CHAPTER 3
THEORY OF CAREER CONCERNS

Optimal Contracts in the Presence of Career Concerns

- This Chapter is based on the work of Gibbons and Murphy (1992).

- See Robert Gibbons and Kevin Murphy (1992), "Optimal Incentive Contracts in the Presence of career Concerns : Theory and Evidence", *Journal of Political Economy*, **100**, p 468-505.

- Of related interest, the pioneering paper (initially circulated in 1982) : Bengt Holmström (1999), "Managerial Incentive Problems : a Dynamic Perspective", *Review of Economic Studies*, **66**, p 169-182.

- Theory of Career Concerns : Concerns about the effects of current performance on future compensation.

- Career concerns occur whenever the (internal or external) labor market uses a worker's current output to *update beliefs* about a worker's ability and base future wages on these updated beliefs.

- The worker will want to take actions to influence beliefs... but the market anticipates these actions...

- Managerial career concerns were first discussed by Eugene Fama (1980), in JPE : incentive contracts are not necessary because managers are disciplined through the managerial labor market...

- Homström (1982, 1999) in a famous paper, shows that the market is not a perfect substitute for contracts in general.

- Managers work too hard in the early years of career and not hard enough in later years.

- Gibbons and Murphy (1992) show that in the presence of career concerns, the optimal contract optimizes *total incentives*, *i.e.*, the combination of explicit and implicit incentives.

The 2-period Model

• A worker works for $T = 2$ periods. In period t, the worker's output is y_t, and we assume,

$$y_t = \eta + a_t + \varepsilon_t$$

where

η denotes *ability*,

$\eta \sim \mathcal{N}(m_0, \sigma_0^2)$, (normally distributed ability)

a_t denotes *effort*.

(The worker controls a stochastic process.)

- The cost of effort is $g(a_t)$.

g is assumed increasing, convex, differentiable, with $g'(0) = 0$, $g'(\infty) = \infty$ and $g''' \geq 0$.

- ε_t is a random noise, assumed normally distributed,

$\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$,

ε_t and η are independent, $t = 1, 2$.

- Employers are assumed risk-neutral.

- The worker is risk averse, with *CARA* utility,

$$U(w_1, w_2, a_1, a_2) = -\exp\{-r[w_1 - g(a_1) + \delta(w_2 - g(a_2))]\}.$$

where w_t is the wage paid in period t and δ is the worker's discount factor, measuring patience.

- *Assumptions about contracts* :

(a) One-period contracts are *linear* in output.

(b) Long-term contracts are not feasible.

- Point *(a)* is for the sake of simplicity. Point *(b)* is equivalent to saying that contracts must be *renegotiation-proof* (inefficient long-term contracts would be renegotiated).

• At he beginning of period $t = 1$, prospective employers simultaneously offer single-period linear wage contracts of the form

$$w_1(y_1) = c_1 + b_1 y_1.$$

• Note : Information is symmetric here, for ability η is known, neither to the worker nor to the employers.

• So w_1 depends on y_1 only ; b_1 and c_1 are parameters, to be determined.

• The worker chooses the most attractive contract and begins production.

• At the end of period $t = 1$, the first-period employer (the firm) and the market (*i.e.*, prospective employers) observe performance y_1 and simultaneously offer linear wage-contracts of the form,

$$w_2(y_2) = c_2 + b_2 y_2.$$

where c_2 and b_2 are the contract parameters.

• Note : The second contract depends on y_1 because y_1 conveys information on η. It follows that c_2 and b_2 will depend on first-period performance y_1. The dependence of w_2 on y_1 is kept implicit here.

- The worker's expected utility is the following,

$$-\mathbb{E}\left\{\exp\left\{-r[c_1 + b_1 y_1 - g(a_1)] - r\delta[c_2 + b_2 y_2 - g(a_2)]\right\}\right\}.$$

- In period $t = 2$, y_1 and a_1 are given, so expected utility can be rewritten,

$$-\exp\left\{-r[c_1 + b_1 y_1 - g(a_1)]\right\}\mathbb{E}\left\{\exp\left\{-r\delta[c_2 + b_2 y_2 - g(a_2)]\right\} \mid y_1\right\}$$

- In period $t = 2$, the worker chooses effort a_2 to maximize

$$-\mathbb{E}\left\{\exp\left\{-r\delta[c_2 + b_2(a_2 + \eta + \varepsilon_2) - g(a_2)]\right\} \mid y_1\right\}$$

or equivalently, maximize,

$$-\exp\left\{-r\delta[c_2 + b_2 a_2 - g(a_2)]\right\}\mathbb{E}\left\{\exp\left\{-r\delta[b_2(\eta + \varepsilon_2)]\right\} \mid y_1\right\}$$

- The problem is then simply to maximize $c_2 + b_2 a_2 - g(a_2)$ with respect to a_2,

- The optimal second-period effort, $a_2^*(b_2)$ solves the FOC,

$$b_2 = g'(a_2). \qquad (IC_2)$$

- Competition among prospective second-period employers implies that the contract accepted in $t = 2$ must earn zero profits.

- With a price of output normalized to 1, c_2 satisfies the zero-profit condition,

$$\mathbb{E}(y_2 - w_2 \mid y_1) = \mathbb{E}(y_2 \mid y_1) - b_2\mathbb{E}(y_2 \mid y_1) - c_2 = 0,$$

or equivalently,

$$c_2 = (1 - b_2)\mathbb{E}(y_2 \mid y_1)$$

where $c_2 = c_2(b_2)$.

- Now, given the assumptions on y_t, we have,

$$\mathbb{E}(y_2 \mid y_1) = \mathbb{E}(\eta \mid y_1) + a_2^*(b_2).$$

since ε_2 is independent of y_1.

- To compute $\mathbb{E}(\eta \mid y_1)$, the market must form a conjecture about the first-period effort choice a_1^*.

- Suppose that this *conjecture* is $\widehat{a}_1$. In *equilibrium*, the conjecture will be correct.

• In the normal case the distribution of η knowing y_1 is given by well-known formulas,

$$\mathbb{E}(\eta \mid y_1) = \mathbb{E}(\eta) + \frac{Cov(\eta, y_1)}{Var(y_1)}(y_1 - \mathbb{E}(y_1)),$$

the theoretical linear regression of η on y_1. So we have

$$\eta = \mathbb{E}(\eta \mid y_1) + \xi$$

with ξ independent of y_1. The variance of the prediction error, that is,

$$Var(\eta \mid y_1) = Var(\xi) = Var\left[\eta - \mathbb{E}(\eta \mid y_1)\right],$$

is given by,

$$Var(\eta \mid y_1) = (1 - \rho^2)Var(\eta),$$

where

$$\rho^2 = \frac{Cov(y_1, \eta)^2}{Var(y_1)Var(\eta)}.$$

• Using the above formula for the *conditional expectation* of a normal random variable with respect to another normal random variable, we find

$$\mathbb{E}(\eta \mid y_1) = \left(\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_0^2}\right) m_0 + \left(\frac{\sigma_0^2}{\sigma_\varepsilon^2 + \sigma_0^2}\right)(y_1 - \widehat{a}_1),$$

a weighted average of the prior mean of ability, *i.e.*, $\mathbb{E}(\eta) = m_0$, and the "signal", *i.e.*, $(y_1 - \widehat{a}_1)$.

• We denote $\mathbb{E}(\eta \mid y_1) = m_1(y_1, \widehat{a}_1)$ because it depends on the conjecture $\widehat{a}_1$.

• The conditional variance, $Var(\eta \mid y_1)$, denoted σ_1^2, is given by the formula,

$$\sigma_1^2 = Var(\eta)\left[1 - \frac{Var(\eta)}{Var(\eta) + Var(\varepsilon)}\right] = \frac{\sigma_0^2 \sigma_\varepsilon^2}{\sigma_0^2 + \sigma_\varepsilon^2}$$

since $Var(y_1) = Var(\eta) + Var(\varepsilon)$, or

$$\frac{1}{\sigma_1^2} = \frac{1}{\sigma_0^2} + \frac{1}{\sigma_\varepsilon^2},$$

Note : In addition,

$$\mathbb{E}(y_2 \mid y_1) = m_1(y_1; \hat{a}_1) + a_2^*(b_2).$$

and

$$Var(y_2 \mid y_1) = Var(\eta + \varepsilon_2 \mid y_1) = \sigma_1^2 + \sigma_\varepsilon^2 = V.$$

Important Remark

• If X is a normal random variable $X \sim \mathcal{N}(\mu, \sigma^2)$, then, $\exp(-rX)$ is log-normal variable and we have the well-known result,

$$\mathbb{E}[\exp(-rx)] = \exp\left[-r\mu + \frac{r^2\sigma^2}{2}\right].$$

• As a consequence, maximizing $-\mathbb{E}[\exp(-rX)]$ is equivalent to maximizing $\mu - \frac{r\sigma^2}{2}$.

It follows that, if we substitute the above results in the worker's second-period utility U_2, we get,

$$\begin{aligned}
U_2 &= -\mathbb{E}\{\exp[-r_2(w_2 - g(a_2^*))] \mid y_1\} \\
&= -\exp\left\{-r_2\left[c_2 + b_2\mathbb{E}(y_2 \mid y_1) - g(a_2^*) - \frac{r_2 b_2^2}{2} Var(y_2 \mid y_1)\right]\right\} \\
&= -\exp\left\{-r_2\left[c_2 + b_2(m_1 + a_2^*) - g(a_2^*) - \frac{r_2 b_2^2 V}{2}\right]\right\}
\end{aligned}$$

and, given that the zero-profit condition implies

$$c_2 = (1 - b_2)\mathbb{E}(y_2 \mid y_1) = (1 - b_2)(m_1 + a_2^*),$$

we finally get,

$$U_2 = -\exp\left\{-r_2\left[m_1 + a_2^*(b_2) - g(a_2^*(b_2)) - \frac{r_2 b_2^2 V}{2}\right]\right\}.$$

This is a function of b_2.

• We assume that competition between employers in period $t = 2$ raises b_2 so as to maximize the second-period expected utility U_2. To find b_2 we solve,

$$\text{Maximize} \quad \left\{ m_1(y_1, \widehat{a}_1) + a_2^*(b_2) - g(a_2^*(b_2)) - \frac{r_2 b_2^2 V}{2} \right\}.$$

• The first-order necessary condition for this problem yields,

$$\frac{da_2^*}{db_2} - g'(a_2^*(b_2)) \frac{da_2^*}{db_2} - r_2 b_2 V = 0.$$

• But since we have $g'(a_2^*) = b_2$, we know that $a_2^*(b_2) = (g')^{-1}(b_2)$ and therefore,

$$\frac{da_2^*}{db_2} = \frac{1}{g''[a_2^*(b_2))]}.$$

• Given this result, the first-order condition for an equilibrium b_2 becomes, $1 - b_2 - r_2 b_2 V g''(a_2^*(b_2)) = 0$, or

$$b_2^* = \frac{1}{1 + r_2 V g''[a_2^*(b_2^*)]}. \qquad (FOC')$$

• Since we assumed $g''' \geq 0$, FOC' is necessary and sufficient., typically yields a unique solution. Clearly, b_2^* decreases with risk aversion r and uncertainty V and $0 < b_2^* < 1$.

• To simplify the presentation, from now on, we assume that $g(a)$ is quadratic, $g(a) = \gamma a^2/2$, so that g'' is a constant and $b_2^* = 1/(1 + r_2 V \gamma)$.

• *Remark :* b_2^* is independent of y_1. By chance (It's a consequence of the normality assumption). This greatly simplifies the analysis. The effect of y_1 on the second-period contract is limited to an effect on $c_2^* = (1 - b_2^*)(m_1(y_1, \hat{a}_1) + a_2^*(b_2^*))$ because of this.

• Given the optimal (c_2^*, b_2^*) derived above, the worker's first-period problem is to choose a_1 so as to maximize U_1, that is,

$$U_1 = -\mathbb{E}\left\{exp\left[-r(c_1 + b_1 y_1 - g(a_1)) - r\delta(c_2^* + b_2^* y_2 - g(a_2^*))\right]\right\}.$$

where $y_2 = \eta + a_2^*(b_2^*) + \varepsilon_2$,

and (substituting $m_1(.)$ into $c_2^* = (1 - b_2^*)(m_1 + a_2^*)$), we have,

$$c_2^* = (1 - b_2^*)\left[\frac{\sigma_\varepsilon^2 m_0 + \sigma_0^2(y_1 - \hat{a}_1)}{\sigma_\varepsilon^2 + \sigma_0^2} + a_2^*(b_2^*)\right].$$

Computation of the first-period effort a_1

• In the expression for U_1 we can separate the terms depending on $(\eta, \varepsilon_1, \varepsilon_2)$ from the non-random terms and the terms depending on effort a_1. Expected utility U_1 has the following form

$$U_1 = -\exp[G(a_1)]\exp(H)\mathbb{E}\{\exp[F(\eta, \varepsilon_1, \varepsilon_2)]\},$$

where, $F(.)$ is the random part that doesn't depend on effort, H is non-random and doesn't depend on a_1, and

$$G(a_1) = -r[c_1 + b_1 a_1 - g(a_1)] - r\delta\left[(1 - b_2^*)\left(\frac{\sigma_\varepsilon^2 m_0 + \sigma_0^2(a_1 - \hat{a}_1)}{\sigma_\varepsilon^2 + \sigma_0^2}\right)\right].$$

• So, maximizing U_1 is equivalent to solving,

$$\text{Maximize} \quad \left\{b_1 a_1 - g(a_1) + \delta(1 - b_2^*)\left(\frac{\sigma_0^2}{\sigma_\varepsilon^2 + \sigma_0^2}\right)(a_1 - \hat{a}_1)\right\}.$$

Computation of the first-period effort a_1, (2)

- This yields the first-order condition,

$$b_1 + \delta(1 - b_2^*)\left(\frac{\sigma_0^2}{\sigma_\varepsilon^2 + \sigma_0^2}\right) = g'(a_1). \qquad (FOC'')$$

- This can be interpreted as,

$$\text{Total} \quad \text{Incentive} = \text{Marginal} \quad \text{Cost} \quad \text{of} \quad \text{Effort}$$

- The total incentive for first-period effort is denoted B_1, and

$$B_1 = b_1 + \delta(1 - b_2^*)\left(\frac{\sigma_0^2}{\sigma_\varepsilon^2 + \sigma_0^2}\right).$$

- B_1 itself is the sum of,

(a), b_1, the *power of the short-term contract*, and

(b), $\delta(1-b_2^*)\left(\frac{\sigma_0^2}{\sigma_\varepsilon^2+\sigma_0^2}\right) > 0$, the *implicit incentives* from the *career effect*.

- The career effect

(i), increases with σ_0^2 (uncertainty about ability η) ;

(ii), decreases with σ_ε^2 (the riskiness of output y_t).

(iii), increases with δ (when the future gains are less discounted).

Equilibrium

• So far, the conjecture $\hat{a}_1$ was taken as given. In equilibrium, the market conjecture must be correct : $\hat{a}_1 = a_1^*(b_1)$, where $a_1^*(b_1)$ solves FOC'' above, that is,

$$a_1^*(b_1) = (g')^{-1}\left[b_1 + \frac{\delta\sigma_0^2(1-b_2^*)}{\sigma_\varepsilon^2+\sigma_0^2}\right].$$

• Again, competition among employers in period $t = 1$ will drive expected profits to zero, that is,

$$c_1 = (1-b_1)\mathbb{E}(y_1) = (1-b_1)(m_0 + a_1^*(b_1)).$$

- the period 1 expected utility U_1 is by definition,

$$U_1 = -\mathbb{E}\left\{\text{exp}\left[-r(c_1 + b_1 y_1 - g(a_1)) - r\delta(c_2^* + b_2^* y_2^* - g(a_2^*))\right]\right\}.$$

where $y_2^* = \eta + \varepsilon_2 + a_2^*(b_2^*)$.

- This expression is of the form $U_1 = -\mathbb{E}\{\text{exp}(Q)\}$ and Q is a linear combination of normal random variables, therefore, it is normal and we have $U_1 = -\text{exp}\{\mathbb{E}(Q) + (1/2)Var(Q)\}$

- We need to compute $\mathbf{M} = \mathbb{E}(Q)$ and $\mathbf{V} = Var(Q)$.

We have,

$$\begin{aligned}\mathbf{M} &= \mathbb{E}[-r(c_1 + b_1 y_1 - g(a_1)) - r\delta(c_2^* + b_2^* y_2^* - g(a_2^*))] \\ &= -r[m_0 + a_1^*(b_1) - g(a_1^*(b_1))] - r\delta([m_0 + a_2^*(b_2^*) - g(a_2^*(b_2^*))]\end{aligned}$$

this is because,

$$\mathbb{E}(c_2^* + b_2^* y_2^*) = (1 - b_2^*)\mathbb{E}[\mathbb{E}(y_2^* \mid y_1)] + b_2^*\mathbb{E}[y_2^*] = \mathbb{E}[y_2^*] = m_0 + a_2^*(b_2^*)$$

To compute the variance of Q, we have,

$$\begin{aligned}
\mathbf{V} &= Var\{-r(c_1 + b_1 y_1 - g(a_1)) - r\delta(c_2^* + b_2^* y_2^* - g(a_2^*))\} \\
&= r^2 Var[b_1 y_1 + \delta(1 - b_2^*)\mathbb{E}(y_2^* \mid y_1) + \delta b_2^* y_2^*] \\
&= r^2 Var\left[b_1 y_1 + \delta(1 - b_2^*)\frac{\sigma_0^2 y_1}{\sigma_0^2 + \sigma_\varepsilon^2} + \delta b_2^*(\eta + \varepsilon_2)\right] \\
&= r^2 Var\left[B_1 y_1 + \delta b_2^*(\eta + \varepsilon_2)\right] \\
&= r^2 B_1^2 Var(y_1) + r^2\delta^2 b_2^{*2} Var(\eta + \varepsilon_2) + 2r^2 B_1 \delta b_2^* Cov(y_1, \eta) \\
&= r^2\left\{(B_1^2 + \delta^2 b_2^{*2})(\sigma_0^2 + \sigma_\varepsilon^2) + 2B_1\delta b_2^*\sigma_0^2\right\}
\end{aligned}$$

This is because

$$\begin{aligned}
Cov(y_1, \eta) &= Cov(\eta + \varepsilon_1, \eta) \\
&= Var(\eta) + 0 \\
&= \sigma_0^2.
\end{aligned}$$

Computation of b_1^*, the Power of the First-period Contract

• Coming back to U_1, we have $U_1 = -\exp\left[\mathbf{M} + \frac{1}{2}\mathbf{V}\right]$ and therefore, U_1 is maximized with respect to b_1 when $\mathbf{M} + \frac{1}{2}\mathbf{V}$ is minimized, or when we maximize,

$$\{m_0 + a_1^*(b_1) - g(a_1^*(b_1)) + \delta[m_0 + a_2^*(b_2^*) - g(a_2^*(b_2^*))] - (1/2r)\mathbf{V}\}.$$

Using the expression for $\mathbf{V}$, and rearranging terms, this is equivalent to maximizing,

$$\left\{m_0 + a_1^*(b_1) - g(a_1^*(b_1)) - \frac{r}{2}[(B_1 + \delta b_2^*)^2(\sigma_0^2 + \sigma_\varepsilon^2) - 2B_1\delta b_2^*\sigma_\varepsilon^2]\right\}$$

Computation of b_1^*, (2)

• We again assume that competition among firms will drive b_1 up in such a way that U_1 is maximized, subject to the zero-profit constraint.

• We can write the first-order condition for the solution b_1^*. First recall that,

$$a_1^*(b_1) = (g')^{-1}(B_1),$$

this implies $da_1^*/db_1 = 1/\gamma$.

Recall that $Var(y_1) = V(y_1) = \sigma_0^2 + \sigma_\varepsilon^2$.

Computation of b_1^*, (3)

We obtain,

$$\frac{1}{\gamma} - \frac{g'(a_1^*(b_1))}{\gamma} - r(B_1 + \delta b_2^*)V(y_1) + r\delta b_2^* \sigma_\varepsilon^2 = 0,$$

Equivalently, using FOC'', that is,

$$g'(a_1^*) = B_1,$$

we obtain,

$$1 - (1 + r\gamma V(y_1))B_1 = r\delta b_2^* \gamma \sigma_0^2$$

Computation of b_1^*, (4)

We finally get,

$$B_1 = \frac{1 - r\delta b_2^* \gamma \sigma_0^2}{1 + rV(y_1)\gamma} \qquad (FOC''')$$

with

$$b_2^* = \frac{1}{1 + r_2 V(y_2 \mid y_1)\gamma},$$

where

$$V(y_2 \mid y_1) = V = \sigma_1^2 + \sigma_\varepsilon^2$$

and

$$B_1 = b_1 + \delta(1 - b_2^*)\frac{\sigma_0^2}{V(y_1)} \qquad (FOC''')$$

Computation of b_1^*, (5)

- We can now rewrite FOC''' as follows,

$$b_1^* = \frac{1}{1 + rV(y_1)\gamma} - \frac{V(y_2 \mid y_1) r_2 \delta \sigma_0^2 \gamma}{V(y_1)} b_2^* - \frac{r\delta\sigma_0^2\gamma}{1 + rV(y_1)\gamma} b_2^*.$$

- This shows how b_1^* is computed in equilibrium.

- We can now provide an interpretation of the result.

Key Result : the Power of Incentives Must Increase with Time

- We have $b_2^* > b_1^*$.

- To see this first note that

$$\sigma_1^2 = \frac{\sigma_0^2 \sigma_\varepsilon^2}{\sigma_0^2 + \sigma_\varepsilon^2} < \sigma_0^2.$$

- Therefore, $V(y_2 \mid y_1) < V(y_1)$. The conditional variance of output y_t declines with time.

- Using FOC''', we can write,

$$b_1^* < \frac{1}{1 + rV(y_1)\gamma} < \frac{1}{1 + rV(y_2 \mid y_1)\gamma} \leq b_2^*$$

using $r_2 \leq r$ (true if $r_2 = r$ or $r_2 = r\delta$).

Interpretation : b_1^* is the Sum of 3 Effects

(1) The Noise Reduction Effect : $\frac{1}{1+r\gamma V(y_1)}$. Note that $b_t^* < 1$, to reduce the employee's exposure to output risk.

(2) The Implicit Career Incentive Term : $\delta(1 - b_2^*)\frac{\sigma_0^2}{V(y_1)}$. Explicit first-period incentives are adjusted downwards in period $t = 1$ to account for the presence of career concerns.

(3) The Human-Capital Insurance Effect : $-\frac{r\delta\gamma\sigma_0^2 b_2^*}{1+r\gamma V(y_1)}$. Workers want to be insured against a low realization of ability η : this can be done by reducing b_1.

Further Remarks :

• It can happen that $B_1 < 0$ is the benefits from insuring the worker against low values of η exceed the benefits of providing effort incentives.

• If the noise on output is negligible, *i.e.*, if $\sigma_\varepsilon^2 \simeq 0$, then, y_1 almost fully reveals ability, and $\sigma_1^2 \simeq 0$. In such a case, we find $b_2^* \simeq 1$, thus $B_1 \simeq b_1^*$ and

$$b_1^* \simeq \frac{1 - r\delta\sigma_0^2\gamma}{1 + r\gamma\sigma_0^2}.$$

• In this case, b_1^* can be negative if $\delta \simeq 1$ and σ_0^2 is large. This poses a problem. We have a corner solution with $a_1^* = 0$ (or we accept negative values of effort...)

Conclusion

“A guy of your age should be paid cash”.

A senior professor cannot increase his salary without accepting piece rates.

The analysis shows that the explicit incentives of older workers should be stronger : there is less uncertainty on output (risk is less costly) and the career (implicit) incentives are weaker.