{\partial \rho_t \left(x^{t-1}, x^*\right)}{\partial x} = \begin{cases} \overline{\rho} > 0 \text{ if } x^* = \overline{x} \\ 0 \text{ otherwise} \end{cases}$$

- In estimation we impose the (theoretical) restrictions that:
  - $\rho_t \left(x^t\right)$ is continuous
  - $\overline{\rho} > 0$ at $x^* = \overline{x}$
- We locally regress separation, $d_t + (1 - d_t)\, d_{0t}$
  - on a common constant
  - and a slope term that depends on whether $x_t \lessgtr x^*$
  - choosing x^* to minimize the quadratic criterion function.

Identification and Estimation

Stage 3: The remaining parameters

- Throughout the analysis $(M_{1,t,j}, M_{2,t,j}, M_{3,t,j})$:
  - represent the prior mean, variance, and skewness of shareholder beliefs about firm value, before executive performance is realized.
  - are assumed sufficient statistics for their (subjective) conditional distribution.
- We parameterize abnormal returns, $f_j(x_t | \theta)$, for both $j \in \{1, 2\}$:
  - with a normal *pdf* $\mathcal{N}(\theta - d_{0t}\Delta, \sigma^2)$ where $\Delta > 0$.
- We also parameterize $h(\theta)$, the match distribution:
  - with a normal *pdf* $\mathcal{N}(\phi, \psi^2)$
  - to exploit the computational properties of a conjugate prior.
- When:
  - $x > \overline{x}$ all separations are quits.
  - $x < \overline{x}$ there is both quitting and firing.
- Given the change points and the preference parameters, the FOCs suffice to identify and estimate the remaining parts of the model.