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:

- During $[0, t^*]$, the agent is offered no information, and is paid $\bar{R}$ if the signal arrives by $t^*$
- During $[t^*, T^*]$, the agent is told as soon as the signal arrives, and is paid a smaller reward at that instant
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 $\Rightarrow$ 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 $\Rightarrow$ 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 concave.

A.2. $F(\cdot)$ is twice differentiable with pdf $f(\cdot)$ and its hazard rate

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

is weakly decreasing. (A.2 implies A.1)

A.3. The function

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

is increasing. (Weak log-concavity of $\lambda(\cdot)$ implies A.3)
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 concave.

**A.2.** $F(\cdot)$ is twice differentiable with pdf $f(\cdot)$ and its hazard rate

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

is weakly decreasing. (*A.2 implies A.1*)

**A.3.** The function

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

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

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

which is non-increasing (akin to a CDF).

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 to recommends to the agent to work or quit at each instant.

For each $t$, define function

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

which is non-increasing (akin to a CDF).

Wolog, ask the agent to work at all $t \leq T$ if the signal has not arrived
A-lá revelation principle, it suffices to consider direct policies in which the principal to recommends to the agent to work or quit at each instant.

For each \( t \), define function

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

which is non-increasing (akin to a CDF).

Wolog, ask the agent to work at all \( t \leq T \) if the signal has not arrived
A-lá revelation principle, it suffices to consider direct policies in which the principal to recommends to the agent to work or quit at each instant.

For each $t$, define function

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

which is non-increasing (akin to a CDF).

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 asked to quit by \( s \).

The agent’s expected payoff when obeying the recommendations is

\[ \int_{0}^{T} R(s)f(s)ds - c \times \int_{0}^{T} p(s)ds \]

expected reward \[ \int_{0}^{T} p(s)ds \] 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 asked to quit by \( s \).

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

\[
\int_0^T R(s)f(s)ds - c \times \int_0^T p(s)ds
\]

- **expected reward**
- **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 asked to quit by \( s \)

The agent’s expected payoff when obeying the recommendations is

\[
\left( \int_0^T R(s)f(s)\,ds \right) - c \times \int_0^T p(s)\,ds
\]

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

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 it satisfies global IC.
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 it satisfies global IC.
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 it satisfies global IC.
Lemma 1.
The reward schedule $R(\cdot)$ and recommendation policy $q(\cdot|\cdot)$ are locally IC if

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

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*: Fixing $t$, raising $R(s)$ for any $s > t$ tightens the $t$-constraint
- Second term measures the change in total effort cost
Proposition 1.

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
- $3^{rd}$ term. “Information rebate” if agent is advised 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 advises 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:

\[ R^*_\text{silence} = \frac{c}{f(T)} \]

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

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

- 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:

\[ R_{\text{silence}}^* = \frac{c}{f(T)} \]

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

\[ R_{\text{pronto}}^* = c \times \frac{1 - F(T)}{f(T)} \]

- 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 2.
The principal’s payoff evaluated at the minimal reward schedule is

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

1$^{st}$ term: Total effort

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

- Note that $\Phi(t)$ initially equals 1 and is increasing by assumption A.3
Main Result

Proposition 2. The following contract is optimal:

- The agent is asked to work throughout $[0, t^*]$ without any feedback.
- From $t^*$ onwards, advised to quit as soon as the signal arrives.
- Reward $\overline{R}$ if signal arrives before $t^*$ and $\underline{R} < \overline{R}$ if it arrives in $(t^*, T^*)$.
- In equilibrium, the agent follows all recommendations.

Optimal contract *frontloads ignorance*:

- By being silent, principal can elicit more effort, but must pay more.
- Due to backward compounding, if rewards are raised at some $t$, they must also be raised at all earlier dates.
- Bringing the silent period forward minimizes backward compounding.
Main Result

Proposition 2. The following contract is optimal:

- The agent is asked to work throughout \([0, t^*]\) without any feedback.
- From \(t^*\) onwards, advised to quit as soon as the signal arrives.
- Reward \(\bar{R}\) if signal arrives before \(t^*\) and \(\bar{R} < \bar{R}\) if it arrives in \((t^*, T^*)\).
- In equilibrium, the agent follows all recommendations.

Optimal contract *frontloads ignorance*:

- By being silent, principal can elicit more effort, but must pay more.
- Due to backward compounding, if rewards are raised at some \(t\), they must also be raised at all earlier dates.
- Bringing the silent period forward minimizes backward compounding.
Proof Sketch: Relaxed Principal’s Problem

Consider maximizing total effort subject only to $p(\cdot)$ being feasible:

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

s.t. $p(t) \in [1 - F(t), 1]$ 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.
Proof Sketch: Relaxed Principal’s Problem

Consider maximizing total effort subject only to \( p(\cdot) \) being feasible:

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

s.t. \( p(t) \in [1 - F(t), 1] \) 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.
Proof Sketch: Solving for the optimal $p(\cdot)$ given a fixed $T$

Fix a $T > 0$ and solve

$$\max_{p(\cdot)} \int_0^T p(t)[1 - c - c\Phi(t)] dt + cT$$

s.t. $p(t) \in [1 - F(t), 1]$ for all $t$ and non-increasing.

- The term in brackets crosses 0 at most once from above ($\Phi(t) \uparrow$)
- Define $t^*$ to be the smallest $t$ such that $1 - c - c\Phi(t) \leq 0$

The following non-increasing function solves the above program:

$$p(t) = \begin{cases} 
1 & \text{if } t \leq \min\{t^*, T\} \\
1 - F(t) & \text{otherwise.}
\end{cases}$$
Proof Sketch: Solving for the optimal $p(\cdot)$ given a fixed $T$

Fix a $T > 0$ and solve

$$\max_{p(\cdot)} \int_0^T p(t)[1 - c - c\Phi(t)]dt + cT$$

s.t. $p(t) \in [1 - F(t), 1]$ for all $t$ and non-increasing.

- The term in brackets crosses 0 at most once from above ($\Phi(t) \uparrow$)
- Define $t^*$ to be the smallest $t$ such that $1 - c - c\Phi(t) \leq 0$

The following non-increasing function solves the above program:

$$p(t) = \begin{cases} 
1 & \text{if } t \leq \min\{t^*, T\} \\
1 - F(t) & \text{otherwise.}
\end{cases}$$
The optimal $p(\cdot)$ is implemented by the following recommendation policy

$$q(s|t) = \begin{cases} 
1 & \text{for all } t \leq s \leq t^* \\
0 & \text{otherwise.}
\end{cases}$$

- Principal advises the agent to work for all $t < t^*$ \textit{no matter what}
- After $t^*$, principal asks agent to work \textit{only if} signal hasn't arrived yet
Proof Sketch: Optimal Terminal Date $T^*$

Substituting $p(\cdot)$ into the objective we have

$$\max_T \int_0^{t^*} [1 - c - c\Phi(t)]dt + \int_{t^*}^{T \lor t^*} [1 - c - c\Phi(T)][1 - F(t)]dt + cT$$

Objective need not single-peaked. Let $T^*$ denote optimal terminal date

Because the first integrand is positive, $T^* > t^*$ (there is always a phase 2)

- With constant hazard rate, $T^* = \infty$
- With “Poisson good news” experimentation, in general, $T^* < \infty$
Proof Sketch: Optimal Terminal Date $T^*$

Substituting $p(\cdot)$ into the objective we have

$$\max_T \int_0^{t^*} [1 - c - c\Phi(t)]dt + \int_{t^*}^{T^\vee t^*} [1 - c - c\Phi(T)][1 - F(t)]dt + cT$$

Objective need not single-peaked. Let $T^*$ denote optimal terminal date

Because the first integrand is positive, $T^* > t^*$ (there is always a phase 2)

- With constant hazard rate, $T^* = \infty$
- With “Poisson good news” experimentation, in general, $T^* < \infty$
Proof Sketch: Optimal Reward Schedule

Per Proposition 1, the *minimal reward schedule* is

\[
R^*(t) = \begin{cases} 
\frac{c}{\lambda(T^*)} + \frac{cF(t^*)}{f(t^*)} & \text{if } t \leq t^* \\
\frac{c}{\lambda(T^*)} & \text{if } t^* < t \leq T^*
\end{cases}
\]

- Agent is paid a time-invariant reward if signal arrives before \( t^* \); and a smaller, time-invariant reward again if it arrives in \((t^*, T^*)\).
Proof Sketch: Global Deviations

During \((t^*, T^*)\):

- The agent earns rents that decrease in time (due to \(\lambda(t)\) declining), and is advised 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 \(F(\cdot)\).
Proof Sketch: Global Deviations

**During \((t^*, T^*)\):**

- The agent earns rents that decrease in time (due to \(\lambda(t)\) declining), and is advised 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 \(F(\cdot)\).
Discussion

Dynamic agency model with joint design of monetary rewards & feedback

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

\( \varphi_1 \). No feedback with “success” accompanied by a relatively large reward

\( \varphi_2 \). 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)