Explicit vs. Implicit Incentives

Margaret A. Meyer
Nuffield College and Department of Economics
Oxford University

2014
Explicit incentives - provided through explicit contractual commitments by principal ($P$)

Implicit incentives - arise when a principal (e.g. a firm) or principals (e.g. firms competing in a labor market) have some ex post discretion how to respond to agent’s (A’s) performance; A has implicit incentives to change performance to influence $P$’s discretionary decision

Two different types of implicit incentives (= two different models of incentives deriving from a concern with “reputation”):

1. Self-enforcing contracts (= relational contracts)

2. Learning about A’s intrinsic characteristics
**Implicit incentives** - arise when principal(s) has some *ex post discretion* how to respond to agent’s (A’s) performance; A has *implicit incentives* to change performance to influence P’s discretionary decision

Applications of models of implicit incentives:

- incentives for employees (within firms)
- incentives for suppliers (between firms or between firms and consumers)
- incentives for public servants or for expert advisors
- incentives for countries to cooperate and avoid conflict
- more generally, incentives for actors to behave in a manner different from that which is in their short-run self-interest

N.B.: Implicit incentives may *encourage* or *discourage* “cooperative” (efficient) behavior
Self-enforcing contracts: this lecture

Learning about intrinsic characteristics: next lecture

Both lectures will

• first explore pure implicit incentives

• then study how implicit and explicit incentives interact
1. **Self-enforcing contracts ( = relational contracts)**

- infinitely-repeated interaction (at least potentially)
- select one equilibrium among many
- can work even if there is no learning about A's intrinsic characteristics

Earliest and most influential applications to:


Outline of presentation of self-enforcing contracts (= relational contracts):


a. Simple model of “pure” implicit incentives for effort

b. Simple model of “pure” but distortionary explicit incentives

c. Interaction between implicit incentives (relational contracts) and explicit incentives (formal contracts)

a. Simple model of “pure” implicit incentives for effort (BGM, 1994)

- risk-neutral agent $A$ privately chooses $e$ each period at cost $c(e)$
- $e$ influences $A$’s contribution to firm value, $y \in \{0, 1\}$: $P(y = 1 \mid e) = e$
- $y$ not verifiable (objectively measurable), so can’t be used in an explicit contract
- $y$ can be observed by both firm and $A$ and used in an implicit contract
- $A$ has reservation utility $w_a \geq 0$
- firm has discount rate $r$ ($A$’s discount rate is irrelevant)
- if payment equals $I$, firm’s payoff $= y - I$, and $A$’s payoff $= I - c(e)$
- first-best benchmark: if effort contractible, optimal $e^*$ satisfies $1 = c'(e^*)$
- contract: $(s, b)$, where $s$ is base salary paid if $A$ turns up and $b$ is bonus to be paid when $y = 1$; $s$ is explicit component (since whether $A$ turns up is verifiable) and $b$ is implicit component
a. Simple model of “pure” implicit incentives for effort (BGM, 1994)

Questions: i) Under what conditions will the firm be willing to pay bonus when \( y = 1 \), i.e. under what conditions will the implicit contract be “self-enforcing”? ii) What is the firm’s profit-maximizing self-enforcing bonus?

- If \( A \) trusts firm to pay bonus when \( y = 1 \), then \( A \) chooses \( e \) s.t. \( b = c'(e) \).

- Given \( b \), and \( e \) induced if \( A \) trusts, firm will choose lowest \( s \) that induces \( A \) to participate (\( = s \) s.t. \( A \)’s IR constraint binds).

- If there were a known finite horizon, relationship would “unravel”: firm would never pay bonus, so \( A \) would never supply effort.

- Consider an infinitely-repeated relationship and cooperation supported by trigger strategies:
  - Firm and worker begin by cooperating (i.e. worker chooses \( e \) s.t. \( c'(e) = b \), and firm pays \( b \) iff \( y = 1 \)). Cooperation continues unless firm ever defects. Following a defection, worker chooses \( e = 0 \) and firm never pays \( b \): these continuation strategies are best responses to each other.
a. Simple model of “pure” implicit incentives for effort (BGM, 1994)

i) Under what conditions will firm be willing to pay bonus? Define $V(b) =$ firm’s expected per-pd. profit in cooperative eqm., given that $s$ is set so $A$’s IR constraint binds. Then the “no reneging constraint” (NRC) for firm is:

$$1 - s - b + \frac{V(b)}{r} \geq 1 - s \iff V(b) \geq rb.$$

ii) What is the profit-maximizing self-enforcing bonus? Firm optimally chooses $b$ to solve max $V(b)$ s.t $V(b) \geq rb$. Denote the solution to this problem by $\tilde{b}$.

Explicit-contract benchmark: Let $b^*$ denote the optimal bonus in an explicit contract, i.e. if $y$ were contractible. With $y$ contractible and $A$ risk-neutral, optimal contract implements the first-best effort, which solves $1 = c'(e^*)$. Hence optimal explicit bonus is $b^* = 1$. 
a. Simple model of “pure” implicit incentives for effort (BGM, 1994)

**Proposition:** The profit-maximizing self-enforcing bonus $\tilde{b}$ is weakly decreasing in $r$, the firm’s discount rate. If $b^* = 1$ violates the NRC $V(b) \geq rb$, then $\tilde{b}$ is the largest $b$ satisfying the NRC.

- $r$ small (e.g. $r_1$): $\tilde{b} = b^*$ is feasible
- $r$ medium (e.g. $r_2$): $0 < \tilde{b} < b^*$
- $r$ large (e.g. $r_3$): the only $\tilde{b}$ satisfying the no-reneging constraint is $\tilde{b} = 0$
a. Simple model of “pure” implicit incentives for effort (BGM, 1994)

Discussion: In this model, only the firm is concerned with its reputation, i.e. with the present value of the ongoing relationship $V(b)/r$. A’s incentives are provided entirely by the short-term (implicit) contract.


- Observes that “firm” may have an incentive to honor the implicit contract even if the individuals making decisions about paying the bonus have finite employment horizons.

- If these indivs. own shares in the firm and can sell their shares to others (e.g. law partnerships), reneging today reduces value of shares.

- Workers, too, can have finite horizons, as long as each worker learns whether the firm ever reneged in the past.
b. Simple model of “pure” but distortionary explicit incentives (BGM (94))

- A’s privately-observed effort stochastically determines a performance measure $p$, which is objectively measurable, so can be used in an explicit contract.

- But $p \in \{0, 1\}$ is an imperfect proxy for $y$, and $A$ observes the ease of influencing $p$ before choosing effort:
  
  $\quad P(p = 1 \mid e) = ue$, where $u$ is ex ante random, $E(u) = 1$, and $u$ is observed by $A$ before choosing $e$ at cost $c(e) = \gamma e^2$.

- Explicit contract: $(s, \beta)$, where bonus $\beta$ is paid iff $p = 1$.

- First-best effort satisfies $1 = c'(e^*)$, so $e^* = (1/2\gamma)$.

- Given $(s, \beta)$ and $u$, $A$ chooses $e$ s.t. $u\beta = c'(e)$, so $e = u\beta/2\gamma$.

  - Thus, the explicit contract induces $A$’s effort to fluctuate with $u$, even though first-best effort $e^*$ is independent of $u$. But fluctuations in effort per se reduce $A$’s expected utility (since $c(e)$ is strictly convex), thus increasing the salary $s$ required to induce $A$ to participate.
b. Simple model of “pure” but distortionary explicit incentives (BGM (94))

**Optimal explicit contract:** Firm chooses \((\beta, s)\) to solve

\[
\max \mathbb{E}[e - \beta ue - s]
\]

subj. to

\[
\mathbb{E}[s + \beta ue - \gamma e^2] \geq w_a \quad (IR)
\]

and

\[
e = u\beta / 2\gamma \quad (IC)
\]

Given \(\beta\), optimal \(s\) makes (IR) bind.

**Proposition:** The optimal explicit bonus, \(\beta^*\), satisfies \(\beta^* = 1/(1 + \text{var}(u)) < 1\) and is decreasing in \(\text{var}(u)\). The firm’s maximized per-period profit, \(V(\beta^*)\), is decreasing in \(\text{var}(u)\) and can be positive or negative (given that \(w_a \geq 0\)).

- Firm trades off strength of incentives against magnitude of (inefficient) fluctuations in effort. The more *distortionary* is the explicit performance measure (larger \(\text{var}(u)\)), the smaller are the optimal explicit incentives.

- cf. “effort substitution problem” in multi-task moral hazard model of Holmstrom and Milgrom (1991)—in both cases, \(A\)’s ability to substitute efforts across tasks or states (values of \(u\)) makes lower-powered incentives optimal.
c. Interaction btw. implicit and distortionary explicit incentives (BGM (94))

- Contract: \((s, b, \beta)\), where implicit bonus \(b\) is to be paid iff \(y = 1\) and explicit bonus \(\beta\) is paid iff \(p = 1\).

- If \(A\) trusts firm to pay \(b\) when \(y = 1\), \(A\) chooses \(e\) s.t. \(b + u\beta = c'(e)\).

Questions: i) Under what conditions is implicit bonus \(b\) self-enforcing? ii) What are optimal \(b\) and \(\beta\)? iii) Are implicit and explicit incentives substitutes or complements for firm? i.e. do optimal \(b\) and \(\beta\) vary with changes in environment, e.g. \(\text{var}(u)\), in opposite directions (substitutes) or same direction (complements)?

- Again, focus on cooperation supported by trigger strategies.

- Now, what happens if firm reneges on implicit component of the contract depends on whether \(V(\beta^*) > 0\) or \(V(\beta^*) \leq 0\):
  - If \(V(\beta^*) > 0\), then if firm reneges, parties revert to the optimal pure explicit contract \((s^*, \beta^*)\).
  - If \(V(\beta^*) \leq 0\), then if firm reneges, parties revert to no further interaction.
i) Under what conditions is implicit bonus $b$ self-enforcing? Define $V(b, \beta) =$ firm's expected per-pd. profit in cooperative eqm., given that $s$ is set so $A$'s IR constraint binds.

- **Case 1:** If $V(\beta^*) > 0$, then NRC is $V(b, \beta) - V(\beta^*) \geq rb$.

- **Case 2:** If $V(\beta^*) \leq 0$, then NRC is $V(b, \beta) \geq rb$.

ii) What are optimal $b$ and $\beta$? Denote these by $(b^{**}, \beta^{**})$.

**Proposition:** In both Cases 1 and 2, for any $b$, the optimal $\beta^{**}(b)$ satisfies

$$\beta^{**}(b) = \frac{(1 - b)}{1 + \text{var}(u)} = (1 - b)\beta^*,$$

where $\beta^* = 1/(1 + \text{var}(u))$ is optimal in a pure explicit contract. If $b = 1$ is not self-enforcing, then given $\beta = \beta^{**}(b)$, the optimal $b$ is the largest one satisfying NRC.

**NB:** $b$ and $\beta$ are substitutes in their effect on $e$: $b + u\beta = c'(e)$, as reflected in $(*)$. 
iii) Are implicit and explicit incentives substitutes or complements as $\text{var}(u)$ changes?

**Case 1:** $\text{var}(u)$ small, so $V(\beta^*) > 0$ (explicit incentives alone are profitable)

Firm chooses $(b, \beta)$ to solve $\max V(b, \beta)$ subj. to $V(b, \beta) - V(\beta^*) \geq rb$ (NRC).

**Prop.:** If $V(\beta^*) > 0$, then as $\text{var}(u) \downarrow$, $b^{**} \downarrow$ and $\beta^{**} \uparrow$.

- So in this case, implicit and explicit incentives are *substitutes*.

- When $\text{var}(u)$ is sufficiently low, then no implicit contract is feasible ($b^{**} = 0$), because firm’s fallback position after reneging, reverting to a pure explicit contract, is too attractive.

- As $\text{var}(u) \downarrow$, both $V(\beta^*)$ and $V(b, \beta^{**}(b))$ rise, but $V(\beta^*)$ rises faster than $V(b, \beta^{**}(b))$, so NRC becomes *tighter*. Hence $b$ must be lowered, and (*) then implies $\beta$ is raised.
\( \beta^{**} = \beta^* \): optimal explicit bonus in the absence of implicit contracts

\( b^{**} \): optimal implicit bonus in the presence of explicit contracts

\( \beta^* \): optimal implicit bonus in the absence of explicit contracts

\( \beta^{**} \): optimal explicit bonus in the presence of implicit contracts

\( V(\beta^*) = 0 \)

\( V(\beta^*) > 0 \) \( \iff \) \( V(\beta^*) < 0 \)

\( \text{var}(u) \)
Case 2: \( \text{var}(u) \) large, so \( V(\beta^*) \leq 0 \) (explicit incentives alone not profitable)

Firm chooses \((b, \beta)\) to solve \( \max V(b, \beta) \) subj. to \( V(b, \beta) \geq rb \) (NRC).

Prop.: If \( V(\beta^*) \leq 0 \), then as \( \text{var}(u) \downarrow \), \( b^{**} \uparrow \) and \( \beta^{**} \) may \( \uparrow \) or \( \downarrow \).

- In Case 2, making the environment more attractive for explicit incentives (\( \downarrow \text{var}(u) \)) allows stronger implicit incentives.

- Now, as \( \text{var}(u) \downarrow \), NRC becomes looser, because \( V(b, \beta) \uparrow \). Hence \( b \) can be raised. But \( \downarrow \text{var}(u) \) directly raises \( \beta^* \), so (*) shows conflicting effects on \( \beta^{**} \).

- When both \( b^{**} \uparrow \) and \( \beta^{**} \uparrow \) as \( \text{var}(u) \downarrow \), then implicit and explicit incentives are complements – a more surprising result.

- This complementarity may help explain empirical findings that subjective bonus plans (implicit) and group piece rates (explicit) are more effective together than singly (Holmstrom-Milgrom, _AER_, 94; Ichniowski et al, _AER_, 97).
Formally similar interactions between relational contracts (informal arrangements) and formal contracts have been examined in other contexts:

- **Development**: publicly-provided social insurance can affect performance of informal family insurance arrangements (Di Tella and MacCulloch, *EJ*, 2002)

- **Macro**: public insurance provided through redistributive taxation may crowd out private insurance arrangements that are subject to limited enforcement (Krueger and Perri, 2010)

- **International**: “globalization”, by expanding (fallback) market alternatives, can potentially impede risk-sharing in long-term employment relationships (McLaren and Newman, 2002; Spagnolo, 2002)

- **Experiments**: interaction between informal and formal arrangements (Lazzarini, Miller, and Zenger, *JLEO*, 2004)
More general analysis of relational contracts: Levin, AER, 2003

- $A$ has i.i.d. (over time) private information about cost of effort (HI) and chooses effort privately (HA)
- more general, time-invariant, distribution of output $y_t$ conditional on effort $e_t$
- general forms of relational contracts allowed
- $P$ and $A$ risk-neutral and have time-invariant outside options
- infinitely-repeated relationship with exponential discounting

Prop.: If output is observable by both $P$ and $A$ (though not verifiable), then among optimal relational contracts (those that maximize expected surplus), there are ones that are stationary, i.e. the effort rule and the compensation schedule are independent of time and history.
More general analysis of relational contracts: Levin, AER, 2003

Prop.: If output is observable by both $P$ and $A$ (though not verifiable), then among optimal relational contracts (those that maximize expected total surplus), there are ones that are stationary, i.e. the effort rule and the compensation schedule are independent of time and history.

- ∃ two ways of giving incentives to $A$: discretionary transfers and future conduct of relationship; for risk-neutral $A$, perfect substitutes in their effect.
- But changing transfer to reward $A$ must harm $P$ and vice versa, while adjusting continuation behavior could punish both $P$ and $A$.
- With output observable by both $P$ and $A$, it is never necessary to punish both $P$ and $A$ by destroying future surplus.
- Hence, all incentives are optimally provided through discretionary transfers to $A$ based on most recent output.
- NB: Availability of monetary transfers distinguishes relational contract analysis from repeated-game analysis, and is key to stationarity result.
More general analysis of relational contracts: Levin, AER, 2003

**Prop.** Optimal stationary contracts maximize expected total surplus subj. to incentive-compatibility constraint for A and subj. to dynamic enforcement constraint

\[
\max \text{payment} - \min \text{payment} \leq \frac{\delta}{1 - \delta} (s - \tilde{s}), \quad (DE)
\]

where \(s\) is expected total surplus and \(\tilde{s}\) is the sum of P’s and A’s outside options.

- (DE) captures the no-reneging/no-quitting constraints for both P and A

**Prop.** Optimal stationary contracts under either hidden information (HI) or hidden action (HA) display "information compression":

i) Under (HI), if first-best not feasible, then effort of every cost-type of A is strictly below first-best, and full separation of cost-types is not optimal.

ii) Under (HA), A receives only 2 possible payments, high (low) if output is above (below) a threshold.

- (HA): For a risk-neutral A, incentives should be as strong as possible subj. to (DE) \(\implies\) one-step payment rule optimal

Now introduce **subjective performance measurement**: output is **privately** observed by $P$

**Prop.** An optimal relational contract in which $P$ reports $y_t$ in every period is a **termination contract** specifying a time-invariant effort and in which, if output is below a threshold, payment is low and relationship is terminated, while if output is above that threshold, payment is high and relationship continues.

- Now it is impossible, with a stationary contract, to provide $P$ with incentives for truthful reporting of output as well as $A$ with incentives for effort.

- So must destroy future surplus along the eqm path when $P$ reports that output is low.
Recent literature studies the **dynamics within relational contracts** when optimal relational contracts are not stationary:

- **e.g.** Li and Matouschek, *AER*, 2013: If firm privately informed about opportunity cost of paying bonus, optimal relational contract generates periodic conflicts during which effort and expected profit fall, but relationship never terminates.

  – cf. Yared, *JET*, 2010: uses a similar model to analyze dynamics of relationship between an aggressive (*A*) and a peaceful (*P*) country, where *A* demands transfers, *P* is privately informed about cost of paying, and *A* can respond to non-payment by going to war.

- **e.g.** Chassang, *AER*, 2010: Analyzes how parties learn to cooperate, given asymmetric information about effects of actions; initially, while learning occurs, there may be inefficient termination, but if not, learning eventually stops and parties adopt a “routine” for cooperating.
Two different types of implicit incentives (＝ two different models of incentives deriving from a concern with “reputation”):

1. **Self-enforcing contracts (＝relational contracts)**

2. **Learning about A’s intrinsic characteristics**

2. Learning about agent’s intrinsic characteristics: can be studied using **signaling (hidden info) or signal-jamming (hidden action)**

   - In both signaling and signal-jamming, *P learns* from *A’s current performance* about *A’s intrinsic characteristics*, which are relevant to *A’s future performance*; *A* has implicit incentives to change current performance to influence *P’s beliefs* about *A’s type and hence *P’s future decisions*.

   - This mechanism can operate even in **finite-horizon** settings, e.g. 2 periods.

     - (How do incentives change as near end of horizon? Depends upon whether *A* can “sell” his reputation and on whether changes in ownership are observable by clients—see Tadelis, *JPE*, 2002; Mailath and Samuelson, *REStud*, 2001; Hakenes and Peitz, *IER*, 2007.)
2a. Learning about A’s intrinsic characteristics: Signaling (hidden info)

Example: Spence (1973) model of education as a signal in labor market

- A’s ability – privately observed by A
- A’s effort (e.g. educational, self-promotion) – public signal
- market pays wages based on estimate of ability
- If MC of effort is lower for higher-ability agents, agents have incentives to use effort to signal higher ability and hence generate higher wage offers.
2b. Learning about A’s intrinsic characteristics: Signal-jamming (hidden action)

*Example:* Holmstrom (1999) model of managerial career concerns

- A’s ability – observed by no one
- A’s effort – privately-observed by A
- A’s output – public signal; depends on ability and effort
- Labor market pays wages based on its estimate of ability
- Incentives for effort derive from possibility of influencing, at the margin, market’s beliefs about ability, since market *can’t distinguish* between effects of ability and effort on output.
2c: Many models of implicit incentives stemming from learning are intermediate between signaling and signal-jamming

- With hidden information, $A$ might take an unobservable action (depending on his type) that generates a noisy signal. Call this “noisy signaling”.

- Noisy signaling models are particularly useful for studying the dynamics of reputational incentives.

Additional examples of implicit incentives stemming from learning:

- ratchet effect in regulation (firm’s type is its intrinsic efficiency)

- incentives for incumbent politicians seeking re-election (politician’s type can be his ability or his preferences/corruptibility)
  - e.g. Coate and Morris (1995) – noisy signaling
Additional examples of implicit incentives stemming from learning:

- incentives for influence activities in firms (employee’s type can be his own quality or quality of his project, division, etc.)
  - e.g. Meyer, Milgrom, and Roberts (1992) – signal-jamming

- incentives for firms to provide high-quality products (firm’s type is its ability to provide high quality)
  - e.g. Tadelis (2002) and Mailath and Samuelson (2001) – noisy signaling

- incentives for experts to provide good advice (expert’s type can be his ability (=quality of his information) or how aligned his preferences are with decision-maker’s)
  - e.g. Scharfstein and Stein (1990), Ottaviani and Sorensen (2006), Prat (2005) – signal-jamming (expert does not know his ability, and his hidden action is his choice of how to map his observation into his advice)