Executive Compensation: Perspectives and Insights from the Theoretical Literature

Pierre Chaigneau[1]

June 2024

NBER Executive Compensation Workshop

Organized by Dirk Jenter and Kelly Shue

[1]Smith School of Business at Queen's University

Incentives

- "The most important development in economics in the last forty years has been the study of incentives." Jean-Jacques Laffont and David Martimort
- "Nothing is more important for a market economy than the structure of incentives for managers and investors." Mihir A. Desai
- "The premise behind modern corporate finance is that corporate insiders need not act in the best interests of the providers of the funds." Jean Tirole
- "The formal foundation for much of [a new understanding of accounting practices and principles from rigorous theoretical models of settings in which accounting information has real economic implications] is the principal-agent model." Thomas Hemmer

In the news

Yesterday's front-page headline in the *Financial Times*

FT Weekend

Updated 10:44am

Executive Pay

US executive pay rises at fastest rate in 14 years

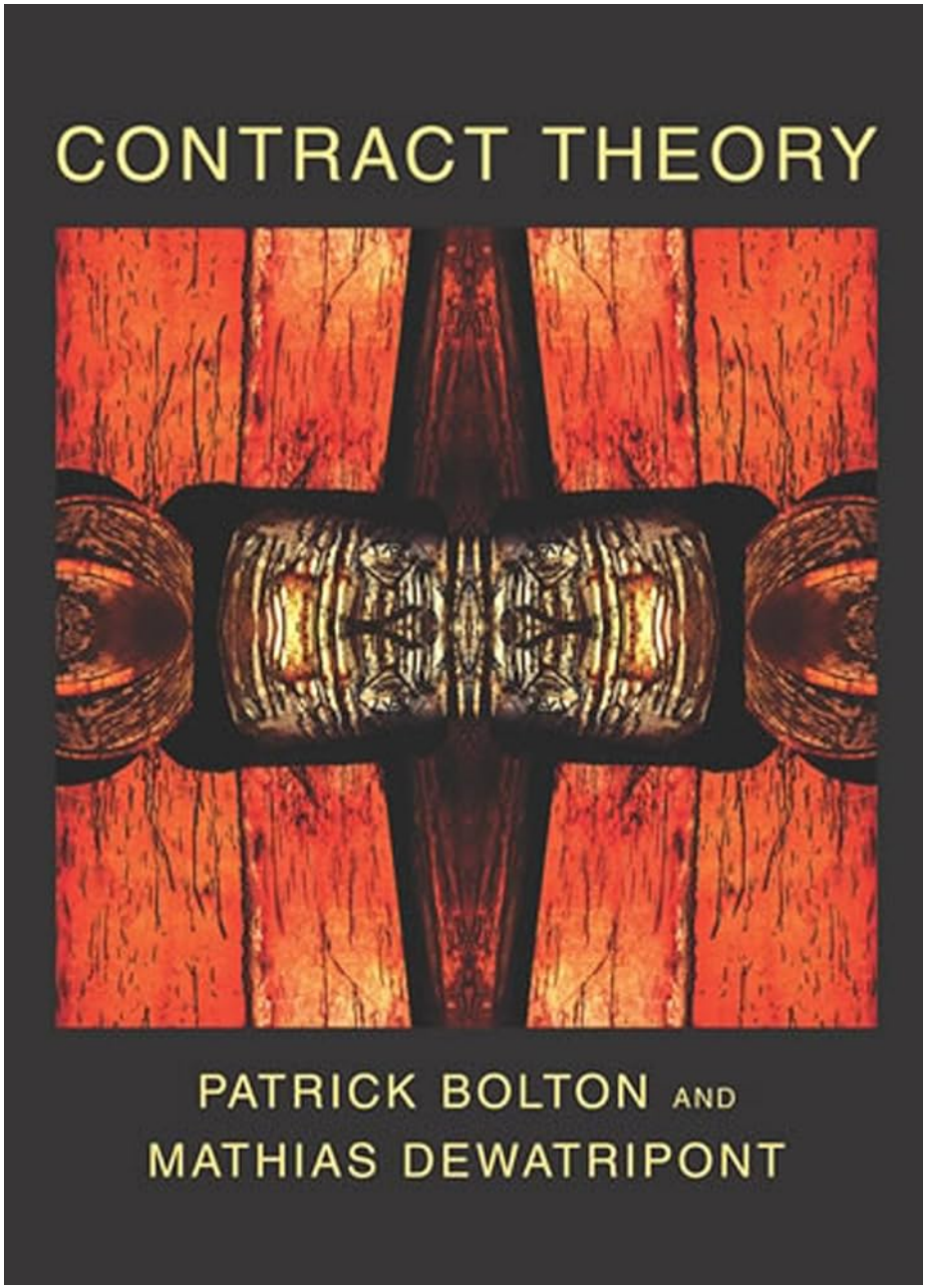
Figures follow Elon Musk's triumph over record $56bn Tesla package

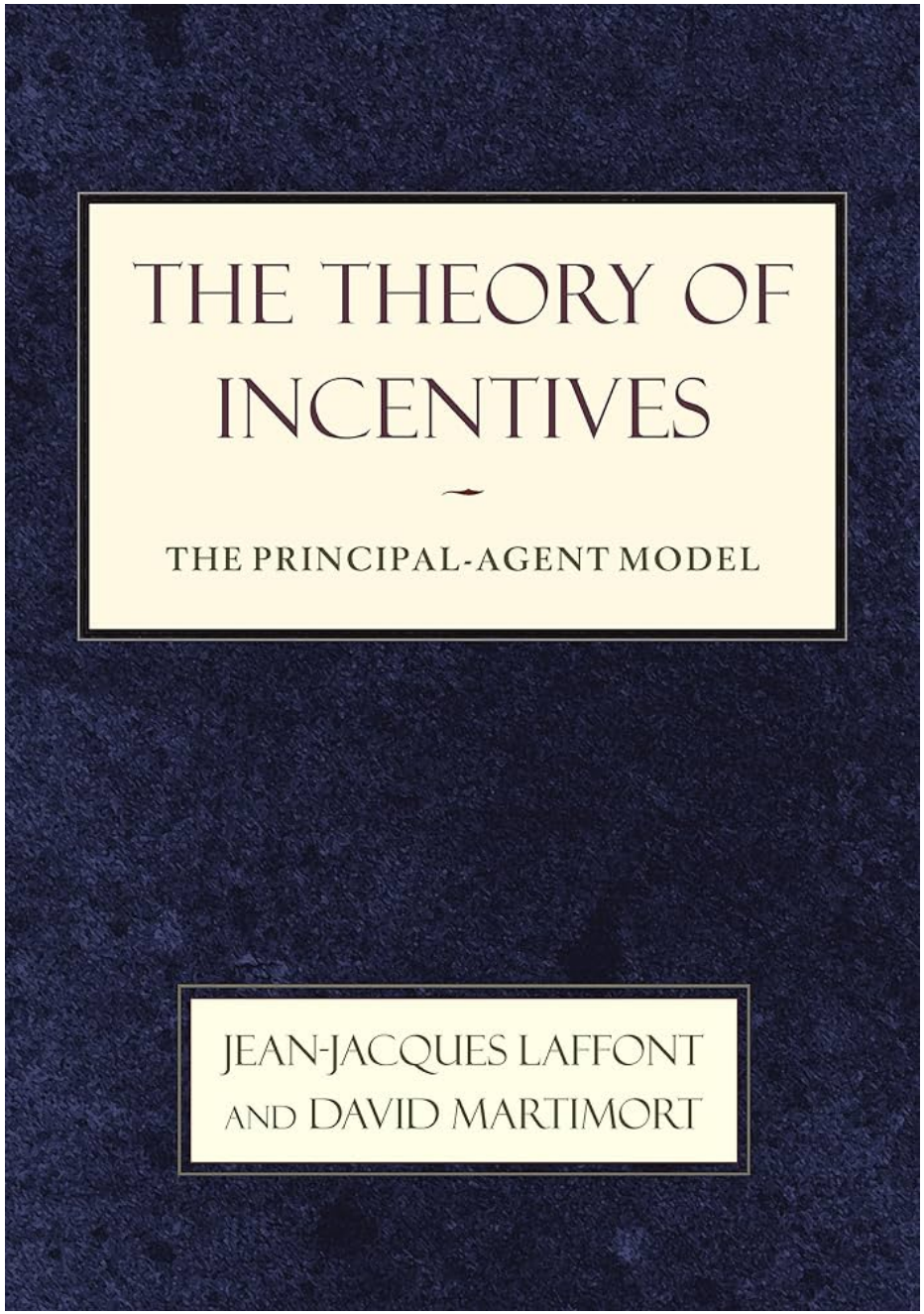
Models

- About Milton Friedman: "He has a model in mind, he's just not showing it to you."
- Even empiricists have a model in mind:
  - People like money: $u' > 0 \Rightarrow$ FOSD, incentives matter ...
  - People are risk averse: $u'' < 0 \Rightarrow$ SOSD, incentives matter ...
- Cross-pollination of theoretical and empirical research

"Economics is a science of thinking in terms of models joined to the art of choosing models which are relevant to the contemporary world."
J.M. Keynes

"All models are wrong, some models are useful (sometimes)"


CONTRACT THEORY
PATRICK BOLTON AND
MATHIAS DEWATRIPONT


THE THEORY OF
INCENTIVES
THE PRINCIPAL-AGENT MODEL
JEAN-JACQUES LAFFONT
AND DAVID MARTIMORT

American Economic Review 2017, 107(7): 1753–1777
https://doi.org/10.1257/aer.107.7.1753

Pay for Performance and Beyond†

By Bengt Holmström*

Incentives are often associated with narrow financial rewards such as bonuses or executive stock options. But in general such rewards are just a small part of the design of incentives. Properly designed incentive systems have to take into account the full portfolio of activities that the agent can engage in, the array of instruments, many nonfinancial, that are available to influence individuals and consider the factors that motivate them in different settings. Thinking about incentives as a system of interacting instruments and influences has been a major advance in the economics of incentives in recent years. In this lecture I will describe the path from pay for performance to the broader view of incentive systems. (*JEL* D21, D82, D86, J33, J41, M12, M52)

In this lecture I will talk about my work on incentive contracts, especially incentives related to moral hazard. I will provide a narrative of my intellectual journey from models narrowly focused on pay for performance to models that see the scope of the incentive problem in much broader terms featuring multitasking employees and firms making extensive use of nonfinancial instruments in designing coherent incentive systems.

I will highlight the key moments of this journey, including misunderstandings as well as new insights. The former are often precursors to the latter. In the process, I hope to convey a sense of how I work with models. There is no one right way about theorizing, but I believe it is important to develop a consistent style with which one is comfortable.

I begin with a brief account of the roundabout way in which I became an economist. It reveals the origins of my interest in incentive problems and accounts for my life long association with business practice, something that has strongly influenced my research and style of work.

I did not plan to become an academic. After graduating from the University of Helsinki, I got a job with Ahlstrom as a corporate planner. Ahlstrom was one of the ten

* Massachusetts Institute of Technology, 77 Massachusetts Avenue, Bldg. E52-520, Cambridge, MA 02139 (email: bengt@mit.edu). I want to thank George Baker, Robert Gibbons, Oliver Hart, Paul Milgrom, Canice Prendergast, John Roberts, and Jean Tirole for years of fruitful discussions on the topic of this lecture and Jonathan Day, Dale Delitis, Robert Gibbons, Gary Gorton, Parag Pathak, Alp Simsek, David Warsh, and especially Iván Werning for comments on various versions of the paper.

† This article is a revised version of the lecture Bengt Holmström delivered in Stockholm, Sweden, on December 8, 2016, when he received the Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel.

CHAPTER 7

Executive Compensation: A Survey of Theory and Evidence*

Alex Edmans[*,†,‡,1], **Xavier Gabaix**[§,¶,†,‡], **Dirk Jenter**[‖,†]
[†]London Business School, London, United Kingdom
[†]CEPR, London, United Kingdom
[‡]ECGI, Brussels, Belgium
[§]Harvard University, Cambridge, MA, United States
[¶]NBER, Cambridge, MA, United States
[‖]London School of Economics, London, United Kingdom
[1]Corresponding author. E-mail: aedmans@london.edu

Contents

* We thank Aubrey Clark, Fabrizio Ferri, Carola Frydman, Tom Gosling, Steve Kaplan, Gaizka Ormazabal, Kelly Shue, Alexander Wagner, David Zhang, and especially Pierre Chaigneau for helpful comments and Janet Chater, Irem Erten, Jesus Gorrin, Joseph Kalmenovitz, and Jiaying Wei for research assistance.

The Handbook of the Economics of Corporate Governance, Volume 1
ISSN 2405-4380, https://doi.org/10.1016/bs.hecg.2017.11.010

Moral hazard models for executive compensation

- Continuous effort, continuous performance, risk averse manager: Holmström BJE 1979 Standard model
- Binary effort, binary performance, risk neutral manager, limited liability: Holmström and Tirole QJE 1997
- Continuous efforts, normally distributed performance, exponential utility, multiple performance measures, linear contracts: Holmström and Tirole JPE 1993
- Continuous time, normally distributed performance, risk averse manager: Sannikov REStud 2008

"The basic moral hazard problem has a fairly simple structure, yet general conclusions have been difficult to obtain. . . . Very few general results can be obtained about the form of the optimal contract."

Bolton and Dewatripont 2005

Solutions:

1. Put more structure on the moral hazard problem, e.g. Dittmann and Maug JF 2007
2. Extend the moral hazard model
3. Consider other models, e.g. learning and retention as in Harris and Holmström REStud 1982

Disclaimer: what follows is an overview of the theoretical literature, not a comprehensive literature review!

MULTITASKING

Multitasking

- Multitasking: the action set of the agent is multidimensional and the principal cares about several actions or "tasks".
- In this sense, there are many models of multitasking, for example:
  - Holmström and Tirole JPE 1993
  - von Thadden REStud 1995
  - Dittmann, Yu, and Zhang RF 2017
- NB: the "multitasking literature" in a narrow sense focuses on the interdependencies between incentive provision on various tasks.

Multitasking

"Knowing the agent's full portfolio of activities–what his authority and responsibilities are–is essential for designing a coherent, balanced solution that takes into account the interdependencies. This is challenging when easy-to-measure and hard-to-measure activities compete for the agent's attention or if the available performance measures are poorly aligned with the principal's objectives."

Holmström AER 2017

- Holmström and Milgrom JLEO 1991: multiple tasks, but imperfect or inexistent performance measurement for some tasks. This may prevent incentive provision on other tasks if efforts on various tasks are complements or substitutes.
- Feltham and Xie TAR 1994: multiple tasks, multiple performance measures, noisy measurement, risk averse agent. The noise in performance measures can induce deviations from the first-best (multidimensional) action to reduce the risk borne by the agent.

Holmström Tirole JPE 1993

- LEN model: linear (contract), exponential (utility), normal (distribution).
  - Replace the agent's nonlinear utility of wealth by a certainty equivalent function which is linear in the expectation and the variance of wealth
- Multiple performance measures, multiple actions.
  - Including performance measure with endogenously determined distribution: stock price
- Include a model à la Kyle ECMA 1985 to study the effect of market liquidity (which depends on ownership concentration) on stock price informativeness
- Implications:
  1. Market monitoring
  2. Incentives of investors to get informed depend on free float
  3. Balance between short-term and long-term incentives

Short-termism

- Managerial short-termism with career concerns: Narayanan JF 1985
  - Asymmetric information in project/action choice
  - **Signal jamming** equilibrium
  - Not robust to optimal contracting: Darrough JF 1987
- Poor firm performance leads to **termination** of the manager or the project, for example via a takeover: Stein (JPE 1988, QJE 1989)
- **Monitoring**: Von Thadden REStud 1995
  - Asymmetric information in project/action choice
  - Allow for long-term contracts, renegotiation, state-contingent monitoring
  - Incomplete information: termination $\Rightarrow$ cannot observe all outcomes
- **Innovation**: Manso JF 2011: $\mathbb{E}[p_2] < p_1 < \mathbb{E}[p_2|S, 2]$
  - Do not punish early failure, reward long-term success
  - Laboratory experiment: Ederer and Manso MS 2013
- Confounding **time dimension**: high output can follow from long-term action yesterday or short-term action today: Zhu REStud 2018
- Managerial short-termism with optimal contracting and **inefficient markets**: Bolton, Scheinkman, and Xiong REStud 2006

Risk taking

- John and John JF 1993
  - Investment 1: safe project with payoff I
  - Investment 2: risky project whose payoff is L with probability $1-p$ and H with probability p, with $L < I < H$
  - John, Saunders, and Senbet RFS 2000: regulation of bank managers' compensation to avoid risk shifting
- Dittmann, Yu, and Zhang RF 2017
  - Risk averse manager needs effort and risk-taking incentives
- Barron, Georgiadis, and Swinkels TE 2020
  - After learning output, agent can costlessly add MPS to output
  - The agent will engage in this type of risk-taking whenever his utility is convex in output
  - Risk averse principal

Executive Pay

How to pay executives in the age of stakeholder capitalism

WSJ

How to Make Climate Progress: Tie it to CEO Pay

More company boards are doing just that. But the key is to do it right.

By Cheryl Winokur Munk

April 17, 2024 at 10:00 am ET

Experts say climate-related goals should be firmly rooted in science, with targets that are clearly achievable and measurable.

ILLUSTRATION: ROB DOBI

WSJ

Why Companies Shouldn't Tie CEO Pay to ESG Metrics

It may sound like a good idea. But it likely won't achieve the results proponents want.

By Alex Edmans

Updated June 27, 2021 at 11:00 am ET

FINANCIAL POST

How do you know a company is accurately rewarding the CEO for climate targets?

Lack of standardized data and disclosure may be opening the way for some to game the system

Barbara Shecter

Published May 10, 2023

86% of investors "believe [that] incorporating non-financial ESG-related metrics into executive compensation programs is an appropriate way to incentivize executives." 2021 Global Benchmark Policy Survey, ISS Governance

ESG scores and ratings

Berg, Kölbel, and Rigobon RF 2022

- **Definition** of ESG ratings and ESG scores:

  "category scores represent a rating agency's assessment of a certain ESG category. They are based on different sets of indicators that each rely on different measurement protocols."

  - Examples of categories: greenhouse gases emissions, workplace safety, board composition, etc.

  "an important part of the service that ESG rating agencies offer is an interpretation of what ESG performance means."

- ESG scores are **standardized** with publicly known formulas:
  - "We are transparent. We publish our S&P Global ESG Score methodology on our website" S&P Global
  - "High-quality, transparent and reliable data is the foundation of our comprehensive approach. (. . .) Bloomberg's proprietary ESG Scores are fully transparent including methodology & company-reported data underlying each score." Bloomberg Professional Services
- Some compensation contracts rely on ESG scores: Cohen et al. JAR 2023

Executive Compensation with Social and Environmental Performance

Chaigneau and Sahuguet 2024

- **Standard arguments**: ESG-based compensation is misguided:
  1. ESG performance is multifaceted
  2. ESG performance is poorly measured – there is even disagreement on ESG measurement!
  3. ESG-based incentives can be gamed: 'hit the target but miss the point'
- **Methodological contribution**: extend a principal-agent model of multitasking to allow the manager to observe the noise in performance measurement at the time of choosing his action $\rightarrow$ captures gaming
- **Results**:
  - ESG-based compensation is optimal only when the level of corporate social responsibility preferred by the board exceeds the one that maximizes the stock price
    $\Rightarrow$ social incentives and socially responsible investors are substitutes
  - Relying on multiple measures based on different methodologies will generally mitigate inefficiencies due to gaming
    $\Rightarrow$ harmonization of social performance measurement can backfire

Executive Compensation with Social and Environmental Performance

Chaigneau and Sahuguet 2024

Stylized principal-agent model of corporate governance with:

1. Socially conscious board that provides incentives
   - Many shareholders value social and environmental factors: Hartzmark and Sussman (2019), Barber et al. (2021), Bauer et al. (2021), Heeb et al. (2023)
2. Measures of financial performance and measures of social performance
   - Profits, stock price, "ESG scores"
3. Timing as in Edmans and Gabaix RFS 2011
   - The manager can anticipate how his social investment decisions will affect the firm's ESG scores $\Rightarrow$ gaming of ESG-based incentive system

Key ingredients: optimal contracting + stock price + gaming

- Other papers have only one of those ingredients: either stock price or optimal contracting

Executive Compensation with Social and Environmental Performance

Chaigneau and Sahuguet 2024

- Effort and investment decisions, risk neutral agent, normal distributions, linear contract.
- Investment y_i in dimension i of social performance is costly but improves social performance by a factor η_i.
- The agent needs financial incentives to exert effort, but these incentives may disincentivize social investment.
- Social performance on dimension i is measured with a contractible measure: $m_i = \varepsilon_i y_i$, where $\varepsilon_i \sim \mathcal{N}(\eta_i, \sigma_\varepsilon^2)$.
- The manager privately observes ε_i before making investment decision y_i.
  - Timing assumption similar to Edmans and Gabaix RFS 2011 which captures gaming

Timeline

BEHAVIORAL

Behavioral factors

- CEO **overconfidence**: Gervais, Heaton, and Odean JF 2011
- CEO **loss aversion**: de Meza and Webb JEEA 2007, Dittmann, Maug, and Spalt JF 2010
- CEO with **fairness** concerns: Chaigneau, Edmans, and Gottlieb 2024

fair CEO pay

ACTION SPACE

Action space

- **Repeated actions**: Holmström Milgrom ECMA 1987
  $\Rightarrow$ linear contract
- **Asymmetric info on technology**: Carroll AER 2015
  $\Rightarrow$ linear contract
- **Flexible technology**: Bonham and Riggs-Cragun 2023
  - Agent chooses the output distribution at a cost
  - The principal wants to reward output not effort

MATCHING

Gabaix and Landier QJE 2008

"the sixfold increase of U.S. CEO pay between 1980 and 2003 can be fully attributed to the sixfold increase in market capitalization of large companies during that period."

- Empirically, the level of CEO pay is strongly related to firm size
- Matching model in which firms compete for managers:
  - CEOs have different talent
  - Firms have different size (market capitalization)
  - CEO talent (T) has multiplicative effect on firms earnings (a): $a_1 = a_0(1 + C \times T)$ or $a_1 = a_0 + Ca_0^{\gamma} T$
  - A firm maximizes its discounted earnings minus CEO wage
  - Pareto firm size distribution: $S(n) = An^{-\alpha}$ with $\alpha \approx 1$
  - Extreme value theory: in the right tail, $T'(x) = -Bx^{\beta-1}$
- The competitive equilibrium involves positive assortative matching
- Pay at reference firm n is proportional to: $(S(n))^{\gamma}$ with $\gamma \approx 1$
  - A small difference in CEO talent implies a large difference in CEO pay
  - Extension with a moral hazard problem to study pay-for-performance and its relation with firm size: Edmans, Gabaix, and Landier RFS 2009
  - Search model: Cao and Wang JF 2013

Besley and Ghatak AER 2005

- A "motivated agent" (exogenous agent type) gets a nonpecuniary benefit from project success if she works with a "mission-oriented" principal (exogenous principal type).
- Model:
  - risk neutrality and limited liability
  - binary outcome ($\Rightarrow$ binary contract)
  - effort = probability of high outcome
  - agent can have three types:
    - type 0: standard pecuniary utility
    - type 1: nonpecuniary benefit from success of type 1 principal
    - type 2: nonpecuniary benefit from success of type 2 principal
- Agent motivation is a substitute for explicit incentives
  $\Rightarrow$ motivated agents receive lower incentives.
- Important to match principals and agents with similar preferences

Bandiera, Guiso, Prat, Sadun JLE 2015

"A firm should select a hiring policy in view of the incentive structure it has in place, and it should select an incentive structure in view of the people it wants to hire."

- Assumptions:
  - Managers have heterogeneous talent and risk aversion
  - Firm owners put different weights on private benefits of control relative to profits
  - Compensation contract fully described by two parameters: fixed wage and slope (PPS)
  - Effort and the distribution of firm performance are endogenously determined
- Matching: risk averse and untalented managers prefer low-powered incentives, and so do firms owners who put little weight on profits
- Empirical implication: firms that offer higher-powered incentives are more profitable (even though all economic agents behave optimally!)

CEO ability, monitoring, and termination

- Hermalin JF 2005: monitoring allows a firm to update its beliefs about CEO ability and dismiss a CEO with low ability
- Chaigneau and Sahuguet JFQA 2018: matching such that firms with larger blockholder ownership hire CEOs with more uncertain ability
  - Firms with a greater ability to monitor are more willing to hire a CEO with more uncertain ability
- Empirical implications:
  1. When the cost of CEO dismissal is high or managerial skills are highly transferable, firms prefer to hire a CEO with a more precisely estimated ability ($\neq$ Hermalin JF 2005)
     * Prediction: CEOs in firms with less concentrated ownership receive higher salaries
  2. Sensitivity of CEO pay to exogenous business conditions relative to firm value ("pay-for-luck") is higher in firms with more dispersed ownership
     * This explains the controversial finding in Bertrand and Mullainathan QJE 2001

INDUSTRIAL ORGANIZATION

Executive compensation and industrial organization

- Old literature: test the informativeness principle by considering relative performance evaluation.

- Antón, Ederer, Giné, Schmalz JPE 2023: embed moral hazard in a standard model of product market competition
  - Moral hazard: effort increases firm productivity
  - Competition: higher productivity $\Rightarrow$ lower product market price $\Rightarrow$ lower profits for all firms in the industry

- Common ownership: owners internalize the effects of increasing productivity in any firm

$\Rightarrow$ provide weaker managerial incentives, tolerate managerial slack

- Contribution: describe a plausible mechanism that explains *how* common ownership affects industry-level outcomes

ATTRACTION AND RETENTION

Oyer JF 2004

- Two periods model without moral hazard
- Exponential utility with ARA r, normal distribution with binary mean, linear contract
- Contracting at the beginning of the first period; costly recontracting/renegotiation at the end of the period
- Market conditions affect:
  1. The distribution of firm performance in the second period, which is $\tilde{\theta} \sim \mathcal{N}(\theta^i, \sigma^2)$ with a mean $\theta^i \in \{\theta^l, \theta^h\}$
  2. The manager's outside option s_i at the end of the first period

$\Rightarrow$ compensation contract is contingent on firm performance to reduce recontracting/renegotiation costs

- Contract linear in θ with fixed wage w and slope $b \Rightarrow$ manager's utility is described by certainty equivalent $w + b\,\theta^i - rb^2\sigma^2$
- Participation constraint: $w + b\,\theta^i - rb^2\sigma^2 \geq s_i$ for $i \in \{s_l, s_h\}$

Chaigneau and Sahuguet JFQA 2018

- Two periods model without moral hazard
- Risk averse managers, normal distributions, one-sided commitment
- Second period reservation wages are endogenously determined in a market equilibrium: they depend on firm performance and market conditions

$\Rightarrow$ optimal compensation contract can be implemented with a fixed wage and indexed options on firm value:

1. Downside protection to insure risk averse managers
2. Upside participation to retain a manager with outside options
   - Managers with more uncertain ability are matched with firms with higher blockholder ownership $\Rightarrow$ compensation of these managers is more sensitive to firm value

- An improvement in the monitoring ability of the worst firms (e.g. diffusion of best monitoring practices) increases CEO pay in all firms
  - Market equilibrium effect when firms compete for CEOs
  - Examples: $\uparrow$ blockholder ownership, or $\downarrow$ cost of monitoring

THE STANDARD MODEL

The principal-agent model

- Assume a risk neutral principal and a risk averse agent with utility function u: $u' > 0$, $u'' < 0$.
- The agent chooses effort $e \in \mathcal{E}$ at personal cost $C(e)$.
  For example: $\mathcal{E} = [0, \overline{e}]$, or $\mathcal{E} = \{0, \overline{e}\}$.
- Optimization problem:

$$\max_{e, w(x)} \mathbb{E}\left[\tilde{x} - w(\tilde{x}) | e\right]$$

$$\text{subject to} \quad \mathbb{E}\left[u(w(\tilde{x})) | e\right] - C(e) \geq U,$$

$$\mathbb{E}\left[u(w(\tilde{x})) | e\right] - C(e) \geq \mathbb{E}\left[u(w(\tilde{x})) | \hat{e}\right] - C(\hat{e}) \quad \forall \hat{e} \in \mathcal{E}.$$

- There can also be constraints on contracting such as:
  1. limited liability for the agent: $w(x) \geq 0$
  2. limited liability for the principal: $w(x) \leq x$
  3. monotonicity constraint: $x - w(x)$ is nondecreasing in x:
     $\frac{\partial}{\partial x}\{x - w(x)\} \geq 0 \Leftrightarrow w'(x) \leq 1$; see Innes JET 1990 back

Two steps approach

Grossman and Hart ECMA 1983

Two-steps approach to the maximization problem:

1. For a given effort $e \in \mathcal{E}$, minimize the cost $V(e)$ of inducing this effort.

   $$V(e) \equiv \mathbb{E}[w(\tilde{x})|e] \text{ for a contract } w \text{ that solves the problem above}$$

2. Optimize over the level of effort:

   $$\max_e \mathbb{E}[\tilde{x}|e] - V(e)$$

Many papers only solve the first step.

Holmström's formulation

Holmström BJE 1979

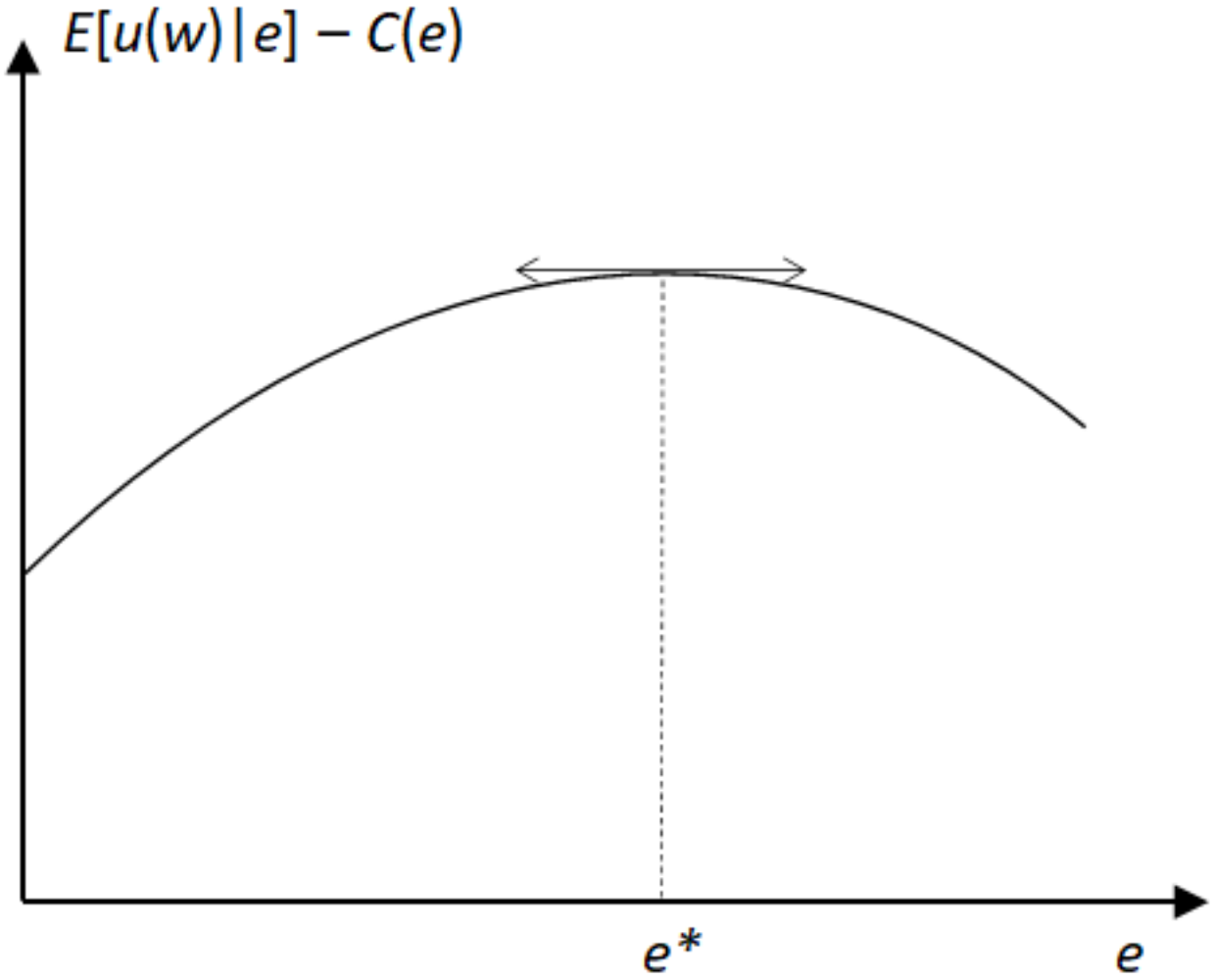
Output $\tilde{x} \in [\underline{x}, \overline{x}]$ with density $f(x|e)$. Effort $e \in [0, \bar{e}]$.
Optimization problem:

$$\max_{e, w(x)} \int_{\underline{x}}^{\overline{x}} (x - w(x))\, f(x|e) dx$$

$$\text{s.t. PC} \quad \int_{\underline{x}}^{\overline{x}} u(w(x)) f(x|e) dx - C(e) \geq U$$

$$\text{IC} \quad \int_{\underline{x}}^{\overline{x}} u(w(x)) f(x|e) dx - C(e) \geq \int_{\underline{x}}^{\overline{x}} u(w(x)) f(x|\hat{e}) dx - C(\hat{e}) \quad \forall \hat{e}$$

First-order approach (FOA): replace IC by $\int_{\underline{x}}^{\overline{x}} u(w(x)) f_e(x|e) dx = C'(e)$



Given a contract w, optimal effort solves:

$$\frac{\partial}{\partial e}\left\{\int_{\underline{x}}^{\overline{x}} u(w(x))f(x|e)dx\right\} - C'(e) = 0$$

First-Order Approach (FOA)

- The first-order condition to the agent's optimization problem gives the optimal level of effort given the contract if the agent's objective function, $\int_{\underline{x}}^{\overline{x}} u(w(x))f(x|e)dx - C(e)$, is globally concave in effort.
- Integrate by parts:

$$\begin{aligned}\int_{\underline{x}}^{\overline{x}} u(w(x))f(x|e)dx &= [u(w(x))F(x|e)]_{\underline{x}}^{\overline{x}} - \int_{\underline{x}}^{\overline{x}} w'(x)u'(w(x))F(x|e)dx \\ &= u(w(\overline{x})) - \int_{\underline{x}}^{\overline{x}} \underbrace{w'(x)}_{>0 \text{ by MLRP}} \underbrace{u'(w(x))}_{>0} F(x|e)dx\end{aligned}$$

- Rogerson (ECMA 1985) condition is $F_{ee}(x|e) > 0$ (CDFC). With MLRP, it holds irrespective of the contract but it imposes a strong condition on the distribution of output.

Other conditions for the FOA:

1. Jewitt ECMA 1988
2. Jung and Kim JET 2015
3. Chaigneau, Edmans, and Gottlieb RFS 2022 in a setting with limited liability

FOA with limited liability

Chaigneau, Edmans, and Gottlieb RFS 2022:

- Sufficient condition for the FOA with log utility, normal distribution, and quadratic effort cost ($C(e) = \frac{\beta}{2}e^2$): $\beta \geq \frac{\bar{e}}{\sigma^2}$
- Sufficient condition for the FOA with utility bounded from above:

$$K_e^+ := \sum_s \int_{\underline{q}}^{+\infty} \max\left\{\frac{\partial^2 f}{\partial e^2}(q,s|e),\, 0\right\} dq, \tag{B.1}$$

$$K_e^- := \sum_s \int_{\underline{q}}^{+\infty} \min\left\{\frac{\partial^2 f}{\partial e^2}(q,s|e),\, 0\right\} dq. \tag{B.2}$$

Lemma B.1 (First-Order Approach). Suppose that

$$K_e^- u(\bar{W}) + K_e^+ \lim_{c \nearrow \infty} u(c) < C''(e) \tag{B.3}$$

for all $e \in (0, \bar{e})$. Then, the FOA is valid.

Example B.1. Let $\bar{W} = 10$, $u(w) = \frac{w^{-2}}{-2}$ (i.e., the manager has CRRA utility with $\gamma = 3$), the cost of effort is $C(e) = \alpha e + \frac{\beta}{2}e^2$, with $\alpha \geq 0$ and $\beta > 0$, the set of possible effort levels is $e \in [0, 10]$, output follows a normal distribution with mean e and standard deviation 2, $S = 1$, and the manager is protected by limited liability. Then we have: $\int_{\underline{q}}^{\infty} \min\left\{\frac{\partial^2 f}{\partial e^2}(q|e,s), 0\right\} dq \approx -0.121$ for all e, $u(\bar{W}) = -\frac{1}{200}$, $C''(e) = \beta$, and the condition in Lemma B.1 is simply $\beta > \frac{0.121}{200}$.

Microfoundation for binary effort model

Many models assume that effort is binary. This is not very realistic but it can be microfounded as in Example 1 in Chaigneau Edmans Gottlieb GEB 2019. The logic is as follows:

- Effort is chosen optimally s.t. $e \in [0, \overline{e}]$
- Integrating by parts shows that, with MLRP and $F_{ee}(x|e) < 0$, the agent's expected utility of wealth is convex w.r.t. effort, i.e. the optimal effort level is either $e = 0$ or $e = \overline{e}$.
- The probability of observing an outcome below x decreases at an increasing rate as effort increases: "stochastic increasing returns to scale"
  - With this type of production technology, there is no point in making a half-hearted effort.
- This justifies solving a simpler ("relaxed") optimization problem with binary effort and a single IC that compares the agent's expected utility under effort $e = \overline{e}$ and effort $e = 0$.

Holmström's formulation

Holmström BJE 1979

Lagrangian:

$$\mathcal{L} = \int_{\underline{x}}^{\overline{x}} (x - w(x))\, f(x|e)dx + \lambda \left(\int_{\underline{x}}^{\overline{x}} u(w(x)) f(x|e)dx - C(e) \right) + \mu \left(\int_{\underline{x}}^{\overline{x}} u(w(x)) f_e(x|e)dx - C'(e) \right)$$

First-order condition (FOC) with respect to $w(x)$ (differentiate pointwise):

$$-f(x|e) + \lambda u'(w(x)) f(x|e) + \mu u'(w(x)) f_e(x|e) = 0$$

Optimal contract that induces a given effort e:

$$\frac{1}{u'(w(x))} = \lambda + \mu \frac{f_e(x|e)}{f(x|e)} \tag{1}$$

MLRP: $\frac{f_e(x|e)}{f(x|e)}$ increasing in $x \Rightarrow w'(x) > 0$ $\left(\frac{\partial}{\partial x} \frac{1}{u'(w(x))} = -\frac{w'(x)u''(w(x))}{(u'(w(x)))^2}\right)$

Structure of the contract

Chaigneau, Sahuguet, and Sinclair-Desgagné EL 2017

- Write the likelihood ratio of x given e as: $LR(x|e) \equiv \frac{f_e(x|e)}{f(x|e)}$.
- The structure of the optimal contract is such that:

$$\frac{w''(x)}{w'(x)} = \frac{LR''(x|e)}{LR'(x|e)} + \underbrace{w'(x)}_{>0}\left[P(w(x)) - 2A(w(x))\right]$$

where $A(w) \equiv -\frac{u''(w)}{u'(w)}$ is the coefficient of absolute risk aversion

$P(w) \equiv -\frac{u'''(w)}{u''(w)}$ is the coefficient of absolute prudence

- Optimal contract is more convex than LR iff $P(w(x)) > 2A(w(x))$.
  - With CRRA utility, $P(w) > 2A(w)$ for any w iff $RRA < 1$.
- With a skew-normal distribution, LR is convex iff skewness is negative, see Chaigneau, Chang, Hillegeist CAR 2024

Two examples

- Assume:
  1. $u(w) = \ln(w) \quad \Rightarrow \quad \frac{1}{u'(w)} = w$
  2. $\tilde{x} \sim \mathcal{N}\left(e, \sigma^2\right)$

     $\Rightarrow \quad f_e(x|e) = \frac{x-e}{\sigma^2} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-e}{\sigma}\right)^2\right\} = \frac{x-e}{\sigma^2} f(x|e)$

     In this case, effort e is a location parameter of the output distribution.

  Then, for a given effort e, equation (1) rewrites as:

$$w(x) \quad = \quad \lambda + \mu \frac{x-e}{\sigma^2}$$

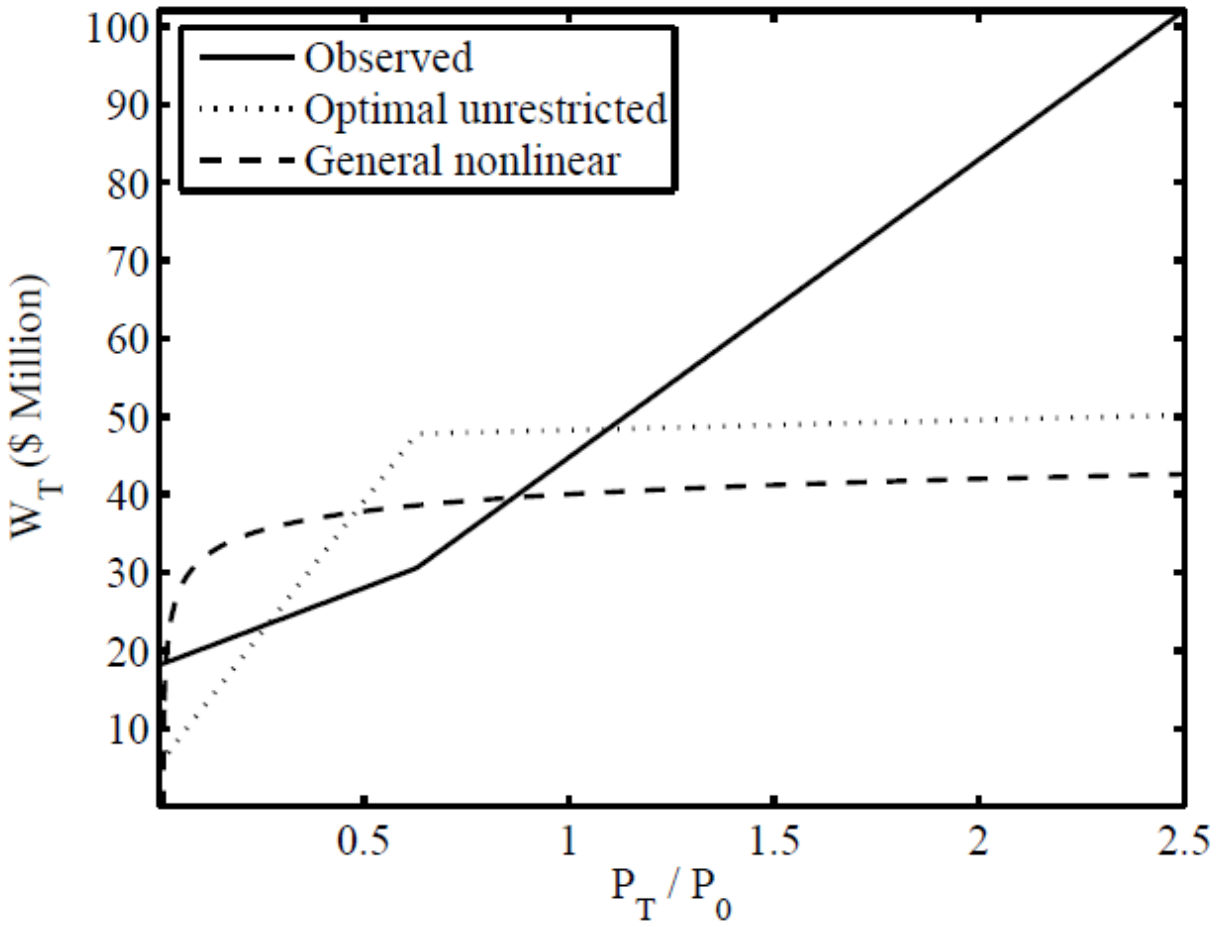
  $\rightarrow$ the payment w is linear in output x.

- Dittmann and Maug JF 2007 calibrate a principal-agent model with CRRA utility (RRA $= \gamma$), limited liability, and a lognormally distributed stock price:

$$w(x) \quad = \quad \begin{cases} \left(\lambda + \mu \frac{\ln(x)-e}{\sigma^2}\right)^{1/\gamma} & \text{if } x \geq \hat{x} \\ 0 & \text{if } x < \hat{x} \end{cases}$$

"We conclude that we need a different contracting model to understand salient features of executive compensation contracts."

Dittmann and Maug JF 2007



Constraints on contracting

This generates realistic contracts under some conditions

- Innes JET 1990: optimal contract with *risk neutral* agent, limited liability, and monotonicity constraint:

$$w(x) = \max\{x - x^*, 0\} \qquad \rightarrow \text{inside equity, outside debt}$$

- Jewitt, Kadan, and Swinkels JET 2008: optimal contract with risk averse agent and limited liability:

$$\frac{1}{u'(w(x))} = \begin{cases} \lambda + \mu \frac{f_e(x|e)}{f(x|e)} & \text{if } \lambda + \mu \frac{f_e(x|e)}{f(x|e)} \geq \frac{1}{u'(0)}, \\ \frac{1}{u'(0)} & \text{if } \lambda + \mu \frac{f_e(x|e)}{f(x|e)} < \frac{1}{u'(0)}, \end{cases}$$

- Chaigneau, Edmans, and Gottlieb 2024a: optimal contract with *risk averse* agent, limited liability, and monotonicity constraint:

$$w(x) = \max\{x - x^*, 0\} \qquad \text{if } \frac{\frac{d}{dx} LR(x|e)}{LR(x|e)} \geq A\left(\bar{W} + x - x^*\right) \ \forall x > x^*$$

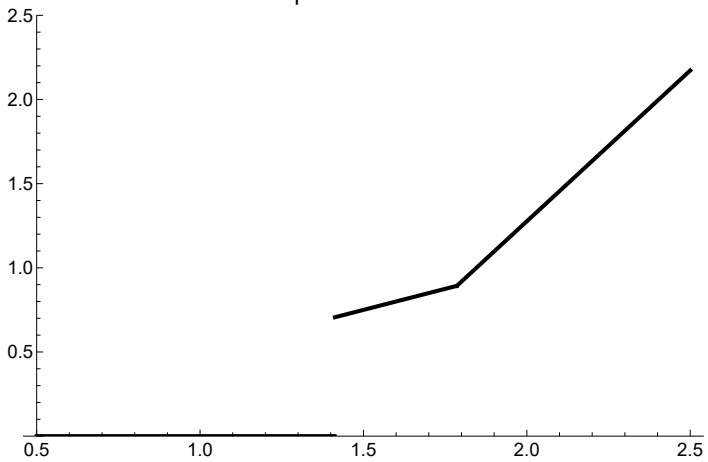
CEOs with fairness concerns

Chaigneau, Edmans, and Gottlieb 2024b behavioral papers

- Motivated by the survey by Edmans, Gosling, and Jenter JFE 2023
- CEO utility function:

$$u(w,x) = \begin{cases} w & \text{if } w(x) \geq w^*(x) \\ w - \gamma\left(w^*(x) - w\right) & \text{if } w(x) < w^*(x) \end{cases}$$

optimal contract vs. observed contract with performance-based vesting:



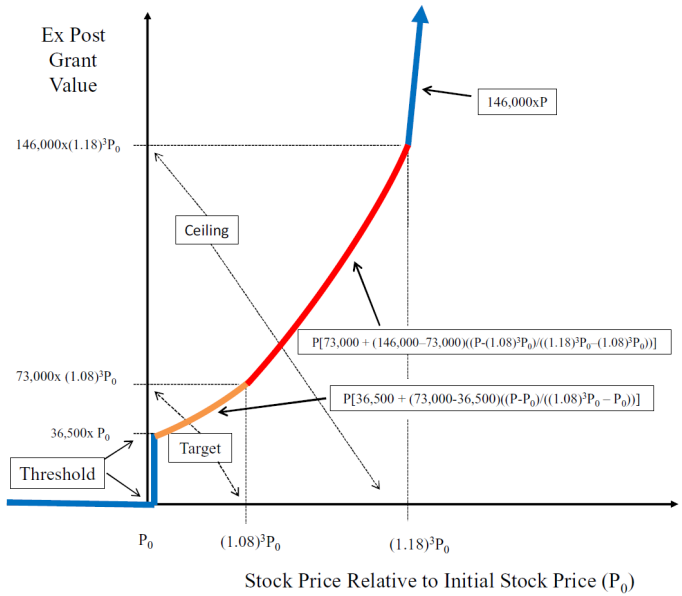


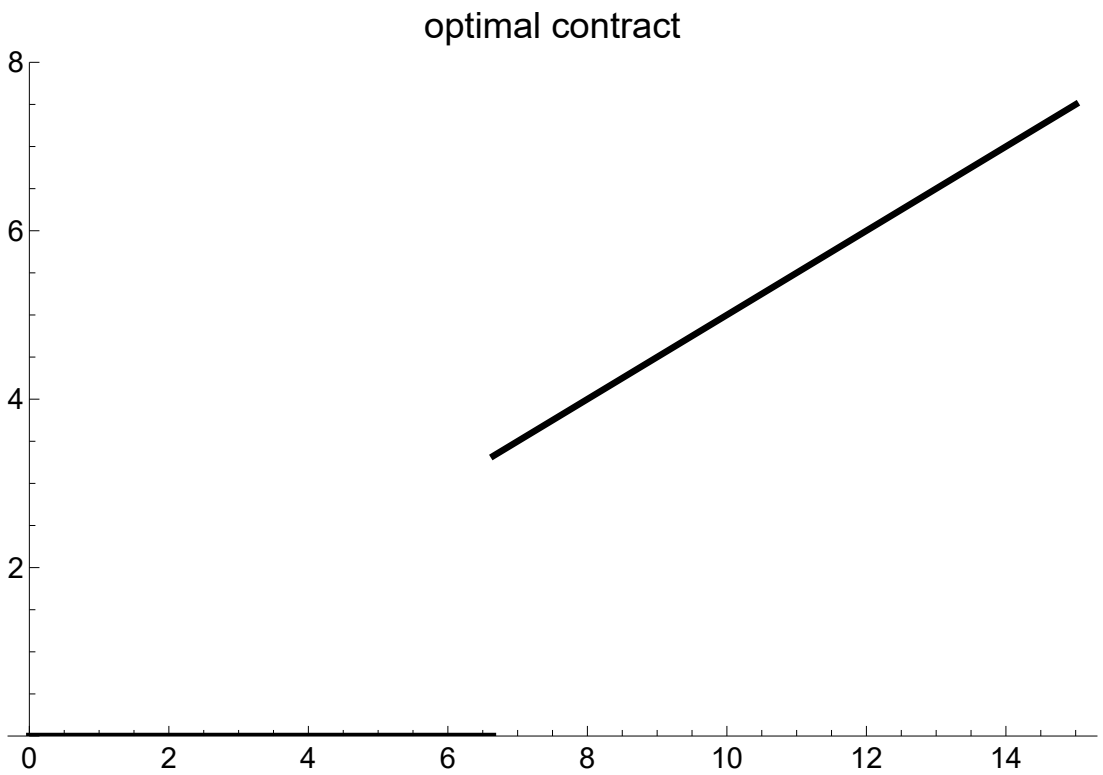
Fig. 2. Ex post value of amgen 2008 grant to CEO of stock with conditional vesting provision.

Source: Bettis, Bizjak, Coles, Kalpathy JAE 2018

CEOs with fairness concerns

Chaigneau, Edmans, and Gottlieb 2024b

- performance-based vesting
- realistic discontinuity
- realistic pay-for-performance and structure
- . . . even for a manager risk averse above the fair wage



The value of information

- *When* to incorporate additional signals of performance in a contract:
  - Holmström BJE 1979: **informativeness principle** ("sufficient statistic" result): filter out exogenous shocks, do not reward for luck, use incrementally informative performance signals.
  - What about when contracting constraints (limited liability, monotonicity, ...) bind? $\rightarrow$ new condition for a signal to be valuable: Chaigneau, Edmans, and Gottlieb [2014]2024a
- *How* to incorporate additional signals of performance in a contract:
  - Chaigneau, Edmans and Gottlieb RFS 2022 impose some structure on the distribution of signals and derive empirical implications
- What is the value of a more precise signal of performance?
  - Chaigneau, Edmans, and Gottlieb JFE 2018 show that a more precise signal can reduce effort incentives
- Progressive learning or attenuation of information asymmetries over time $\Rightarrow$ **deferred compensation**, e.g. time-based vesting
  - Chaigneau FRL 2018
  - Hoffmann, Inderst, and Opp REStud 2021

Mirrlees mechanism

- A simple mechanism can solve the agency problem at zero agency cost under some assumptions:
  1. the distribution of $\tilde{x}$ has a location parameter, e, i.e. $\tilde{x} = e + \tilde{\epsilon}$
  2. $\tilde{x}$ has bounded support, denoted by $[\underline{x} + e, \overline{x} + e]$
     $\Rightarrow$ effort moves the support
  3. utility function u is unbounded from below
  4. no constraints on contracting

  Mechanism: fixed wage and arbitrarily large punishment for $x < \underline{x} + e^*$, where e^* is the effort level to be induced.
  - This allows to achieve first-best risk sharing and effort inducement.

- See Mirrlees 1975 for the case with unbounded support (e.g. normal distribution) when the optimal contract does not exist.

Formulating the problem

Grossman and Hart ECMA 1983

Rewrite the optimization problem with "utils":
$u_x \equiv u(w(x)) \quad \Rightarrow \quad w(x) = u^{-1}(u(w(x))) = u^{-1}(u_x)$.
Assuming $u' > 0$ and $u'' < 0 \quad \Rightarrow \quad u^{-1\prime\prime} > 0$.

$$\max_{e,u_x} \mathbb{E}\left[\tilde{x} - u^{-1}(\tilde{u}_x)|e\right]$$

$$\text{subject to PC} \qquad \mathbb{E}\left[\tilde{u}_x|e\right] - C(e) \geq U,$$

$$\text{IC} \qquad \mathbb{E}\left[\tilde{u}_x|e\right] - C(e) \geq \mathbb{E}\left[\tilde{u}_x|\hat{e}\right] - C(\hat{e}) \quad \forall \hat{e} \in \mathcal{E}.$$

Formulating the problem

- Assume $\tilde{x} \equiv e + \tilde{\epsilon}$, where $\tilde{\epsilon}$ is independent of e and has density g such that $f(x|e) = g(x - e)$.
- Mirrlees/Holmström formulation:

$$\mathbb{E}[u(w(x))] = \int_{\underline{x}}^{\overline{x}} u(w(x)) f(x|e) dx \tag{2}$$

- State space formulation:

$$\mathbb{E}[u(w(x))] = \int_{\underline{\epsilon}}^{\overline{\epsilon}} u(w(e + \epsilon)) g(\epsilon) d\epsilon \tag{3}$$

  - $g(\epsilon)$ is independent of effort whereas $f(x|e)$ is not
  - But $w(\cdot)$, which is not exogenously given, may not be differentiable
- For an application see Poblete and Spulber RAND 2012