Estimating Models of Optimal Contracting

Robert A. Miller

DSE 2022 MIT

August 2022

A Pure Moral Hazard Model

- Gross revenue to a *risk neutral* principal is a random variable x.
- The probability distribution for x depends on choices by a *risk averse* agent.
- The principal proposes a *compensation plan* to the agent, denoted by $w(x)$.
- The agent an *employment* choice:
  - *rejecting* the principal's offer in favor of an outside option $(l_0 = 1)$
  - *accepting* the principal's offer $(l_0 = 0)$.
- If $l_0 = 0$ the agent makes an *effort* choice:
  - *working*, pursuing value maximization $(l = 1)$.
  - *shirking*, his optimal action if paid a fixed wage $(l = 0)$.
- The principal observes l_0 whether the offer is accepted, but not the agent's work effort l.
- After revenue is realized, the principal pockets $x - w(x)$ as profit.

- Denote by $f(x)$ the pdf for x conditional on the agent working.
- Let $f(x)g(x)$ denote the pdf for x when the agent shirks.
- The expected value of x from work exceeds the expected value from shirking:

$$E[xg(x)] \equiv \int xf(x)g(x)\,dx < \int xf(x)\,dx \equiv E[x]$$

- The inequality reflects the preference of principal for working over shirking.
- We assume the likelihood ratio $g(x) \to 0$ as $x \to \infty$: intuitively, shirking hardly ever produces truly extraordinary results.

- The agent is an expected utility maximizer with utility:

$$-l_0 - \alpha E\left[e^{-\rho w(x)}\right] l - \beta E\left[e^{-\rho w(x)} g(x)\right](1-l)$$

where:

  - utility is normalized to the outside option (when $l_0 = 1$).
  - ρ is the coefficient of absolute risk aversion.
  - α is a utility parameter with consumption equivalent $-\rho^{-1}\ln(\alpha)$ measuring the distaste for working.
  - β is a utility parameter with consumption equivalent $-\rho^{-1}\ln(\beta)$ measuring the distaste for shirking.

- A conflict of interest arises between the principal and the agent because:
  - the agent prefers shirking, meaning $\beta < \alpha$.
  - yet the principal prefers working since $E[xg(x)] < E[x]$.

- To induce the agent to accept employment and shirk, it suffices to propose a contract giving the agent an expected utility of at least minus one:

$$\beta E\left[e^{-\rho w(x)} g(x)\right] \equiv \beta E\left[v(x) g(x)\right] \leq 1$$

- To elicit work from the agent as well, the principal must offer a contract that gives the agent a higher expected utility than:
  1. the outside option (the *participation constraint*):

$$\alpha E\left[e^{-\rho w(x)}\right] \equiv \alpha E\left[v(x)\right] \leq 1$$

  2. the utility from shirking (the *incentive compatibility constraint*):

$$\alpha E\left[e^{-\rho w(x)}\right] \equiv \alpha E\left[v(x)\right] \leq \beta E\left[v(x) g(x)\right] \equiv \beta E\left[e^{-\rho w(x)} g(x)\right]$$

Solving the Pure Moral Hazard Model

- In the transformed problem of $v(x) \equiv \exp\left[-\rho w\left(x\right)\right]$ the principal maximizes the strictly concave objective function with linear constraints:

$$E\left\{\log\left[v(x)\right]\right\} + \eta_0 E\left[1 - \alpha_2 v(x)\right] + \eta_1 E\left[\beta g\left(x\right) v(x) - \alpha v(x)\right]$$

to obtain:

$$w^o\left(x\right) \equiv \rho^{-1} \ln \alpha + \rho^{-1} \ln\left[1 + \eta\left(\frac{\alpha}{\beta}\right) - \eta g(x)\right]$$

where η is the unique positive solution to the equation:

$$E\left[\frac{g\left(x\right)}{\alpha + \eta[(\alpha/\beta) - g\left(x\right)]}\right] = E\left[\frac{(\alpha/\beta)}{\alpha + \eta[(\alpha/\beta) - g\left(x\right)]}\right]$$

Measuring the Importance of Moral Hazard

Three measures

1. With *perfect monitoring* the principal would pay the agent a fixed wage of $\rho^{-1} \ln \alpha_2$ so the maximal amount prinicpal would pay for a perfect monitor is:

$$\tau_1 \equiv E_t \left[w^o(x) - \rho^{-1} \ln \alpha \right] = \rho^{-1} E \left\{ \ln \left[1 + \eta \left(\frac{\alpha}{\beta} \right) - \eta g(x) \right] \right\}$$

2. The agent's *nonpecuniary benefits from shirking* is:

$$\tau_2 \equiv \rho^{-1} \ln \left(\alpha / \beta \right)$$

3. The *gross loss* the principal incurs from the agent shirking instead of working is:

$$\tau_3 \equiv E \left[x - xg(x) \right]$$

Identification

Model primitives and the data generating process

- The model is defined by:
    - $f(x)$ the probability density function of x from working
    - $g(x)$ the likelihood ratio for shirking versus working
    - α distaste for working relative to outside option
    - β distaste for shirking relative to outside option
    - ρ risk-aversion parameter.
- The panel data set is $\{x_n, w_n\}_{n=1}^{N}$ where $w(x) = E[w_n | x_n]$.
- Thus $f(x)$ and $w(x)$ are identified.
- This leaves only $g(x)$ plus (α, β, ρ) to identify.

Identification

What if the risk parameter is known?

- The FOC for the Lagrangian can be expressed as:

$$v(x)^{-1} = \alpha \left[1 + \eta \left(\alpha / \beta\right) - \eta g(x)\right] = \overline{v}^{-1} - \alpha \eta g(x)$$

where:

$$\lim_{x \to \infty} \left[g(x)\right] = 0 \Rightarrow \lim_{x \to \infty} \left[v(x)^{-1}\right] = \alpha \left[1 + \eta \left(\alpha / \beta\right)\right] \equiv \overline{v}^{-1}$$

- These equalities imply:

$$g(x) = \frac{\overline{v}^{-1} - v(x)^{-1}}{\alpha \eta} = \frac{\overline{v}^{-1} - v(x)^{-1}}{\overline{v}^{-1} - E\left[v(x)^{-1}\right]} \tag{1}$$

- Also since both pariticpation and incentive compatibility constraints bind:

$$\begin{aligned} \alpha &= E\left[v(x)\right]^{-1} &(2)\\ \beta &= E\left[v(x) g\left(x\right)\right]^{-1} &(3) \end{aligned}$$

Identification

The identified set (Gayle and Miller, 2015)

- Noting $v(x) = e^{-\rho w(x)}$ and $\overline{v} \equiv e^{-\rho \overline{w}}$ equations (1), (2) and (3) imply:

$$\alpha(\rho) = E\left[e^{-\rho w^o(x)})\right]^{-1}$$

$$\beta(\rho) = \frac{1 - E\left[e^{\rho w^o(x) - \rho \overline{w}}\right]}{E\left[e^{-\rho w^o(x)}\right] - e^{-\rho \overline{w}}}$$

$$g(x, \rho) = \frac{e^{\rho \overline{w}} - e^{\rho w^o(x)}}{e^{\rho \overline{w}} - E\left[e^{\rho w^o(x)}\right]}$$

- Finally since paying $w^o(x)$ is more profitable than paying $\gamma^{-1} \ln(\beta)$:

$$\begin{aligned} 0 &\leq E[x] - E\left[w^o(x)\right] - E[xg(x)] + \rho^{-1} \ln(\beta) \\ &= \frac{cov\left(x, e^{\rho w^o(x)}\right)}{e^{\rho \overline{w}} - E\left[e^{\rho w^o(x)}\right]} - E\left[w^o(x)\right] + \rho^{-1} \ln\left(\frac{1 - E\left[e^{\rho w^o(x) - \rho \overline{w}}\right]}{E\left[e^{-\rho w^o(x)}\right] - e^{-\rho \overline{w}}}\right) \end{aligned}$$

- The set of ρ satisfying this inequality is *sharp* and *tight*, and the model is rejected if this set is empty.

Identification

Introducing dynamics

- Adding dynamics to this model further restricts the set of *observationally equivalent* parameterizations:
  1. In a multiperiod model where the agent can borrow and save *changes in the interest rate* (or bond price) affect the value of (smoothing) an extra dollar, leading to adjustments in the incentive compatibility and participation constraints (Gayle and Miller, 2009)
  2. Letting the agent choose between alternative sources of employement offering *different contracts* with uncertain compensation (like different lotteries), also shrinks the set (Gayle and Miller, 2015).
- We now extend the prototype to a generalized *Roy model*:
  - in a *dynamic* setting
  - where there is one *principal* and *multiple agents*
  - each one having *several employment choices*
  - and accumulating *human capital*
- We apply this model to managerial compensation:
  - estimate the three measures of moral hazard defined above.
  - explain why executives in large firms are paid more than those in small firms.

Data

Sources and summary statistics (Gayle, Golan and Miller, 2012)

- Data taken from ExecuComp for the S&P 1500 and COMPUSTAT were matched with data from Who's Who for the years 1992-2006.
- The matching algorithm yielded 16,300 executives (from 30,614) in 2100 firms (from 2818) yielding 59,066 observations.
- Data on executives include: compensation, title, including interlock status, and background, including age, gender, education, annual transitions by title and firm.
- Data on firms include annual return, size (large, medium, small) and sector (primary, service, consumer).
- Summarizing:
  1. The exit rate is between 12% and 18% per year.
  2. Turnover is about 2% to 3% per year.
  3. Executives average between 51 and 54 years old.
  4. On average executives have about 13 to 14 years firm tenure.
  5. They average about 17 years executive experience.
  6. About 80% graduated from college and about 20% have an MBA.
  7. Total compensation averages between $1.5 and $4.5 million.
  8. Compensation increases with firm size.

Data

Compensation, education and tenure by firm size (Figures 1 and 2, GGM 2015, pages 2302-2303)

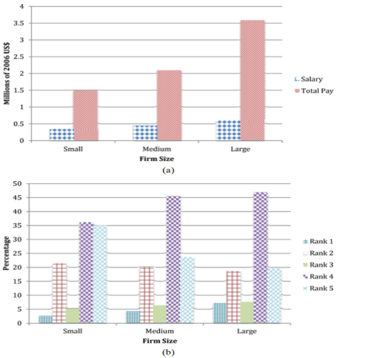

FIGURE 1.—Pay and hierarchy by firm size. (a) Firm size pay premium, (b) hierarchy by firm size.

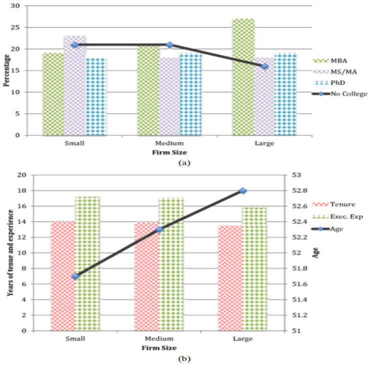

FIGURE 2.—Education and experience by firm size. (a) Education and firm size, (b) experience and firm size.

Data

What explains firm-size pay premium in the market for top executives?

- There are three basic factors that might be playing a role:
  1. Human Capital:
     1. Executives in large firms are older, more educated, but have less executive experience and less tenure than those in smaller firms; presumably human capital of the kind described by Mincer (1974) is playing a role.
     2. Working as executives in more firms increases an executive compensation at higher ranks in the hierarchy. This is a form of productivity enhancing on-the-job experience.
  2. Moral Hazard:
     1. Top executives are paid a significant portion of their total compensation in stock and options.
     2. The composition of firm denominated securities varies substantially across ranks and executives at different points in their lifecycle.

- Executive chooses job k in firm j, by setting indicator variable $d_{jkt} = 1$, and effort level $l_t \in \{0, 1\}$ where:
  - $j = j_1 \otimes j_2 \in \{0, 1\} \otimes \{1, 2..., J_2\}$
  - $j_1 \in \{0, 1\}$ denotes moving to a new firm ($d_{jt}^{(1)} = 1$) or not ($d_{jt}^{(1)} = 0$).
  - $j_2 \in \{1, 2..., J_2\}$, denotes firm size and industrial sectors
  - he retires by setting $d_{0t} = 1$ and:

$$d_{0t} + \sum_{j=1}^{J} \sum_{k=1}^{K} d_{jkt} = 1$$

- Human capital is given by $h_t \equiv (t, h_1, h_{2t})$ where h_1 are fixed demographics and $h_{2t} \equiv (h_{2t}^{(1)}, h_{2t}^{(2)}, h_{2t}^{(3)})$, where:
  - $h_{2t}^{(1)}$ is tenure with current firm and $h_{2,t+1}^{(1)} = 1 + \left(1 - d_{jt}^{(1)}\right) h_{2t}^{(1)}$
  - $h_{2t}^{(2)}$ is years of executive experience and $h_{2,t+1}^{(2)} = h_{2t}^{(2)} + 1$
  - $h_{2t}^{(3)}$ is the number of firms employed as an executive and $h_{2,t+1}^{(3)} = h_{2t}^{(3)} + d_{jt}^{(1)}$
- For notational convenience, let $H_{jk}(h_t)$ denote h_{t+1} when $d_{jkt} = 1$.

Model

Preferences and budget constraint

- Executives get utility from current consumption c_t.
- Executives have absolute risk aversion parameter ρ.
- Utility depends on h_t where h_1 includes education and gender.
- Jobs, firms, and effort level give nonpecuniary utility through functions $\beta_{jk}(h_t)$ (shirking) and $\alpha_{jk}(h_t)$ (working), where:

$$\alpha_{jk}(h_t) > \beta_{jk}(h_t) > 0$$

- An *iid* firm-job privately observed taste shock ε_{jkt} also affects utility.
- Lifetime utility is parameterized as:

$$-\sum_{t=1}^{\infty} \sum_{j=0}^{J} \sum_{k=1}^{K} \delta^t e^{-\rho c_t - \varepsilon_{jkt}} d_{jkt} \left[\alpha_{jk}(h_t) l_t + \beta_{jk}(h_t)(1 - l_t) \right]$$

where we abbreviate by setting $d_{0kt} \equiv d_{0t}$ for all k.

- There are complete markets for all publicly disclosed events, but no borrowing against future executive compensation.

- Denoting calendar time by τ, firm production is:

$$\sum_{k=1}^{K} F_{jk\tau}\left(h_{t(\tau)}^{(k)}\right) + e_{j\tau}\left(\pi_{\tau+1} - 1\right) + e_{j\tau}\pi_{j,\tau+1}$$

where for expositional ease, each executive holds a distinct position and:

  - $t\left(\tau\right)$ is the age of executive at calendar time τ.
  - $h_t^{(k)}$ denotes the human capital of the executive in position k.
  - $F_{jk,t(\tau)}\left(h_t\right)$ denote the individual contribution of k to the firm.
  - $e_{j\tau}$ denotes the value of firm j at the beginning of calendar time τ.
  - $\pi_{\tau+1}$ denotes the gross returns to the market portfolio.
  - $\pi_{j,\tau+1}$, denotes abnormal return to the firm before executive compensation.

- We assume the probability density for $\pi_{j,\tau+1}$ is:
  - $f_j(\pi_{j,\tau+1})$ when all K executives work
  - $f_j(\pi_{j,\tau+1})g_{jk}(\pi_{j,\tau+1}\left|h_t\right.)$ when all executives but k work.

- The gross expected return to a firms are higher if everybody works:

$$\int \pi f_j\left(\pi\right) d\pi > \int \pi f_j\left(\pi\right) g_{jk}(\pi\left|h_t\right.) d\pi$$

Model

Timing, information, and overview

- Each executive knows his h_t and privately observes realization of ε_{jkt}.
- He selects a firm and position, and submits a compensation proposal, w_{jkt+1}, to shareholders represented by a board.
- If his demand is not approved, the executive retires.
- If the board approves his proposed compensation plan, the executive privately chooses consumption c_t and effort l_t.

- Recursively define $A_t(h)$ an index of human capital by:

$$
\begin{aligned}
A_t(h) &= p_{0t}(h) E\left[\exp\left(-\varepsilon^*_{0t}/b_t\right)\right] \\
&+ \sum_{j=1}^{J}\sum_{k=1}^{K}\left(\begin{array}{c} p_{jkt}(h)\left[\alpha_{jkt}(h)\right]^{\frac{1}{b_t}} E\left[\exp\left(-\varepsilon^*_{jkt}/b_t\right)\right] \\ \times\left\{A_{t+1}\left[H_{jk}(h)\right] E_t\left[\upsilon_{jk,t+1}\right]\right\}^{1-\frac{1}{b_t}} \end{array}\right)
\end{aligned}
$$

where:

- b_t is the bond price at t.
- $\upsilon_{jk,t+1} = \exp\left(-\rho w_{jk,t+1}/b_{t+1}\right)$
- ε^*_{jkt} is the value of the private disturbance ε_{jkt} conditional on $d_{jkt} = 1$.
- $p_{jkt}(h)$ is the CCP for choosing rank k in firm j, period t.

- Lower values of $A_t(h)$ are associated with higher values of human capital.
- Defining $\Gamma[\cdot]$ as the complete gamma function, if ε_{jkt} is distributed T1EV then:

$$A_t(h) = p_{0t}(h)\,\Gamma\left[1 + \frac{1}{b_{t+1}}\right] \tag{4}$$

Firm and Job Choices

Optimization (Theorem 4.2 of GGM 2015)

- The value function is derived in two steps, solving for:
  1. optimal consumption given any career path
  2. the optimal career path.
- In the second step jobs are chosen to maximize:

$$\sum_{j=0}^{J} \sum_{k=0}^{K} d_{jkt} \left\{ \varepsilon_{jkt} - \ln \alpha_{jkt}(h) - (b_t - 1) \left(\ln A_{t+1}(H_{jk}(h)) + \ln E_t[v_{jk,t+1}] \right) \right\} \tag{5}$$

- Executives trade off jobs based on three dimensions:
  1. nonpecuniary benefit, $\alpha_{jkt}(h)$;
  2. human-capital accumulation, $H_{jk}(h) - h$;
  3. expected utility from compensation, $E_t[v_{jk,t+1}]$.

Cost Minimization

Participation constraint

- By the inversion theorem there exists $q(p)$ to R^{JK} such that:

$$q_{jk}\left[p_t(h)\right] = \ln\left[\alpha_{jkt}(h)\right] + (b_t - 1)\left\{\ln A_{t+1}\left(H_{jk}(h)\right) + \ln E_t[\upsilon_{jk,t+1}]\right\} \tag{6}$$

where $q_{jk}\left[p_t(h)\right] \equiv \varepsilon'_{jkt} - \varepsilon'_{0t}$, for all shock pairs $\left(\varepsilon'_{0t}, \varepsilon'_{jkt}\right)$ making the executive indifferent between retiring and (j,k).

- Define $w^*_{jk,t+1}(h)$ as the certainty equivalent wage to a executive indifferent between $(j.k)$ and retirement given CCPs $p_t(h)$:

$$q_{jk}\left[p_t(h)\right] = \ln\alpha_{jkt}(h) + (b_t - 1)\left\{\begin{array}{l}\ln A_{t+1}\left(H_{jk}(h)\right) \\ +\ln E_t[\exp\left(-\rho w^*_{jk,t+1}(h)/b_{t+1}\right)]\end{array}\right\}$$

- Solving for $w^*_{jk,t+1}(h)$ gives the participation constraint:

$$w^*_{jk,t+1}(h) = \frac{b_t}{\rho}\left\{\frac{1}{(b_t-1)}\ln\alpha_{jkt}(h) + \ln A_{t+1}\left[H_{jk}(h)\right] - \frac{1}{(b_t-1)}q_{jk}\left[p_t(h)\right]\right\}$$

Cost Minimization

Incentive compatibility constraint

- One-period (short term) contracts are optimal in this model. (See Fudenberg, Holmstrom and Milgrom, 1990.)
- In this model the firm can deter shirking in a one-period contract by offering a compensation schedule that satisfies the incentive-compatibility constraint:

$$\left[\frac{\alpha_{jkt}(h)}{\beta_{jkt}(h)}\right]^{1/(b_t-1)} \leq \frac{E_t\left[v_{jk,t+1}g_{jkt}(\pi \mid h)\right]}{E_t\left[v_{jk,t+1}\right]}. \tag{7}$$

- Paying the manager a constant wage, such as $w^*_{jk,t+1}(h)$, simplifies the right side of the (7) to:

$$\frac{\exp\left(-\rho w^*_{jk,t+1}(h)/b_{t+1}\right) E_t\left[g_{jkt}(\pi \mid h)\right]}{\exp\left(-\rho w^*_{jk,t+1}(h)/b_{t+1}\right)} = 1, \tag{8}$$

- Since $\alpha_{jkt}(h) > \beta_{jkt}(h)$, the inequality given by (7) is violated: paying a constant wage guarantees shirking in this model.

Cost Minimization

Optimal Contract (Theorem 4.3 of GGM 2015)

- Minimizing The cost minimizing contract is:

$$\begin{aligned} w_{jk,t+1}(h,\pi) &= w^{*}_{jk,t+1}(h) + r_{jk,t+1}(h,\pi) \\ &\equiv \Delta^{\alpha}_{jkt}(h) + \Delta^{A}_{jkt}(h) + \Delta^{q}_{jkt}(h) + r_{jk,t+1}(h,\pi) \end{aligned}$$

1. $\Delta^{\alpha}_{jkt}(h) \equiv \rho^{-1}\left(b_t - 1\right)^{-1} b_{t+1} \ln \alpha_{jkt}\left(h\right)$ is the systematic component of non-pecuniary utility of (j,k)
2. $\Delta^{A}_{jkt}(h) \equiv \rho^{-1} b_{t+1} \ln \left\{ A_{t+1}\left[H_{jk}\left(h\right)\right]\right\}$ is the investment value of (j,k).
3. $\Delta^{q}_{jkt}(h) \equiv \rho^{-1}\left(b_t - 1\right)^{-1} b_{t+1} q_{jk}\left[p_t\left(h\right)\right]$ are the idiosyncratic values making executive in fractal $p_{jkt}\left(h\right)$ indifferent between (j,k) and retirement.
4. $\Delta^{r}_{jkt}(h)$ is the risk premium defined as:

$$\Delta^{r}_{jkt}(h) \equiv E\left[r_{jk,t+1}(h,\pi)\right] = \frac{b_{\tau+1}}{\rho} E\left[\ln\left\{1 - \eta g_{jkt}\left(\pi \,|\, h\right) + \eta \left[\frac{\alpha_{jkt}(h)}{\beta_{jkt}(h)}\right]^{1/(b_{\tau}-1)}\right\}\right.$$

with η the unique positive root to:

$$\int \left\{ \eta^{-1} + \left[\frac{\alpha_{jkt}\left(h\right)}{\beta_{jkt}\left(h\right)}\right]^{1/(b_t - 1)} - g_{jkt}\left(\pi \,|\, h\right)\right\}^{-1} f_j\left(\pi\right) d\pi = 1$$

- Free entry by firms implies:

$$F_{jkt}\left(h\right)=E\left[w_{jk,t+1}^{*}(h)+r_{jk,t+1}(h,\pi)\right]$$

- This entry condition essentially completes the equilibrium.

Identification and Estimation

Compensating differentials and risk aversion

- Note that (5) is a dynamic discrete choice problem.
- Appealing to Arcidiacono and Miller (2020), $\alpha_{jkt}(h)$ and ρ are identified up the distribution of ε_t.
- Intuitively both are identified off from the different characteristics their job choices, inducing executives to reveal their attitude towards risk, the value they place on nonpecuniary features of the job, and their investment value.
- Assuming ε_t is T1EV, (4) and (6) imply the participation constraint can be expressed as:

$$\begin{aligned} \ln\left(\frac{p_{jkt}(h)}{p_{0t}(h)}\right) &= -\ln\alpha_{jkt}(h) - \frac{b_t - 1}{b_{t+1}} \ln p_{0,t+1}\left[H_{jk}\left(h_t\right)\right] \\ &\quad -(b_t - 1)\ln\Gamma\left[1 + \frac{1}{b_{t+1}}\right] - (b_t - 1)\ln E_t[v_{jk,t+1}] \end{aligned} \tag{9}$$

- Sample analogs were constructed for the CCPs, compensation schedule, and conditional and unconditional densities of the abnormal return.
- A GMM estimator can be constructed from moment conditions using (9).

Estimates from the Structural Model

Figure 3 from GGM 2015, page 2345

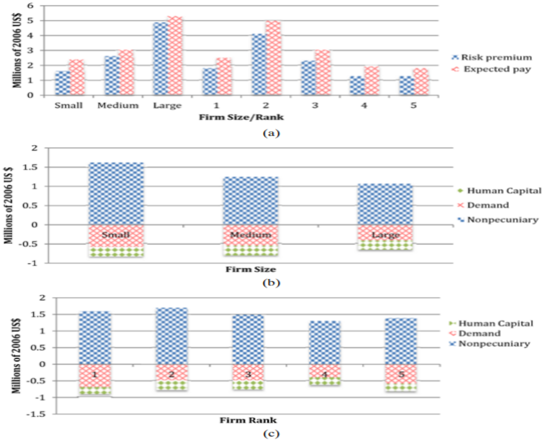

FIGURE 3.—Rank and firm-size pay decomposition. (a) Risk premium, (b) decomposition of certainty-equivalent pay, (c) decomposition of certainty-equivalent pay. *Note*: The certainty equivalent is the sum of human capital, demand, and nonpecuniary compensating differentials.

Estimates from the Structural Model

Figure 4 from GGM 2015, page 2352

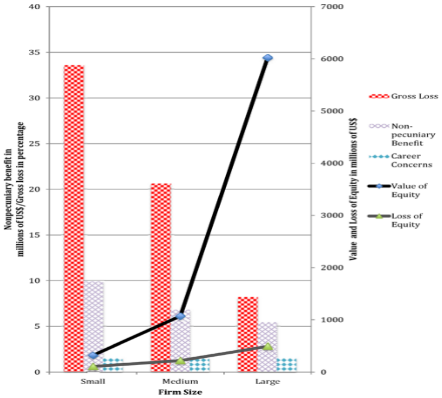

FIGURE 4.—Agency cost decomposition. Sources of agency cost by firm size. *Note*: Gross loss is the percentage of the firm value lost if an executive shirks instead of working. Loss of equity is the firm value lost if an executive shirks instead of working. Nonpecuniary benefit is the value to an executive of shirking relative to working. Career concerns measures the extent to which career concerns ameliorate the agency problem.

Estimates from the Structural Model

Figure 5 from GGM 2015, page 2354

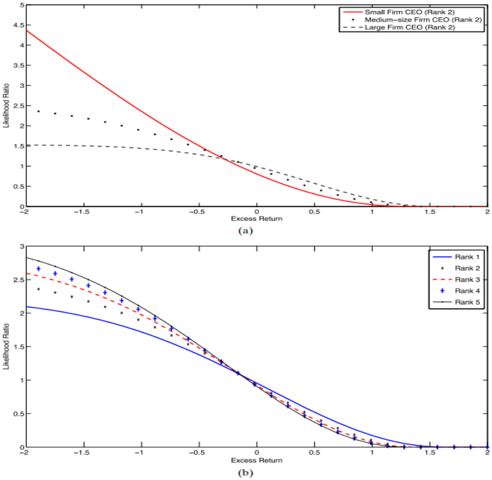

FIGURE 5.—Likelihood ratio. (a)Likelihood ratio by firm size for a CEO, (b) Likelihood ratio by rank 1 for a medium size firm. *Note*: Likelihood ratios are calculated at the average of the sample for the appropriate groups.

Estimates from the Structural Model

Three factors explain the firm-size executive pay premium

1. Large firms employ more talented executives.
2. There is no support for the hypothesis that executives prefer working in small firms; they are willing to work in a large firm for less pay.
3. There is no firm-size premium for human capital. Education and experience gained from different firms are individually significant, but collectively the firm-size pay differentials net out.
4. 80% of the firm-size total-compensation gap comes from the risk premium. Signal quality about effort is unambiguously poorer in larger firms, and this fully explains the larger risk premium. Larger firms having more supervisory positions and accountability is more difficult.
5. The remaining 20% comes from demand. Large firms pay a premium to meet demand because their bigger resource base amplifies the marginal productivity of their executives.