Contracting with Disagreement about Performance Evaluation and Compensation

Anqi Li

Department of Economics
Washington University in St. Louis

March 2015
Motivation

Organizational frictions from disagreements about employee performance evaluation and compensation:

- **Self-serving bias**: Babcock et al. (1996);
- **Subtle information**: Bracken et al. (2001);
- **Intertemporal comparison**: Endlich (2000);
- **Interpersonal comparison**: Frank (1984, 1985).
Modeling: a general framework for analyzing the organizational implications of various kinds of disagreements about employee performance evaluation and compensation.

Result: seemingly rigid policies are robust tools for providing incentives and managing disagreements simultaneously:

- Long-term performance appraisal; compressed compensation scheme, efficiency wage, seniority-based promotion, etc.
Agenda

1. Baseline Model
   - Setup
   - Static case
   - Dynamic case
2. Optimal features
3. Extensions
4. Literature
5. Conclusion
Agenda

1. Baseline Model
   - Setup
   - Static case
   - Dynamic case

2. Optimal features

3. Extensions

4. Literature

5. Conclusion
Setup

**Players:** a principal and an agent: risk-neutral, zero outside option, no discounting over $t = 1, \cdots, T$

**Production technology:** in period $t$, the agent privately exerts an effort $a_t \in \mathcal{A} = \{0, 1\}$ that incurs a cost $c(a_t)$ to himself and generates an expected payoff $a_t$ to the principal:

- $1 > c(1) = c > c(0) = 0$; $a = 1$ is efficient.

**Monitoring technology:** at the end of period $t$, a signal $X_t$ is publicly realized where

$1 > \mathbb{P}(X_t = H|a_t = 1) = p(1) > \mathbb{P}(X_t = H|a_t = 0) = p(0) > 0$.

**Incentive contract** $\left\{w_T, \left\{\psi_{t,T}(\cdot)\right\}_{t=1}^T\right\}$:

- $w_T \in \mathbb{R}$: baseline wage;
- $\psi_{t,T} : \Omega^t \to \mathbb{R}$: period-$t$ incentive payment.
Disagreement Point

Disagreement point $S_t \in \Omega$: a non-contractible signal about the period-$t$ performance evaluation that the agent thinks he deserves:

- Self-serving bias: $S_t \equiv H$;
- Subtle (private) information;
- Intertemporal comparison: $S_t \sim F(\cdot | a_s, X_s, s \leq t)$;
- Interpersonal comparison: $S_{i,t} \sim F_i(\cdot | a_{i,t}, a_{-i,t}, X_{i,t}, X_{-i,t})$.
- Expectation: $S_t = F^{-1}(Q|a = 1)$

Monitoring and Disagreement technology

$$I(T) = \left\{ \left\{ X_1, S_1, \ldots, X_T, S_T : a^T \right\} : a^T \in A^T \right\}$$
Disagreement Cost

**Period-\(t\) disagreement cost** If \(\psi_{t,T}(S^t) > \psi_{t,T}(X^t)\), then each player \(i = p, a\) bears a disagreement cost \(\Lambda_i(\chi_t(X^t, S^t)) \geq 0\) where

\[
\chi_t(X^t, S^t) = [\psi_{t,T}(S^t) - \psi_{t,T}(X^t)]^+
\]

**Assumptions:**
- The principal’s participation is essential;
- Asymmetry between gain and loss;
- Disagreement about internal matters; contract as a frame;
- (For the time being) no anticipatory conflicts.
Agenda

1. Baseline Model
   - Setup
   - Static case
   - Dynamic case

2. Optimal features

3. Extensions

4. Literature

5. Conclusion
Static Case: Tension Between Incentive and Disagreement

**Observation 1.**

*The optimal contract that elicits the low effort satisfies* \( w = 0 \) *and* \( \psi(X) \equiv 0 \) *and incurs no disagreement cost.*

**Observation 2.**

*The optimal contract that elicits the high effort solves*

\[
\min_{w, \psi(\cdot)} 1 - c - \rho(1)\Lambda_p(\psi(H) - \psi(L)),
\]

\[
s.t. \quad \psi(H) - \psi(L) - \frac{\rho(1) - \rho(0)}{p(1) - p(0)}\Lambda_a(\psi(H) - \psi(L)) \geq \frac{c}{p(1) - p(0)}
\]

*and* \( w + p\psi(H) + (1 - p)\psi(L) - \rho(1)\Lambda_a(\psi(H) - \psi(L)) \geq c \)

*where* \( \rho(a) = \mathbb{P}(S = H, X = L|a; l(1)) \)
Static Case: Tension Between Incentive and Disagreement

Assumption 1.

\[ z - \frac{\rho(1)-\rho(0)}{p(1)-p(0)} \Lambda_a(z) \text{ single-crosses } \frac{c}{p(1)-p(0)} \text{ from below at } z^* > 0 \]

Lemma 1.

Under Assumption 1,

(i) The optimal profit from eliciting the high effort is

\[ 1 - c - \rho(1)\Lambda_a(z^*) \]

(ii) High effort is more profitable than low effort if and only if

\[ 1 - c \geq \rho(1)\Lambda_a(z^*) \]
Agenda

1. Baseline Model
   - Setup
   - Static case
   - Dynamic case

2. Optimal features

3. Extensions

4. Literature

5. Conclusion
Efficiency Wage Contract

A fixed wage $w_T$ at each $t = 1, \cdots, T$. At $t = T$, test if

$$\varphi_T(X^T) = \frac{1}{T} \sum_{t=1}^{T} X_t > \mathbb{E}[X|a = 1] - b_T$$

where

$$b_T = T^{-\frac{1}{2} + \varepsilon} \text{ for some arbitrary } \varepsilon \in \left(0, \frac{1}{2}\right)$$

- If “$>$” then the agent passes and earns $\overline{\psi}_T$;
- If “$\leq$” then he fails and earns $\underline{\psi}_T$:

$$\overline{\psi}_T - \underline{\psi}_T = B_T = \alpha c T \text{ for some arbitrary } \alpha > 1$$
Graphical Illustration

Average Output; a=1
Average Output; a=0
$E[Y|a=1] - b_T$
$E[Y|a=1] + b_T$
$E[Y|a=1]$
$T$
Define $\lambda_{i,T} = \frac{\Lambda_i(B_T)}{T}$ for each $i = p, a$, $\mu_T = \exp\left(-\frac{2Tb_T^2}{(H-L)^2}\right)$ and $\kappa = \left[(p(1) - p(0))(H - L)\right]^{-1}$.

**Assumption 2.**

$$\max\{\lambda_{a,T} \mu_T, b_T\} \cdot (1 + \lambda_{p,T}) \to 0 \text{ as } T \to \infty.$$
Analytical Performance Bound

**Theorem 1.**

Fix $\varepsilon$ and $\alpha$. For each $T \in \mathbb{N}$ and $I(T)$, every BNE $\sigma_T^*$ satisfies

(i) $\mathbb{P}(\mathcal{F}|\sigma_T^*; I(T)) \leq \pi_T + 2\mu_T$ where

$$\pi_T \leq \frac{((\alpha + 2) + \lambda_{a,T}/c)\mu_T + 2\kappa b_T(1 - 2\mu_T)}{\alpha - 1 + 2\kappa b_T}$$

(ii) The per-period expected profit is at least

$$
\underbrace{(1 - \pi_T - 2\mu_T)(1 - 2\kappa b_T)}_{\text{Lower bound for the principal's expected payoff}}
- \underbrace{c + (\alpha c + \lambda_{a,T})\mu_T}_{\text{Upper bound for the expected payment}}
- \underbrace{\lambda_{p,T}\mathbb{P}(\mathcal{F}|\sigma_T^*; I(T))}_{\text{Upper bound for the principal's disagreement cost}}
$$
Asymptotic Efficiency

Theorem 1 (continued).

Take an arbitrary sequence $\{I(T)\}_{T=1}^{\infty}$. Under Assumption 2, as $T \to \infty$,

1. $\mathbb{P}(F|\sigma_T^*; I(T)) \sim O(\max\{\lambda_a, T \mu_T, b_T\})$;

2. The per-period expected profit is $1 - c - O(\max\{\lambda_a, T \mu_T, b_T\}(1 + \lambda_{p,T}))$. 
Intuition: Bound the Principal’s Disagreement Cost

Disagreement arises from the feeling of being underpaid ⇒ the prob. of disagreement ≤ the prob. of failure for all $I(T)$.

Thus

$$\frac{1}{T} \mathbb{E} \left[ \Lambda_p \left( \chi_T \left( X^T, S^T \right) \right) \mid \sigma^*_T; I(T) \right] \leq \lambda_{p,T} \mathbb{P}(\text{Failure} \mid \sigma^*_T; I(T))$$
Intuition: Bound the Prob. of Failure

Limit the gain from two types of deviations:

- **Concentration of measure**: pass by sheer luck;
- **Informativeness**: fine-tune the effort choice with past outcomes.

**Lemma 2 (Concentration of Measure).**
\[
\mathbb{P}\left(\left|\varphi_T(X^T) - \mathbb{E}[\varphi_T(X^T)|a^T]\right| > b_T |a^T\right) < 2\mu_T \text{ for all } T \in \mathbb{N} \text{ and } a^T \in \mathcal{A}^T.
\]

**Lemma 3 (Informativeness).**
\[
|\varphi_T(X^T) - \mathbb{E}[\varphi_T(X^T)|a^T]| < b_T, \quad \varphi_T(X^T) > \mathbb{E}[X|a = 1] - b_T
\]
implies \[
\frac{1}{T} \sum_{t=1}^{T} a_t \geq 1 - 2\kappa b_T \text{ and } \frac{1}{T} \sum_{t=1}^{T} c(a_t) \geq c(1 - 2\kappa b_T).
\]
Agenda

1. Baseline Model
2. Optimal features
3. Extensions
4. Literature
5. Conclusion
Short-term bonus contract $\{w_T, \psi_t, T(X_t)\}_{t=1}^T$.

**Assumption 3.**

$S_t$ depends only on $(a_t, X_t)$.

**Corollary 1.**

*Under Assumption 3, the per-period expected disagreement cost incurred by the short-term bonus contract is bounded below by*

$$\frac{1}{T} \mathbb{E} \left[ \sum_{t=1}^{T} \mathbb{1}_{\psi_t, T(H) - \psi_t, T(L) \geq z^*_t} \cdot \rho_t(1) \Lambda(z^*_t) \sigma^*_T \mid I(T) \right]$$
Compressed Compensation Scheme

**Review contract:** \{w_T, D_T, \Psi_T\}

- \(w_T: \) baseline wage;
- \(D_T \subset \{1, 2, \ldots, T\}: \) dates for performance appraisals;
- \(\Psi_T = \{\psi_{t,T}: \Omega^{t-t'+1} \to \mathbb{R}: t \in D_T\}\) where \(t' = \max\{s < t: s \in D_T \cup \{0\}\}\).

A max-min game between the principal, the adversary and the agent:

- The principal commits to \(\{w_T, D_T, \Psi_T\}\) first.
- The adversary then specifies \(I(T)\) to minimize the principal's expected profit.
- The agent observes \(I(T)\) and makes the optimal effort choice.
The Principal’s Problem

Let $\lambda_{a,T} = 0$ and write the principal’s problem as follows:

$$\max_{\{w_T, D_T, \psi_T\}} \min_{I(T)} \mathbb{E} \left[ \sum_{t=1}^{T} X_t - \psi_{t,T}(X_{t'}^T) - \Lambda_p(\chi_T(X^T, S^T))|\sigma^m_T; I(T) \right]$$

s.t. $\forall I(T), \sigma^m_T \in \arg \max_{\sigma_T} \mathbb{E} \left[ \sum_{t=1}^{T} \psi_{t,T}(X_{t'}^T) - c(a_t)|\sigma_T \right]$ and $Tw_T + \mathbb{E} \left[ \sum_{t=1}^{T} \psi_{t,T}(X_{t'}^T) - c(a_t)|\sigma^m_T \right] \geq 0$
Compressed Payment Scheme

For each $Q \in [0, 1]$ and $t = 1, \cdots, T$, define

$$\psi_{t,T}^m(Q) = \max \left\{ x \in \mathbb{R} : \frac{\mathbb{P}(\psi_{t,T}^m(X_t^T) \leq x | \sigma_T^m; I_m(T)) \geq Q}{\mathbb{P}(\psi_{t,T}^m(X_t^T) < x | \sigma_T^m; I_m(T)) \leq Q} \right\}$$

**Proposition 1.**

For each $T \in \mathbb{N}$ and $Q \in [0, 1]$,

$$\frac{1}{T} \mathbb{E} \left[ \sum_{t \in D_T} \Lambda_p \left( [\psi_{t,T}^m(Q) - \psi_{t,T}^m(X_t^T)]^+ \right) | \sigma_T^m; I_m(T) \right] \leq 1 - c - \text{Theorem 1 (ii)}$$
The optimal compensation scheme exhibits local bunching for many disagreement technologies of interest:

- Self-serving bias: $Q = 1$;
- Average payment of many colleagues: $Q = \frac{1}{2}$;
- Expectation: $Q \in [0, 1]$.

Compression becomes increasingly global as the set of feasible disagreement technologies enlarges.
Testable Implications

Compressed compensation scheme is common if employees

- Have heterogeneous views about how they should be evaluated and compensated because of the distinct socio-economic backgrounds and (or) prior experiences;
- Have access to the information that facilitates disagreement formation, e.g., colleagues’ pay.
- Can effectively penalize the principal if they are retained in the firm.
Distinguish Our Theory from Alternative Explanations for Compressed Compensation Schemes

- Second-order risk aversion: e.g., Spear and Srivastava (1987):
  - Disclose information that facilitates disagreement formation while fixing the agent’s risk attitude \(\Rightarrow\) our contract becomes more profitable than the short-term bonus contract;
  - First-order risk aversion \(\Rightarrow\) our contract is more profitable than the short-term bonus contract even if the outcome volatility is small.

- Subjective evaluation, e.g., Fuchs (2007), Maestri (2012):
  - The implication of self-assessment.
Agenda

1. Baseline Model
2. Optimal features
3. Extensions
   - Anticipatory conflicts
   - Promotion system design
   - Endogenous disagreement point formation
   - General model and simulation results
4. Literature
5. Conclusion
Anticipatory Conflict

Fix the aforementioned contract and take an arbitrary effort strategy \( \sigma_T \).

At the end of \( t = 1, \cdots, T \), a conflict state variable \( C_t \in \{0, 1\} \) is publicly realized.

If \( C_t = 1 \), then conflict arises and each player \( i \) bears a disagreement cost \( \Lambda_i(h^t; \sigma_T) \).
Assumption 4.

(i) For all $t, h^t, \sigma_T$,

(a) $\Lambda_a(h^t; \sigma_T) \leq \hat{\Lambda}(\mathbb{E}[\chi_T(X^T, S^T) | h^t; \sigma_T])$ for some weakly concave function $\hat{\Lambda} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$;

(b) There exists $\beta \in (0, 1)$ s.t. $\Lambda_p(h^t; \sigma_T) \leq \beta \Lambda(h^t; \sigma_T)$.

(ii) The two terms below vanish as $T \rightarrow \infty$:

$$m_T = \hat{\Lambda}(B_T \cdot \mu_T) \mathbb{E} \left[ \frac{1}{T} \sum_{t=1}^{T} C_t | 1^T \right]$$

$$\rho_T = \frac{1}{T} \mathbb{E} \left[ \sum_{t=1}^{T} \operatorname{Cov} \left( \hat{\Lambda}(\mathbb{E}[\chi_T(X^T, S^T) | h^t; 1^T]), C_t | h^t, 1^T \right) | 1^T \right]$$
Proposition 2.

Under Assumption 4,

(i) \( \mathbb{P}(\mathcal{F}|\sigma_{T}^*; I(T)) \leq 2\mu_{T} + \pi_{T} \) where \( \pi_{T} \) satisfies

\[
\pi_{T} \leq \frac{(\alpha + 2)\mu_{T} + 2\kappa b_{T}(1 - 2\mu_{T}) + (m_{T} + \rho_{T})/c}{\alpha - 1 + 2b_{T}}
\]

(ii) The per-period expected profit is bounded below by

\[
\mathcal{L}_{\pi} = \left(1 - \pi_{T} - 2\mu_{T}\right)(1 - 2\kappa b_{T})
\]

\[
- \left[ \frac{\beta}{1 - \beta} \left[ (\alpha + 2)\mu_{T} + \pi_{T} + 2\kappa b_{T}(1 - \pi_{T} - 2\mu_{T}) + (m_{T} + \rho_{T})/c \right] \right]
\]

Lower bound for the principal's expected payoff

Upper bound for the expected compensation

Upper bound for the principal's disagreement cost
Two junior positions with no payment and one senior position.

Once every $T$ periods, the senior position becomes vacant and is filled by either an internal junior agent or an external candidate. The second option costs $K_T \sim O(T)$.

The junior labor market is competitive and replacement is costless.

Each agent can work at the junior position for at most $2T$ periods.

From a junior agent’s viewpoint, promotion conveys a benefit $B_T = 2\alpha cT, \alpha > 1$. 
Define merit-based promotion policies \( \{\rho_1, \rho_2\} \) by

\[
\rho_i : \Omega^{2T} \rightarrow [0, 1], \quad i = 1, 2,
\]

\[
s.t. \quad \sum_i \rho_i \left( x_1^T, x_2^T \right) \leq 1 \quad \forall x_1^T, x_2^T \in \Omega^T \quad \text{(Feasibility)}
\]

and \( \rho_1 \left( x^T, x'^T \right) = \rho_2 \left( x'^T, x^T \right) \quad \forall x^T, x'^T \in \Omega^T \quad \text{(Anonymity)} \)

If one agent wins and the other loses then the social disagreement cost is at least \( \Lambda(B_T - cT) \).
Label the two junior positions as the *entry-level position* and the *junior-level position*.

An agent works at the entry level for $T$ periods; moves to the junior level if $\frac{1}{T} \sum_{t=1}^{T} X_t > \mathbb{E}[X|a = 1] - b_T$ and gets fired otherwise.

Works for another $T$ periods before he is promoted or fired.

Agents do not envy the outcome of colleagues at different organizational hierarchies (even if they do, the prob. is small).
Proposition 3.

(i) The per-period efficiency loss of the merit-based policy is at least

\[ \frac{1}{T} \left\{ K_T (1 - \sum_i P(\rho_i > \rho_{-i}) \sigma^*_T) \right. \\
+ \Lambda(B_T - cT) \sum_i P(\rho_i > \rho_{-i}) \sigma^*_T \\
+ \text{efficiency loss due to shirking} \right\} \]

(ii) The per-period efficiency loss of the seniority-based policy is at most

\[ K_T O(b_T)/T + O(b_T) \sim O(b_T) \]

(iii) When \( T \) is large, (ii) is smaller than (i) if \( \frac{\Lambda(B_T - cT)}{B_T} \gg \Theta(b_T) \) as \( T \to \infty \).
Simulation

Poisson signals $\Omega = \{0, 1\}$. General production technology and sophisticated preferences over risks and time.

Treat one month as a period and assume that:

- Standard annual discount rate 5%;
- $\lambda_{a,T} = 0$, $\varepsilon = .45$, $\alpha = 2, 3$;
- Conservative estimate of the informativeness parameter: $c_T \in [\frac{1}{15}, \frac{1}{8}]$.

Estimated prob. of failure is below 10% after 5 months.

Attain near-efficiency over a relatively short time span under reasonable assumptions about (1) the production technology $\frac{w(a^*)}{c(a^*)}$ (2) the informativeness parameter $w_T$ and (3) the disagreement technology $(\lambda_{a,T}, \lambda_{p,T})$. 
**Literature**

**Contracting with reference-dependent preference:** Herweg et al. (2010), Eliaz and Spiegler (2014).

**Contracting with social preference:** Akerlof (1982), Moldovanu et al. (2007), Dubey and Geanakoplos (2010).

**Hart and Moore (2008)**


**Efficiency in agency games and mechanisms:** Rubinstein and Yaari (1983), Radner (1981, 1985), Jackson and Sonnenschein (2007), Li (2014 (a)).
Summary

Formalize various kinds of disagreements about performance evaluation and compensation; examine their organizational implications.

Open questions:
- Contract renegotiation;
- Learning the disagreement point process.