

The (No) Value of Commitment

Nathan Hancart

February, 2025

Abstract

I provide a sufficient condition called stable preferred strategies under which a principal does not benefit from committing to a mechanism in economic models represented by a maximisation problem under constraints. These problems include mechanism design, principal-agent models or sender-receiver games. In principal-agent problems, this condition holds if the agent has a finite strategy space and the principal's value function is continuous in the mechanism.

1 Introduction

Commitment plays an important role in many economic models. The general insight of economic theory is that the value of commitment is positive: if a principal has commitment, he can replicate any action he would play without commitment. Moreover, commitment plays a key role in many standard tools used in economic theory such as the revelation principle (Myerson, 1982; Bester and Strausz, 2000; Doval and Skreta, 2022). However, commitment is usually a strong assumption and is sometimes hard to justify. Even if we are willing to assume commitment, knowing that best-replies to the strategy of the agent restricts the set of strategies the modeller has to look at.

To fix ideas, I introduce the setup I will be looking at directly, then will discuss the results.

A principal chooses a mechanism $\alpha \in A$. The principal can also choose a strategy $\sigma \in \Sigma$, subject to the constraint that $\sigma \in C(\alpha) \subseteq \Sigma$. The strategy σ could be the strategy of agents in mechanism design problems or a receiver in a sender-receiver game. Denote by $v(\alpha, \sigma)$ the principal's payoffs.

When the principal commits to α , it gets

$$V(\alpha) = \max_{\sigma} v(\alpha, \sigma) \quad \text{s.t. } \sigma \in C(\alpha). \quad (\mathcal{V}(\alpha))$$

The principal does not benefit from commitment if there is $\alpha^* \in \arg \max_{\alpha} V(\alpha)$ and $\sigma^* \in C(\alpha^*)$ such that $v(\alpha^*, \sigma) \geq v(\alpha, \sigma)$ for all $\alpha \in A$.

In this note, I show that there is no value of commitment when a condition I call stable preferred strategies hold at the optimal mechanism. It holds if for every nearby mechanism, one of the strategies that attains the principal's optimal payoffs is still feasible. In the context of principal-agent problems, where the constraint set $C(\alpha)$ is the best-reply correspondence of the agent, I show that the continuity of the value function $V(\alpha)$ implies the stable preferred strategies property. This result generalises the result from Ben-Porath et al. (2023) who show that there is no value of commitment whenever the principal is indifferent amongst the best-replies of the agent at the optimal mechanism: when this condition is satisfied, the value function is continuous. I also link this result to Lipnowski's (2020) result that also shows that continuity of the value function implies the equivalence between cheap-talk and Bayesian persuasion in sender-receiver games.

In an example, I also show that the simple type dependence assumption on the agent's preferences is key to the result of no value of commitment in games with evidence disclosure introduced in Ben-Porath et al. (2019). In particular, it is not true that in evidence games where the principal's payoff is $v(a, \theta) = \nu(\theta)u(a, \theta) + \psi(a)$ where a is the action taken by the principal, θ the type of the agent and $u(a, \theta)$ his payoffs, the principal does not benefit from commitment.

2 Results

I assume throughout that A is a compact convex subset of \mathbb{R}^n . The principal can also choose a strategy $\sigma \in \Sigma$, also a closed convex subset of \mathbb{R}^n , subject to the constraint that $\sigma \in C(\alpha) \subseteq \Sigma$, where $C(\cdot)$ is a non-empty, closed, convex and upper hemicontinuous correspondence.

The constraint $C(\alpha)$ can be a best-reply correspondence, a set of correlated equilibria or some other constraint set. However, it cannot be a set of Nash Equilibria as these are not necessarily convex. I also

note that σ can incorporate some elements of the mechanism the principal has the power to commit to. The mechanism α can be a strategy of the principal in a mechanism design problem or an information structure in a sender-receiver game.

I assume that $v(\alpha, \sigma)$ is concave in α and jointly continuous.

I denote by $V^* = \max_{\alpha} V(\alpha)$ the maximal value attainable by the principal.

The principal has *stable preferred strategies* at α^* if there is $\epsilon > 0$ such that for all $\alpha \in B_{\epsilon}(\alpha^*)$,

$$\text{there is } \sigma \in C(\alpha) \text{ such that } v(\alpha^*, \sigma) = V^*. \quad (1)$$

In words, the principal has stable preferred strategies at α^* if for every nearby mechanism, one of the strategies that attains the principal's optimal payoffs is still feasible.

Proposition 1. *Let $\alpha^* \in \max V(\alpha)$. If the principal has stable preferred strategies at α^* , there is no value of commitment.*

The proof builds on and generalises the proof of Ben-Porath et al. (2023), Theorem 4.

Proof. Let \tilde{A} a closed ball of A containing α^* where condition (1) holds. The set \tilde{A} is compact and convex.

Let $\Sigma^* = \text{conv}\{\sigma : \sigma \in C(\alpha) \text{ for some } \alpha \in \tilde{A} \text{ and } v(\alpha^*, \sigma) = V^*\}$. By condition (1), Σ^* is non-empty. Because C is upper hemicontinuous and v is continuous in σ , Σ^* is also closed and therefore compact. By definition, $\Sigma^* \cap C(\alpha)$ is non-empty for all $\alpha \in \tilde{A}$ and is also closed and convex. Similarly, $\arg \max_{\alpha \in \tilde{A}} v(\alpha, \sigma)$ is non-empty, closed and convex valued. Moreover, both $\Sigma^* \cap C(\alpha)$ and $\arg \max_{\alpha \in \tilde{A}} v(\alpha, \sigma)$ are upper hemicontinuous in α and σ respectively.

By the Kakutani fixed-point theorem, there exists a fixed point of the correspondence $C(\alpha) \cap \Sigma^* \times \arg \max_{\alpha \in \tilde{A}} v(\alpha, \sigma)$, $(\hat{\alpha}, \hat{\sigma})$. We want to show that $\hat{\sigma} \in C(\hat{\alpha})$, $\hat{\alpha} \in \arg \max_{\alpha \in A} v(\alpha, \hat{\sigma})$ and $V^* = v(\hat{\alpha}, \hat{\sigma})$.

By definition, $\hat{\sigma} \in C(\hat{\alpha})$.

Because $\hat{\sigma} \in \Sigma^* \cap C(\hat{\alpha})$, we have $V^* = v(\alpha^*, \hat{\sigma}) \geq v(\hat{\alpha}, \hat{\sigma})$. Because $\alpha^* \in \tilde{A}$, we also have

$v(\hat{\alpha}, \hat{\sigma}) \geq v(\alpha^*, \hat{\sigma})$. Therefore, $v(\hat{\alpha}, \hat{\sigma}) = v(\alpha^*, \hat{\sigma})$.

For any $\alpha \in A$, $(1 - \epsilon)\alpha^* + \epsilon\alpha \in \tilde{A}$ for $\epsilon > 0$ small enough. Because $\hat{\alpha} \in \arg \max_{\alpha \in \tilde{A}} v(\alpha, \hat{\sigma})$, we get

$$v(\alpha^*, \hat{\sigma}) = v(\hat{\alpha}, \hat{\sigma}) \geq v((1 - \epsilon)\alpha^* + \epsilon\alpha, \hat{\sigma}).$$

Using the concavity of v and rearranging, we obtain $v(\alpha^*, \hat{\sigma}) = v(\hat{\alpha}, \hat{\sigma}) \geq v(\alpha, \hat{\sigma})$. Therefore, $\hat{\alpha} \in \arg \max_{\alpha \in A} v(\alpha, \hat{\sigma})$. \square

Proposition 1 gives us a tool to verify whether the principal has no value of commitment.

3 Principal-agent problems

Suppose the situation we want to model is a principal-agent problem. Let $u(\alpha, \sigma)$ the utility function of the agent and $BR(\alpha)$ be the set of best-replies to α . Here, we have $BR(\alpha) = C(\alpha)$. I also assume that $\Sigma = \Delta(S)^n$ for some finite set S . I assume that $u(\alpha, \sigma)$ and $v(\alpha, \sigma)$ are linear in σ and jointly continuous. If these assumptions are satisfied, we are in a *principal-agent problem*.

Proposition 2. *Suppose we are in a principal-agent problem. Let $\alpha^* \in \arg \max_{\alpha} V(\alpha)$. If $V(\alpha)$ is continuous at α^* , then the principal has stable preferred strategies at α^* .*

Proof. Let $\Sigma^* = \{\sigma : v(\alpha^*, \sigma) = V^*\}$.

Suppose it is not the case, i.e., for all $\epsilon > 0$, there is $\alpha_\epsilon \in B_\epsilon(\alpha^*)$ such that

$$\text{for all } \sigma \in \Sigma^*, u(\alpha_\epsilon, \sigma^*) < u(\alpha_\epsilon, \sigma), \text{ for some } \sigma \notin \Sigma^*.$$

Take a sequence of α_ϵ and a principal-preferred best-reply $\sigma_\epsilon \in BR(\alpha_\epsilon)$. By hypothesis, $\sigma_\epsilon \notin \Sigma^*$. Without loss, we can take σ_ϵ in pure strategy and because it does not belong to Σ^* , the pure strategy σ_ϵ does not belong to the support of Σ^* . This implies that $\sigma_\epsilon \rightarrow \tilde{\sigma} \notin \Sigma^*$ as $\epsilon \rightarrow 0$ (possibly taking a convergent subsequence and using that there is a finite number of pure strategies). By continuity of u , we have $\tilde{\sigma} \in BR(\alpha^*)$. But then by continuity of V , we have $V(\alpha_\epsilon) = v(\alpha_\epsilon, \sigma_\epsilon) \rightarrow V^* = v(\alpha^*, \tilde{\sigma})$ for $\sigma \notin \Sigma^*$. A contradiction. \square

Continuity of the value does not generally imply that the stable preferred strategy condition holds. A key assumption we used in the proof here is that we can focus without loss of generality on pure strategies and that converging sequences of pure strategies have to be constant at some point in the sequence. That would not be the case in an infinite strategy space.

Lipnowski (2020) (conjectured by Forges (2020)) shows that in sender-receiver games with a finite number of actions and states, an optimal information structure can be implemented by cheap-talk when the value function that takes as input posterior beliefs is continuous. The model presented here allows for modelling sender-receiver games. Let Θ be a set of state, M a message space and S a finite action space. The strategy of the principal is an information structure $\alpha : \Theta \rightarrow \Delta M$ and the strategy of the agent is $\sigma : M \rightarrow \Delta S$. The posterior distribution given message m varies continuously in the information structure α . Therefore the two results are equivalent when restricting attention to sender-receiver games.

Ben-Porath et al. (2023) show that whenever at the optimal mechanism α , the principal is indifferent across all the best-replies of the agent, then the principal does not benefit from commitment. If a game has this property, they call this game *aligned*. The following result shows that if the game is aligned, then the value function is continuous.

Proposition 3. *Suppose we are in a principal-agent problem. If for all $\sigma \in BR(\alpha)$, $v(\alpha, \sigma) = V(\alpha)$, then $V(\alpha)$ is continuous at α .*

Proof. By the upper hemicontinuity of the best-reply correspondence and the fact that $\Sigma = (\Delta S)^n$ and $u(\alpha, \cdot)$ is linear, there is a open ball around α , $B_\epsilon(\alpha)$ such that for all $\alpha' \in B_\epsilon(\alpha)$, $BR(\alpha') \subseteq BR(\alpha)$. Therefore, for any sequence (α_n, σ_n) with $\alpha_n \in B_\epsilon(\alpha)$ and $\sigma_n \in BR(\alpha_n)$, we have (possibly taking a convergent subsequence), $\sigma_n \rightarrow \sigma \in BR(\alpha)$. Therefore, $v(\alpha_n, \sigma_n) \rightarrow V(\alpha)$. \square

This observation was also made by Lipnowski (2020) in the context of sender-receiver games. Most results on the value of commitment in the mechanism design with evidence literature have games that are aligned. Some examples are Glazer and Rubinstein (2004); Glazer and Rubinstein (2006), Ben-Porath et al. (2019) or Hart et al. (2017) (see Ben-Porath et al. (2023) for a discussion).

The conditions of aligned games and continuous value function differ when there is a best-reply of the agent, say σ' that makes the principal strictly worse off at the optimal mechanism α^* . To guarantee

that the value function is continuous at α^* , it must also be the case that σ' is not strictly preferred if we perturb α^* a little bit.

Below, I provide an example of a mechanism design problem with evidence and moral hazard where the value function is continuous and therefore the principal does not benefit from commitment. Noteworthy features of this examples are that the agent's payoff is non-linear in the mechanism and the principal is not indifferent across best-replies of the agent at the optimum. I then show an example of a game where there is a discontinuity in the value function and there is value to commitment. At the same time, I show that the result on the value of commitment in Ben-Porath et al. (2019) does not extend beyond simple type dependence.¹

Example 1. This example is a variation over mechanism with Dye evidence (Dye, 1985). The principal must make an investment decision, $a \in [-2, 2]$. The agent can make an investment in their skills $x = 0, 1$. In addition, the agent has a private information about their productivity $\theta \in \Theta \subset [-1, 1]$. With some probability, the agent can certify his productivity. With complement probability, he can only disclose that he did not receive evidence, denoted by \emptyset . Each type is pair of productivity and evidence to disclose, either $\{\{\theta\}, \emptyset\}$ or $\{\emptyset\}$.² Therefore a type is an element of $\Theta \times \{e_1, e_0\}$, where e_1 indicates that the agent can produce evidence and e_0 that he cannot. The set of evidence to disclose is $M = \Theta \cup \{\emptyset\}$. The prior distribution on types is denoted by $\mu \in \Delta(\Theta \times \{e_1, e_0\})$.

The payoff of the principal is $v(a, x, \theta) = \theta \cdot (1 + x) \cdot a - \frac{a^2}{2}$. The payoff of the agent is $u(a, x) = xa + (1 - x)(\mathbb{1}[a < 0] \cdot \frac{a}{2} + \mathbb{1}[a \geq 0] \cdot a)$. Therefore when $a \geq 0$, the agent is indifferent between action $x = 1$ and $x = 0$. When $a < 0$, the agent strictly prefers action $x = 0$. The agent always strictly prefers a higher a .

The interpretation is that the principal can invest in the production of two goods where positive a indicates the first good and negative a the second good. The agent wants the principal to invest as much as possible in the first good as little as possible in the second. Positive productivity θ is associated with positive productivity in the first good and negative θ is associated in positive productivity in the second good. On top of that, the agent can choose whether to improve his productivity in the dimension he is good at by choosing $x = 1$ or $x = 0$. He strictly prefers to not invest if he is assigned to the task he

¹In a principal-agent problem with private information $\theta \in \Theta$, a payoff function satisfies simple type dependence if for all $\theta \in \Theta$, $u(a, \theta) = u(a)$ or $= -u(a)$ for some function $u(a)$.

²Given the payoffs we will assume, cheap-talk cannot help the principal in this mechanism design problem.

does not like ($a < 0$) but is otherwise indifferent.

A strategy for the agent is a mapping $\sigma : \Theta \times \{e_1, e_0\} \rightarrow \Delta(M \times \{0, 1\})$ with the restriction that $\sigma(\theta, x|\theta, e) > 0 \Rightarrow e = e_1$ and $\sigma(\theta', x|\theta, e) = 0$ for all $\theta' \neq \theta$. That is, only types that have evidence can produce evidence and evidence has to be correct.

A strategy for the principal is a mapping $\alpha : M \rightarrow [-2, 2]$. Note that I assume that the investment decision x is not observed by the principal.

For a given evidence e , the principal's preferred best-reply is if $\theta > 0$, to play $x = 1$ whenever $\alpha(m) \geq 0$ and $x = 0$ otherwise. If $\theta < 0$, to play $x = 0$ for all $\alpha(m)$.

Given α , the agent always chooses the message that maximises his probability of being accepted.

For type (θ, e_1) with $\theta > 0$, the value to the principal is

$$\theta \max\{\alpha(\theta), \alpha(\emptyset)\} (1 + \mathbb{1}[\max\{\alpha(\theta), \alpha(\emptyset)\} \geq 0]) - \frac{\max\{\alpha(\theta), \alpha(\emptyset)\}^2}{2}.$$

For type (θ, e_1) with $\theta < 0$, the value to the principal is

$$\theta \max\{\alpha(\theta), \alpha(\emptyset)\} - \frac{\max\{\alpha(\theta), \alpha(\emptyset)\}^2}{2}.$$

For type (θ, e_0) with $\theta > 0$, the value to the principal is

$$\theta \alpha(\emptyset) (1 + \mathbb{1}[\alpha(\emptyset) \geq 0]) - \frac{\alpha(\emptyset)^2}{2}.$$

For type (θ, e_0) with $\theta < 0$, the value to the principal is

$$\theta \alpha(\emptyset) - \frac{\alpha(\emptyset)^2}{2}.$$

All these functions are continuous in α and therefore for any distribution μ , the value function $V(\alpha)$ is continuous.

The following example considers a generalisation of the payoff function introduced in Ben-Porath et al. (2019) that does not assume simple type dependence.

Example 2. Let $\Theta = \{\theta_1, \theta_2\}$, $M = \{m_1, m_2\}$, $M(\theta_1) = \{m_1\}$, $M(\theta_2) = \{m_1, m_2\}$ and $A = \{1, 2, 3\}$. Take some $\epsilon \in (\frac{2}{9}, \frac{1}{3})$. The prior probability on type θ_1 is μ .

The utility of the agent is $u(a, \theta_1) = a$ and

$$u(a, \theta_2) = \begin{cases} 0 & \text{if } a = 1 \\ 1 - \epsilon & \text{if } a = 2 \\ 1 & \text{if } a = 3 \end{cases}.$$

The utility of the principal is $v(a, \theta) = \nu(\theta)u(a, \theta) + \psi(a)$ with $\nu(\theta_1) = 1$, $\nu(\theta_2) = -3$ and $\psi(1) = 1$ and $\psi(2) = \psi(3) = 2$. The principal's payoff is increasing in a when $\theta = \theta_1$ and decreasing when $\theta = \theta_2$.

A strategy for the agent is a mapping $\sigma : \Theta \rightarrow \Delta M$ with $\sigma(m|\theta) > 0 \Rightarrow m \in M(\theta)$. Only type θ_2 has a real choice of strategy and we will denote by σ the probability of choosing message m_2 . A strategy for the principal is a mapping $\alpha : M \rightarrow \Delta A$.

Claim: In the game described above, there is value to committing to α .

Suppose there is no value of commitment. At the optimum, if $\sigma(m_2|\theta_2) > 0$, then the best-reply to m_2 is to play $a = 1$ as only θ_2 can send message m_2 . Given this best-reply, type θ_2 is only willing to send m_2 with positive probability if $\alpha(1|m_1) = 1$, otherwise type θ_2 strictly prefers to send m_1 . Therefore the principal plays the same action independently of the message received.

We conclude that without commitment, the best the principal can do is to use the same action independently of the message received.

Note that given the assumption on ϵ , there is always a prior μ such that playing $a = 2$ is a best-reply ex-ante for the principal:

$$\begin{aligned} 4\mu + (1 - \mu)(-1 + 3\epsilon) &\geq 2\mu + (1 - \mu), \\ 4\mu + (1 - \mu)(-1 + 3\epsilon) &\geq 5\mu + (1 - \mu)(-1). \end{aligned}$$

Now I show that there is a strategy α that can do better than that. Let $\alpha(1|m_1) = \epsilon$, $\alpha(3|m_1) = 1 - \epsilon$

and $\alpha(2|m_2) = 1$. The strategy of the agent is $\sigma(m_2|\theta_2) = 1$.

This strategy gives strictly higher profits to the principal and is incentive-compatible for the agent.

The IC constraint of θ_2 must satisfy

$$1 - \epsilon \geq \epsilon \cdot 0 + (1 - \epsilon) \cdot 1.$$

I show that the profits must be higher than playing $a = 2$ with probability one:

$$\begin{aligned} \mu\epsilon \cdot 2 + \mu(1 - \epsilon) \cdot 5 + (1 - \mu)(-1 + 3\epsilon) &> \mu 4 + (1 - \mu)(-1 + 3\epsilon) \\ \Leftrightarrow \epsilon &< 1/3. \end{aligned}$$

Since playing $a = 2$ is strictly optimal ex-ante for some μ , there is a mechanism where the principal does not best-reply and gives strictly higher profits against all the strategies where the principal best-replies.

To see there is a discontinuity, observe that if the principal decreases $\alpha(2|m_1)$ and increases $\alpha(1|m_2)$, then type θ_2 has a strict incentive to choose m_1 leading to a discrete drop in profits.

References

- Ben-Porath, E., Dekel, E. and Lipman, B. L. (2019), ‘Mechanisms with evidence: Commitment and robustness’, *Econometrica* **87**(2), 529–566.
URL: <https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA14991>
- Ben-Porath, E., Dekel, E. and Lipman, B. L. (2023), ‘Mechanism design for acquisition of/stochastic evidence’.
- Bester, H. and Strausz, R. (2000), ‘Imperfect commitment and the revelation principle: the multi-agent case’, *Economics Letters* **69**(2), 165–171.
- Doval, L. and Skreta, V. (2022), ‘Mechanism design with limited commitment’, *Econometrica* **90**(4), 1463–1500.
- Dye, R. A. (1985), ‘Disclosure of nonproprietary information’, *Journal of accounting research* pp. 123–145.
- Forges, F. (2020), ‘Games with Incomplete Information: From Repetition to Cheap Talk and Persuasion’, *Annals of Economics and Statistics* (137), 3–30.
- Glazer, J. and Rubinstein, A. (2004), ‘On optimal rules of persuasion’, *Econometrica* **72**(6), 1715–1736.
- Glazer, J. and Rubinstein, A. (2006), ‘A study in the pragmatics of persuasion: a game theoretical approach’, *Theoretical Economics* **1**(4), 395–410.
URL: <https://econtheory.org/ojs/index.php/te/article/view/20060395/0>
- Hart, S., Kremer, I. and Perry, M. (2017), ‘Evidence games: Truth and commitment’, *American Economic Review* **107**(3), 690–713.
- Lipnowski, E. (2020), Equivalence of cheap talk and bayesian persuasion in a finite continuous model, Technical report, Working paper.
- Myerson, R. B. (1982), ‘Optimal coordination mechanisms in generalized principal–agent problems’, *Journal of mathematical economics* **10**(1), 67–81.