Robust Contracts: A Revealed Preference Approach

Nemanja Antic and George Georgiadis

Northwestern Kellogg

Introduction

Virtually every firm has an incentive program for at least some employees

Proper design of incentive programs is crucial

- In the *management practices* literature high scores in the "incentives" category correlate with better performance (Bloom et al., 2007)
- At Safelite Autoglass, switching from hourly wages to piece rates led to a 44% productivity increase (Lazear, 2000)
- Many other success stories—see the references in Lazear (2018)
- But poorly designed incentives can have dire consequences

  e.g., multitask problems incl. gaming (Jensen, 2002), excessive risk-taking (Rajan, 2011), fraud (Wells Fargo & VW scandals), etc

Motivation

- Imagine designing an incentive plan for a dealership's salespeople
- To simplify matters, focus on the pay-for-performance relationship
- One approach is to adopt industry best practices (Zoltners et al, '06)
- Another option is to take guidance from contract theory.
- Std. models are difficult to operationalize as they impose implausible assumptions about principal's knowledge of production environment
- Misra and Nair (2011) offer (and implement) the following approach:
  i. Make parametric assumptions about the production environment
  ii. Exploit variation in offered incentives to recover unknown parameters
  iii. Find an optimal contract

Motivation

- Imagine designing an incentive plan for a dealership's salespeople
- To simplify matters, focus on the pay-for-performance relationship
- One approach is to adopt industry best practices (Zoltners et al, '06)
- Another option is to take guidance from contract theory.
- Std. models are difficult to operationalize as they impose implausible assumptions about principal's knowledge of production environment
- Misra and Nair (2011) offer (and implement) the following approach:
  - i. Make parametric assumptions about the production environment
  - ii. Exploit variation in offered incentives to recover unknown parameters
  - iii. Find an optimal contract

Motivation

- Imagine designing an incentive plan for a dealership's salespeople
- To simplify matters, focus on the pay-for-performance relationship
- One approach is to adopt industry best practices (Zoltners et al, '06)
- Another option is to take guidance from contract theory.
- Std. models are difficult to operationalize as they impose implausible assumptions about principal's knowledge of production environment
- Misra and Nair (2011) offer (and implement) the following approach:
  i. Make parametric assumptions about the production environment
  ii. Exploit variation in offered incentives to recover unknown parameters
  iii. Find an optimal contract

A Simple Approach ("A/B Contracts", Georgiadis & Powell)

- If you restrict to linear contracts, you can write expected profit as

$$\Pi = (m - \alpha)a$$

where m is profit margin, α is piece rate, and a are mean sales (effort)

$$\Rightarrow \text{First-order condition: } \epsilon(\alpha^*) =: \frac{\alpha^*}{Q}\frac{dQ}{d\alpha} = \frac{\alpha^*}{m - \alpha^*}$$

- Given an A/B test of contracts, you can estimate ϵ and compute α^*
- Angrist et al. (2021) estimate at UBER, $\epsilon \simeq 1.2$ implying $\alpha^* \simeq 0.55m$

This implicitly assumes that (1) the agent's action is one-dimensional, (2) elasticity is constant, and (3) the agent has specific preferences over money

This paper.

What if the principal is reluctant to make such (strong) assumptions?

A Simple Approach ("A/B Contracts", Georgiadis & Powell)

- If you restrict to linear contracts, you can write expected profit as

$$\Pi = (m - \alpha)a$$

where m is profit margin, α is piece rate, and a are mean sales (effort)

$$\Rightarrow \text{First-order condition: } \epsilon(\alpha^*) =: \frac{\alpha^*}{Q}\frac{dQ}{d\alpha} = \frac{\alpha^*}{m - \alpha^*}$$

- Given an A/B test of contracts, you can estimate ϵ and compute α^*
- Angrist et al. (2021) estimate at UBER, $\epsilon \simeq 1.2$ implying $\alpha^* \simeq 0.55m$

This implicitly assumes that (1) the agent's action is one-dimensional, (2) elasticity is constant, and (3) the agent has specific preferences over money

This paper.

What if the principal is reluctant to make such (strong) assumptions?

A Simple Approach ("A/B Contracts", Georgiadis & Powell)

- If you restrict to linear contracts, you can write expected profit as

$$\Pi = (m - \alpha)a$$

where m is profit margin, α is piece rate, and a are mean sales (effort)

$$\Rightarrow \text{First-order condition: } \epsilon(\alpha^*) =: \frac{\alpha^*}{Q}\frac{dQ}{d\alpha} = \frac{\alpha^*}{m - \alpha^*}$$

- Given an A/B test of contracts, you can estimate ϵ and compute α^*
- Angrist et al. (2021) estimate at UBER, $\epsilon \simeq 1.2$ implying $\alpha^* \simeq 0.55m$

This implicitly assumes that (1) the agent's action is one-dimensional, (2) elasticity is constant, and (3) the agent has specific preferences over money

This paper.

What if the principal is reluctant to make such (strong) assumptions?

A Simple Approach (“A/B Contracts”, Georgiadis & Powell)

- If you restrict to linear contracts, you can write expected profit as

$$\Pi = (m - \alpha)a$$

where m is profit margin, α is piece rate, and a are mean sales (effort)

$$\Rightarrow \text{First-order condition: } \epsilon(\alpha^*) =: \frac{\alpha^*}{Q}\frac{dQ}{d\alpha} = \frac{\alpha^*}{m - \alpha^*}$$

- Given an A/B test of contracts, you can estimate ϵ and compute α^*
- Angrist et al. (2021) estimate at UBER, $\epsilon \simeq 1.2$ implying $\alpha^* \simeq 0.55m$

This implicitly assumes that (1) the agent's action is one-dimensional, (2) elasticity is constant, and (3) the agent has specific preferences over money

This paper.

What if the principal is reluctant to make such (strong) assumptions?

Framework in a Nutshell

- We consider a contracting game between a principal and an agent
- The principal is oblivious to the agent's action set *and* their costs
  - An "action" is a probability distribution over outcomes
- The principal however knows the agent's best response to K contracts
  - E.g., has outcome data under each of these K exogenous contracts
- The principal's objective is to maximize her worst-case profit
  - As if *nature* chooses agent's set of actions and their costs to minimize the principal's profit subject to a set of revealed preference constraints

Overview of Takeaways

i. Either the most profitable of the known contracts or a mixture of the known contracts and a linear one is optimal.

ii. Straightforward to implement given outcome data, which we demonstrate using data from an experiment of diff. incentive schemes

iii. Practically the best of known contracts provides max. profit guarantee

* Optimality of linear contracts with maxmin preferences (Carroll, 2015) relies on the principal knowing the costs of some of the agent's actions

Overview of Takeaways

i. Either the most profitable of the known contracts or a mixture of the known contracts and a linear one is optimal.

ii. Straightforward to implement given outcome data, which we demonstrate using data from an experiment of diff. incentive schemes

iii. Practically the best of known contracts provides max. profit guarantee

* Optimality of linear contracts with maxmin preferences (Carroll, 2015) relies on the principal knowing the costs of some of the agent's actions

Overview of Takeaways

i. Either the most profitable of the known contracts or a mixture of the known contracts and a linear one is optimal.

ii. Straightforward to implement given outcome data, which we demonstrate using data from an experiment of diff. incentive schemes

iii. Practically the best of known contracts provides max. profit guarantee

* Optimality of linear contracts with maxmin preferences (Carroll, 2015) relies on the principal knowing the costs of some of the agent's actions

Overview of Takeaways

i. Either the most profitable of the known contracts or a mixture of the known contracts and a linear one is optimal.

ii. Straightforward to implement given outcome data, which we demonstrate using data from an experiment of diff. incentive schemes

iii. Practically the best of known contracts provides max. profit guarantee

* Optimality of linear contracts with maxmin preferences (Carroll, 2015) relies on the principal knowing the costs of some of the agent's actions

Model – The Setting

Timing:

1. A principal offers a "contract" $w : [0, \overline{x}] \to \mathbb{R}_+$ to an agent
2. The agent chooses an "action" $F \in \mathcal{F} \subseteq \Delta([0, \overline{x}])$
3. Output $x \sim F$ is drawn and payoffs are realized

Agent's payoff (for a given action):

$$\int w(x) dF(x) - C(F)$$

Principal's profit (for a given contract and action):

$$\int [mx - w(x)] dF(x),$$

where $m > 0$ is the marginal gross profit

Model – The Setting

Timing:

1. A principal offers a "contract" $w : [0, \overline{x}] \to \mathbb{R}_+$ to an agent
2. The agent chooses an "action" $F \in \mathcal{F} \subseteq \Delta([0, \overline{x}])$
3. Output $x \sim F$ is drawn and payoffs are realized

Agent's payoff (for a given action):

$$\int w(x) dF(x) - C(F)$$

Principal's profit (for a given contract and action):

$$\int [mx - w(x)] dF(x),$$

where $m > 0$ is the marginal gross profit

Model – Principal's Knowledge

The principal does **NOT** know the agent's action set $\mathcal{F}$ or any costs $C(F)$

However, she knows the agent's best response to K "known" contracts; i.e., for each $k \in \{1, \dots, K\}$, she knows that

$$F_k \in \arg\max_{F \in \mathcal{F}} \int w_k(x) dF(x) - C(F)$$

In addition, the agent is known to be able to costlessly produce 0 output

She also does not have a prior over $\mathcal{F}$ or C

Instead, she aims to maximize her *worst-case profit*

This is *as if*, after offering a contract, an adversarial 3^{rd} party chooses $\mathcal{F} \supseteq \{F_0, \dots, F_K\}$ and C subject to the K revealed preference constraints

Model – Principal's Knowledge

The principal does **NOT** know the agent's action set $\mathcal{F}$ or any costs $C(F)$

However, she knows the agent's best response to K "known" contracts; i.e., for each $k \in \{1, \ldots, K\}$, she knows that

$$F_k \in \arg\max_{F \in \mathcal{F}} \int w_k(x) dF(x) - C(F)$$

In addition, the agent is known to be able to costlessly produce 0 output

She also does not have a prior over $\mathcal{F}$ or C

Instead, she aims to maximize her *worst-case profit*

This is *as if*, after offering a contract, an adversarial 3^{rd} party chooses $\mathcal{F} \supseteq \{F_0, \ldots, F_K\}$ and C subject to the K revealed preference constraints

Model – Principal's Knowledge

The principal does **NOT** know the agent's action set $\mathcal{F}$ or any costs $C(F)$

However, she knows the agent's best response to K "known" contracts; i.e., for each $k \in \{1, \ldots, K\}$, she knows that

$$F_k \in \arg\max_{F \in \mathcal{F}} \int w_k(x) dF(x) - C(F)$$

In addition, the agent is known to be able to costlessly produce 0 output

She also does not have a prior over $\mathcal{F}$ or C

Instead, she aims to maximize her *worst-case profit*

This is *as if*, after offering a contract, an adversarial 3^{rd} party chooses $\mathcal{F} \supseteq \{F_0, \ldots, F_K\}$ and C subject to the K revealed preference constraints

Problem Formulation

Given a contract w, the adversarial 3^{rd} party, *nature* solves

$$\Pi(w) := \inf \int [mx - w(x)] dF(x)$$

$$\text{s.t. } F \in \arg\max_{\widetilde{F} \in \mathcal{F}} \int w(x) d\widetilde{F}(x) - C(\widetilde{F}) \quad \text{(IC)}$$

$$F_k \in \arg\max_{\widetilde{F} \in \mathcal{F}} \int w_k(x) d\widetilde{F}(x) - C(\widetilde{F}) \quad \text{(RP)}$$

$$\mathcal{F} \supseteq \{F, F_0, \ldots, F_K\} \text{ and } C(\widetilde{F}) \geq 0 \ \forall \widetilde{F} \in \mathcal{F} \text{ with } C(F_0) = 0.$$

The principal solves

$$\sup \Pi(w)$$

$$\text{s.t. } w : [0, \overline{x}] \to \mathbb{R}_+$$

Problem Formulation

Given a contract w, the adversarial 3^{rd} party, *nature* solves

$$\Pi(w) := \inf \int [mx - w(x)] dF(x)$$

$$\text{s.t. } F \in \arg\max_{\widetilde{F} \in \mathcal{F}} \int w(x) d\widetilde{F}(x) - C(\widetilde{F}) \qquad \text{(IC)}$$

$$F_k \in \arg\max_{\widetilde{F} \in \mathcal{F}} \int w_k(x) d\widetilde{F}(x) - C(\widetilde{F}) \qquad \text{(RP)}$$

$$\mathcal{F} \supseteq \{F, F_0, \ldots, F_K\} \text{ and } C(\widetilde{F}) \geq 0 \ \forall \widetilde{F} \in \mathcal{F} \text{ with } C(F_0) = 0.$$

The principal solves

$$\sup \Pi(w)$$

$$\text{s.t. } w : [0, \overline{x}] \to \mathbb{R}_+$$

Assumptions

We impose 3 assumptions on the K *known* contract–action pairs:

A.1. At least one contract delivers a strictly positive profit. Without loss, we adopt the convention that w_1 delivers the largest profit.

A.2. Each contract has $w_k(0) = 0$ and the makes smallest payment at $x = 0$

- If $w_k(0) > 0$, the contract can trivially be improved by a downward shift

A.3. The agent's best responses can be rationalized; *i.e.*, for all i and j

$$\int w_i(x)dF_i(x) + \int w_j(x)dF_j(x) \geq \int w_i(x)dF_j(x) + \int w_j(x)dF_i(x)$$

Otherwise, the revealed preference constraints cannot be met

Simplifying the Problem

Consider the following simpler max-min problem:

$$\sup_{w_{K+1}} \inf_{F_{K+1},\mathbf{c}} \int [mx - w_{K+1}(x)] dF_{K+1}(x) \qquad \text{(P)}$$

$$\text{s.t. } \int w_i(x) dF_i(x) - c_i \geq \int w_i(x) dF_j(x) - c_j \text{ for all } i, j$$

$$w_{K+1}(\cdot) \geq 0 \ , \ F_{K+1} \in \Delta([0, \overline{x}]), \text{ and } \mathbf{c} \in \mathbb{R}_+^{K+1}$$

where $c_i := C(F_i)$ and $\mathbf{c} = \{c_1, \dots, c_{K+1}\}$.

Here nature, instead of $\mathcal{F}$, chooses *one* action and the corresponding costs

Lemma 1.

A contract w_{K+1} solves the principal's problem if and only if it solves (P)

Adding extra actions increases the #(RP) constraints to principal's benefit

Simplifying the Problem

Consider the following simpler max-min problem:

$$\sup_{w_{K+1}} \inf_{F_{K+1},\mathbf{c}} \int [mx - w_{K+1}(x)] dF_{K+1}(x) \tag{P}$$

$$\text{s.t.} \int w_i(x) dF_i(x) - c_i \geq \int w_i(x) dF_j(x) - c_j \text{ for all } i, j$$

$$w_{K+1}(\cdot) \geq 0 \text{ , } F_{K+1} \in \Delta([0, \overline{x}]), \text{ and } \mathbf{c} \in \mathbb{R}_+^{K+1}$$

where $c_i := C(F_i)$ and $\mathbf{c} = \{c_1, \ldots, c_{K+1}\}$.

Here nature, instead of $\mathcal{F}$, chooses *one* action and the corresponding costs

Lemma 1.

A contract w_{K+1} solves the principal's problem if and only if it solves (P)

Adding extra actions increases the #(RP) constraints to principal's benefit

Connection to Carroll (AER, 2015)

- Our problem is conceptually similar to Carroll (2015)
- In his model, principal knows some of agent's actions *and their costs*
- In that setting, a linear contract is optimal
- In our model, the principal does not know the cost of *any* action, but the revealed preference constraints bound the rationalizable costs
- The motivation is that practically, costs are **not** observable
- And the upshot is that in general, the optimal contract is **not** linear

Connection to Carroll (AER, 2015)

- Our problem is conceptually similar to Carroll (2015)
- In his model, principal knows some of agent's actions *and their costs*
- In that setting, a linear contract is optimal

- In our model, the principal does not know the cost of *any* action, but the revealed preference constraints bound the rationalizable costs
- The motivation is that practically, costs are **not** observable
- And the upshot is that in general, the optimal contract is **not** linear

One known contract ($K = 1$): A Benchmark

Theorem 1.

With one known contract, w_1 maximizes the principal's worst-case profit

- With only one known contract it's a case of *better the devil you know*, no matter how *irregular* w_1 is; *e.g.*, non-monotone, kinks, etc!

Proof Sketch: Consider any $w \neq w_1$

- If $\int w dF_1 < \int w_1 dF_1$, then nature can induce the agent to choose F_0 by setting $c_1 > \int w dF_1$, delivering a negative profit to the principal
- If $\int w dF_1 > \int w_1 dF_1$, then nature can induce the agent to choose F_1, in which case the principal's profit decreases vis-a-vis w_1
- If $\int w dF_1 = \int w_1 dF_1$ yet $w(\cdot) \neq w_1(\cdot)$, then nature can ensure that the principal obtains a vanishingly small profit

One known contract ($K = 1$): A Benchmark

Theorem 1.

With one known contract, w_1 maximizes the principal's worst-case profit

- With only one known contract it's a case of *better the devil you know*, no matter how *irregular* w_1 is; *e.g.*, non-monotone, kinks, etc!

Proof Sketch: Consider any $w \neq w_1$

- If $\int w dF_1 < \int w_1 dF_1$, then nature can induce the agent to choose F_0 by setting $c_1 > \int w dF_1$, delivering a negative profit to the principal
- If $\int w dF_1 > \int w_1 dF_1$, then nature can induce the agent to choose F_1, in which case the principal's profit decreases vis-a-vis w_1
- If $\int w dF_1 = \int w_1 dF_1$ yet $w(\cdot) \neq w_1(\cdot)$, then nature can ensure that the principal obtains a vanishingly small profit

Two known contracts: Characterization

- For each i and j define

$$v_{ij} := \int w_i(x) dF_j(x) \text{ and } \mu_j := \int x dF_j(x),$$

let $\phi := v_{11} + v_{22} - v_{12} - v_{21}$ (≥ 0 by A.3), and for each j define

$$w_j^*(x) := \rho_j w_j(x) + (1 - \rho_j) mx \text{ where } \rho_j := 1 - \sqrt{\phi / (m\mu_{-j} - v_{j,-j})}$$

- *Note:* w_j^* is a mixture of w_j and the "residual claimant" contract

Theorem 2. Suppose there are two known contracts; i.e., $K = 2$.

i. If $\sqrt{m\mu_2 - v_{12}} - \sqrt{\phi} > \sqrt{m\mu_1 - v_{11}}$, then w_1^* is optimal;

ii. Else if $\sqrt{m\mu_1 - v_{21}} - \sqrt{\phi} > \sqrt{m\mu_1 - v_{11}}$, then w_2^* is optimal;

iii. Otherwise, w_1 is optimal.

Two known contracts: The agent's best response

Corollary 1. Suppose the *new* contract w_j^* is optimal

In response, the agent chooses action $F_j^*(x) = \rho_j F_{-j}(x) + (1 - \rho_j) F_0(x)$.

Moreover, $C(F_j^*) \leq \rho_j C(F_{-j})$

e.g., if w_1^* is optimal, then the agent chooses a mixture of F_2 and F_0

- Note that the agent prefers F_j^* to randomizing between F_{-j} and F_0
- If w_1 is optimal, then the agent best-responds with action F_1

Two known contracts: Some Intuition

Suppose condition (i) of Theorem 2:

Theorem 2. Suppose there are two known contracts; i.e., $K = 2$.

i. If $\underbrace{\sqrt{m\mu_2 - v_{12}} - \sqrt{\phi}}_{=\sqrt{\Pi(w_j^*)}} > \underbrace{\sqrt{m\mu_1 - v_{11}}}_{=\sqrt{\Pi(w_1)}}$, then w_1^* is optimal;

- Here, the principal prefers that the agent chooses F_2 in response to w_1
- Of course this is not IC so she must adjust incentives appropriately
- ϕ relates to the profit she must give up to appropriately adjust incentives (since by RP, the agent prefers F_1 when w_1 is offered)

Corollary 2.

If both known contracts are linear, then w_1 is optimal.

Two known contracts: Some Intuition

Suppose condition (i) of Theorem 2:

Theorem 2. Suppose there are two known contracts; i.e., $K = 2$.

i. If $\underbrace{\sqrt{m\mu_2 - v_{12}} - \sqrt{\phi}}_{=\sqrt{\Pi(w_j^*)}} > \underbrace{\sqrt{m\mu_1 - v_{11}}}_{=\sqrt{\Pi(w_1)}}$, then w_1^* is optimal;

- Here, the principal prefers that the agent chooses F_2 in response to w_1
- Of course this is not IC so she must adjust incentives appropriately
- ϕ relates to the profit she must give up to appropriately adjust incentives (since by RP, the agent prefers F_1 when w_1 is offered)

Corollary 2.

If both known contracts are linear, then w_1 is optimal.

$K \geq 3$ known contracts

A full characterization of optimal contract doesn't appear feasible for $K \geq 3$

Theorem 3.

Every optimal contract takes the form

$$w^*(x) = \sum_{k=1}^{K} \rho_k w_k(x) + \left(1 - \sum_{k=1}^{K} \rho_k\right) mx$$

i.e., it is a mixture of the known contracts and the residual claimant one

- Obtaining the optimal contact entails solving a non-convex program
- As long as K is not too large, this is achievable numerically

Dataset (DellaVigna and Pope, 2018)

- *Goal:* Demonstrate the applicability of our methodology
- Real-effort experiment on M-Turk: Subjects press a-b keys for 10 min
- 7 treatments with different monetary incentives (+participation fee):

	Contract (in ¢)	Avg. #points (x)	N
No incentives	$\pi^1(x) = 0$	1521	540
Piece-rate	$\pi^2(x) = 0.001x$	1883	538
	$\pi^3(x) = 0.01x$	2029	558
	$\pi^4(x) = 0.04x$	2132	566
	$\pi^5(x) = 0.10x$	2175	538
Bonus	$\pi^6(x) = 40\,\mathbb{I}_{\{x\geq 2000\}}$	2136	545
	$\pi^7(x) = 80\,\mathbb{I}_{\{x\geq 2000\}}$	2187	532

- Each subject participated in a single treatment, once.

Empirical Exercise

For each subset $\mathcal{W} \subseteq \{\pi^1, \ldots, \pi^7\}$ take this to be the "known" contracts

i. Letting $K = |\mathcal{W}|$, define the known contracts $w_1, \ldots, w_K$

ii. For each k, use the outcome data to compute the ecdf F_k

*We abstract away from statistical error & unobserved heterogeneity

iii. Compute the optimal contract (given an assumption about m)

Optimal Contract ($K = 1, 2$)

Case 1: $K = 1$ *(7 combinations)*

- Theorem 1: The single known contract provides max. profit guarantee

Case 2: $K = 2$ *(21 combinations)*

- Check which of the conditions in Theorem 2 are met
- We did so for $m \in [0.05, 1]$ on a fine grid
- In every instance, the more profitable of the known contracts delivers the largest profit guarantee

Optimal Contract ($K = 1, 2$)

Case 1: $K = 1$ *(7 combinations)*

- Theorem 1: The single known contract provides max. profit guarantee

Case 2: $K = 2$ *(21 combinations)*

- Check which of the conditions in Theorem 2 are met
- We did so for $m \in [0.05, 1]$ on a fine grid
- In every instance, the more profitable of the known contracts delivers the largest profit guarantee

Optimal Contract ($K \geq 3$)

Case 3: $K \in \{3, \dots, 7\}$ *(99 combinations)*

- Per Theorem 3, it suffices to solve the dual program

$$\max_{\boldsymbol{\lambda} \geq 0} \frac{\sum_{j=1}^{K} \lambda_{K+1,j} \left(m\mu_j + \sum_{k=1}^{K} \lambda_{k,K+1} v_{kj} \right)}{1 + \sum_{j=0}^{K} \lambda_{K+1,j}} + \text{linear terms}(\boldsymbol{\lambda})$$
$$\text{s.t. Linear constraints}(\boldsymbol{\lambda})$$

where λ_{kj} is the dual multiplier corresp. to kj^{th} IC/RP constraint

- We solve this (non-convex) program in two steps:
  - i. Fix $\{\lambda_{K+1,0}, \dots, \lambda_{K+1,K}\}$, solve LP, and denote objective by $\widetilde{\Pi}(\boldsymbol{\lambda}^{K+1})$
  - ii. Maximize w.r.t remaining multipliers (using simulated annealing)
- Again, **most profitable of the known contracts is always optimal!**

Optimal Contract ($K \geq 3$)

Case 3: $K \in \{3, \ldots, 7\}$ *(99 combinations)*

- Per Theorem 3, it suffices to solve the dual program

$$\max_{\boldsymbol{\lambda} \geq 0} \frac{\sum_{j=1}^{K} \lambda_{K+1,j} \left(m\mu_j + \sum_{k=1}^{K} \lambda_{k,K+1} v_{kj} \right)}{1 + \sum_{j=0}^{K} \lambda_{K+1,j}} + \text{linear terms}(\boldsymbol{\lambda})$$
$$\text{s.t. Linear constraints}(\boldsymbol{\lambda})$$

  where λ_{kj} is the dual multiplier corresp. to kj^{th} IC/RP constraint

- We solve this (non-convex) program in two steps:
  - i. Fix $\{\lambda_{K+1,0}, \ldots, \lambda_{K+1,K}\}$, solve LP, and denote objective by $\widetilde{\Pi}(\boldsymbol{\lambda}^{K+1})$
  - ii. Maximize w.r.t remaining multipliers (using simulated annealing)

- Again, **most profitable of the known contracts is always optimal!**

Summary & Takeaways

Agency model with *minimal assumptions* about the prod. environment

- The principal knows only the agent's best response to K contracts, and designs a contract to maximize worst-case profit

What to take away?

i. Theoretically, it is sometimes possible to improve the profit guarantee with a mixture of the known contracts and a linear one

ii. Straightforward to implement given outcome data

iii. Practically, best of known contracts provides max. profit guarantee

- Explains firms' aversion to experimenting with incentive schemes
- Predicts path dependence—firms find smth that works and stick to it

* Opt. of linear contracts relies on principal knowing some actions' costs

Summary & Takeaways

Agency model with *minimal assumptions* about the prod. environment

- The principal knows only the agent's best response to K contracts, and designs a contract to maximize worst-case profit

What to take away?

i. Theoretically, it is sometimes possible to improve the profit guarantee with a mixture of the known contracts and a linear one

ii. Straightforward to implement given outcome data

iii. Practically, best of known contracts provides max. profit guarantee
  - Explains firms' aversion to experimenting with incentive schemes
  - Predicts path dependence—firms find smth that works and stick to it

* Opt. of linear contracts relies on principal knowing some actions' costs

Further Robustness

Our principal still makes some assumptions about the prod. environment:

a. Agent has additively separable preferences with known utility functions

b. Ignores unobserved heterogeneity—in practice, outcome data would be aggregated across several agents, who may choose different actions

Allowing robustness along these dimensions as well would make it only more likely that one of the known contracts is optimal

On the other hand, incorporating estimation error would make it more likely that a new contract provides a larger profit "guarantee"

Further Robustness

Our principal still makes some assumptions about the prod. environment:

a. Agent has additively separable preferences with known utility functions

b. Ignores unobserved heterogeneity—in practice, outcome data would be aggregated across several agents, who may choose different actions

Allowing robustness along these dimensions as well would make it only more likely that one of the known contracts is optimal

On the other hand, incorporating estimation error would make it more likely that a new contract provides a larger profit "guarantee"