Robust Risk Sharing Contracts with Costly Signaling

Laurent Lamy & **Clément Leblanc**

Ecole National des Ponts et Chaussées – CIRED (Paris)
Universidad Carlos III de Madrid – EnergyEcoLab (Madrid)

June 2nd, 2025

Motivation

Delegating a **risky project** where only the **outcome is contractable**, a **risk-neutral principal** faces a trade-off between:

- Incentivizing the agent to take actions that induce the most favorable outcome $\rightarrow$ **Moral Hazard**
- Taking the risk for herself to reduce the risk premium of the possibly risk-averse agent $\rightarrow$ **Risk Sharing**

Holmström (1979)

Canonical framework for the **agency under moral hazard problem**

1. The principal offers a **contract** w depending on the agent's **ex-post performance** q
2. The agent chooses a **costly action** a that affects the distribution of q
3. q is realized and the **principal pays** $w(q)$ to the agent

Principal's problem:

$$\max_{w(\cdot)} \int (q - w(q))dF(q \mid a^*)$$

$$\text{s.t.} \quad a^* \in \operatorname*{Arg\,max}_{a \geq 0} \int U(w(q))dF(q \mid a) - C(a) \quad \text{(IC)}$$

$$\int U(w(q))dF(q \mid a^*) - C(a^*) \geq 0 \quad \text{(IR)}$$

Limits of the existing literature

Existing literature relies on

- Structure on the **space of actions**
  - e.g., one dimensional with $a \in \mathbb{R}$
- **Common knowledge** assumptions
  - about **the distribution** $F(\cdot \mid a)$
  - about **the agent's risk-preferences** U

Predicts sophisticated contracts **not observed in practice**

This paper – Robust Approach

- No **parametric restrictions**
  - Large set of potential actions and risk distributions associated *(all distributions on $\mathbb{R}_+$)*
  - Large set of potential risk preferences *(all increasing concave functions)*
- No **prior on the characteristics** of the agent *(no distribution of "types")*
- **Criterion** = **No-regret** (*Pareto-improvement*) compared to a *marginal reward* (MR) contract *(→ optimal under a risk-neutrality)*
- Better contracts are allowed by some **costly signaling** by the agent

Related Literature

- **Contract Design** – Laffont, Martimort (2002), Armstrong, Sappington (2007), Georgiadis (2022)
  - **Under Risk-Neutrality** – Rogerson (1992, RES), Hatfield, Kojima, Kominers (2015, wp)
  - **Under Risk-Aversion** – Weitzman (1980, QJE), Mirlees (1999, RES), McAfee, MacMillan (1986, RAND), Holmstrom, Milgrom (1987, Ecma), Engel, Fischer, Galetovic (2001, JPE)
- **Robust Contract Design** – Bergemann, Schlag (2008, JEEA), Chassang (2013, Ecma), Carroll (2015, AER), Walton, Carroll (2019, wp)
- **Lying/Misreporting Costs** – Lacker, Weinberg (1989, JPE); Maggi, Rodriguez-Clare (1995, RAND); Kartik (2009, RES); Martimort, Poudou, Sand-Zantmann (2010, JIE)

Model

Model

A **risk-neutral principal** delegates a risky project to an **agent** through a contract w, where the payment $w(q, s)$ depends on:

- A **production** $q \in \mathbb{R}_+$ realized ex-post following a distribution F chosen ex-ante by the agent
- A **signal** $s \in \mathbb{R}_+^*$ sent ex-ante by the agent, said **truthful** if

$$s = \sigma(F) \equiv q_F := \mathbb{E}_F[q]$$

(σ denotes the expected value function)

The agent I

The agent is characterized by

- Any **set of potential actions** $\mathcal{A} \subseteq \mathcal{F} \times \mathbb{R}_+^*$
  - including a **project** $F \in \mathcal{F} \equiv$ the set of **all probability distributions** on a compact subset of $\mathbb{R}_+$ (excluding the Dirac in zero)
  - and a **signal** $s \in \sigma(\mathcal{F}) = \mathbb{R}_+^*$
  - a **truthful signal is always available**: $\forall (F, s) \in \mathcal{A}, (F, \sigma(F)) \in \mathcal{A}$
- Any **cost function** $C : \mathcal{A} \to \mathbb{R}_+$
  - such that sending **a truthful signal is always a least costly option**: $\forall (F, s) \in \mathcal{A}, C(F, s) \geq C(F, \sigma(F))$
  - where the **cost of misreporting admits a lower bound** μ *(next slide)*
- Any **utility function** $U : \mathbb{R} \to \mathbb{R}$
  - $U \in \bar{\mathcal{U}} \equiv$ the set of all increasing and weakly concave utility functions

The agent II

$\mu : (\mathbb{R}_+^*)^2 \to \mathbb{R}_+$ denotes a **minimum lying cost** function

- $\mu(s, \tilde{q})$ is a lower bound on the additional cost $C(F, s) - C(F, \tilde{q})$ for the agent to **report** s **instead of** $\tilde{q}$, where $\tilde{q}$ is the truthful signal
- $\forall s \in \mathbb{R}_+^*, \mu(s, s) = 0$
- $\forall s, \tilde{q} \in \mathbb{R}_+^*, \mu(s, \tilde{q}) \geq 0$

μ defines the agent's **potential cost functions** $C \in \mathcal{C}(\mu)$ where

$$\mathcal{C}(\mu) \equiv \{C : \mathcal{F} \times \mathbb{R}_+^* \to \mathbb{R}_+ \mid \forall (F, s), C(F, s) - C(F, \sigma(F)) \geq \mu(s, \sigma(F))\}$$

Ω denotes the **set of potential agents** $(\mathcal{A}, C, U) \in \Omega$

Timing of the game

The principal and the agent **agree on a contract** $w : \mathbb{R}_+ \times \mathbb{R}_+^* \to \mathbb{R}$, before the following steps:

1. The agent **learns his characteristics** $(\mathcal{A}, C, U) \in \Omega$ and decides whether to opt out with payoff zero
2. The agent **chooses an action** $(F^*, s^*) \in \mathcal{A}$ at cost $C(F^*, s^*)$ with

$$(F^*, s^*) \in \operatorname*{Arg\,max}_{(F,s)\in\mathcal{A}} \mathbb{E}_F[U(w(q,s) - C(F,s))]$$

3. The **production** q **is drawn** from distribution F^* and the principal pays $w(q, s^*)$ to the agent

Risk-neutrality and Marginal Rewards Contracts (Reminder)

- A contract w_0 is said to provide **marginal reward** (MR contract) if

$$\forall q \in \mathbb{R}_+, w_0(q) = q + w_0(0)$$

- When the agent is known to be **risk-neutral**, a contract **is Pareto-optimal for any characteristics of the agent iff it provides marginal reward**

- When the agent can be **risk-averse**, **no contract is Pareto-optimal for any characteristics of the agent**

$\Rightarrow$ *We take a MR contract as benchmark that we want to improve upon*

Criterion = Robust Pareto-dominance

Definition: Robust Pareto-dominance

- $w \succ_\Omega w'$: A contract w strictly robustly Pareto-dominates another contract w' over a set of potential agents Ω if $w \succeq_\Omega w'$ and the Pareto-dominance is strict for some agent $(\mathcal{A}, C, U) \in \Omega$

Impossibility Results:

- A **simple contract** $w(q)$ *(that does not depend on the signal)* cannot strictly robustly Pareto-dominates a MR contract
- If the agent might be able of **cheap talk** $(\forall s, \tilde{q}, \mu(s, \tilde{q}) = 0)$, no contract strictly robustly Pareto-dominates a MR contract

Characterizing robustly Pareto-dominant contracts

Characterizing robustly Pareto-dominant contracts

Outline

1. Example 1: $\mu(s, q_F) = \alpha|s - q_F|$ with $\alpha \in \mathbb{R}_+$
2. Example 2: $\mu(s, q_F) = c \cdot \mathbf{1}\{s \neq q_F\}$ with $c \in \mathbb{R}_+$
3. General characterization for any $\mu : (\mathbb{R}_+^*)^2 \to \mathbb{R}_+$

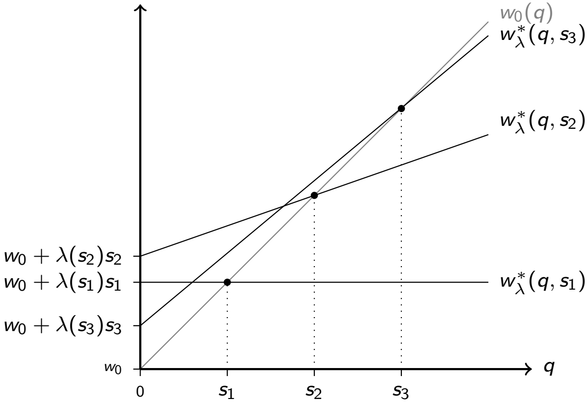
Necessary Condition: Linear contracts

For any $w \succeq_\Omega w_0$, w_0 a marginal reward contract, then w is such that $\forall q \in \mathbb{R}_+, s \in \mathbb{R}_+^*$

$$w(q,s) = \lambda(s)s + (1-\lambda(s))q + w_0 \quad \text{with} \quad \lambda(s) \in [0,1]$$

- Only form that **preserves the expected payment** $w_0(q_F)$ for any distribution $F \in \mathcal{F}$, provided that the signal is truthful ($s = \sigma(F)$)
- $\lambda(s)$ is a **share of the risk that is transferred** to the principal

Necessary Condition: Linear contracts



Necessary Condition: Deter overstatement

- Expected payment by the principal to the agent:

$$\mathbb{E}_F[w(q,s)] = w_0 + q_F + \lambda(s)(s - q_F)$$

$\Rightarrow$ If the agent overstates in equilibrium, switching from w_0 to w is **detrimental to the principal**

$$s > q_F \Rightarrow \mathbb{E}_F[w(q,s)] > w_0 + q_F > \mathbb{E}_F[w_0(q)]$$

$\Rightarrow$ w **must deter any agent** from overstating his expected production

Example 1: $\mu(s, q_F) = \alpha|s - q_F|$ with $\alpha \in \mathbb{R}_+$

Restrict attention to contracts $w_\lambda(q, s) = \lambda \cdot s + (1 - \lambda) \cdot q + w_0$

When deciding whether to overstate $s > q_F$, the agent compares

- His benefit from misreporting: $w_\lambda(q, s) - w_\lambda(q, q_F) = \lambda(s - q_F)$
- His cost of misreporting: $C(F, s) - C(F, q_F) \geq \alpha(s - q_F)$

$\Rightarrow$ Any contract w_λ with $\lambda \in]0, \alpha]$

- Avoids any overstatement by the agent $\rightarrow$ Makes the **principal weakly better-off**
- Lower the risk for the agent with the same expected payment $\rightarrow$ Makes a **risk-averse agent strictly better off**

Example 1: $\mu(s, q_F) = \alpha|s - q_F|$ with $\alpha \in \mathbb{R}_+$

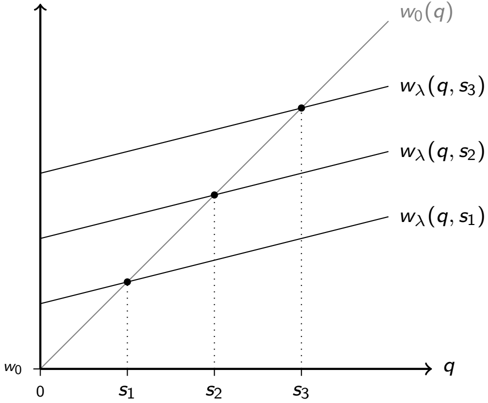


Figure: Example – Contract w_λ for $\lambda = 0.75$

Example 2: $\mu(s, q_F) = c \cdot \mathbf{1}\{s \neq q_F\}$ with $c \in \mathbb{R}_+$

If we keep the restriction to contracts with a fixed λ

- No agent decides to overstate $s > q_F$ iff

$$w_\lambda(q, s) - w_\lambda(q, q_F) = \boxed{\lambda(s - q_F) \leq c} \leq C(F, s) - C(F, q_F)$$

- Worst (limit) case: $q_F = 0$ and $s \to +\infty$

$$\lambda \leq \frac{c}{s - q_F} \leq \frac{c}{s} \xrightarrow{s\to\infty} 0$$

But we can relax the fixed λ restriction and **have a variable** $\lambda(s)$

$$w(q, s) = \lambda(s)s + (1 - \lambda(s))q + w_0 \quad \text{with} \quad \lambda(s) = c/s$$

Example 2: $\mu(s, q_F) = c \cdot \mathbf{1}\{s \neq q_F\}$ with $c \in \mathbb{R}_+$

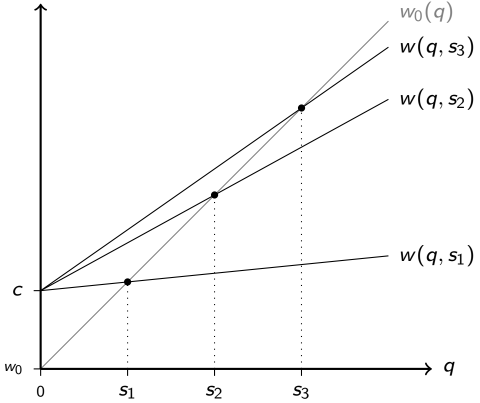


Figure: Example – Contract w with $\lambda(s) = c/s$

General Characterization of RPD contracts

For w_0 a MR contract, $w \succeq_\Omega w_0$ if and only if

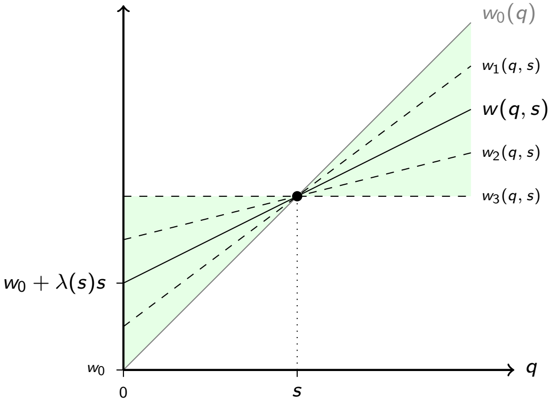
$$\forall q \in \mathbb{R}_+, s \in \mathbb{R}_+^*, \quad w(q,s) = \lambda(s) \cdot s + (1 - \lambda(s)) \cdot q + w_0$$

with

- A $\forall s \in \mathbb{R}_+^*$, $\lambda(s) \in [0,1]$;
- B $\forall \tilde{q}, s \in \mathbb{R}_+^*$ with $s > \tilde{q}$, $\quad w(\tilde{q}, s) - w_0(\tilde{q}) \leq \mu(s, \tilde{q})$;
- C $\forall \tilde{q}, s \in \mathbb{R}_+^*$ with $s > \tilde{q}$, $\quad w(0, s) - w(0, \tilde{q}) \leq \mu(s, \tilde{q})$.

General Characterization – Condition A

$$\forall s \in \mathbb{R}_+^*, \ \lambda(s) \in [0, 1]$$



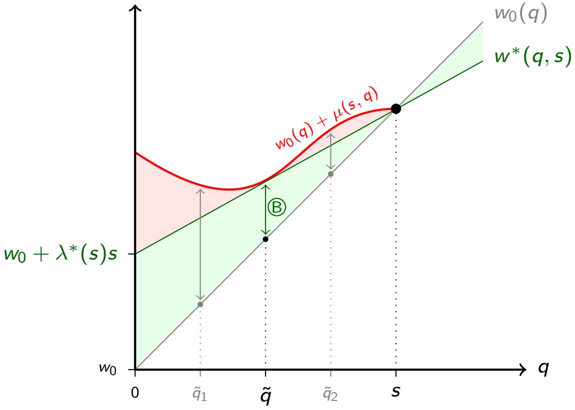
General Characterization – Condition B

$$\forall \tilde{q}, s \in \mathbb{R}_+^* \text{ with } s > \tilde{q},$$

$$w(\tilde{q}, s) - w_0(\tilde{q}) \leq \mu(s, \tilde{q})$$

i.e.,

$$\forall s, \lambda(s) \leq \inf_{\tilde{q}<s} \frac{\mu(s, \tilde{q})}{s - \tilde{q}}$$

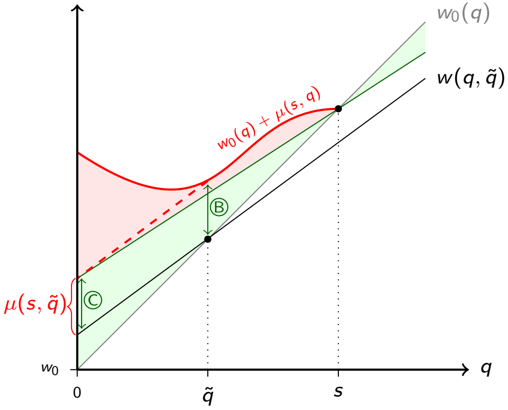


General Characterization – Condition C (binding)

$$\forall \tilde{q}, s \in \mathbb{R}_+^* \text{ with } s > \tilde{q},$$

$$w(0,s) - w(0,\tilde{q}) \leq \mu(s,\tilde{q})$$

$$\Leftrightarrow \lambda(s)s - \lambda(\tilde{q})\tilde{q} \leq \mu(s,\tilde{q})$$

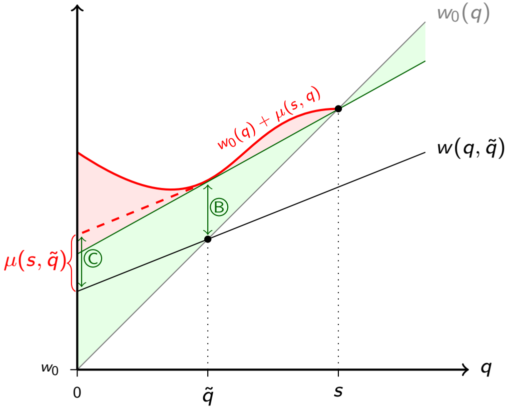


General Characterization – Condition C (not binding)

$\forall \tilde{q}, s \in \mathbb{R}_+^*$ with $s > \tilde{q}$,

$$w(0, s) - w(0, \tilde{q}) \leq \mu(s, \tilde{q})$$

$$\Leftrightarrow \lambda(s)s - \lambda(\tilde{q})\tilde{q} \leq \mu(s, \tilde{q})$$



Are these contracts truthful?

- A **truthful contract** always induces a truthful signal in equilibrium
- **RPD contracts are not all truthful**: some agents may understate their expected production
  - The agent would do so to mitigate risk (get a "flatter" payment)
  - The principal would be better off as it reduces the expected payment
- Contract w **is truthful if a stricter version of Condition** (C) is met

(C) $\forall \tilde{q}, s \in \mathbb{R}_+^*$ ~~with $s > \tilde{q}$~~, $\quad w(0,s) - w(0,\tilde{q}) \leq \mu(s,\tilde{q})$

Ranking RPD contracts

The set of contracts that RPD w_0 over Ω

- Is only **partially ordered** by $\succ_\Omega$
- The partial order is straightforward for **truthful contracts** (only)

> For w_1, w_2 two truthful contracts with $w_1 \succeq_\Omega w_0$, $w_2 \succeq_\Omega w_0$, and λ_1, λ_2 the corresponding defining functions,
>
> - $w_1 \succeq_\Omega w_2$ if and only if $\forall s \in \mathbb{R}_+^*$, $\lambda_1(s) \geq \lambda_2(s)$,
> - $w_1 \succ_\Omega w_2$ if in addition $\exists s \in \mathbb{R}_+^*$ with $\lambda_1(s) > \lambda_2(s)$.

- Implies that the poset of truthful contracts that RPD w_0 is a **lattice**, and thus contains a **unique robustly-undominated contract** w^*

Identify w^* while ignoring (C)

Under some conditions, w^* can be identified

- by saturating Condition (B)
- while ignoring Condition (C)

If the minimum lying cost function μ is such that:

- $\bar{\lambda}(s) \equiv \inf_{q \in [0,s)} \frac{\mu(s,q)}{s-q}$ is weakly decreasing in s,
- but $\bar{\lambda}(s)s$ is weakly increasing in s,

then the contract w^* defined by $\lambda^*(s) = \min\{\bar{\lambda}(s), 1\}$ is the unique robustly undominated contract among $RPD_{\Omega}^{T}(w_0)$.

Lying cost depending on the magnitude of the lie

In particular, these conditions are met by **minimum lying cost functions** in the form

$$\mu(s, \tilde{q}) = d(s - \tilde{q}) \quad \text{or} \quad \mu(s, \tilde{q}) = d\left(\frac{s - \tilde{q}}{s}\right)$$

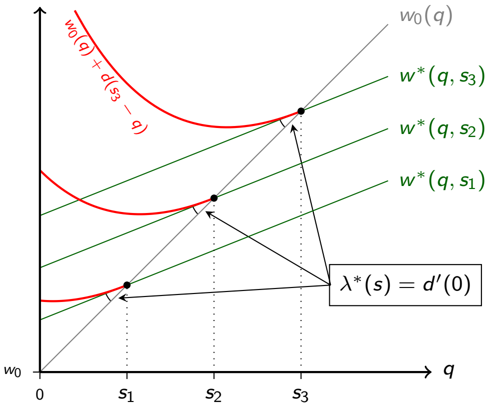
where $d : \mathbb{R} \to \mathbb{R}$ is

- weakly increasing on $\mathbb{R}_+$
- weakly decreasing on $\mathbb{R}_-$

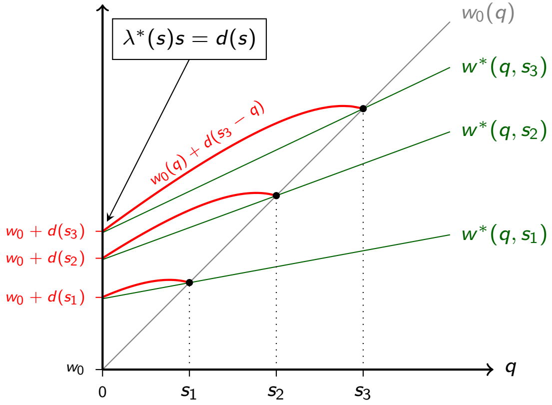
Leads to a **straightforward solution** when:

- d is **convex**: $\lambda^*(s) = \min\{d'(0), 1\}$
- d is **concave**: $\lambda^*(s) = \min\{d(s)/s, 1\}$ or $\lambda^*(s) = \min\{d(1)/s, 1\}$

Convex lying cost



Concave lying cost



Takeaways

Risk-sharing contracts can be designed

- While relying on **no common knowledge assumption** about the agent's technology or risk preferences
- At **no loss in comparison to the marginal reward contract**, that is Pareto-optimal in the risk neutral case
- Insofar as **partially verifiable information** on the effort provided by the agent can be collected

$\Rightarrow$ We characterize **robustly optimal contracts** depending on how costly it is for the agent to misrepresent his action to the principal

$\Rightarrow$ Robust risk-sharing motivates the use of **linear contracts** where some share of the risk is taken by the principal

Thank you for your attention.

Contact: cl.clement.leblanc@gmail.com