

Stochastic mechanisms and quasi-linear preferences*

Christoph Schottmüller and Jan Boone[†]

June 12, 2012

Abstract

Many optimal contracting papers use quasi-linear preferences. To exclude stochastic mechanisms they impose a (sufficient) condition on how the curvature of an agent's objective function varies with type. We show with quasi-linear preferences that an optimal deterministic outcome without bunching implies that stochastic mechanisms are not optimal (without any additional assumptions).

Keywords: stochastic mechanisms, contract theory, quasi-linear preferences

JEL classification: D82, H21

*Financial support from the Dutch National Science Foundation (VICI 453.07.003) is gratefully acknowledged.

[†]Schottmüller: Department of Economics, Tilc and CentER, Tilburg University, email: christophschottmueller@gmail.com. Boone: Department of Economics, University of Tilburg, P.O. Box 90153, 5000 LE, Tilburg, The Netherlands; Tilc, CentER and CEPR, email: j.boone@uvt.nl.

1. Introduction

Since the introduction of adverse selection by Akerlof (1970) and Mirrlees (1971), an extensive literature on this topic has emerged. The overwhelming majority of contributions on this topic analyzes deterministic contracts although it is commonly believed that random contracts can be efficient, see Bolton and Dewatripont (2005, ch. 2.3) and Laffont and Martimort (2002, ch. 2.13) for textbook accounts of this. The idea is that randomness in contracts could be used to relax incentive compatibility constraints. However, random contracts are hardly observed in practice. Hence, there seems to be a mismatch between theory and real world and several more or less plausible explanations have been brought forward to explain it, see for example Arnott and Stiglitz (1988).

This paper shows that the intuition explaining why random contracts dominate deterministic contracts is slightly flawed. This intuition suggests that already in fairly standard models, random contracts improve on deterministic mechanisms. The incorrect intuition originates from an error repeatedly made in the small literature on optimal random contracts. We show that random contracts cannot improve upon the optimal deterministic contract in a standard principal agent framework with quasi-linear preferences. Therefore, the non-existence of random contracts in practice is less surprising and the contributions on optimal deterministic contracts with quasi-linear preferences are more general than previously thought.

The paper is structured as follows. The following section explains the argument in favor of random contracts and the flaw in this reasoning. Section 3 shows that random contracts are never optimal in a two type principal agent model. Section 4 generalizes the result to a continuous type framework.

2. The standard intuition

To fix ideas think of a two type principal agent model. Call the types θ^h and θ^l . Under the optimal contracting scheme the incentive compatibility constraint of one of the types, say of θ^h will bind, *i.e.* θ^h is indifferent between his contract and the contract intended for θ^l . The standard explanation why random contracts can be optimal is along the following lines (see Bolton and Dewatripont (2005, ch. 2.3.2) or Laffont and Martimort (2002, ch.

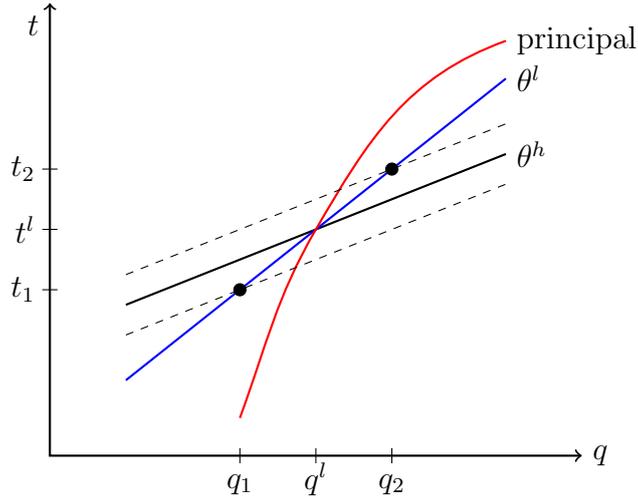


Figure 1: indifference curves through θ^l contract

2.13) for textbook expositions and Arnott and Stiglitz (1988) or Brito et al. (1995) for research papers). Start from the optimal deterministic contract. Now consider giving θ^l a random contract which (for simplicity) randomizes between contracts that are close to his deterministic contract. If θ^h is more risk averse than θ^l (this boils down to an assumption on the third derivative $c_{qq\theta}$ in our framework below), this randomization relaxes the binding incentive compatibility constraint. If the difference in risk aversion is big enough, the relaxation of the incentive constraint is big enough to more than compensate for the loss due to randomization (after all, θ^l is risk averse as well). Therefore, randomization can improve upon deterministic contracts whenever the types differ sufficiently in risk aversion.

This reasoning originates in the literature on optimal income taxation where a similar argument is stated slightly more formal, see for example theorem 2 in Brito et al. (1995). This argument suggests that small changes in a standard set up can make stochastic contracts optimal. Take figure 1 as illustration. Figure 1 depicts the indifference curves of the two types and the principal through the θ^l contract. The contract consists of a transfer t^l and a decision q^l . Now the idea is the following: Instead of the deterministic (q^l, t^l) , assign θ^l a random contract consisting of a random mix between (q_1, t_1) and (q_2, t_2) . As both are on the same indifference curve, the utility of θ^l does not change. Furthermore, one can choose the probability weights such that the expected utility of the principal remains unaffected as well. Therefore, this random contract will improve

on the deterministic one if and only if the incentive constraint is relaxed.

Now consider a concave transformation of the utility of θ^h , *i.e.* the utility function of θ^h is no longer $u^h(q, t)$ but $F(u^h(q, t))$ where F is an increasing and concave function.¹ Note that such a transformation does not change the shape of the indifference curves of type θ^h . It only changes the cardinal utility differences between contracts on different indifference curves. The concavity of F means that θ^h cares less about the utility difference between (q^l, t^l) and (q_2, t_2) compared to the utility difference between (q^l, t^l) and (q_1, t_1) . By choosing F sufficiently concave, the above suggested random contract will relax the incentive constraint for θ^h . To illustrate, one can think of the extreme function $F(u) = u$ if $u < u^h(q^l, t^l)$ and $F(u) = u^h(q^l, t^l)$ if $u \geq u^h(q^l, t^l)$ which guarantees that the incentive constraint is relaxed. By continuity, one can then also choose a strictly increasing F that still relaxes the binding incentive constraint. The result is that for concave enough u^h functions random contracts are always optimal.

What is the flaw in this reasoning? The problem is that as soon as one transforms u^h the optimal deterministic contract changes. The reasoning above only shows that the proposed random contract is better than the initial deterministic contract. However, the optimal deterministic contract under the transformed utility function differs from the initial deterministic contract.

Intuitively, the concave transformation is designed to make θ^h more indifferent between (q^l, t^l) and slightly “higher” contracts, *i.e.* contracts in the direction of (q_2, t_2) . But this implies that the standard distortion for rent extraction reasons becomes less relevant for the principal: the θ^l contract is distorted away from first best to extract more rents from the θ^h type. If this θ^h type cares less about utility differences for slightly “higher” contracts, it is efficient to reduce the distortion. The easiest way to see this is again the extreme F described above: If these are θ^h 's preferences, one can assign θ^l his first best contract without affecting incentive compatibility.

A second story why stochastic mechanisms are optimal is that they help to convexify the domain. We explain why this intuition does not carry through in our setup in the supplementary material to this paper.

¹We depart here from the quasi-linear utility setup we will analyze in the paper. This allows to see the problem with the standard intuition more clearly.

3. Non-optimality in a principal agent framework

Take a principal agent model. The principal has utility $S(q) - t$ of some decision q and transfer t where $S(q)$ is twice differentiable, increasing and concave. The agent has utility $u(q, t, \theta) = t - c(q, \theta)$ where θ is his type which is private information. The cost function is assumed to be twice differentiable, increasing and convex in q with $c(0, \theta) = 0$.² Both principal and agent maximize expected utility.

To simplify things as much as possible, this section assumes that there are only two types θ^h and θ^l . The principal's prior assigns probability f^i to type θ^i with $f^h = 1 - f^l$.

To rule out countervailing incentives it is assumed that $c(q, \theta^h) < c(q, \theta^l)$ for $q > 0$, *i.e.* the high type is better absolutely. In addition, an agent has an outside option normalized to zero, *i.e.* the offered contracts have to satisfy the individual rationality constraint $t^i - c(q^i, \theta^i) \geq 0$. Note that we do not make assumptions on $c_{qq\theta}$ which are generally used to rule out the optimality of stochastic contracts (Laffont and Martimort, 2002, pp. 65/6).

Finally, the incentive compatibility constraint

$$t^i - c(q^i, \theta^i) \geq t^j - c(q^j, \theta^i) \quad \text{with } i, j \in \{h, l\} \text{ and } i \neq j \quad (\text{IC})$$

has to be satisfied. The principal's program is

$$\begin{aligned} \max_{t^i, t^h, q^h, q^l} \quad & f^h [S(q^h) - t^h] + f^l [S(q^l) - t^l] \\ \text{s.t. :} \quad & t^i - c(q^i, \theta^i) \geq 0 \quad \text{with } i \in \{h, l\} \\ & t^i - c(q^i, \theta^i) \geq t^j - c(q^j, \theta^i) \quad \text{with } i, j \in \{h, l\} \text{ and } i \neq j \end{aligned}$$

With the assumptions above the following lemma is standard in the literature. As shown in the proof (in the appendix), this result also holds when we allow for random contracts.

Lemma 1. *Under the optimal contract scheme, the individual rationality constraint of θ^l and the incentive compatibility constraint of θ^h are binding. The individual rationality constraint of θ^h and the incentive compatibility constraint for θ^l are lax. Furthermore, q^h is the first best decision, *i.e.* $q^h = \operatorname{argmax}_q S(q) - c(q, \theta^h)$.*

²The concavity of S and convexity of c is only used to establish existence of a deterministic first best decision and existence of a solution to (1). As long as these two exist, also non-concave S and non-convex c can be allowed.

Using lemma 1, the principal's program can be rewritten as an unconstrained optimization over q^l

$$\max_{q^l} f^l [S(q^l) - c(q^l, \theta^l)] + f^h [S(q^{h*}) - c(q^{h*}, \theta^h) + c(q^l, \theta^h) - c(q^l, \theta^l)] \quad (1)$$

where q^{h*} is the first best decision of type θ^h . Because of the incentive compatibility constraint, this program is not necessarily concave. However, it is easy to see that the objective is approaching $-\infty$ for $q^l \rightarrow \infty$. Using the Weierstrass theorem (or extreme value theorem), the maximum for this problem is well defined. The (deterministic) maximum is achieved either at an interior point or at $q^l = 0$.

Now turn to random decisions. Lemma 1 still applies and also implies that the decision of θ^h is non-random (it is the deterministic first best decision). As utilities are quasilinear in transfers, randomization over t^i is immaterial and only randomization over q^l has to be considered. In the literature, two types of arguments are used to show the optimality of stochastic contracts: local and non-local. We consider each in turn and local randomizations are considered first.³

Proposition 1. *Local randomizations around the optimal deterministic contract of θ^l cannot be optimal.*

Proof. Let \tilde{q}^l denote the deterministic optimum of q^l and consider a random contract where θ^l is assigned $\tilde{q}^l - \varepsilon_1$ with probability p and $\tilde{q}^l + \varepsilon_2$ with probability $1 - p$.⁴ Hence,

³Readers familiar with Arnott and Stiglitz (1988, propositions 11 and 12) or Brito et al. (1995, theorem 3) may be surprised to see that this local argument rules out stochastic contracts. Indeed, these authors use this type of argument to prove the optimality of stochastic contracts. In these papers, the condition that u^h has to be "concave enough" at the optimal deterministic q^l implies a certain convexity of the principal's program at the optimal deterministic contract. Viewing the principal's optimization problem as a function of q^l shows the tension in this argument. Optimality of the best deterministic contract requires that the principal's objective is locally concave at the optimal deterministic q^l . Randomizing with a concave objective function, however, reduces the objective value. Randomization can only work if the objective function is locally convex.

As shown in an example in Brito et al. (1995), the conditions can nevertheless be met. Loosely speaking, this happens when the variable in which the principal's objective is convex is fixed by constraints. Hence, the principal has no real choice and therefore his objective does not have to be locally concave at the optimal deterministic contract. In our standard setting with quasilinear preferences this situation cannot emerge.

⁴The proof holds also for more elaborate randomization schemes as long as all q^l in the support are close enough to \tilde{q}^l , *i.e.* the randomization is local.

the principal's payoff is

$$\begin{aligned}
W &= f^l [pS(\tilde{q}^l - \varepsilon_1) - pc(\tilde{q}^l - \varepsilon_1, \theta^l) + (1-p)S(\tilde{q}^l + \varepsilon_2) - (1-p)c(\tilde{q}^l + \varepsilon_2, \theta^l)] \\
&\quad + f^h [S(q^{h*}) - c(q^{h*}, \theta^h) + pc(\tilde{q}^l - \varepsilon_1, \theta^h) - pc(\tilde{q}^l - \varepsilon_1, \theta^l) \\
&\quad\quad + (1-p)c(\tilde{q}^l + \varepsilon_2, \theta^h) - (1-p)c(\tilde{q}^l + \varepsilon_2, \theta^l)].
\end{aligned}$$

Denote the principal's deterministic objective by $\tilde{W}(q^l)$. Since the focus is on local stochastic contracts around \tilde{q}^l , we have for some $\tilde{q}_1^l \in [\tilde{q}^l - \varepsilon_1, \tilde{q}^l]$ and $\tilde{q}_2^l \in [\tilde{q}^l, \tilde{q}^l + \varepsilon_2]$:

$$W = \tilde{W}(\tilde{q}^l) + \frac{d\tilde{W}(\tilde{q}^l)}{dq^l}(-p\varepsilon_1 + (1-p)\varepsilon_2) + \frac{1}{2} \left(p \frac{d^2\tilde{W}(\tilde{q}_1^l)}{dq^{l2}} \varepsilon_1^2 + (1-p) \frac{d^2\tilde{W}(\tilde{q}_2^l)}{dq^{l2}} \varepsilon_2^2 \right) \leq \tilde{W}(\tilde{q}^l)$$

for $\varepsilon_{1,2} > 0$ close to 0. The inequality follows from the deterministic optimality of \tilde{q}^l : The first order condition requires $d\tilde{W}/dq^l = 0$ and $d^2\tilde{W}/dq^{l2} \leq 0$ at q close to $\tilde{q}^l > 0$. If $\tilde{q}^l = 0$ is a corner solution of the deterministic problem, incentive compatibility is equivalent to the individual rationality constraint of θ^h . Random q^l cannot relax this constraint. \square

After ruling out the optimality of local randomization around the deterministic optimum, one might wonder whether non-local randomization can be optimal.

Proposition 2. *No random contract can improve on the optimal deterministic contract.*

Proof. Suppose that in the optimal contract q^l was random and distributed according to a probability distribution $P(q^l)$. Using lemma 1, the principal's payoff can be written as

$$\begin{aligned}
W &= \int_{q^l} f^l [S(q^l) - c(q^l, \theta^l)] + f^h [S(q^{h*}) - c(q^{h*}, \theta^h) + c(q^l, \theta^h) - c(q^l, \theta^l)] dP(q^l) \\
&= \int_{q^l} \tilde{W}(q^l) dP(q^l) \leq \int_{q^l} \tilde{W}(\tilde{q}^l) dP(q^l) = \tilde{W}(\tilde{q}^l)
\end{aligned}$$

where $\tilde{W}(q^l)$ is the principal's payoff under the deterministic contract q^l . The inequality follows from the definition of \tilde{q}^l as optimal deterministic q^l . \square

The results above are driven by two assumptions: First, the assumption of quasilinear utility and, second, the assumption that the high type has lower costs for all levels of q . Both assumptions are commonplace in the contract theory literature. These assumptions lead in lemma 1 to the result that θ^h 's individual rationality constraint is not binding. Without these two assumptions stochastic contracts can help to implement first best

decisions where both types individual rationality constraints are binding. To illustrate, assume now that θ^h has the following utility function:

$$\hat{u}(q, t, \theta^h) = \log(t^h) - c(q, \theta^h) \quad (2)$$

The utility function of θ^l and the principal are the same as before. The intuition is now that θ^h never wants to misrepresent as θ^l if t^l is 0 with some probability. Therefore, the following first best contract is implementable for θ^l : Assign q^{l*} to θ^l accompanied by transfer $2c(q^{l*}, \theta^l)$ with probability 1/2 and 0 with probability 1/2. If the incentive compatibility constraint of θ^h is binding under the optimal deterministic contract, the stochastic contract improves strictly on the deterministic outcome.

4. Continuum of types

The previous section dealt with a two type model. This section shows that some results carry over to a standard model where the type space is a continuum. The principal still maximizes the expected value of $S(q) - t$ with $q \in \mathbb{R}_+$. An agent of type θ has utility $u(q, t, \theta) = t - c(q, \theta)$ as in section 3. Principal and agent maximize expected utility. The type space, however, is now $\Theta = [\underline{\theta}, \bar{\theta}]$. The principal has a prior distribution $F(\theta)$ with strictly positive density $f(\theta)$ over Θ . The functions S and c are assumed to be twice continuously differentiable. Furthermore, it is assumed that a finite and positive first best decision $q^{fb}(\theta) = \operatorname{argmax}_{q \geq 0} S(q) - c(q, \theta)$ exists.

The outside option is normalized to 0 for all types. To rule out individual rationality constraints binding at interior types, $c_\theta < 0$ is assumed, *i.e.* higher types have lower costs. For simplicity, full participation is assumed; *i.e.* $S(q) - c(q, \underline{\theta})$ is high enough so that the principal does not want to exclude some types.

By the quasi-linearity of the utility functions, randomization over t is again pointless and is therefore not considered in the remainder. \mathbb{E} is used to denote the expectation of a lottery over q . As every type θ could face a different lottery over q , the expectation over the lottery intended for type θ will be noted \mathbb{E}_θ . The lottery over q for type θ will be denoted by the distribution $G(q, \theta)$.

With a slight abuse of notation, denote the agent's expected rent under the imple-

mented mechanism by $u(\theta)$. The envelope theorem yields

$$u_\theta(\theta) = \mathbb{E}_\theta[-c_\theta(q(\theta), \theta)]. \quad (3)$$

By the assumption $c_\theta < 0$, the participation constraint can only bind for the lowest type $\underline{\theta}$. Therefore, the rent function can be written as

$$u(\theta) = \int_{\underline{\theta}}^{\theta} \mathbb{E}_t[-c_\theta(q(t), t)] dt. \quad (4)$$

The principal's program is

$$\max_{G(q, \theta)} \int_{\underline{\theta}}^{\bar{\theta}} \int_q (S(q) - c(q, \theta) - u(\theta)) dG(q, \theta) dF(\theta) \quad (5)$$

$$s.t. : \quad u(\theta) \geq u(\hat{\theta}) + \mathbb{E}_{\hat{\theta}}[c(q, \hat{\theta}) - c(q, \theta)] \quad (\text{IC}') \quad (5)$$

$$u(\underline{\theta}) = 0 \quad (\text{IR}') \quad (5)$$

where $t(\theta) = u(\theta) - \mathbb{E}_\theta c(q, \theta)$ is used to rewrite the objective and the incentive compatibility constraint (IC'). Using (4), the objective can be rewritten as

$$\begin{aligned} & \int_{\underline{\theta}}^{\bar{\theta}} \int_q (S(q) - c(q, \theta)) dG(q, \theta) - \int_{\underline{\theta}}^{\theta} \int_q -c_\theta(q(t), t) dG(q, t) dt dF(\theta) \\ &= \int_{\underline{\theta}}^{\bar{\theta}} \int_q \left(S(q) - c(q, \theta) + \frac{1 - F(\theta)}{f(\theta)} c_\theta(q, \theta) \right) dG(q, \theta) dF(\theta) \end{aligned} \quad (6)$$

where integration by parts is used to get from the first to the second line.

Next the *relaxed deterministic solution* $(q^{rd}(\theta), u^{rd}(\theta))$ is defined: q^{rd} is the solution of the maximization problem

$$\max_q S(q) - c(q, \theta) + \frac{1 - F(\theta)}{f(\theta)} c_\theta(q, \theta) \quad (\text{RD})$$

and u^{rd} is defined by $u^{rd}(\theta) = \int_{\underline{\theta}}^{\theta} -c_\theta(q^{rd}(t), t) dt$.

Proposition 3. *If the relaxed deterministic solution is implementable, it is the solution to (5); i.e. random contracts cannot improve on deterministic contracts.*

Proof. The principal's program is a maximization of (6) subject to (IC') and (IR'). By the definition of q^{rd} ,

$$\begin{aligned} & \int_q \left(S(q) - c(q, \theta) + \frac{1 - F(\theta)}{f(\theta)} c_\theta(q, \theta) \right) dG(q, \theta) \\ & \leq S(q^{rd}(\theta)) - c(q^{rd}(\theta), \theta) + \frac{1 - F(\theta)}{f(\theta)} c_\theta(q^{rd}(\theta), \theta) \end{aligned}$$

for any type θ and any distribution $G(q, \theta)$. Consequently, no distribution can lead to a higher objective (6) than the relaxed deterministic solution. The relaxed deterministic solution satisfies (IR') by construction and (IC') by assumption. Hence, it is the solution of (5). \square

The following corollary states standard conditions under which $q^{rd}(\theta)$ is implementable. As these assumptions are often invoked in principal agent models, stochastic contracts cannot improve the outcome (irrespective of the sign of $c_{qq\theta}$).

Corollary 1. *The following conditions are jointly sufficient for the optimality of deterministic contracts:*

- *single crossing: $c_{q\theta} < 0$*
- *monotone q^{rd} , i.e. $q_{\theta}^{rd}(\theta) \geq 0$ for all $\theta \in [\underline{\theta}, \bar{\theta}]$*

Proof. It is well known that under single crossing non-local incentive constraints are not binding in deterministic solutions, see Bolton and Dewatripont (2005, ch. 2.3) for a textbook exposition. Local first order incentive compatibility is satisfied by construction while monotonicity of q^{rd} ensures second order local incentive compatibility (see again Bolton and Dewatripont (2005, ch. 2.3) for a textbook exposition). \square

Single crossing is a standard assumption in most of the literature. Standard properties that lead to a monotone q^{rd} are the monotone hazard rate assumption on the type distribution and $c_{q\theta\theta} \geq 0$. Most contract theory papers make these assumptions anyway in order to be able to use the first order approach. The corollary then says that whenever the first order approach gives the optimal deterministic contract, no stochastic contract can improve upon the optimal deterministic contract.

5. Conclusion

The dominant view in the literature is that random contracts can improve upon deterministic contracts in a wide class of problems by relaxing incentive compatibility constraints (unless an assumption is made on the sign of $c_{qq\theta}$). This paper shows that this is not true in a principal agent setting that satisfies commonly made regularity conditions like quasi-linear utility, single crossing and type independent participation constraints.

Consequently, economic models are closer to reality—where random contracts are rarely observed—than previously thought.

Appendix

Proof of lemma 1: As this will be useful later on, lemma 1 is proven allowing for random contracts. Since utilities are linear in transfers, randomization over t^i is immaterial and therefore neglected. Assume that the contract has q^i being distributed according to some distribution $P^i(q^i)$.

The individual rationality constraint of θ^h is implied by the incentive compatibility constraint of θ^h , the assumption $c(q, \theta^l) \geq c(q, \theta^h)$ and the individual rationality constraint of θ^l :

$$t^h - \int_{q^h} c(q^h, \theta^h) dP^h(q^h) \geq t^l - \int_{q^l} c(q^l, \theta^h) dP^l(q^l) \geq t^l - \int_{q^l} c(q^l, \theta^l) dP^l(q^l) \geq 0$$

Note that the second inequality holds strictly whenever P^l does not put all probability mass on 0.

Next it is argued that the incentive compatibility constraint (IC^h) of θ^h has to bind. Suppose it did not and for now concentrate on the case where $q^i > 0$ with positive probability for both types. The principal could then increase his payoff by marginally decreasing t^h and this decrease would be feasible: It relaxes the incentive constraint of θ^l . It tightens the incentive and individual rationality constraint of θ^h but those do not bind and are therefore irrelevant for small enough changes of t^h . Hence, the incentive constraint of θ^h has to bind under the optimal contract if $q^i > 0$.

If $q^l = 0$ with probability 1, then the same argument (decreasing t^h) shows that the individual rationality constraint of θ^h has to be binding. As $q^l = 0$ implies $t^l = 0$ ⁵, the individual rationality constraint of θ^h is equivalent to the incentive compatibility constraint of θ^h and therefore IC^h binds as well.

If $q^h = 0$ with probability 1, then either IC^h binds (and that is what we want to show) or $t^h = 0$ (otherwise reducing t^h is feasible and increases the principal's payoff). $t^h = 0$ implies then $q^l = 0$ with probability 1 as well. Otherwise, IC^h is violated because of $c(q, \theta^l) > c(q, \theta^h)$ for $q > 0$. Hence, both types get the same contract, again IC^h binds trivially.

Consequently, IC^h will be binding.

⁵ IC^l cannot be binding in this case as jointly reducing t^l and t^h by the same $\varepsilon > 0$ would improve the principal's payoff without harming any constraint.

Now it is shown that the individual rationality constraint of θ^l has to be binding. If not, decreasing t^l and t^h by the same (small) $\varepsilon > 0$ would increase the principal's payoff without harming any binding constraint.

The next step is to show that IC^l is not binding whenever $q^h > 0$ with positive probability. Since $c(q, \theta^l) > c(q, \theta^h)$ for $q > 0$, a binding IC^l would imply that IC^h is lax. But it was just shown that IC^h has to bind.

Solving the binding constraints for t^h and t^l and plugging them into the principal's objective gives:

$$W = f^l \left[\int_{q^l} S(q^l) - c(q^l, \theta^l) dP^l(q^l) \right] + f^h \left[\int_{q^h} S(q^h) - c(q^h, \theta^h) dP^h(q^h) \right] \\ + f^h \left[\int_{q^l} c(q^l, \theta^h) - c(q^l, \theta^l) dP^l(q^l) \right]$$

With respect to P^h , this term is clearly maximized by choosing $P^h(q^h)$ such that all probability is put on the first best decision q^{h*} .

If S is not concave and/or c not convex, the first best decision might not be unique. Note that the principal can—even in this case—still maximize his payoff with a deterministic q^h (*i.e.* putting all probability mass of P^h on one of the first best decisions). \square

References

- Akerlof, G. (1970). The market for “lemons”: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics* 84(3), 488–500.
- Arnott, R. and J. E. Stiglitz (1988). Randomization with asymmetric information. *The RAND Journal of Economics* 19(3), pp. 344–362.
- Bolton, P. and M. Dewatripont (2005). *Contract theory*. Cambridge: The MIT Press.
- Brito, D., J. Hamilton, S. Slutsky, and J. Stiglitz (1995). Randomization in optimal income tax schedules. *Journal of Public Economics* 56(2), 189–223.
- Laffont, J. and D. Martimort (2002). *The theory of incentives: the principal-agent model*. Princeton University Press.
- Mirrlees, J. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies* 38(2), 175–208.

Supplementary material: Convexifying the domain

Another explanation for the optimality of random contracts which is sometimes mentioned is that they convexify the principal's domain. In this supplementary material we explain why this intuition cannot lead to optimality of random contracts in our quasi-linear setting.

The idea is that—using lemma 1—the principal's problem can be written as an optimization over q^l only. The incentive compatibility constraint of θ^h requires a minimum rent $u^h(q^l)$ that has to be given to type θ^h depending on the choice of q^l . This minimum rent is given by

$$u^h(q^l) = u^l + c(q^l, \theta^l) - c(q^l, \theta^h) = c(q^l, \theta^l) - c(q^l, \theta^h)$$

where the second equality follows from θ^l 's individual rationality constraint. In figure 2, this means that for all (q^l, u^h) combinations above the IC^h -curve incentive compatibility of θ^h is satisfied. As depicted in the figure, this set can be non-convex.⁶ The figure also depicts an indifference curve of the principal which is tangent to the implementable set at the points A and B.

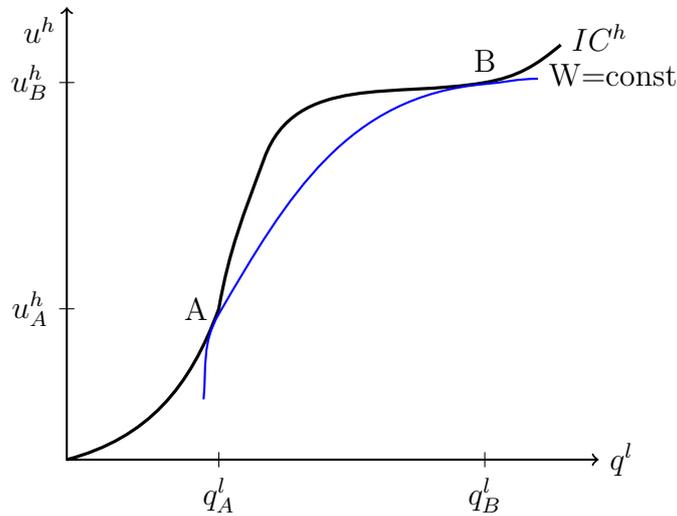


Figure 2: non-convex domain

At first, one could think that random contracts allow convex combinations between the contracts A and B which are strictly preferred by the principal to A and B, *i.e.*

⁶A sufficient condition for this set to be convex is $c_{qq\theta} \leq 0$. However, we argue that random contracts cannot be optimal in our framework without an assumption on the sign of $c_{qq\theta}$.

stochastic mechanisms allow to convexify the domain. However, this is not at all clear; and we know from the analysis in the main text that it is actually not the case for quasi-linear preferences.

As utilities are quasilinear, randomization in t^i is pointless. From lemma 1, we know that the decision of the high type is his first best decision. Hence, the only variable we can randomize is q^l . For illustration, let us check that randomizing between q_A^l and q_B^l does not raise the principal's utility. Take a convex combination putting probability α on q_A^l and $1 - \alpha$ on q_B^l . As IR^l is binding, $t^l = \alpha c(q_A^l, \theta^l) + (1 - \alpha)c(q_B^l, \theta^l)$. We know that the incentive compatibility constraint of θ^h will bind. Hence, the rent of θ^h is $u^h = t^l - \alpha c(q_A^l, \theta^h) - (1 - \alpha)c(q_B^l, \theta^h) = \alpha[c(q_A^l, \theta^l) - c(q_A^l, \theta^h)] + (1 - \alpha)[c(q_B^l, \theta^l) - c(q_B^l, \theta^h)]$. The principal's utility can then be written as

$$\begin{aligned} W &= f^h(S(q^h) - c(q^h, \theta^h) - u^h) + f^l(\alpha S(q_A^l) + (1 - \alpha)S(q_B^l) - \alpha c(q_A^l, \theta^l) - (1 - \alpha)c(q_B^l, \theta^l)) \\ &= \alpha \{f^l(S(q_A^l) - c(q_A^l, \theta^l)) + f^h(S(q^h) - c(q^h, \theta^h) - [c(q_A^l, \theta^l) - c(q_A^l, \theta^h)])\} \\ &+ (1 - \alpha) \{f^l(S(q_B^l) - c(q_B^l, \theta^l)) + f^h(S(q^h) - c(q^h, \theta^h) - [c(q_B^l, \theta^l) - c(q_B^l, \theta^h)])\} \\ &= \alpha W_A + (1 - \alpha)W_B \end{aligned}$$

where W_A (W_B) is the principal's utility of offering the deterministic contract that gives q_A^l (q_B^l) to the θ^l agent. Since both A and B are on the same indifference curve for the principal in figure 2, $W_A = W_B = W$ which means that randomization does neither increase nor decrease the principal's utility. If, however, the deterministic contracts A and B would lead to different utility levels for the principal, the randomization would lead to a worse outcome than the best deterministic contract.

The problem with the convexification intuition is the following. Indeed, the principal prefers a *deterministic* contract with $(q^l, u^h) = (1/2q_A^l + 1/2q_B^l, 1/2u_A^h + 1/2u_B^h)$ over the deterministic contracts A or B but this is not a feasible (incentive compatible) contract. The random version of this contract, however, cannot generate a higher utility for the principal than the deterministic contracts A or B.