Feedback Design in Dynamic Moral Hazard

Jeffrey Ely, George Georgiadis and Luis Rayo

Northwestern University

Introduction

We study the joint design of monetary and informational incentives in a dynamic agency model under moral hazard.

Framework in a Nutshell

- In continuous-time, the agent chooses to work or shirk
- Working generates a binary signal, privately observed by the principal
- The principal designs a reward schedule and a feedback policy to maximize total effort net of payments to the agent

Application

- Consider an associate at a professional partnership
  - Can be a “high” or a “low” type
- Signal represents conclusive evidence that the associate’s type is high; e.g., has the necessary skills to be partner
  - The firm is better informed whether the associate is “partner material”
- Firm values effort (billable hours). Decides *when* to tell the agent.

Main Result

The optimal incentive scheme comprises two phases:

- A *silent* phase lasting until t^* during which the agent is provided no information, and is paid $\overline{R}$ if the signal arrives by t^*; and
- A *pronto* phase lasting from t^* until T^* during which the agent is told as soon as the signal arrives, and is paid a smaller reward

Related Literature

Dynamic contract theory.

- *Poisson output.* Mason and Välimäki (2015); Green and Taylor (2016)
- *Good-news experimentation.* Halac, Kartik and Liu (2016)

Dynamic Information design. Information about payoff-relevant variable

- *Receiver behaves myopically.* Ely (2017) and Renault et al. (2017),...
- *Strategic receiver.* Ely and Szydlowski (2020), Orlov et al. (2020), Smolin (2021), Ball (2022),...
- *Moral hazard.* Varas et al. (2020), and Hörner and Lambert (2021), Kaya (2022), Ely et al. (Forthcoming) ...

Our model is (*one of?*) the first to study dynamic information design with a fully endogenous payoff structure

Model – The Setting

Agent makes effort decision in continuous time to generate a binary signal

- Effort is binary (work / shirk) with flow cost $c > 0$
- Signal arrives stochastically as a function of accumulated effort
- Agent privately observes effort; principal privately observes the signal

* Signal structure ⇒ agent quits upon learning that the signal has arrived
  - e.g., at that moment, the job market also learns he is a "high" type and the principal must pay him his marginal output going forward

Model – The Setting

Agent makes effort decision in continuous time to generate a binary signal

- Effort is binary (work / shirk) with flow cost $c > 0$
- Signal arrives stochastically as a function of accumulated effort
- Agent privately observes effort; principal privately observes the signal

⋆ Signal structure ⇒ agent quits upon learning that the signal has arrived

- e.g., at that moment, the job market also learns he is a "high" type and the principal must pay him his marginal output going forward

Model – Principal's Design Problem

At time 0, the principal chooses a "contract", which comprises

1. A **terminal date** T;
2. A **reward schedule** $R(t)$ specifying a reward for the agent as a function of the arrival date of the signal; and
3. A **feedback policy** that specifies messages at each instant as a function of whether the signal has arrived and past messages.

The principal maximizes total effort net of monetary payments

Wolog, focus on designs that motivate working continuously until quitting

Model – Principal's Design Problem

At time 0, the principal chooses a "contract", which comprises

1. A **terminal date** T;
2. A **reward schedule** $R(t)$ specifying a reward for the agent as a function of the arrival date of the signal; and
3. A **feedback policy** that specifies messages at each instant as a function of whether the signal has arrived and past messages.

The principal maximizes total effort net of monetary payments

Wolog, focus on designs that motivate working continuously until quitting

Assumptions

Denote by $F(t)$ the probability that the signal arrives by t (assuming the agent has worked continuously until that time)

A.1. $F(\cdot)$ is twice differentiable with pdf $f(\cdot)$ and its hazard rate

$$\lambda(t) = \frac{f(t)}{1 - F(t)}$$

is weakly decreasing. (*Implies that $F(\cdot)$ is concave.*)

A.2. The function

$$\Phi(t) = F(t)\frac{d}{dt}\frac{1}{f(t)}$$

is increasing. (*Weak log-concavity of $\lambda(\cdot)$ implies A.2*)

Assumptions

Denote by $F(t)$ the probability that the signal arrives by t (assuming the agent has worked continuously until that time)

A.1. $F(\cdot)$ is twice differentiable with pdf $f(\cdot)$ and its hazard rate

$$\lambda(t) = \frac{f(t)}{1 - F(t)}$$

is weakly decreasing. (*Implies that $F(\cdot)$ is concave.*)

A.2. The function

$$\Phi(t) = F(t)\frac{d}{dt}\frac{1}{f(t)}$$

is increasing. (*Weak log-concavity of $\lambda(\cdot)$ implies A.2*)

$F(\cdot)$: Two Examples that satisfy Assumptions A.1-3

I. *Poisson signal*

- Conditional on working, signal arrives with constant hazard rate

II. *Poisson "good news" experimentation*

- Agent's type is either "high" or "low"
- Type is ex-ante unknown and the players share a common prior
- Conditional on the agent being a high type and working, the signal arrives according to a Poisson process
- If the agent is a low type, the signal never arrives

The Challenge

The space of message/feedback policies is immense. For example:

- No messages at all (*silence*)
- Inform the agent as soon as the signal arrives (*pronto*)
- Inform with a delay
- Inform probabilistically
- Message at t can condition on messages at prior dates

Simplifying the Problem

A-lá revelation principle, it suffices to consider direct policies in which the principal recommends to the agent to work or quit at each instant.

For each t, define the (survival) function

$$q(s|t) := \Pr\{\text{ask agent to work at least until } s \mid \text{signal arrived at } t\},$$

which must be non-increasing.

Wolog, ask the agent to work at all $t \leq T$ if the signal has not arrived

Simplifying the Problem

A-lá revelation principle, it suffices to consider direct policies in which the principal recommends to the agent to work or quit at each instant.

For each t, define the (survival) function

$$q(s|t) := \Pr\{\text{ask agent to work at least until } s \mid \text{signal arrived at } t\},$$

which must be non-increasing.

Wolog, ask the agent to work at all $t \leq T$ if the signal has not arrived

Simplifying the Problem

A-lá revelation principle, it suffices to consider direct policies in which the principal recommends to the agent to work or quit at each instant.

For each t, define the (survival) function

$$q(s|t) := \Pr\{\text{ask agent to work at least until } s \mid \text{signal arrived at } t\},$$

which must be non-increasing.

Wolog, ask the agent to work at all $t \leq T$ if the signal has not arrived

Payoffs

Let

$$p(s) := 1 - F(s) + \int_0^s q(s|t)f(t)dt$$

be the unconditional probability the agent has **not** been told to quit by s

The **agent's expected payoff** when obeying the recommendations is

$$\underbrace{\int_0^T R(s)f(s)ds}_{\text{expected reward}} - c \times \underbrace{\int_0^T p(s)ds}_{\text{expected effort}}$$

The **principal's objective** is to maximize

$$\int_0^T p(s)ds - \int_0^T R(s)f(s)ds$$

subject to the IC constraint that the agent obeys recommendations

Payoffs

Let

$$p(s) := 1 - F(s) + \int_0^s q(s|t) f(t) dt$$

be the unconditional probability the agent has **not** been told to quit by s

The **agent's expected payoff** when obeying the recommendations is

$$\underbrace{\int_0^T R(s) f(s) ds}_{\text{expected reward}} - c \times \underbrace{\int_0^T p(s) ds}_{\text{expected effort}}$$

The **principal's objective** is to maximize

$$\int_0^T p(s) ds - \int_0^T R(s) f(s) ds$$

subject to the IC constraint that the agent obeys recommendations

Payoffs

Let

$$p(s) := 1 - F(s) + \int_0^s q(s|t)f(t)dt$$

be the unconditional probability the agent has **not** been told to quit by s

The **agent's expected payoff** when obeying the recommendations is

$$\underbrace{\int_0^T R(s)f(s)ds}_{\text{expected reward}} - c \times \underbrace{\int_0^T p(s)ds}_{\text{expected effort}}$$

The **principal's objective** is to maximize

$$\int_0^T p(s)ds - \int_0^T R(s)f(s)ds$$

subject to the IC constraint that the agent obeys recommendations

Towards Incentive Compatibility: Impact of Deviations

Fix a contract (terminal date, reward schedule & recommendation policy)

We will focus on *local* incentive compatibility whereby the agent obeys recommendations to work except during some interval $(t, t + dt)$

This deviation has three effects:

1. The agent misses on any rents during this interval;
2. The arrival rate of the signal from t onwards changes; and
3. Total effort changes (since recommendations depend on signal arrival)

We establish a condition that deters instantaneous pauses

We characterize the profit-maximizing contract subject to local IC. Then we verify that this contract also deters all global deviations.

Towards Incentive Compatibility: Impact of Deviations

Fix a contract (terminal date, reward schedule & recommendation policy)

We will focus on *local* incentive compatibility whereby the agent obeys recommendations to work except during some interval $(t, t + dt)$

This deviation has three effects:

1. The agent misses on any rents during this interval;
2. The arrival rate of the signal from t onwards changes; and
3. Total effort changes (since recommendations depend on signal arrival)

We establish a condition that deters instantaneous pauses

We characterize the profit-maximizing contract subject to local IC. Then we verify that this contract also deters all global deviations.

Towards Incentive Compatibility: Impact of Deviations

Fix a contract (terminal date, reward schedule & recommendation policy)

We will focus on *local* incentive compatibility whereby the agent obeys recommendations to work except during some interval $(t, t + dt)$

This deviation has three effects:

1. The agent misses on any rents during this interval;
2. The arrival rate of the signal from t onwards changes; and
3. Total effort changes (since recommendations depend on signal arrival)

We establish a condition that deters instantaneous pauses

We characterize the profit-maximizing contract subject to local IC. Then we verify that this contract also deters all global deviations.

Local Incentive Compatibility

Proposition 1.

The reward schedule $R(\cdot)$ and recommendation policy $q(\cdot|\cdot)$ are locally IC if

$$R(t)f(t) - cp(t) \geq \int_t^T R(s)|f'(s)|ds - c \times \int_t^T \dot{p}(s|t)ds \text{ for all } t,$$

where $\dot{p}(s|t)$ is the marginal change in $p(s)$ following the deviation at t.

LHS measures the agent's on-path flow rents

RHS represents impact of pause on the $\mathbb{E}[\text{rents}]$ going forward

- First term captures a *backward compounding effect*: For fixed t, raising any future $R(s)$ makes delaying effort at t more tempting
- Second term measures the change in total effort cost

A key observation: Other things equal, raising $R(s)$ tightens IC for all $t < s$

Local Incentive Compatibility

Proposition 1.

The reward schedule $R(\cdot)$ and recommendation policy $q(\cdot|\cdot)$ are locally IC if

$$R(t)f(t) - cp(t) \geq \int_t^T R(s)|f'(s)|ds - c \times \int_t^T \dot{p}(s|t)ds \text{ for all } t,$$

where $\dot{p}(s|t)$ is the marginal change in $p(s)$ following the deviation at t.

LHS measures the agent's on-path flow rents

RHS represents impact of pause on the $\mathbb{E}[\text{rents}]$ going forward

- First term captures a *backward compounding effect*: For fixed t, raising any future $R(s)$ makes delaying effort at t more tempting
- Second term measures the change in total effort cost

A key observation: Other things equal, raising $R(s)$ tightens IC for all $t < s$

Minimal "Implementing" Reward Schedule

Proposition 2.

Fix a recommendation policy $q(\cdot|\cdot)$. The reward schedule

$$R^*(t) = c\left[\frac{p(t)}{f(t)} - \int_t^T \frac{f'(s)}{f(s)^2} p(s) ds - \int_t^T 1 - q(s|t) ds\right]$$

satisfies IC and is pointwise smaller than any other implementing $R(\cdot)$

- 1^{st} *term.* Reward that gives zero flow rent at time t
- 2^{nd} *term.* Backward compounding effect – modulated by the speed at which $1/f$ grows
- 3^{rd} *term.* "Information rebate" if agent is asked to quit before T

Minimal Reward Schedule for 2 Polar Cases

A. Principal always asks the agent to work:

$$R^*_{silence} = \frac{c}{f(T)}$$

B. Principal asks the agent to quit as soon as the signal arrives:

$$R^*_{pronto} = c \times \frac{1 - F(T)}{f(T)}$$

- To give 0 flow rents, $R(t)$ must increase in t. But this would lead the agent to pause, so R^* is closest to the ideal while meeting IC.
- Feedback reduces cost p.u effort but imposes a bound on max. effort.

Minimal Reward Schedule for 2 Polar Cases

A. Principal always asks the agent to work:

$$R^*_{silence} = \frac{c}{f(T)}$$

B. Principal asks the agent to quit as soon as the signal arrives:

$$R^*_{pronto} = c \times \frac{1 - F(T)}{f(T)}$$

- To give 0 flow rents, $R(t)$ must increase in t. But this would lead the agent to pause, so R^* is closest to the ideal while meeting IC.
- Feedback reduces cost p.u effort but imposes a bound on max. effort.

Minimal Reward Schedule for 2 Polar Cases

A. Principal always asks the agent to work:

$$R^*_{silence} = \frac{c}{f(T)}$$

B. Principal asks the agent to quit as soon as the signal arrives:

$$R^*_{pronto} = c \times \frac{1 - F(T)}{f(T)}$$

- To give 0 flow rents, $R(t)$ must increase in t. But this would lead the agent to pause, so R^* is closest to the ideal while meeting IC.
- Feedback reduces cost p.u effort but imposes a bound on max. effort.

Simplifying the Principal's Objective

Using $R^*(\cdot)$ we can write the principal's objective as a function of just $p(\cdot)$, the unconditional probability the agent has **not** been asked to quit by s

Lemma 1.

The principal's payoff evaluated at the minimal reward schedule is

$$\int_0^T p(t)dt - c\int_0^T p(t)\left[1+\Phi(t)\right] - \left[1-p(t)\right] dt$$

1^{st} *term:* Total effort

2^{nd} *term:* True cost of effort + information rents

- Note that $\Phi(t)$ initially equals 0 and is increasing by assumption A.2

A Relaxed Problem

$$\max_{T \geq 0,\, p(\cdot)} \int_0^T p(t)[1 - 2c - c\Phi(t)]dt + cT \tag{RP}$$

$$\text{s.t. } 1 - F(t) \leq p(t) \leq 1 \text{ for all } t \text{ and non-increasing.}$$

This is a relaxed problem because:

i. No guarantee that a recommendation policy implementing $p(\cdot)$ exists;

ii. Local IC does not generally imply global incentive compatibility.

Proposition 3.

Let t^* be earliest time when $1 - 2c - c\Phi(t) \leq 0$. (RP) is solved by setting

$$p(t) = \begin{cases} 1 & \text{if } t \in [0, t^*) \\ 1 - F(t) & \text{if } t \in [t^*, T^*] \text{ for some } T^* > t^*. \end{cases}$$

A Relaxed Problem

$$\max_{T \geq 0,\, p(\cdot)} \int_0^T p(t)[1 - 2c - c\Phi(t)]dt + cT \tag{RP}$$

s.t. $1 - F(t) \leq p(t) \leq 1$ for all t and non-increasing.

This is a relaxed problem because:

i. No guarantee that a recommendation policy implementing $p(\cdot)$ exists;

ii. Local IC does not generally imply global incentive compatibility.

Proposition 3.

Let t^* be earliest time when $1 - 2c - c\Phi(t) \leq 0$. (RP) is solved by setting

$$p(t) = \begin{cases} 1 & \text{if } t \in [0, t^*) \\ 1 - F(t) & \text{if } t \in [t^*, T^*] \text{ for some } T^* > t^*. \end{cases}$$

A Relaxed Problem

$$\max_{T \geq 0,\, p(\cdot)} \int_0^T p(t)[1 - 2c - c\Phi(t)]dt + cT \qquad \text{(RP)}$$

$$\text{s.t. } 1 - F(t) \leq p(t) \leq 1 \text{ for all } t \text{ and non-increasing.}$$

This is a relaxed problem because:

i. No guarantee that a recommendation policy implementing $p(\cdot)$ exists;

ii. Local IC does not generally imply global incentive compatibility.

Proposition 3.

Let t^* be earliest time when $1 - 2c - c\Phi(t) \leq 0$. (RP) is solved by setting

$$p(t) = \begin{cases} 1 & \text{if } t \in [0, t^*) \\ 1 - F(t) & \text{if } t \in [t^*, T^*] \text{ for some } T^* > t^*. \end{cases}$$

Main Result

Theorem 1. Every optimal contract comprises at most two phases:

φ.1. Until t^*, the agent is asked to work under silence; i.e., $q(s|t) \equiv 1$

φ.2. During $[t^*, T^*]$, the agent is told to quit as soon as the signal arrives; i.e., $q(s|t) \equiv 0$ (a.k.a *pronto* feedback)

Reward $\overline{R}$ if signal arrives by t^* and $\underline{R} < \overline{R}$ if it arrives in $(t^*, T^*]$.

- Pronto feedback is always used because it minimizes the cost p.u of effort. But it bounds the total effort that can be elicited.
- To elicit extra effort, the agent must be sometimes kept in the dark.
- This requires larger rewards. And because rewards are compounded backwards, it is optimal to maximally frontload the silent period.

Main Result

Theorem 1. Every optimal contract comprises at most two phases:

φ.1. Until t^*, the agent is asked to work under silence; i.e., $q(s|t) \equiv 1$

φ.2. During $[t^*, T^*]$, the agent is told to quit as soon as the signal arrives; i.e., $q(s|t) \equiv 0$ (a.k.a *pronto* feedback)

Reward $\overline{R}$ if signal arrives by t^* and $\underline{R} < \overline{R}$ if it arrives in $(t^*, T^*]$.

- Pronto feedback is always used because it minimizes the cost p.u of effort. But it bounds the total effort that can be elicited.
- To elicit extra effort, the agent must be sometimes kept in the dark.
- This requires larger rewards. And because rewards are compounded backwards, it is optimal to maximally frontload the silent period.

Tying the loose ends: Optimal Reward Schedule

The rewards schedule is computed per Proposition 2 such that IC binds:

- If the signal arrives before t^*, then the agent is paid

$$R(t) = \overline{R} := \frac{c}{\lambda(T^*)} + \frac{cF(t^*)}{f(t^*)}$$

- If the signal arrives during $(t^*, T^*]$, then the agent is paid

$$R(t) = \underline{R} := \frac{c}{\lambda(T^*)}.$$

Tying the loose ends: Global Deviations

During $(t^*, T^*]$**:**

- The agent earns rents that decrease in time (due to $\lambda(t)$ declining), and is asked to work only if the signal hasn't yet arrived.
- So following the recommendations is a dominant strategy.

During $[0, t^*]$**:**

- Because the reward schedule is time-invariant and $p(t) = 1$, it suffices to check that shirking for any $s \in (0, t^*]$ units of time is not profitable.
- This follows from the concavity of $F(\cdot)$.

Tying the loose ends: Global Deviations

During $(t^*, T^*]$**:**

- The agent earns rents that decrease in time (due to $\lambda(t)$ declining), and is asked to work only if the signal hasn't yet arrived.
- So following the recommendations is a dominant strategy.

During $[0, t^*]$**:**

- Because the reward schedule is time-invariant and $p(t) = 1$, it suffices to check that shirking for any $s \in (0, t^*]$ units of time is not profitable.
- This follows from the concavity of $F(\cdot)$.

Practical Implications

- Obfuscating information can be a useful tool for eliciting more effort.
- Hiding information is costly & the cost depends on *when* information is withheld. In our setting, it is ideal to do so early on.
- A final phase of full transparency is desirable as it elicits additional effort at minimal cost.
- If the hazard rate falls sufficiently over time, it is best to give up on the employee; i.e., terminate the relationship at $T^* < \infty$.
- Adopting a decreasing reward schedule helps prevent pauses and minimizes backward compounding.

Remarks

I. Optimal terminal date T^* is finite if the hazard rate falls sufficiently over time. If the hazard rate is constant, then $T^* = \infty$.

II. Theorem 1 true if the hazard rate is hump-shaped with $\lambda(t^*) \geq \lambda(T^*)$

e.g., with good-news experimentation where the "good" type "succeeds" with hazard rate that increases in cumulative effort.

III. Theorem 1 true (albeit with a different T^*) if the principal or the agent value the signal in and of itself.

IV. If the agent can succeed repeatedly, then the principal optimally keeps the agent fully apprised and rewards them with c/λ for each success.

Discussion

Dynamic agency model with joint design of monetary rewards & feedback

In our setting, the optimal incentive scheme comprises two phases:

φ1. No feedback with "success" accompanied by a relatively large reward

φ2. Full transparency with success accompanied by a smaller reward

Broader agenda: Paying with information + money

- More general signal processes; e.g., multiple or continuous signals
- Many agents; e.g., contests (for experimentation)

Robustness and Extensions

- If the reward schedule is given...

  the optimal recommendation policy is generally more complicated; *e.g.*, $q(s|t)$ may depend on s and t in intricate ways.

- If the hazard rate is increasing...

  the optimal contract features a silent phase, followed by a partial disclosure phase, and possibly a *pronto* phase at the end.

- With many agents and given a fixed prize...

  the principal administers a *two-phase, weighted-egalitarian* contest: Each agent who succeeds during the first (silent) phase earns one share; each who succeeds during the second (*pronto*) phase earns $\alpha < 1$ shares, and the prize is split according to the awarded shares.