Università Commerciale "Luigi Bocconi"
PhD School

PhD program in Economics and Finance
Cycle: 31^{st}
Disciplinary Field (code): SECS-P/01

Positioning choice problems and their applications to the theory of Moral Hazard

Advisor: Nenad Kos
Co-Advisor: Massimo Marinacci

PhD Thesis by
Jean-Gabriel Lauzier
ID number: 3002666

Academic Year: 2019/20

Contents

Chapter 1

Introduction

Optimisation is vital to modern economics, however tooling is non existent for the exploration optimisation problems for which the objective functions are discontinuous. This is unfortunate as there are interesting economic problems which can be modeled as discontinuous functions. This thesis by publication aims to develop and demonstrate some uses for such tooling. Of the three papers contained in this thesis, one develops tooling while the other two show their usage in the context of ex-post moral hazard problems.

The first article, *Positioning choice problems: the mathematics* develops tooling for a class of optimisation problems which we define as positioning choice problems. This is a class of optimisation problems defined in finite-dimensional Euclidean spaces for which the value function is always almost everywhere continuous even when the objective function is discontinuous.

For this new class of problem, we show that we can state an equivalent of first-order conditions in terms of Dini supergradients. This allows us to prove that the value function is Lipschitz continuous, which implies by Rademacher's theorem that it is almost everywhere differentiable. This implication is essentially an ad-hoc envelope theorem for positioning choice problem. The proof's use of Rademacher's theorem also yields a fuller characterization of the derivative of positioning choice problems, in a way similar to Danskin's theorem. We conclude the paper by suggesting that positioning choice problems are likely to be generalized on other spaces as well.

The first application we explore in *Ex-Post moral hazard and manipulation-proof contracts* discusses ex-post moral hazard in the context of entrepreneurial financing. We examine the trade-off between the provision of incentives to exert costly effort (ex-ante moral hazard) and the incentives needed to prevent the agent from manipulating the profit observed by the principal (ex-post moral hazard). Formally, we build a model of two stage hidden actions where the agent can both influence the expected revenue of a business and manipulate its observed profit. We focus our efforts in the analysis of bonus contracts as tools to provide incentives.

As the manipulation stage of the game is a positioning choice problem, and we can thus directly import results from Lauzier (2020d). We show that the optimal contract is manipulation-proof whenever the manipulation technology is linear. This linearity of the manipulation technology entirely drives the result because it allows us to replace any contract by the value function of the optimisation problem it defines in the manipulation stage without changing the incentives to exert effort.

However, convex manipulation technologies sometimes lead to contracts for which there are manipulations in equilibrium. Whenever the distribution satisfies the monotone likelihood ratio property, we can always find a manipulation technology for which this is the case. This results obtains because bonus contracts have the desirable property of incentivizing hard work while maintaining the expected losses to (acceptable) manipulations low.

The next application we turn to is the design of insurance. *Insurance design and arson-type risks*. In *Insurance design and arson-type risks*, arson-type risks are modeled similarly to the linear manipulation technologies of Lauzier (2020a). Therefore in addition to the tooling we developed in Lauzier (2020d), we will additionally be able to import some of the proof technique developed Lauzier (2020a).

Various authors obtain optimal insurance contracts which are at odds to the types of contracts offered by the industry. Huberman et al. (1983)'s completely disappearing deductible and Picard (2000)'s discontinuous contract serve as our main examples. The authors observe that these contracts are not optimal when there are arson-type risks.

Similarly to Lauzier (2020a) the manipulation stage of the game is a positioning choice problem. We can therefore, using our tooling, show that the optimal insurance contract with arson-type risks must be continuous and have a bounded slope.

The economic intuition gathered from Lauzier (2020a) is reversed while dealing with insurance. In Lauzier (2020a) there is a trade-off in the provision of incentivization dealing with ex-ante and ex-post moral hazards and the optimal contract sometimes entails manipulations in equilibrium.

Insurance contracts never entail manipulations in equilibrium and all contracts mixing coinsurance and deductibles are robust to arson-type risks. However, these types of contracts are also used to mitigate other ex-ante agency problems such as adverse selection and ex-ante moral hazard. As the contracts used for the prevention of arson-type manipulations are needed for the resolution of other agency problems we intuit that here are no trade-offs as in "securities design".

Chapter 2

Positioning choice problems: the mathematics

Abstract

This article examines differentiability properties of the value function of *positioning choice problems*, a class of optimisation problems in finite-dimensional Euclidean spaces. We show that the value function of positioning choice problems is always almost everywhere differentiable even when the objective function is discontinuous. To obtain this result we first show that the Dini superdifferential is always well-defined for the maxima of positioning choice problems. This last property allows to state first-order necessary conditions in term of Dini supergradients. We then prove our main result, which is an ad-hoc envelope theorem for positioning choice problems. Lastly, after discussing necessity of some key assumptions, we conjecture that similar theorems might hold in other spaces as well.

Keywords

Envelope theorem, Optimisation, Discontinuous optimisation, Danskin's theorem, Rademacher's theorem, Lipschitz continuity, Positioning choice problems

2.1 Introduction

An envelope theorem is a statement about the derivative of value functions. Envelope theorems have foundational applications in several fields of mathematical analysis, notably in the calculus of variations and optimal control. As such, they are also fundamental to the microeconomic analysis of consumer and producer problems (Mas-Colell et al. (1995)).

Envelope theorems and their generalizations, as they are currently formulated and used, rely on a fundamental hypothesis: the continuity of the objective function. However, this assumption is problematic in applications where discontinuities are meaningful. Notably, discontinuous functions are essential to modeling pivotal economic phenomena such as executive bonuses, tax brackets and indivisible capital investments such as power plants.

Optimisation problems with discontinuities are not differentiable and thus standard first-order necessary conditions cannot be used to find maxima at such points. Consequently, prior envelope theorems cannot be used to characterize the behaviour of their value function.

In this article, to tackles this issue we define a class of optimisation problems in finite-dimensional Euclidean spaces for which the value function is almost everywhere differentiable even when the objective function is discontinuous. These optimisation problems are dubbed *positioning choice problems* since they have a straightforward geometrical interpretation as a choice of position. To the best of our knowledge this is the first article to explicitly examine the value function of optimisation problems with discontinuous objective functions.

As mentioned, the maxima of discontinuous functions are not differentiable and so standard first-order necessary conditions theorems cannot be used. We propose a generalisation of first-order conditions using the Dini superdifferential. Dini supergradients are well-defined at the maxima of positioning choice problems and so is the Dini superdifferential. To the best of our knowledge this property is not shared by other notions of superdifferentials in the literature.

We then prove an ad-hoc envelope theorem for positioning choice problems along the lines of Danskin's theorem (Danskin, 1967). A deeper look at the theorem shows that it relies extensively on Euclidean spaces being a (Dedekind-)complete ordered field, an observation worth investigating.

We set to do so with examples aimed at breaking down our main result. We first present an optimisation problem for which the maxima are isolated points but where the value function is still almost everywhere differentiable. This highlights that first-order conditions are not essential to the characterization of the derivative of value functions. The second example shows how differentiability depends on the domain of the value function. Piecing those observations together we suggest that envelope theorems might be obtained on other spaces which are isomorphic to the reals. We conclude with an example on a Riemannian manifold which supporting this conjecture.

The presentation is as follows. We first give a brief overview of the literature. We further detail important concepts in the beginning of section 1. Section 1 also defines positioning choice problems. It contains two examples that help build the intuition behind our ad-hoc envelope theorem. Variations of these examples will be used while discussing our findings.

Theorem 2, which provides a generalized notion of first-order necessary conditions for a maximum in the context of positioning choice problems, and theorem 3, the ad-hoc envelope theorem, are presented in section 2 which serves the core of the article.

Section 3 discuss the main results. The presentation is less formal as we use examples to highlight certain properties of positioning choice problems. We show how continuity and differentiability properties of value functions are tightly connected to the completeness of Euclidean spaces. We conjecture that envelope theorems similar to theorem 3 might be stated for any spaces which are isomorphic to the reals.

Avenues for further research are identified in conclusion. Appendix A contains the definitions omitted in the text and a relevant statement of Rademacher's theorem, while Appendix B contains the omitted proofs.

Literature review

The essential property of positioning choice problems is that the value function is always a locally Lipschitz map between two finite-dimensional Euclidean spaces, even when the objective function of the optimisation problem is discontinuous. By Rademacher's theorem, this implies that the value function is almost everywhere differentiable (on open sets, with regard to the Lebesgue measure) and therefore relatively "well-behaved".

The literature on envelope theorems is well-established. It is largely focused on providing increasingly general conditions on objective functions as to characterize the derivative of value functions. Recent examples are Morand et al. (2015) and Morand et al. (2018). The authors scrutinize what they call Lipschitz programs, a large class of parametric optimisation problems where the objective function satisfies Lipschitzianity. They show that under weak assumptions the Lipschitz property is inherited by the value function. Our approach is different but complementary to theirs as we do not study parametric problems instead lifting the Lipschitz assumption on the objective function while still obtaining Lipschitzianity of the value function.

Both Morand et al. (2015) and Morand et al. (2018) derive many results of Milgrom and Segal (2002) as special cases since the latter assumes the absolute continuity of the objective function. Our findings complements the results of Milgrom and Segal (2002) as well as we consider discontinuous objective functions. More details on Milgrom and Segal (2002) are provided in the next section.

To the best of our knowledge this paper is the first to explicitly examine the value function of optimisation problems with discontinuous objective functions. Thus, we prioritise clarity and keep notation, assumptions and proofs as simple as possible. This comes at the cost of generality. The sequel paper, (Lauzier, 2020c), explores positioning choice problems in greater generality by showing how we can relax some assumptions in order to model different economic problems.

2.2 Positioning choice problems

Preliminaries

Let Θ be a set of parameters and let $Y(\theta)$ be a choice set given parameter θ. Let the function $h : Y(\Theta) \times \Theta \to \mathbb{R}$ be the *objective function.* The problem

$$\sup_{y(\theta)\in Y(\theta)} h(y(\theta);\theta)$$

is an *optimisation problem.* The *optimal choice correspondence* $\sigma : \Theta \rightrightarrows 2^{Y(\theta)}$ is

$$\sigma(\theta) = \arg\max_{y(\theta)\in Y(\theta)} h(y(\theta);\theta)$$

and the *value function* $V : \Theta \to \mathbb{R}$ is

$$V(\theta) = h(y^*(\theta);\theta)$$

for $y^*(\theta) \in \sigma(\theta)$. An *envelope theorem* is a statement about the rate of change of $V(\theta)$ when θ changes, often stated in terms of the derivative

$$\frac{\partial V(\theta)}{\partial \theta}.$$

Typically, envelope theorems rely building a *continuous selection*[1] $y^*(\theta)$ of $\sigma(\theta)$ in order to use the chain rule and write this derivative as

$$\frac{\partial V(\theta)}{\partial \theta} = \frac{\partial h(y^*(\theta);\theta)}{\partial \theta} = \frac{\partial h(y^*(\theta);\theta)}{\partial y^*(\theta)} \frac{\partial y^*(\theta)}{\partial \theta}.$$

This approach relies on the family

$$\{h(\cdot;\theta) : \theta \in \Theta\}$$

being composed of functions satisfying some form of continuity. The simplest envelope theorem such as the one found in Mas-Colell et al. (1995) assumes that this family consists exclusively of differentiable functions. The basic statement of Milgrom and Segal (2002) assumes that this family consists of absolutely continuous functions[2]. However, continuity can be problematic in application. This paper shows that it is sometime possible to make statements about the derivative of a value function even when the objective function is discontinuous.

Definitions

Consider a slightly different notation to avoid confusion. Let $1 \leq n < \infty$ and let $C \subsetneq \mathbb{R}^n$ be a closed box in $\mathbb{R}^n$ which contains zero. Let the function $f : C \to \mathbb{R}$ be the "payoff function" of being at position $y \in C$. Let the function $g : C \times C \to \mathbb{R}$ be the "cost of movement function" of moving from being at $x \in C$ to y.

Assume that g is continuous, that $\sup_{y \in C} |f(y)| < +\infty$ and that $\max_{(x,y) \in C \times C} |g(x,y)| < +\infty$. Further assume that the function g is always a metric on $\mathbb{R}^n$ (see appendix A). This assumption is not essential to obtain the main results of this paper, but it greatly streamlines the proofs and thus the exposition.

Let $x \in C$ be given and let $h(x,y) = f(y) - g(x,y)$. The optimisation problem

$$\sup_{y \in C} h(x,y)$$

is a **positioning choice problem**.

---

[1]A selection f of a correspondence F is a function such that for every $x \in domain(F)$ it is $f(x) \in F(x)$.
[2] Lipschitzianity is assumed in Morand et al. (2015), Morand et al. (2018) and Clarke (2013).

Further, assume that f is almost everywhere continuous except on a set of points where it satisfies either

$$\limsup_{x \in C,\, x \to y} f(x) = f(y)$$

or

$$\liminf_{x \in C,\, x \to y} f(x) = f(y).$$

As previously, for $x \in C$ given, $h(x, y)$ is called the objective function and the correspondence $\sigma : C \rightrightarrows 2^C$ defined as

$$\sigma(x) = \arg\max_{y \in C} h(x, y)$$

is called the optimal choice correspondence. A function $y^* : C \to C$ is a selection of σ if for every $x \in C$, $y^*(x) \in \sigma(x)$. Letting $y^*(x) \in \sigma(x)$, the value function $V : C \to \mathbb{R}$ is

$$V(x) = h(x, y^*(x)) = f(y^*(x)) - g(x, y^*(x)).$$

Example: the structure of the line at *n=1*

Let $n = 1$, $C = [0, 2]$, $f(x) = x \cdot 1_{x \geq 1}$ and $g(x, y) = \alpha|y - x|$ for $1_{x \geq 1}$ the indicator function, $\alpha > 1$ and $| \cdot |$ the usual Euclidean distance on $\mathbb{R}$. Notice that since f is discontinuous the objective function also is. In other words, the family of functions

$$\{h(x, \cdot) : x \in C\}$$

consists exclusively of functions which are discontinuous. Thus, typical envelope theorems cannot be used to characterize the derivative of the value function V

$$\frac{\partial V(x)}{\partial x}.$$

However, notice that by the projection theorem there exists at least one pair (x, y), $y > x$, for which we have

$$f(x) - g(x, x) = f(y) - g(y, x).$$

Since $\alpha > 1$ this pair is unique and is $y = 1$ and $x = 1 - 1/\alpha$. The optimal choice correspondence is

$$\sigma(x) = \begin{cases} \{x\} & \text{if } x \in [0,\, 1 - 1/\alpha) \cup [1, 2] \\ \{1\} & \text{if } x \in (1 - 1/\alpha, 1) \\ \{x, 1\} & \text{if } x = 1 - 1/\alpha \end{cases}$$

and so the value function is

$$V(x) = \begin{cases} 0 & \text{if } x \in [0,\, 1 - 1/\alpha] \\ 1 - \alpha(1 - x) & \text{if } x \in (1 - 1/\alpha, 1) \\ x & \text{if } x \in [1, 2]. \end{cases}$$

Clearly, $V(x)$ is almost everywhere differentiable on $(0, 2)$ except at $x \in \{1 - 1/\alpha, 1\}$ and has derivative given by

$$V'(x)|_{x \in (0,2) \setminus \{1 - 1/\alpha, 1\}} = \begin{cases} 0 & \text{if } x \in (0,\, 1 - 1/\alpha) \\ \alpha & \text{if } x \in (1 - 1/\alpha, 1) \\ 1 & \text{if } x \in (1, 2) \end{cases}$$

Moreover, $V(x)$ is kinked at $x = 1 - 1/\alpha$ with subdifferential given by

$$\underline{\partial} V(1 - 1/\alpha) = [0, \alpha]$$

and at $x = 1$ with superdifferential

$$\overline{\partial} V(1) = [1, \alpha].$$

This examples shows that even when the objective function is not particularly "well-behaved" the value function is. Implicitly, this example uses the structure of the line to guarantee that σ is upper hemicontinuous which guarantees a simple optimisation problem. Variations of it will be used in section 3 for further discussions.

Example: *n=2* and an intuitive argument

The previous example can be generalized to additional dimensions. The next constructive example aims to provide the reader with more intuition about the main results of the article.

Let $n = 2$, $C = [0,2]^2$, $g(x,y) = \alpha||y - x||$, $\alpha > 1$ and

$$f(x_1, x_2) = \begin{cases} 1 & \text{if simultaneously } x_1 \geq 1 \text{ and } x_2 \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Notice that f is almost everywhere continuous except on the set

$$Dis_f = \{(x,y) : x = 1,\, y \in [1,2]\} \cup \{(x,y) : x = [1,2],\, y = 1\}$$

where it is upper semicontinuous. By the projection theorem to every point

$$(x,y) \in Dis_f \setminus \{(1,1)\}$$

there exists a unique $(a,b) \neq (x,y)$ such that

$$V\left((a,b)\right) = f\left((a,b)\right) - g\left((a,b),(a,b)\right) = 0 = f\left((x,y)\right) - g\left((x,y),(a,b)\right).$$

This (a,b) corresponds exactly to the point $x = 1 - 1/\alpha$ in the previous 1-dimensional example so that if $(x,y) = (1,y)$ then $(a,b) = (1-1/\alpha, y)$ and vice-versa for $(x,y) = (x,1)$. Let K be the set of points that satisfies the previous condition.

The arc of circle with center $(1,1)$ defined by

$$L := \{(x,y) : x \leq 1, y \leq 1, (x,y) \text{ belongs to a circle with center } (1,1) \text{ and radius } 1-1/\alpha\}$$

is also a set of points that satisfies the previous equality. That is, for every $(a,b) \in L$ it is

$$V((a,b)) = f((a,b)) - g((a,b),(a,b)) = 0 = f((1,1)) - g((1,1),(a,b)).$$

Notice that the set

$$Kink_V := Dis_f \cup K \cup L$$

contains all the (interior) points of non-differentiability of V and is a set of measure zero with regard to the Lebesgue measure on $\mathbb{R}^2$. Let *int* denotes the interior of a set. It is clear that V is differentiable on $\mathcal{M} := int(C) \setminus Kink_V$, with $\nabla V(x) = (0,0)$ for every point (x,y) for which $\sigma((x,y)) = \{(x,y)\}$. Furthermore, whenever $(x,y) \in \mathcal{M}$ and

$$\sigma((x,y)) = \{(a,b)\}, \quad (a,b) \neq (x,y)$$

then $(a, b) \in Dis_f$ and the directional derivative of V at (x, y) in the direction of $(a - x, b - y)$ is[3]

$$V'((x, y); (a - x, b - y)) = \alpha||(a - x, b - y)||.$$

Constructing this example bare hands is tedious but insightful. It shows that there are higher dimensional environments where the value functions are "well-behaved" even when the objective functions are not. In other words, it is clear that the same problem could be cast in $n = 3, 4, ..., N < +\infty$ dimensions and the value function would still be almost everywhere differentiable.

2.3 Statements

Let us start with a lemma that allows for simpler notation. Since C is closed, f is bounded and g is continuous it is easy to show the following:

Lemma 1 *Under the assumption made in section 1 there always exists a correspondence* $\overline{f} : C \rightrightarrows 2^{\mathbb{R}}$ *and a selection* $\overline{f}^*$ *of* $\overline{f}$ *such that*

$$f = \overline{f} = \overline{f}^* \qquad \textit{almost everywhere}$$

and for $\overline{h}(x, y) = \overline{f}^*(y) - g(x, y)$ *it holds*

$$\sup_{y \in C} h(x, y) = \sup_{y \in C} \overline{h}(x, y) = \max_{y \in C} \overline{h}(x, y).$$

The proof follows from setting

$$\overline{f}(y) = \left\{ z \in \mathbb{R} : \text{ either } \limsup_{x \in C,\, x \to y} f(x) = z \text{ or } \liminf_{x \in C,\, x \to y} f(x) = z \text{ or both} \right\}$$

and observing that for every $x \in C$ it is possible to build a selection for which the problem $\sup_{y \in C} \overline{h}(x, y)$ attains its supremum. In light of the lemma, the rest of the presentation uses max instead of sup and drops the subscript $x \in C$ while taking limits.

Consider a positioning choice problem

$$\max_{y \in C} h(x, y)$$

satisfying the assumptions made in section 1. Let x be given and let $y^* \in cor(C)$ be a maximizer of $h(x, \cdot)$, where cor denotes the algebraic interior of a set[4] . Since the

[3] See section 2 for a definition of the directional derivative.

[4] Recall that $cor(A) = int(A)$ whenever $A \subset \mathbb{R}^n$. See appendix A for more details

objective function h is discontinuous it is possible that the gradient $\nabla h(x, \cdot)$ is ill-defined at y^*. Thus, the standard first-order necessary condition

$$0 = \nabla h(x, y^*)$$

is not necessarily meaningful. There is, however, a suitable substitute notion.

The *upper Dini derivative*[5] of $h(x, \cdot)$ in the direction of $v \in \mathbb{R}^n$ is

$$dh(x, y; v) = \limsup_{\substack{w \to v \\ t \downarrow 0}} \frac{h(x, y + tw) - h(x, y)}{t} \tag{UDD}$$

and the *Dini supperdifferential* of $h(x, \cdot)$ at y, denoted by $\partial_d h(x, y)$, is the set of $\zeta \in \mathbb{R}^n$ such that

$$dh(x, y; v) \leq \langle \zeta, v \rangle \quad \forall v \in \mathbb{R}^n. \tag{DS}$$

Each $\zeta \in \partial_d$ is a *Dini supergradient* and if $h(x, \cdot)$ is differentiable at $y^* \in \sigma(x)$ then

$$\partial_d h(x, y) = \{\nabla h(x, y^*)\}.$$

Theorem 2 (First-order necessary conditions for an interior solution on C) *Under the assumptions of Section 1, for every given $x \in C$, if $y^* \in cor(C)$ and $y^* \in \sigma(x)$ then*

$$0 \in \partial_d h(x, y^*) \tag{CFOC}$$

Moreover, if both f and g are differentiable at y^* then

$$0 = \nabla h(x, y^*) = \nabla(f(y^*) - g(x, y^*)) = \nabla f(y^*) - \nabla g(x, y^*). \tag{SFOC}$$

Proof. We only have to show the case when h is discontinuous at y^* as the other cases follow. Recall that by Lemma (1), we can assume without lost of generality that

$$\limsup_{y \to y^*} h(x, y) = h(x, y^*).$$

Lemma: $\partial_d h(x, y^*)$ is well-defined
Proof of the lemma: By definition it is

$$h(x, y^*) = f(y^*) - g(x, y^*) = \limsup_{y \to y^*} f(y^*) - g(x, y^*).$$

[5] The definitions used here follow Clarke (2013) chapter 11.

For every direction $v \in \mathbb{R}^n$ for which there is a discontinuity it holds that

$$dh(x, y^*; v) = \limsup_{\substack{w \to v \\ t \downarrow 0}} \frac{h(x, y^* + tw) - h(x, y^*)}{t} = -\infty$$

because $h(x, y^* + tw) - h(x, y^*) \to K < 0$, K constant. Since f is almost everywhere continuous there exists a direction v for which $|dh(x, y^*; v)|$ is finite. As $y^* \in cor(C)$ and $y^* \in \sigma(x)$ it is

$$\partial_d h(x, y^*) \neq \emptyset$$

and $\partial_d h(x, y^*)$ is well-defined. □

It follows immediately from $y^* \in \sigma(x)$ that for every $v \in \mathbb{R}^n$ it holds

$$dh(x, y^*; v) \leq 0 = \langle 0, v \rangle$$

and

$$0 \in \partial_d h(x, y^*),$$

as desired. ■

The proof is almost identical to the proof of necessity of first-order conditions in books on convex analysis and non-smooth optimisation such as Borwein and Lewis (2010). The difference is the use of the Dini superdifferential, which is well-defined at discontinuity points.

To simplify the presentation the remainder of this section always consider positioning choice problems satisfying the assumptions of section 1 unless specified otherwise. The next statement is the ad-hoc envelope theorem. It uses the *Fréchet* derivative defined in appendix A. The *Fréchet* derivative is a generalisation of the standard gradient ∇ and the two notions coincides on finite-dimensional euclidean spaces.

Theorem 3 ("Ad-Hoc" envelope theorem for positioning choice problems) *Positioning choice problems always have almost everywhere Fréchet differentiable value functions V.*

The logic of the proof is instructive. Since C is closed, f is bounded and g is continuous, it suffice to prove the continuity of V to show that it is Lipschitz and obtain, by Rademacher's theorem, that it is almost everywhere *Fréchet.* Proving continuity of V is relatively simple but a bit tedious. By the way of contradiction it is assumed that $V(x)$ is discontinuous at a point. Since g is continuous this implies that there exists a converging sequence $x_n \to x$ along which (a) either $V(x) > V(x_n)$ or $V(x_n) > V(x)$ so (b) the set of maximizers $\sigma(x_n)$ and $\sigma(x)$ are disjoint. For this to happen it must be the case that x is a discontinuity point of f. WLOG let $V(x) > V(x_n)$ and consider $y^*(x) \in \sigma(x)$. Since g is a metric the projection theorem guarantees that we can find a neighborhood around x such that $y^*(x)$ strictly dominates every points of $\sigma(x_n)$, a contradiction.

Using Rademacher's theorem also ensure that every points of V which are not *Fréchet* differentiable have well-defined directional derivatives[6] in every direction. This property is summarized in the next corollary.

Corollary 4 (Properties of the derivative) *The value function of a positioning choice problem always satisfies the following:*

1. *Every point of non-differentiability in the sense of Fréchet have finite directional derivative in every directions;*
2. *if $\sigma(x) = \{x\}$ then $V(x) = f(x)$ and $\nabla V(x) = \nabla f(x)$ whenever $\nabla f(x)$ is well-defined;*
3. *if $g(x, y) = \alpha||y - x||$ for $\alpha \geq 1$ then to every $x \neq y$, $x \in int(C)$, $y \in C$, it holds that $|V'(x; y - x)| \leq \alpha||y - x||$. Equality holds whenever $y \in \sigma(x)$.*

The corollary is an immediate consequence of Rademacher's theorem and its proof does not provide further insights. Intuitively, theorem 3 and its corollary provides a limited version of Danskin's theorem (Danskin (1967)) in the context of positioning choice problems.

The rest of this section gathers results which are useful in application. Following Milgrom and Segal (2002), it is sometimes handy to write the value function as an integral.

Fact 5 (Integral representation of the value function) *Let $n = 1$ and assume $C = [0, M]$ for $0 < M < +\infty$. There exists a continuous function $v_0 : C \to \mathbb{R}$ such that for every $\tilde{x} \in C$ it holds*

$$V(\tilde{x}) = V(0) + \int_0^{\tilde{x}} v_0(x)dx. \tag{IRVF}$$

This representation is a simple application of the Second fundamental theorem of calculus. The statement can be generalized to higher dimensions, but doing so does not provide

[6]Recall that the directional derivative of a function $h : C \to \mathbb{R}$ at $x \in C$ in the direction of $y \in C$ is

$$h'(x; y) := \lim_{t \downarrow 0} \frac{h(x + ty) - h(x)}{t}$$

when the limit exists.

further insights on the behaviour of positioning choice problems.

The assumption that g is a metric is useful because it allows to streamline proofs by using the projection theorem. However, it is restrictive.

Proposition 6 (Sufficient conditions for monotonicity of the value function) *Let g be*

$$g(x, y) := \sum_{i=1}^{n} g_i(y_i - x_i)$$

for $g_i(y_i - x_i)$ continuous, non-decreasing and almost everywhere differentiable except maybe at 0. Then the value function $V(x)$ is almost everywhere differentiable.

Moreover, if g is positive and for every $y_i < 0$ it holds $g_i(y_i) = 0$ then V is non-decreasing: if $y >> x$ then $V(y) \geq V(x)$.

Proposition 6 is an example of a class of functions g which are useful in economics. Lauzier (2020a) considers the case when $n = 1$ and $g(x, y) = \max\{0, y - x\}$ is interpreted as a manipulation technology of the observable profit of a firm. A simplified version of the narrative goes as follows. A business owner hires a manager to take care of the company. The real profit x randomly realizes at the end of a quarter. The manager observes x and reports y, namely, the accounting profit of the quarter. The owner observes only the accounting profit y. Before reporting, the manager can manipulate the accounting profit by either burning some money (if $y < x$ then $g(x, y) = 0$) or injecting liquidities (if $y > x$ then $g(x, y) = y - x$). Proposition 6 implies that the manager's pay must be non-decreasing in observable profit y because the value function of the optimisation problem defined by this simple game is monotonic[7].

Many optimisation problems which are important in application are defined on $\mathbb{R}^n$ and not on a closed box C. Extending Theorem 3 to such setting can be challenging. Second-order derivatives are not defined for discontinuous functions, and second-order sufficient conditions for optimality cannot be used (i.e. Karush-Kuhn-Tucker theorem). Proposition 8 shows how to extend Theorem 3 to such setting. The next useful intermediate result is stated independently as a lemma for ease of presentation.

Lemma 7 (A plane is a plane) *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a plane and let $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ be a positive and convex transformation of the Euclidean distance, i.e $g(x, y) = \tilde{g}(||y - x||)$*

[7] Lauzier (2020c) shows how different functions g can be used to model economic problems.

for some positive and strictly convex function $\tilde{g} : \mathbb{R}_+ \to \mathbb{R}_+$. Then for every finite x the optimisation problem

$$\max_{y \in \mathbb{R}^n} h(x, y) = \max_{y \in \mathbb{R}^n} \{f(y) - g(x, y)\}$$

attains a finite maximum. Moreover, the value function $V(x)$ is a plane parallel to $f(x)$ and hence continuously differentiable with $\nabla V = \nabla f$.

The proof of the lemma is straightforward and need not be covered.

Proposition 8 (Positioning choice problems and interior solutions) *Let f and g be as in lemma 7 with f non-decreasing. Let $\tilde{f} : \mathbb{R}^n \to \mathbb{R}$ be a non-decreasing function which is bounded by f. Then the value function*

$$\tilde{V}(x) = \tilde{h}(x, \tilde{y}^*(x)) = \tilde{f}(\tilde{y}^*(x)) - g(x, \tilde{y}^*(x)) \quad \textit{for} \quad \tilde{y}^*(x) \in \tilde{\sigma}(x) = \arg\max_{y \in \mathbb{R}^n} \tilde{h}(x, y)$$

is almost everywhere Fréchet differentiable and is bounded by the function

$$V(x) = h(x, y^*(x)) = f(y^*(x)) - g(x, y^*(x)) \quad \textit{for} \quad y^*(x) \in \sigma(x) = \arg\max_{y \in \mathbb{R}^n} h(x, y).$$

The intuition behind proposition 8 is as follows. For $\tilde{f}$ a discontinuous function of interest is found a function f which bounds it. The latter is taken continuous and locally bounded so that standard second-order conditions can be used solving problem $\max\{f(y) - g(x, y)\}$. The function g must be "sufficiently convex" so that "g crosses both f and $\tilde{f}$ from below". This is enough to guarantee that

$$\tilde{\sigma}(x) \subset \{y \in \mathbb{R}^n : f(y) - g(x, y) \geq 0\}$$

and conclude that $\tilde{V}(x)$ is finite and bounded by $V(x)$. Similarly, the local Lipschitzianity of $\tilde{V}$ is insured by the local Lipschitzianity of V. Rademacher's theorem then guarantees that $\tilde{V}$ is almost everywhere *Fréchet*. The sequel Lauzier (2020c) introduces set orders to study further properties of positioning choice problems.

2.4 Discussion

We start by providing a simple example of a positioning choice problem where theorem 2 does not hold but where the value function can be characterised in a way paralleling theorem 3. The example emphasis the role of $\mathbb{R}^n$ being a *Dedekind-complete ordered field.* It also suggests that statements similar to Theorem 3 can be obtained for optimisation problems defined on other spaces as well, for example for optimisation problems defined on Riemannian manifolds. We then conclude with an example where this "conjecture" is correct.

Pathological cases

Positioning choice problems are defined on finite-dimensional Euclidean spaces. This choice is consequential for three reasons. First, finite-dimensionality allows to define maxima in straightforward ways. Second, it allows to use Rademacher's theorem. Third, theorem 3 implicitly exploits the properties of Euclidean spaces being *Dedekind-complete ordered fields*. The next example shows why this is important.

Let $n = 1$, $C = [0, 2]$ and consider the following functions:

$$f(y) = \begin{cases} 1 & \text{if } x = 1 \\ 0 & \text{otherwise,} \end{cases}$$

and $g(x, y) = |x - y|$. For $x \in [0, 2]$ given the optimisation problem is

$$\max_{y \in [0,2]} h(x, y) = \max_{y \in [0,2]} \{f(y) - g(x, y)\}.$$

It is easy to check that the optimal choice correspondence is $\sigma(x) = \{1\}$ and thus

$$V(x) = 1 - |1 - x|.$$

The value function V is almost everywhere differentiable. The derivative is

$$V'(x)|_{x \in (0,2) \setminus \{1\}} = \begin{cases} 1 & \text{if} \quad 0 < x < 1 \\ -1 & \text{if} \quad 1 < x < 2 \end{cases}$$

Two observations are immediate.

First, the assumption that for every $z \in dom(f)$ it is

$$\limsup_{y \to z} f(y) \geq f(z)$$

is necessary to write down theorem 1. Otherwise the Dini superdifferential of $h(x, y^*)$ might not be well-defined for y^* a maximizer. The problem is that no superdifferentials are well-defined for isolated points (the points $h(x, 1)$ in the example). However the assumption is not necessary to obtain differentiable value functions.

Second, the derivative $V'(x)$ is directly computed using the chain rule (setting $y^*(x) = 1$). In other words, it is possible to characterize the derivative of value functions with the chain rule even for problems where the objective function fails all continuity properties traditionally assumed in the literature[8].

Those observations demonstrate that properties of value functions such as continuity and differentiability are not necessarily "inherited" from objective functions. Such properties are tighten to the space on which the value function is defined.

Let us highlight further the role of Euclidean spaces being Dedekind-complete ordered fields. Consider again the two functions f and g defined above but now assume that $x \in X = \{0, 1, 2\}$. Theorem 3 cannot be used to characterize further the value function of this problem, nor does any other envelope theorem we are aware of.

However, the value function is easily seen as

$$V(x) = \begin{cases} 1 & \text{if } x = 1 \\ 0 & \text{if } x \in \{0, 2\} \end{cases}.$$

This function is defined on the discrete space X so its derivative cannot be defined. This difficulty is handled implicitly in positioning choice problems by insuring that the space of "parameters" X is a (subset of a) Dedekind-complete ordered field. Precisely, the problem stems from the fact that when $X = \{0, 1, 2\}$ it is

$$cor(X) = \emptyset,$$

while the set $cor(X)$ should be non-empty[9].

Riemannian manifolds

The previous two examples are pathological but informative. Let X, Y be arbitrary sets, let $f : Y \to \mathbb{R}$ be a function and $g : X \times Y \to \mathbb{R}$ be an induced metric[10] on $X \times Y$. Set

[8]For instance, the main envelope theorem of Milgrom and Segal (2002) relies on the family

$$\{h(x, y) : x \in [0, 2] \text{ and } y \in [0, 2]\}$$

consisting of functions that are absolutely continuous in y. The problem naturally fails this assumption.

[9]This example also shed lights on the role of finite-dimensionality. In the setting of this paper, the interior *int* and the algebraic interior *cor* of a set coincides. This is not true anymore in infinite-dimensional spaces where the two notions can differ dramatically. It might be possible to prove results similar to theorem 3 in some infinite-dimensional spaces such as the ones defined in dynamic programming.

[10]This assumption is to rule-out discontinuous functions like the discrete metric.

$h(x, y) = f(y) - g(x, y)$ and consider the optimisation problem

$$\sup_{y \in Y} h(x, y)$$

for $x \in X$ given. The value function is now

$$V(x) = h(x, y^*(x))$$

for

$$y^*(x) \in \sigma(x) = \arg\max_{y \in Y} h(x, y).$$

What are the minimal assumptions which must be imposed on the set X and Y in order to obtain statements similar to theorem 3?

The examples showed that one such minimal assumption is the algebraic interior of the set X being non-empty. Moreover, the examples also showed that the characterization of maxima in theorem 1 is not necessary to write down theorem 3. In other words, *theorem 3 does not rely in any fundamental ways on topological properties of the Euclidean space.*

Thus, we believe that envelope theorems can be obtained for positioning choice problems defined on many other spaces. This "conjecture" is supported by the following example of a positioning choice problem defined on a Riemmanian manifold.

Let M be the ring (circle) obtained by defining the one-point closure of the $(0, 2)$ interval. Let again $f(x) = x \cdot 1_{x \geq 1}$, $g(x, y) = |y - x|$ and $h(x, y) = f(y) - g(x, y)$. For $x \in M$ given the optimisation problem is

$$\max_{y \in M} h(x, y).$$

Notice that by the definition of the one-point closure it is

$$\limsup_{z \in M, z \to 2} f(z) = 2 = \limsup_{z \in M, z \to 0} f(y).$$

The optimal choice correspondence is

$$\sigma(x) = \begin{cases} \{0\} & \text{if} \quad 0 \leq x < 1 \\ [x, 2] & \text{if} \quad 1 \leq x < 2 \end{cases}$$

and so the value function is

$$V(x) = 1 + |1 - x|.$$

This function is almost everywhere differentiable with derivative

$$V'(x)|_{M\setminus\{0,1\}} = \begin{cases} -1 & \text{if} \quad 0 < x < 1 \\ 1 & \text{if} \quad 1 < x < 2. \end{cases}$$

2.5 Conclusion

Envelope theorems are currently formulated to rely on the assumption of continuity of the objective function. In cases where discontinuities need to be analysed this assumption, fundamental to current envelope theorems, forces the development of different approaches.

Accordingly we defined a class of optimisation problems called positioning choice problems for which the value function is almost everywhere differentiable, even when the objective function is discontinuous.

Discontinuous functions do not have well-defined derivatives at their discontinuity points. Thus, standard first-order necessary conditions theorems cannot be used for solving maximization problems with discontinuous objective functions. We have shown that the Dini superdifferential is always well-defined at the maxima of positioning choice problems, even when the objective function is discontinuous. This allows us to state first-order necessary conditions in terms of Dini supergradients in theorem 2.

This characterisation of the maxima lets us examine further properties of the value function of positioning choice problems. We have shown in theorem 3 that these value functions are always almost everywhere *Fréchet* differentiable. *Fréchet* differentiability obtains by Rademacher's theorem because the value functions of positioning choice problems are always a locally Lipschitz map between two Euclidean spaces.

We have found interesting properties of positioning choice problems but further study is required as the argument "Lipschitzianity of the value function implies almost everywhere differentiability" should hold in a more general context. We provided many examples supporting this conjecture, including one on a Riemannian manifold.

We have emphasized clarity over generality and have not modeled positioning choice problems as parametric optimisation problems. In other words, we have not explicitly modeled sets of constraints, which are of great interest in optimisation theory. However, the definition of positioning choice problems can be extended to accommodate such feature. We are planning to do so in further work.

2.6 Appendix A: Omitted definitions

The text always considers the measure spaces $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda_n)$ of finite-dimensions $1 \leq n < \infty$, where $\mathcal{B}(\mathbb{R}^n)$ is the Borel σ-algebra of $\mathbb{R}^n$ and λ_n is the n-dimensional Lebesgue measure. This appendix always considers functions $f : \mathbb{R}^n \to \mathbb{R}$.

A property P of f is said to **hold almost everywhere** if there exists a set $A \in \mathcal{B}(\mathbb{R}^n)$ such that $\lambda_n(A) = 0$ and P is true on $\mathbb{R}^n \setminus A$.

A function $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is **a metric on** $\mathbb{R}^n$ if for every $x, y, z \in \mathbb{R}^n$ it satisfies the following conditions:

- *non-negativity*: $g(x, y) \geq 0$;
- *identity of indiscernibles*: $g(x, y) = 0 \iff x = y$;
- *symmetry*: $g(x, y) = g(y, x)$ and
- *subadditivity* or the *triangle inequality*: $g(x, z) \leq g(x, y) + g(y, z)$.

The notation $||\cdot||$ denotes the usual Euclidean metric.

The function f is said to be **Lipschitz** if there exists a $0 \leq K < \infty$ such that for every $x, y \in \mathbb{R}^n$ it holds

$$|f(x) - f(y)| \leq K||x - y||.$$

Let V be a normed space. A point $\theta_0 \in \Theta \subset V$ is in the **algebraic interior** of Θ, denoted by $cor(\Theta)$, if for every $v \in V$ there exists $\eta_\theta > 0$ such that $\theta_0 + tv \in \Theta$ for all $0 \leq t < \eta_\theta$. If $V = \mathbb{R}^n$ then $cor(\Theta) = int(\Theta)$, where int denotes the interior of a set.

Let V and W be normed vector spaces with $||\cdot||_V$ and $||\cdot||_W$ the induced metric of V and W respectively and let $O \subset V$ be open. The function $f : O \to W$ is ***Fréchet* differentiable** at $x \in O$ if there exists a bounded linear operator $A : V \to W$ such that

$$\lim_{||h||_V \to 0} \frac{||f(x+h) - f(x) - Ah||_W}{||h||_V} = 0.$$

If $V = \mathbb{R}^n$ and $W = \mathbb{R}$ then *Fréchet* differentiability is the "usual" differentiability and A is the gradient $\nabla f(x)$ of f at x.

A proof of Rademacher's theorem requiring minimal knowledge of measure theory can be found in Borwein and Lewis (2010) (Theorem 9.1.2). It is stated here for completeness.

Theorem 9 (Rademacher's theorem) *Let $O \subset \mathbb{R}^n$ be an open subset of $\mathbb{R}^n$ and let $f : O \to \mathbb{R}$ be Lipschitz. Then the function f is almost everywhere Fréchet differentiable.*

2.7 Appendix B: Omitted proofs

Proof of Lemma (1)**:** Recall that f is almost everywhere continuous except on a set of points where it satisfies

$$\limsup_{x \in C,\, x \to y} f(x) = f(y) \quad \text{or} \quad \liminf_{x \in C,\, x \to y} f(x) = f(y)$$

and set the correspondence $\overline{f} : C \rightrightarrows 2^{\mathbb{R}}$ as

$$\overline{f}(y) = \left\{ z \in \mathbb{R} : \text{ either } \limsup_{x \in C,\, x \to y} f(x) = z \text{ or } \liminf_{x \in C,\, x \to y} f(x) = z \text{ or both} \right\}$$

The set

$$D := \{ y \in C : \limsup_{x \in C,\, x \to y} f(x) \neq \liminf_{x \in C,\, x \to y} f(x) \}$$

is a set of measure zero and $\overline{f} = f$ on $C \setminus D$ so the two are almost everywhere equal.

Let x be given. The problem

$$\sup_{y \in C} h(x, y)$$

attains a finite supremum because C is a closed box and it is assumed that

$$\sup_{y \in C} |f(y)| < +\infty \quad \text{and} \quad \max_{(x,y) \in C \times C} |g(x, y)| < +\infty.$$

If this supremum is attained for some $y \in C$ then we are done.

Suppose it is not. Then $\sigma(x) := \arg\max_{y\in C} h(x,y) \subset D$. Pick a $y^* \in \sigma(x)$ and observe that since g is continuous it holds

$$\sup_{y\in C} h(x,y) = \limsup_{y\in C, y\to y^*} (f(y) - g(x,y)) = \limsup_{y\in C, y\to y^*} f(y) - \limsup_{y\in C, y\to y^*} g(x,y) = \limsup_{y\in C, y\to y^*} f(y) - g(x,y^*).$$

Let

$$\overline{f}^*(y) = \begin{cases} f(y) & \text{if } y \notin D \\ \max \overline{f}(y) & \text{if } y \in D. \end{cases}$$

The function $\overline{f}^*$ is a selection of $\overline{f}$ for which

$$f = \overline{f} = \overline{f}^* \qquad \text{almost everywhere}$$

and for $\overline{h}(x,y) = \overline{f}^*(y) - g(x,y)$ it holds

$$\sup_{y\in C} \overline{h}(x,y) = \max_{y\in C} \overline{h}(x,y) = \sup_{y\in C}, h(x,y)$$

as desired. ■

Proof of Theorem 3:
It suffice to show that V is continuous, as it follows immediately by observing that since C is a closed box, g is a metric and f is bounded the value function is also Lipschitz and hence, by Rademacher's theorem, almost everywhere differentiable.

Suppose, by the way of contradiction, that V is discontinuous at $x \in C$. Then there exists an $\alpha > 0$ such that for every sequences $(x_n)_{n\in\mathbb{N}} \subset C$ converging to x, $x_n \to x$, it is

$$|V(x) - V(x_n)| \geq \alpha.$$

By definition this means that

$$|f(y^*(x)) - g(x, y^*(x)) - f(y^*(x_n)) + g(x_n, y^*(x_n))| \geq \alpha$$

for every selection $y^*(x) \in \sigma(x)$ and every selection $y^*(x_n) \in \sigma(x_n)$. Since g is continuous, the previous implies that for n large it is

$$\sigma(x) \cap \sigma(x_n) = \emptyset.$$

Without loss of generality suppose that $V(x) > V(x_n)$ and pick a $y^*(x) \in \sigma(x)$. Since g is continuous it holds that

$$|V(x) - f(y^*(x)) - g(x_n, y^*(x))| \to 0.$$

Thus, there exists a $\beta > 0$ such that for every x_n satisfying

$$||x - x_n|| < \beta$$

it is

$$f(y^*(x)) - g(x_n, y^*(x)) > f(y^*(x_n)) + g(x_n, y^*(x_n)),$$

for every $y^*(x_n) \in \sigma(x_n)$, a contradiction. ■

Proof of fact 5:
Equality (IRVF) follows immediately from the Second fundamental theorem of calculus by observing that since V is Lipschitz it is almost everywhere differentiable (by Rademacher theorem) and there exists a Riemann integrable function v_0 such that for every $a, b \in C$, $a < b$, it holds

$$V(b) - V(a) = \int_a^b v_0(x)dx.$$

Continuity of v_0 is also ensured by the Lipschitzianity of V. ■

Proof of Proposition 6:
As C is closed, f is bounded and g_i are continuous it holds that V is Lipschitz and almost everywhere differentiability follows from Rademacher's theorem.

The last statement is trivial if f is non-decreasing. If f is decreasing somewhere and if $g_i(y) = 0$ for every $y < 0$ then the last statement follows by observing that there always exists an open ball $B(x)$ such that there exists a $y << x$, $y \in B(x)$, for which

$$f(y) = f(y) - 0 = f(y) - g(y, x) > f(x) - g(y, x) = f(x)$$

and V is non-decreasing. ■

Proof of Lemma 7:
Let $\tilde{x} \in \mathbb{R}^n$ be given. Under the assumption that g is a positive and convex transformation

of the Euclidean distance $||\cdot||$ it holds that the optimisation problem

$$\sup_{y\in\mathbb{R}^n} h(\tilde{x}, y) = f(y) - g(\tilde{x}, y)$$

attains a finite supremum. Let $y^* \in \sigma(\tilde{x})$ and define $z = y^* - \tilde{x}$. Since the Euclidean distance is translation invariant g also is. Since f is a plane it holds that for every $x \in \mathbb{R}^n$

$$x + z \in \sigma(x).$$

Hence, V is a plane parallel to f. ■

Proof of Proposition 8: Let x be given and consider the sets

$$Z(x) := \{y \in \mathbb{R}^n : f(x, y) - g(x, y) \geq 0\} \text{ and } \tilde{Z}(x) := \{y \in \mathbb{R}^n : \tilde{f}(x, y) - g(x, y) \geq 0\}.$$

Since f dominates $\tilde{f}$ it holds that $\tilde{Z}(x) \subset Z(x)$. Thus, the problem

$$\max_{y\in\mathbb{R}^n} \tilde{h}(x, y)$$

attains an interior and by definition it holds that $\tilde{\sigma}(x) \subset \tilde{Z}(x) \subset Z(x)$.

This implies that the value function V dominates the value function $\tilde{V}$ and since the former is a plane, the latter is a locally Lipschitz map between two Euclidean spaces. By Rademacher's theorem V is almost everywhere *Fréchet* differentiable, as desired. ■

Chapter 3

Ex-post moral hazard and manipulation-proof contracts

Abstract

We examine the trade-off between the provision of incentives to exert costly effort (ex-ante moral hazard) and the incentives needed to prevent the agent from manipulating the profit observed by the principal (ex-post moral hazard). Formally, we build a model of two stage hidden actions where the agent can both influence the expected revenue of a business and manipulate its observed profit.

We use a novel proof technique to show that manipulation-proofness is sensitive to the interaction between the manipulation technology and the probability distribution of the stochastic output. The optimal contract is manipulation-proof whenever the manipulation technology is linear. However, a convex manipulation technology sometimes lead to contracts for which there is manipulation in equilibrium. Whenever the distribution satisfies the monotone likelihood ratio property we can always find a manipulation technology for which this is the case.

Keywords

Moral hazard, hidden actions, monotone likelihood ratio, security design, fraud, earnings management, window dressing, costly state falsification, positioning choice problem, acceptable manipulations

3.1 Introduction

Ex-post moral hazard arguments have been widely used to rationalize features of real-world contracts. For instance, the earlier literature on financial contracts considers simple models where a borrower can lie about the real profit of a business while hiding money from the lender. Such manipulations provide a theoretical foundation for simple or collateralized debt contracts as optimal securities, as these contracts minimize the incentives to lie [(Attar and Campioni, 2003), (Lacker, 2001)]. Well-known macroeconomic models use a similar argument to microfound a borrowing constraint for the representative firm, as seen in Kiyotaki and Moore (1997)'s and Bernanke et al. (1999)'s famous credit rationing.

Recent literature suggests that the rise of performance-based executive compensation is linked an explosion of accounting scandals during the early twenty-first century, such as Nortel Telecom's. Intuitively, the more CEOs are incentivized with bonuses, shares and options, the more incentives they have to use accounting techniques to make reported profits look higher than they are (Crocker and Slemrod, 2007). In fact, the empirical literature on earnings management consistently observes a positive correlation between earnings management and CEOs' incentive pay. However this correlation may be driven by optimal contracting and is thus likely to be efficient as seen in Sun (2014) and Beyer et al. (2014).

The idea that the optimal contract strictly trades-off between opposite incentives has also found ground in recent literature on securities design. The entrepreneurial financing model of Koufopoulos et al. (2018) shows that bonus contracts, even while inducing manipulation in equilibrium, sometimes dominate debt contracts. Intuitively, debt contracts prevent manipulation perfectly while being incapable of separating entrepreneurial types when there is adverse selection. Bonus contracts do the exact opposite and are thus optimal when separating types is sufficiently valuable.

In other words many strands of literature suggest that it is not always optimal to perfectly prevent the manipulation of observed profit. The implicit implication of such a statement is that unethical behaviours, such as defrauding, are to be expected in a well-functioning economy. That is, if such argument is to be true, then the unintended consequence of high-powered incentives is also to incentivize manipulation, and not much can or should be done to prevent this type of unethical behaviour.

To the best of our knowledge, no previous articles provide a set of general conditions for which the optimal contract entails manipulation in equilibrium. We are not aware of any general treatment which identifies the conditions under which the possibility of ex-post moral hazard is problematic or not. This situation is unfortunate given the strong normative implications of such models.

This articles aims to identify such a set of general conditions. Specifically, we build a general model of two stage hidden actions and try to identify assumptions under which the optimal contract entails manipulation in equilibrium. We interpret the model as a model of entrepreneurial financing where the entrepreneur can burn business's money while having access to hidden borrowing[1].

Following the results of Koufopoulos et al. (2018), we examine in greater details the use of bonuses for the incentivization of hard work. Bonus contracts can be represented as discontinuous functions. Such contracts set the rules of the our two-stage game. The manipulation stage of the game is by definition an optimisation problem for which the first-order conditions are not well-defined for bonus contracts. We circumvent this technical difficulty by using a novel approach developed in previous work (Lauzier, 2020d). We obtain two main results:

1) The optimal contract must be manipulation-proof whenever the manipulation technology is linear. This holds for any distribution of profits. We interpret the linearity of the manipulation technology as a situation where there are no frictions on the hidden borrowing market. This result implies that when the profit can take a continuum of values then the optimal contract is a generalized debt contract with a bounded slope.

2) The optimal contract can entail manipulation in equilibrium when the manipulation technology is convex. When the distribution of profit satisfies the monotone likelihood ratio property and another technical condition we can always find a manipulation technology for which the optimal contract is not manipulation-proof. The convex manipulation technology we consider in the main text can be interpreted as situations where there are frictions on the hidden borrowing market.

Intuitively, contracts with manipulation in equilibrium are optimal when they allow the entrepreneur to commit to a high(er) level of effort and the *expected* waste is small. The manipulation technologies we consider are particularly wasteful. During a manipulation the amount of resources wasted to manipulation can be very large in regard to the total profit made. However, such manipulations are infrequent when the effort is "productive enough", a difficult notion to pin down mathematically. We show that the monotone likelihood ratio implicitly makes the entrepreneurs' efforts "very productive"

[1] The possibility of hidden borrowing is sometimes called *window dressing*.

and often leads to optimal contracts for which there is manipulation in equilibrium. This is important because the monotone likelihood ratio is often assumed in application as a mathematical convenience, and our results imply that it is with loss of generality to ignore the possibility of manipulation in such applications.

We proceed with an extensive literature review. Then section 1 aims to provide the reader with intuition by presenting simple models with three states and two levels of effort.

However, these simple models are not well-suited to stating or proving our results. We therefore present in section 2 a full model with a continuum of states and effort levels. We conclude with a short discussion on the difficulties of evaluating empirically our model's predictions. Appendix A is a primer on stochastic orders, Appendix B contains the proofs omitted in the text while Appendix C further explains the equilibrium concept we use.

3.1.1 Literature review

Many papers have examined the agency problem which arises from an agent's ability to manipulate what is observed by the principal. The precise definition of a manipulation is thus important for the understanding and categorization of the literature. In this article we consider an entrepreneurial financing model where the entrepreneur can burn the enterprise's profit while having access to hidden borrowing while the financier can observe the final profit at zero cost but can never observe the state. This type of manipulation is different to a situation in which the entrepreneur sends a message about the state.

When the entrepreneur can send any message at zero cost he always has an incentive to declare a lower profit while keeping the money, essentially stealing from the business. This possibility leads to a complete failure of the lending market if there are no other mechanisms helping to mitigate the ex-post moral hazard problem[2].

[2]Two main mechanisms are explored in the literature:
Observing the state at a cost: Papers belonging to the Costly state verification (CSV henceforth) literature spawned by Townsend (1979) assume that the financier can observe a business' real profit by conducting an audit. However, auditing is costly and the optimal contract minimizes the expected cost of verification as it is wasteful. Attar and Campioni (2003) provides a complete survey of the CSV literature and its relation to the optimality of debt contracts. The paper also discusses the role of financial intermediaries as delegated monitors in the sense of Diamond (1984) and its relation to the CSV framework.
Collateralization: Another way to prevent the entrepreneur to always leave with the money is collateralizing the loan. That is, the investor uses the entrepreneur's asset as collateral to prevent him from declaring bankruptcy. See Lacker (2001) for an extensive study of the idea. This mechanism is also the micro foundation of the borrowing constraints (credit rationing) that can be found in Kiyotaki and Moore (1997) and Bernanke et al. (1999).

The possibility of stealing the money entails difficulties which are irrelevant to the point we want to make. We thus assume that the entrepreneur can burn the business' money. We show that this simple assumption guarantees that the optimal contract must be non-decreasing[3].

The manipulation technology we consider is similar to the one found in the Costly state falsification literature spawned by Lacker and Weinberg (1989) and Maggi and Rodriguez-Clare (1995). This literature generally assumes that the entrepreneur can send a message about the business' profit while the state is not verifiable by the principal. Lying is costly, but small lies are inexpensive. Crocker and Slemrod (2007)'s model shows that perfectly preventing manipulation represents a prohibitively expensive opportunity cost. Intuitively, this is because manipulation-proof contracts are completely flat and do not allow incentivization of hard work. The optimal contract is thus strictly trading-off between the provision of incentives to exert effort and the prevention of manipulation.

Our results complement the findings of Crocker and Slemrod (2007). We show that the optimal contract might entail manipulation in equilibrium even when the cost of manipulation is "high". This possibility is mainly driven by the necessity to incentivize effort and hence by the assumptions on the distribution of profit. In other words, we show that the results of Crocker and Slemrod (2007) are not driven by the specific manipulation technology they consider but by the assumption that the output satisfies the monotone likelihood ratio property (MLRP henceforth).

Many articles investigate the link between the provision of incentives needed for a CEO to exert effort and earnings management[4]. Sun (2014) argues that the empirical positive correlation between managers' incentive pay and earnings management is likely due to optimal contracting and does not necessarily reflect market inefficiencies. Beyer et al. (2014) analyses in details the optimal contract under earnings manipulation and relates the shape of the contract to the quality of the business' governance. The manipulation technology considered in this literature is akin to the one found in Crocker and Slemrod (2007). Thus, we do not directly address these models, although we conclude with a discussion on the difficulty of empirically evaluating the welfare due to ex-post moral hazard. We refer to Beyer et al. (2014) for an extensive literature review on the

[3]This result is already known in the literature as "Free disposal of output by the seller [entrepreneur] is a common assumption in security design problems, and is used to justify restricting the set of securities to designs for which the seller's [entrepreneur's] payoff is weakly increasing in the asset value."Hébert (2017), square brackets added.

[4]The term "earnings management" broadly refers to the possibility that a business' CEO can use accounting techniques to make a business' profit report appear better than it is.

link between CEOs incentives pay and earnings management.

Our model can be interpreted as a security design problem. The formalization of ex-ante moral hazard borrows from Innes (1990), with the notable difference that we consider general distributions of profit. However, ours is not a "proper" model of security design because we do not explicitly consider the competitive environment in which contracting happens.

More recently, Koufopoulos et al. (2018) shed some new light on manipulation's role in the design of securities. The article considers a model with both ex-ante hidden information (adverse selection) and ex-post moral hazard. The authors ask whether the assumption that returns to the lender must be monotonic is without loss of generality[5]. They call manipulation-proof such types of contracts. They show that conditions exist under which bonus contracts are optimal even though they are not manipulation-proof. This is because bonus contracts allow for separating good and bad types of entrepreneurs.

Our results similarly compliment Crocker and Slemrod (2007) and Koufopoulos et al. (2018). Koufopoulos et al. (2018) assumes that the probability distribution of profit satisfies the MLRP in types. We show that bonus contracts implement a higher level of effort than manipulation-proof contracts in a similar way that they allow the separation of types. In other words, we show that their arguments on manipulation-proof contracts hold in the context of ex-ante moral hazard. Again, this property is driven by the MLRP assumption and not by the particular manipulation technology considered in the various papers.

This article indirectly relates to the literature on the first-order approach's validity and the monotonicity of the optimal contract while the approach is valid. Broadly speaking, this literature shows how the MLRP assumption is instrumental to modelling ex-ante moral hazard problems. This single assumption guarantees that both the underlying optimisation problem is easier to solve and that the optimal contract is monotonic. The ensuing literature made wide use of it to simplify many applied problems. We refer Ke and Ryan (2018a) and Ke and Ryan (2018b) for a recent survey of the literature on the first-order approach and the importance of the MLRP assumption.

We show that if the distribution of profit satisfies the MLRP in effort then there are many cases for which the optimal contract entails manipulations in equilibrium. Our results thus suggest that the MLRP is a stronger assumption than previously thought.

[5]The paper contains a thorough literature review discussion of this monotonicity assumption.

The MLRP precisely captures the notion that an "effort is productive enough" for the acceptance of such unethical behaviours in an well-functioning economy.

The optimisation problem at the manipulation stage of the game does not have well-defined first-order conditions when the contract is a discontinuous function of the profit, i.e. when the contract has bonuses. This features poses many technical challenges while solving for the optimal contract. However in previous work Lauzier (2020d) we defined this type of optimization problem as a positioning choice problem. We have shown in the paper that positioning choice problems have desirable properties; their value function are always Lipschitz and thus almost everywhere differentiable. We will use many of these properties amongst others in section 2 to simplify our main statements' proofs. However, we do not fully introduce the mathematical apparatus needed to solve positioning choice problems and we simply refer the reader to the relevant theorems of Lauzier (2020d).

3.2 Basic models

We present simple models with three states and two levels of effort to provide intuition. We first show how much structure linear manipulation technologies impose on the optimal contract. With these types of technologies the optimal contract must be manipulation-proof, which implies that it is non-decreasing and has a bounded slope. However, manipulation-proofness is not a desirable features in of itself, and we explain how this result is solely driven by the peculiar structure of the manipulation technologies.

We then show that manipulation-proofness is too limiting on the entrepreneur's ability to commit to a high level of effort when the manipulation technology is convex. In a nutshell, it is sometimes better to have manipulation in equilibrium provided that the probability of a manipulation is sufficiently low.

3.2.1 Linear manipulation technologies

An entrepreneur needs to raise capital $Q > 0$ in order to finance a project. The project profit is stochastic; let it be a discrete random variable taking value $0 \leq x_l < x_m < x_h = M$. The entrepreneur can take a costly action $e \in E := \{e_l, e_h\}$ which augments the expected profit of the project, i.e. $\mathbb{E}[X(e_h)] > \mathbb{E}[X(e_l)]$, where $X(e)$ is the stochastic profit given effort level e. Exerting effort is costly, the cost c of effort being non-negative, increasing and convex, with $c(0) = 0$. Further, we assume that every random variable $X(e)$ has full support.

The hidden action $e \in E$ is chosen before the realization of the profit. Then, in stage 2, the entrepreneur observes the realization $x \in \{x_l, x_m, x_h\}$ of the profit and can take a hidden action $z \in \mathbb{R}$. This second hidden action modifies the profit $\overline{x} := x + z$ observed by the financier. Finally, the financier observes $\overline{x}$ and the contract is implemented without renegotiation.

The cost of the hidden action z is parametrized by a function $g(z)$. Consider for the moment that $g(z) = (1 + r)max\{0, z\}$ for $r \geq 0$. We interpret this function as the following manipulation technology: (A) when $z < 0$ then $g(z) = 0$ and the entrepreneur burns the business' money and (B) the entrepreneur can borrow the amount $z > 0$ at the interest rate of r and inject the liquidities into the business therefore inflating the business' observed profit.

Let us assume that the financier's upfront payment is always Q. The contract is therefore a vector $(y_i)_{i=l,m,h}$, where $y_i := y(\overline{x}_i)$ is the entrepreneur's state-contingent share of the profit upon the financier observing $\overline{x}_i$. The financier keeps $\overline{x}_i - y_i$.

The entrepreneur is either risk-neutral or risk-averse, with standard Bernoulli utility u twice differentiable and weakly concave. The entrepreneur also has both outside utility and limited liability normalized to zero. The financier is risk-neutral with opportunity cost of investment $1 + r$, where $r \geq 0$ is the interest rate of the economy. Let us also assume that the financier will never pay more than the maximum profit realization of the project, i.e. that $y_i \leq M$.

At time zero the financier makes a take-it-or-leave-it offer[6]. Denote by $\mathbb{P}[x_i|e]$ the conditional probability of x_i given effort level e. The financier's maximization problem is

$$\max_{(y(x_i+z_i))_{i=l,h,m},e\in E} \sum_i (x_i + z_i - y(x_i + z_i))\,\mathbb{P}[x_i|e] - Q \qquad \text{(Problem Discrete)}$$

$$s.t.\ 0 \leq y(x_i + z_i) \qquad \text{(LL-D)}$$

$$y(x_i + z_i) \leq M \qquad \text{(B-D)}$$

$$\sum_i u\left(y(x_i + z_i) - g(z_i)\right)\mathbb{P}[x_i|e] - c(e) \geq 0 \qquad \text{(IR-E-D)}$$

$$\sum_i (x_i + z_i - y(x_i + z_i))\,\mathbb{P}[x_i|e] \geq (1+r)Q \qquad \text{(IR-F-D)}$$

$$e \in arg\max_{\hat{e}\in E}\left\{\sum_i u\left(y(x_i + z_i) - g(z_i)\right)\mathbb{P}[x_i|\hat{e}] - c(\hat{e})\right\} \qquad \text{(IC-D)}$$

$$\forall x_i,\ z_i \in arg\max_z\{y(x_i + z) - g(z)\} \qquad \text{(IIC-D)}$$

where (LL-D) is the limited liability constraint, (B-D) is the boundedness constraint, (IR-E-D) is the individual rationality constraint of the entrepreneur, (IC-D) is the incentive compatibility constraint defined by stage 2 and (IIC-D) is the interim incentive compatibility constraint defined by stage 3. Let us assume that the individual rationality constraint of the financier (IR-F-D) is never binding so that we can drop it.

Consider the following distribution:

Table 3.1: A distribution satisfying FOSD but not MLRP

	$\mathbb{P}[x_l\|e]$	$\mathbb{P}[x_m\|e]$	$\mathbb{P}[x_h\|e]$
e_l	0.5	0.49995	0.0005
e_h	0.5	0.0005	0.49995

This distribution satisfies the assumption of first-order stochastic dominance (FOSD) in effort but does not satisfy the monotone likelihood ratio property (MLRP). Absent ex-post moral hazard if the entrepreneur is risk-averse and if the financier wants to implement the high level of effort e_h then the optimal contract is non-monotonic, i.e. $y_m < y_l < y_h$. However, this contract is not optimal if we consider the possibility of manipulating the observed profit.

[6] The solution concept is a weak Perfect Bayesian Equilibrium where the entrepreneur takes the action the most favoured by the financier whenever indifferent

Property 1 - monotonicity

Suppose that the optimal contract when there is only ex-ante moral hazard is such that

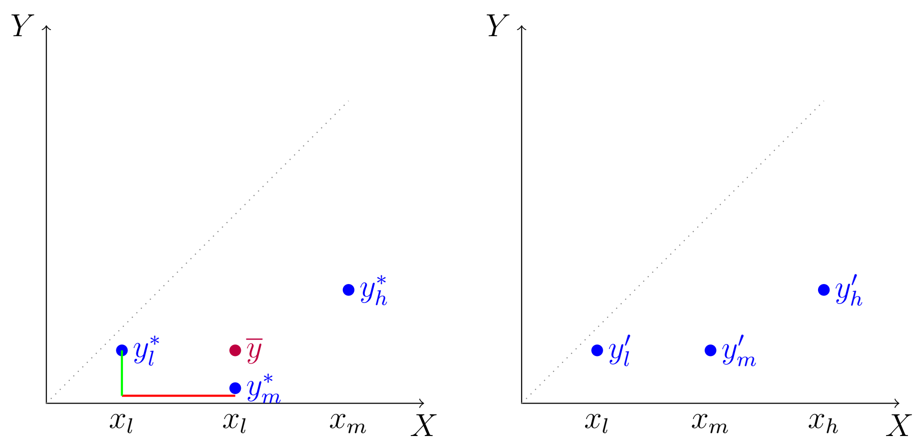
$$y_m < y_l < y_h \text{ and } y_h - y_m < (1+r)(x_h - x_m).$$

This contract is not optimal if the entrepreneur can manipulate the observed profit.

Upon realization of profit x_m the entrepreneur burns amount $x_m - x_l$ and receives $y_l > y_m$. Consider an alternative contract where $y'_m = y_l = y'_l$ and $y'_h = y_h$. This new contract strictly dominates the original contract as it does not induce wasteful manipulations and does not change the incentives to exert effort as the entrepreneur receives state-by-state the same amount with both contracts.

In other words, the value function of the optimisation problem defined at the manipulation stage of the game is non-decreasing whenever the entrepreneur can freely burn money. This implies that any contract which is decreasing somewhere is dominated by a monotonic contract, since replacing the former by by its monotone envelope does not change incentives.

Figure 3.1: Property 1 - monotonicity



Property 2 - bounded slope

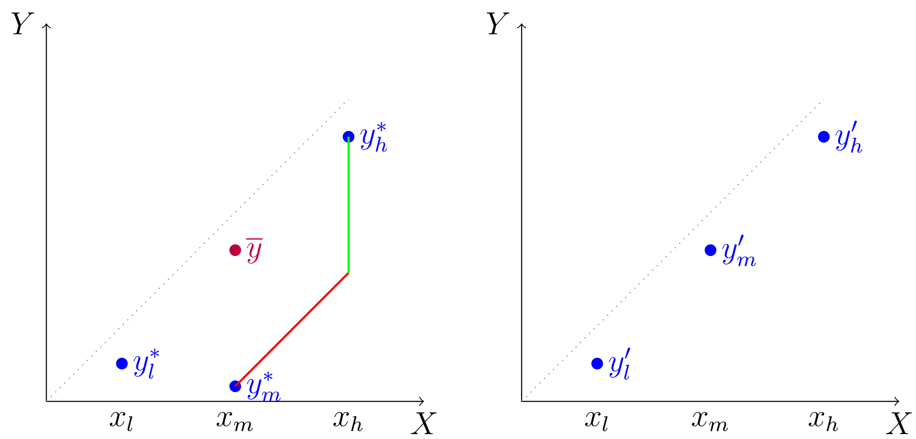
Now suppose that the contract with ex-ante moral hazard only satisfies the following:

$$
\begin{aligned}
&y_m < y_l < y_h\,;\\
&y_h - y_m > (1+r)(x_h - x_m)\,;\\
&y_h - y_m > y_l \quad \text{ and}\\
&y_h - y_l \leq (1+r)(x_h - x_l).
\end{aligned}
$$

Again, this contract is not optimal when the entrepreneur can manipulate the profit.

Updating the contract with the optimisation program's value function at the manipulation stage eliminates wasteful manipulations while leaving incentives intact. This implies that the optimal contract has a bounded slope. However, this is only true when $g(z)$ is linear as this is the only type of manipulation technology for which the entrepreneur receives state-by-state the same amount in both contracts.

Figure 3.2: Property 2 - bounded slope



3.2.2 Convex manipulation technologies

Let us craft an example where the optimal contract entails manipulation. Let

$$
g(z) = \begin{cases} 0 & \text{if } z \leq 0 \\ \tilde{g}(z) & \text{if } z > 0 \end{cases}
$$

for $\tilde{g}$ a strictly convex function satisfying

$$\lim_{z \downarrow 0} \tilde{g}(z) = 0 \quad \text{and} \quad \inf_{z>0} \left\{ \frac{\partial g(z)}{\partial z} \right\} \geq 1.$$

This assumption guarantees that g is continuous and that lying is always "expensive".

Now, for the sake of simplicity suppose that the entrepreneur is risk-neutral and that he makes the take-it or leave-it offer at the initial stage of the game. Let $q \in (0, 0.45)$ and assume the following distribution:

Table 3.2: Another distribution which satisfies FOSD but not the MLRP

	$\mathbb{P}[x_l\|e]$	$\mathbb{P}[x_m\|e]$	$\mathbb{P}[x_h\|e]$
e_l	0.1	0.9 - q	q
e_h	0.1	q	0.9 - q

Best manipulation-proof contract

Let us first consider a manipulation-proof contract $Y = \{y_l, y_m, y_h\}$. The upward interim incentive compatibility constraints are

$$y_m - g(x_m - x_l) \leq y_l \qquad (IIC_{l,m})$$

$$y_h - g(x_h - x_m) \leq y_m \qquad (IIC_{m,h})$$

$$y_h - g(x_h - x_l) \leq y_l. \qquad (IIC_{l,h})$$

Consider the contract for which $IIC_{l,m}$ and $IIC_{m,h}$ are binding, i.e.

$$IIC_{l,m}: \quad y_m - g(x_m - x_l) = y_l$$

$$IIC_{m,h}: \quad y_h - g(x_h - x_m) = y_m.$$

This is the manipulation-proof contract that maximizes the spread between y_l and y_h therefore maximizes the incentives the exertion of high effort[7]. The Individual Rationality constraint of the financier reads

$$\mathbb{E}[X - Y|e_h] \geq (1 + r)Q$$

or

$$\mathbb{E}[X|e_h] - 0.9\,(g(x_m - x_l) + g(x_h - x_m)) + qg(x_h - x_m) \geq (1 + r)Q + y_l. \qquad (IR_F^{MP})$$

[7]By convexity of g the constraint $(IIC_{l,h})$ is slack.

Denote by $y_l(q)$ the highest feasible value of y_l. Since $\mathbb{P}[x_l|e_h] = \mathbb{P}[x_l|e_l]$, the first stage Incentive Compatibility constraint $\mathbb{E}[Y|e_h] - c(e_h) \geq \mathbb{E}[Y|e_l]$ boils down to

$$0.9g(x_h - x_m) - 2q(g(x_h + x_m) + g(x_m - x_l)) \geq c(e_h). \qquad (IC^{MP})$$

Contract with manipulation

Let us consider a contract $Y^M = \{y_l^M, y_m^M, y_h^M\}$ for which $IIC_{l,h}$ is an equality. By the strict convexity of g we have

$$y_h^M = g(x_h - x_l) + y_l^M > g(x_h + x_m) + g(x_m - x_l) + y_l^M$$

which implies that $IIC_{m,h}$ is violated whenever $IIC_{l,m}$ is satisfied. Suppose $IIC_{l,m}$ is satisfied. In equilibrium, x_m is never seen by the financier and its Individual Rationality constraint reads

$$\mathbb{P}[x_l|e_h]x_l + (\mathbb{P}[x_m|e_h] + \mathbb{P}[x_h|e_h])(x_h - g(x_h - x_l)) \geq (1 + r)Q + y_l^M. \qquad (IR_F^M)$$

Denote by $y_l^M(q)$ the highest feasible value of y_l^M that satisfies the previous equality. The first stage Incentive Compatibility constraint becomes

$$0.9g(x_h - x_m) - 2qg(x_h - x_l) \geq c^e(e_h). \qquad (IC^M)$$

Optimal contract

We are now left to show that there exists distributions for which the entrepreneur would prefer proposing Y^M to Y. By construction both contracts satisfy the Individual Rationality constraint of the financier.

The entrepreneur would be better-off proposing the contract Y^M if

$$\mathbb{E}[Y^M|e_h] \geq \mathbb{E}[Y|e_h],$$

i.e. if

$$\begin{aligned} &0.1y_l^M(q) + q(g(x_h - x_l) - g(x_h - x_m)) + (0.9 - q)g(x_h - x_l) \geq \\ &\quad 0.1y_l(q) + qg(x_m - x_l) + (0.9 - q)(g(x_h - x_m) + g(x_m - x_l)). \end{aligned}$$

Suppose that (IR_F^{MP}) and (IR_F^M) holds at equality. The previous equation becomes

$$0.81g(x_h - x_l) - 0.1qg(x_h - x_m) \geq 0.81[g(x_h - x_m) + g(x_m - x_l)].$$

As $q \downarrow 0$, the inequality is strict due to the convexity of g. Moreover, taking $q \downarrow 0$ in (IC^M) and (IC^{MP}) shows that the Incentive Compatibility constraints converge, and are therefore both satisfied provided that $0.9g(x_h - x_m) \geq c(e_h)$. Thus, there exists distributions for which Y^M strictly dominates Y.

3.2.3 Preliminary discussion

These simple examples shed light on the mechanics at play. Manipulations waste valuable resources and avoiding them imposes a lot of structure on contracts. However, this does not imply that the optimal contract necessarily avoids manipulations in equilibrium. This feature depends on the interplay between the "productivity of effort" and the expected waste in equilibrium.

The possibility of burning the money unambiguously implies that the contract is non-decreasing. Burning money is wasteful and no gains in incentives can be made by allowing it. Clearly this feature does not depend on the particular probability distribution of profit we assumed. Theorem 12 of section 2.1 formalizes these observations.

When the manipulation technology is linear the optimal contract does not entail manipulation in equilibrium. However, this feature is entirely driven by the linearity assumption. In this case, any contract can be replaced by the value function of the optimisation problem it defines in the manipulation stage of the game without changing the incentives to exert effort. In other words, any implementable level of effort is implementable with a contract that perfectly prevents manipulation. This is true with any distribution of profit we are considering. Theorem 16 of section 2.2 formalizes this statement, while corollary 17 shows that the optimal contract is a generalized debt contract with bounded slope.

But manipulation-proofness is not necessarily a desirable feature when the manipulation technology is convex. This is because manipulation-proof contracts are sometimes too restrictive upon the incentives to hard work. Under certain circumstances there exist high levels of effort which can only be implemented with a contract which entails manipulation in equilibrium. Contracts with bonuses have the desirable feature of reducing the probability of a manipulation while keeping high powered incentives to hard work.

Intuitively, the effort must be "productive enough" for this to happen, an elusive notion when considering stochastic output. Theorem 18 of section section 3 shows that the MLRP is a somewhat "sufficient condition" for the optimal contract to induce manipulations in equilibrium, although it is not necessary as shown in the example in section 1.2.

3.3 Full model

We now present the full fledged model with a continuum of states and effort levels to prove our main results. We show in section 2.1 that the optimal contract is non-decreasing whenever the entrepreneur can burn the business' money. We also state in this section two lemmas inspired by our previous work, Lauzier (2020d), which help to solve the full fledged model.

We then show in section 2.2 that manipulation-proofness obtains whenever the manipulation technology is linear. As a corollary, we obtain that the optimal contract is a generalized debt contract with a bounded slope. Finally, we show in section 2.3 that when the distribution of profit satisfies the MLRP and another technical condition then we can always find a convex manipulation technology for which the optimal contract entails manipulation in equilibrium.

Let the set of efforts be $E = [0, e_{max}]$ for $e_{max} > 0$ large. The business' profit is a family of continuous random variables[8] $(X(e))_{e \in E}$ with common and full support $[0, M]$. Exerting effort augments the expected profit of the project so that $e > e'$ implies $\mathbb{E}[X(e)] > \mathbb{E}[X(e')]$. Exerting effort is costly, with the cost $c : E \to \mathbb{R}_+$ being increasing, differentiable, (weakly) convex and satisfying $c(0) = 0$.

Let $S = [0, M]$ be a set of states of the world and let $\mathcal{B}(S)$ be the Borel sigma-algebra of S. The family of random variables defined above is thus a family of $X(s, e) : S \times E \to [0, M]$ such that for every $e \in E$ it is

$$\min_{s \in S} X(s, e) = 0 < M := \max_{s \in S} X(s, e) < +\infty.$$

This abstract environment usefully keeps the notation compact.

The entrepreneur needs to borrow the amount $Q > 0$ before starting the project. Let us assume again that the financier is risk-neutral and never pays more than Q. The contract is a transfer function

$$Y = I \circ X \in B_+(\mathcal{B}([0, M]))$$

which depends only on the observed realization of profit $\overline{x}(s)$, where $B_+(\mathcal{B}([0, M]))$ denotes the Banach space (sup-norm) of non-negative bounded functions which are measurable with regard to $\mathcal{B}([0, M])$. This function Y represents the amount received by the en-

[8] The vector notation $(\cdot)$ is used instead of the general $\{\cdot\}$ to emphasize that we will mostly consider families of random variables ordered in at least one of the two stochastic orders defined in appendix A.

trepreneur, with the financier keeping the amount $\overline{x} - Y$.

The entrepreneur first chooses the level of effort and then observes the state $s \in S$. The entrepreneur can then take another hidden action $z \in [-M, M]$ in order to manipulate the profit $\overline{x} = X(s, e) + z$ as observed by the financier. The cost of this hidden action is parameterized by a function $g : [-M, M] \to \mathbb{R}_+$ which represents the manipulation technology. We assume throughout that

$$g(z) = \begin{cases} 0 & \text{if } z \leq 0 \\ \tilde{g}(z) & \text{if } z > 0 \end{cases}$$

for $\tilde{g}$ a (weakly) convex function which is differentiable on $(0, M)$ and which satisfies

$$\lim_{z \downarrow 0} \tilde{g}(z) = 0 \quad \text{and} \quad \inf_{z > 0} \left\{ \frac{\partial g(z)}{\partial z} \right\} \geq 1. \qquad \text{(assumption 1)}$$

The first part of assumption 1 is to guarantee that the cost of a manipulation is a continuous function, while the second is to guarantee that inflating the observed profit is "expensive". We will further discuss the interpretation of this manipulation technology in section 3.

The entrepreneur is either risk-neutral or risk-averse, with Bernoulli utility u weakly concave and differentiable. The entrepreneur also has outside utility $\overline{u} \geq 0$ and limited liability standardized to zero so that $Y \geq 0$. Similarly, we will assume that $Y \leq M$. This boundedness constraint states that the financier never pays the entrepreneur more than the maximum amount which the business can make.

The financier makes a take-it or leave-it offer at the initial stage of the game. The solution concept is a weak Perfect Bayesian equilibrium where the entrepreneur takes the action most favoured by the financier when indifferent[9]. By backward induction the

[9]We consider weak Perfect Bayesian Equilbria in pure strategies (Mas-Colell et al., 1995) with the assumption that the entrepreneur chooses the highest level of effort whenever indifferent and takes the action most favoured by the financier whenever indifferent. See appendix C for more details.

optimal contract satisfies the optimisation program

$$\sup_{Y\in B_+(\mathcal{B}([0,M])),\overline{e}\in E} \int X(s,\overline{e}) + z(s) - Y(X(s,\overline{e}) + z(s))d\mathbb{P} - Q \tag{Problem F}$$

$$s.t.\ 0 \leq Y \tag{LL}$$

$$Y \leq M \tag{B}$$

$$\int u\left(Y(X(s,\overline{e}) + z(s)) - g(z(s)) - c(\overline{e})\right) d\mathbb{P} \geq \overline{u} \tag{E-IR}$$

$$\int X(s,\overline{e}) + z(s) - Y(X(s,\overline{e}) + z(s)))d\mathbb{P} \geq (1+r)Q \tag{F-IR}$$

$$\overline{e} \in \arg\max_e \left\{ \int u(Y(X(s,e) + z(s)) - g(z(s)) - c(e))d\mathbb{P} \right\} \tag{IC}$$

$$\forall s \in S, \quad z(s) \in \arg\max_z \{Y(X(s,\overline{e}) + z) - g(z)\} \tag{IIC}$$

where (E-IR) is the entrepreneur's participation constraint, (F-IR) is the financier's participation constraint, (IC) is the incentive compatibility constraint imposed by stage one and (IIC) is the interim incentive compatibility constraint imposed by stage two. We will assume without loss of generality that the constraint F-IR is redundant.

3.3.1 Monotonicity of the optimal contract

Each contract Y defines a sequential choice of effort e and then of manipulation z. At the manipulation stage both e and s are given so we can write the *optimal choice correspondence* of this stage as

$$\sigma(Y,e,s) = \arg\max_{z\in[-M,M]} \{Y(X(s,e) + z) - g(z)\}.$$

The *value function* of the manipulation stage of the game is

$$V(s;Y,e) = Y(X(s,e) + z(s)) - g(z(s))$$

for $z(s)$ a selection[10] of $\sigma(Y,e,s)$. Intuitively, we want to allow Y being discontinuous because we interpret the upward "jumps" as bonuses. However, this makes characterizing the optimal choice correspondence and the value function much harder.

Fortunately, the optimisation problem of the manipulation stage is a positioning choice problem, a class of optimisation problems which we have defined and examined in detail in previous work. We will simply need to adapt some of the results obtained in Lauzier

[10]A selection f of a correspondence F is a function such that for every $x \in domain(F)$ it is $f(x) \in F(x)$.

(2020d) to our particular setting in order to simplify the treatment.

Without loss of generality we will exclusively consider contracts which are almost everywhere continuous and which satisfy the following technical assumption:

$$\text{for every } x \in [0, M] \text{ it is } \limsup_{x' \to x} Y(x') = Y(x). \qquad \text{(assumption 2)}$$

We can now state two lemmas which help solve problem Problem F. We do not prove them in the current article as they are straightforward applications of our previous work.

Lemma 10 (Continuity of the value function) *Let the function g satisfy assumption 1. Then for every given Y and e the value function*

$$V(s; Y, e)$$

is Lipschitz continuous and thus almost everywhere differentiable.

This lemma follows immediately from Theorem 3 of Lauzier (2020d).

Recall that we currently aim to prove that the possibility of burning the money implies that the optimal contract must be monotonic. The next lemma is a useful intermediate step.

Lemma 11 (Monotonicity of the value function) *Let the function g satisfy assumption 1. Then for every given Y and e the value function*

$$V(s; Y, e)$$

is non-decreasing.

This lemma is Proposition 6 of Lauzier (2020d) applied to the problem at hand. We can now state our first theorem.

Theorem 12 (Monotonicity of the optimal contract) *Any contract solving Problem F is non-decreasing.*

The proof is instructive and will be done carefully. It uses the following notion.

Definition 13 (Monotone envelope) *Let Y satisfy assumption 2. The monotone envelope of the function Y is the smallest non-decreasing function $\overline{Y}$ such that $Y \leq \overline{Y}$.*

Proof of theorem 12: Suppose by contraposition that the contract Y is decreasing somewhere and let e_Y be a level of effort that contract Y implements. By lemma 10 and 11 the value function

$$V(s; Y, e_Y)$$

is continuous and non-decreasing. By assumption 1 it is also the case that $V(s; Y, e_Y) \geq Y$. Consider the alternative contract $\overline{Y}$ defined by the monotone envelope of Y.

Lemma 14 *The contract $\overline{Y}$ implements e_Y and is such that*

$$V(s; Y, e_Y) = V(s; \overline{Y}, e_Y).$$

Proof of lemma 14: Consider $x \in [0, M]$ and redefine the choice correspondence as

$$\sigma(Y, x) = \arg \max_{z \in [0,M]} \{Y(x + z) - g(z)\}$$

and the value function as

$$V(x; Y) = Y(x + z(x)) - g(z(x))$$

for $z(x) \in \sigma(Y, x)$. By definition if $0 \in \sigma(Y, x)$ then $0 \in \sigma(\overline{Y}, x)$ and $V(x; Y) = V(x; \overline{Y})$. It remains to show the cases when $0 \notin \sigma(Y, x)$.

Downward manipulation: If there exists a $z \in \sigma(Y, x)$ such that $z < 0$ then

$$V(x; Y) = Y(x + z) - g(z) = Y(x + z) = \overline{Y}(x)$$

by definition of the monotone envelope. Thus, $0 \in \sigma(\overline{Y}, x)$ and $V(x; Y) = V(x; \overline{Y})$.

Upward manipulation: If every $z \in \sigma(Y, x)$ are such that $z > 0$ then $\sigma(Y, x) = \sigma(\overline{Y}, x)$ and $V(x, Y) = V(x, \overline{Y})$.

We have just shown that for every given level of effort the value functions of the manipulation stage of the game are equal under both contracts. Thus, $\overline{Y}$ implements effort e_Y .$\square$

By construction the contract $\overline{Y}$ satisfies constraint (E-IR) if the contract Y does. Let

$$z(s) \in \sigma(Y, e_Y, s) \quad \text{and} \quad \overline{z}(s) \in \sigma(\overline{Y}, e_Y, s)$$

be selections. The contract $\overline{Y}$ dominates the contract Y since the latter induces downward manipulation which implies that

$$\int X(s, e_Y) + \overline{z}(s) - Y(X(s, e_Y) + \overline{z}(s))d\mathbb{P} > \int X(s, e_Y) + z(s) - Y(X(s, e_Y) + z(s))d\mathbb{P},$$

and Y is not optimal. ■

The critical steps are in lemma 14. Virtually all this article's proofs rely on comparing two contracts and verifying whether or not they implement the same level of effort. Theorem 12 compares a contract to its monotone envelope because lemma 11 implicitly guarantees that they implement the same level of effort.

This monotonicity property is entirely driven by the manipulation technology and does not rely on properties of the distribution of profit. We now show that a similar result is true for linear manipulation technologies.

3.3.2 Linear manipulation technologies and manipulation-proofness

Let us now assume that

$$g(z) = (1 + r) \max\{0, z\} \quad \text{for} \quad r \geq 0. \qquad \text{(Assumption 3)}$$

From corollary 4 of Lauzier (2020d) the following ancillary lemma immediately obtains:

Lemma 15 *Let g satisfy Assumption 3. For every given Y and e the value function*

$$V(s; Y, e)$$

has a slope lesser or equal to $1 + r$.

Lemma 15 allows us to essentially repeat the proof of theorem 12 while using the value function defined by a contract Y instead of its monotone envelope.

Theorem 16 (Manipulation-proof contracts) *Let g satisfy Assumption 3. Then the optimal contract Y is manipulation-proof: for every $x \in [0, M]$ it holds that*

$$0 \in \arg\max_{z \in [-M,M]} \{Y(x+z) - g(z)\}.$$

Moreover, it is continuous, non-decreasing and has a slope lesser or equal to $1 + r$, i.e.

$$0 \leq \frac{\partial Y(x)}{\partial x} \leq 1 + r.$$

whenever this derivative is well-defined.

As mentioned, the proof of theorem 16 is almost identical to the proof of theorem 12. For the sake of brevity we omit it in the main text and refer the reader to appendix B.

It is worth emphasizing again that the proof does not rely on properties of the distribution of profit. The linearity of the manipulation technology entirely drives the result because this is what allows us to replace any contract by the value function of the optimisation problem it defines in the manipulation-stage of the game. Doing so does not change the incentives to exert effort, and we therefore deduce that the optimal contract is manipulation-proof. However, manipulation-proofness is not obtained because a manipulation is a "bad thing" *per see*, but simply because there are no losses in perfectly preventing it.

Some readers might have further interest in the shape of the optimal contract. We conclude this section with a simple corollary which helps characterize it further. Since any continuous, non-negative and non-decreasing function can be written as a maximum we deduce the following:

Corollary 17 *Let g satisfy assumption 3. Then the optimal contract can be written as a generalized debt contract*

$$Y(x) = \max\{0, \alpha(x)x - d\} + w$$

where $d \geq 0$ is a threshold of debt, $w \geq 0$ is a flat wage and $\alpha(x)$ is a non-decreasing and continuous function with slope $\leq 1 + r$.

3.3.3 Convex manipulation technologies

In the last two sections we aimed to characterize the optimal contract in its greatest generality and thus we tried to impose as few assumptions as possible. Our goal now is to show that convex manipulation technologies sometimes lead to contracts which induce manipulation in equilibrium. This is an existence statement, which will allow us to make stronger assumptions in order to highlight the mechanics at play.

Let us assume that the entrepreneur is risk-neutral and that he makes the take-it or leave-it offer at the initial stage of the game. Let the manipulation technology be

$$g(z) = \begin{cases} 0 & \text{if } z \leq 0 \\ \tilde{g}(z) & \text{if } z > 0 \end{cases}$$

for $\tilde{g}$ a strictly convex function which is differentiable on $(0, M)$ and which satisfies

$$\lim_{z \downarrow 0} \tilde{g}(z) = 0 \quad \text{and} \quad \inf_{z>0} \left\{ \frac{\partial g(z)}{\partial z} \right\} = 1. \qquad \text{(Assumption 4)}$$

Finally, let us also assume that the financier will never agree to give the entrepreneur more than the (state-by-state) profit of the business, i.e. that $Y \leq X$.

The optimal contract solves the following optimisation problem

$$\sup_{Y \in B_+(\mathcal{B}([0,M])), \bar{e} \in E} \int Y(X(s,\bar{e}) + z(s)) - g(z(s)) d\mathbb{P} - c(\bar{e}) \qquad \text{(Problem Entrepreneur)}$$

$$s.t.\ 0 \leq Y \leq X \qquad \text{(Feasibility)}$$

$$\int X(s,\bar{e}) + z(s) - Y(X(s,\bar{e}) + z(s))) d\mathbb{P} \geq (1+r)Q \qquad \text{(IR)}$$

$$\bar{e} \in \arg\max_{e} \left\{ \int Y(X(s,e) + z(s)) - g(z(s)) d\mathbb{P} - c(e) \right\} \qquad \text{(IC)}$$

$$\forall s \in S, \quad z(s) \in \arg\max_{z} \{ Y(X(s,\bar{e}) + z) - g(z) \} \qquad \text{(IIC)}$$

Our current goal is showing that probability distributions and manipulation technologies exist for which the optimal contract entails manipulation in equilibrium. We will do so by using bonus contracts, which will define a partition

$$\mathcal{M} = \{[0,a), [a,b), [b,M]\}$$

of $[0, M]$ for which manipulation will be restricted to the middle interval $[a, b)$. Intuitively, these intervals correspond to the three states x_l, x_m and x_h of the example of section 1.2. We will be done by showing that situations where the probability of middle interval is small and the bonus contract implements a strictly higher level of effort than the best-manipulation proof contract can exist simultaneously.

First we claim that the constraint (IR) is binding at any solution Y of Problem Entrepreneur. This claim is standard and we do not prove it in the main text.

Let e_{MP} be the highest level of effort which is implementable with a manipulation-proof contract and let Y^{MP} implement e_{MP}. We want to know if we can find an alternative contract Y and a level of effort e_Y such that simultaneously e_Y cannot be implemented with a manipulation-proof contract and

$$\int Y(X(s,e_Y)+z(s)) - g(z(s))d\mathbb{P} - c(e_Y) > \int Y^{MP}(X(s,e_{MP}))d\mathbb{P} - c(e_{MP}).$$

Assumption 4 implicitly guarantees that every manipulation-proof contract must be continuous and have a slope ≤ 1. The manipulation-proof contract which implement the highest level of effort is thus a simple debt contract representable by the function

$$Y^{MP}(x) = \max\{0, x-d\}$$

for $d \in (0,M)$ solving constraint (IR) at equality.

We want to show that there exists a bonus contract Y^{Bonus} which dominates the contract Y^{MP}. Thus, consider the contract

$$Y^{Bonus}(x) = \begin{cases} 0 & \text{if } x < d' \\ x-\beta & \text{if } x \geq d' \end{cases}$$

for $0 < \beta < d'$ and $d < d' < M$. The value $b = \beta - d' > 0$ is the amount of bonus the entrepreneur keeps upon a realization of profit greater than d'.

Since $\tilde{g}$ is strictly convex the contract Y^{Bonus} defines a partition

$$\mathcal{M}_g = \{[0, d' - \tilde{g}^{-1}(b)), [d' - \tilde{g}^{-1}(b), d'), [d', M]\}$$

for which there is manipulation in the middle interval $[d' - \tilde{g}^{-1}(b), d')$. That is, the function

$$z(x) = \begin{cases} 0 & \text{if } x \in \mathcal{M}_g \setminus [d' - \tilde{g}^{-1}(b), d') \\ d' - x & \text{if } x \in [d' - \tilde{g}^{-1}(b), d') \end{cases}$$

is a selection of the optimal choice correspondence $\sigma(Y,x)$.

As mentioned, these three intervals intuitively correspond to the three states x_l, x_m and x_h we had in section 1.2. Thus, it suffices to show that we can find situations where

$$\mathbb{P}[X(e) \in [d' - \tilde{g}^{-1}(b), d')|e] \to 0, \tag{3.1}$$

a convergence which intuitively corresponds to the limit $q \downarrow 0$ in the example of section 1.2. We can show this by finding a "sequence of increasingly steeper functions" $\tilde{g}$ so that the interval $[d' - \tilde{g}^{-1}(b), d')$ converges to the singleton $\{d'\}$. The assumption that the family $(X(e))_{e\in E}$ consists exclusively of continuous random variables then guarantees the convergence in (3.1).

If Y^{Bonus} implements a higher level of effort than Y^{MP} and if the effort is "productive enough" then we have just shown that the former contract dominates the latter. The notion of an effort level being "productive enough" is elusive. The MLRP is enough for the argument above to be correct, although the example in section 1.2 show that it is not necessary.

Theorem 18 *Let the family $(X(e))_{e\in E}$ be ordered in the likelihood ratio and let $e_{MP} < e_{2nd}$, where e_{MP} is the highest level of effort implementable with a manipulation proof contract and e_{2nd} is the highest level of effort when there is only ex-ante moral hazard.*

Then we can always find a manipulation technology g satisfying assumption 4 for which the solution to problem (Problem Entrepreneur) *entails manipulation in equilibrium: there exists profit realizations $x \in [0, M]$ such that $z(x) > 0$.*

3.4 Discussion

Discussions about theorem 8 will be split, moving at a tutorial pace. First, theorem 8's features will be expanded upon and mapped to various models. These models will then be interpreted in relationship to the various assumptions made. We then provide our own interpretation of the model's normative implications. Discussions will be concluded by explaining the difficulties of assessing the empirical validity of the model.

The possibility that the optimal contract entails manipulation in equilibrium is sensitive to the interplay between the manipulation technology and the stochastic output and thus to the assumptions we have made. It is worth taking a closer look at the proof of theorem 18 to better understand this sensitivity. Theorem 8 relies on two key moving parts, the manipulation technology and the distribution of output. A thorough understanding of both of these moving parts is useful to the interpretation of the model.

First, we want to emphasize that the assumptions we made about the manipulation technologies are quite strong and were in fact binding our hands. Incidentally this discussion will show that theorem 18 is more general than a first glance can tell.

The proof of 18 relies on comparing the best manipulation-proof contract to a slightly modified version of itself. Intuitively, this is a variational argument, as the bonus contract we consider is essentially the best manipulation-proof contract to which we added an upward jump at a well-chosen point. If the new contract implements a higher level of effort and keeps the expected loss of manipulation low then we are done.

We assumed a manipulation technology for which the cost of small upward manipulation is "large", which implied that the best manipulation-proof contract still incentivizes working hard. However, the literature also considers technologies for which the cost of small upward manipulations is "low", for instance by assuming that

$$g(z) = g(\overline{x} - x) = (\overline{x} - x)^2$$

where x is once again the realized profit and $\overline{x}$ is the profit as declared by the entrepreneur. Such type of manipulation technology is used in articles like Crocker and Slemrod (2007) and Sun (2014), where it is interpreted as a situation where the manager can manipulate the firm's accounting profit.

Taking derivative around $z = 0$ we see that

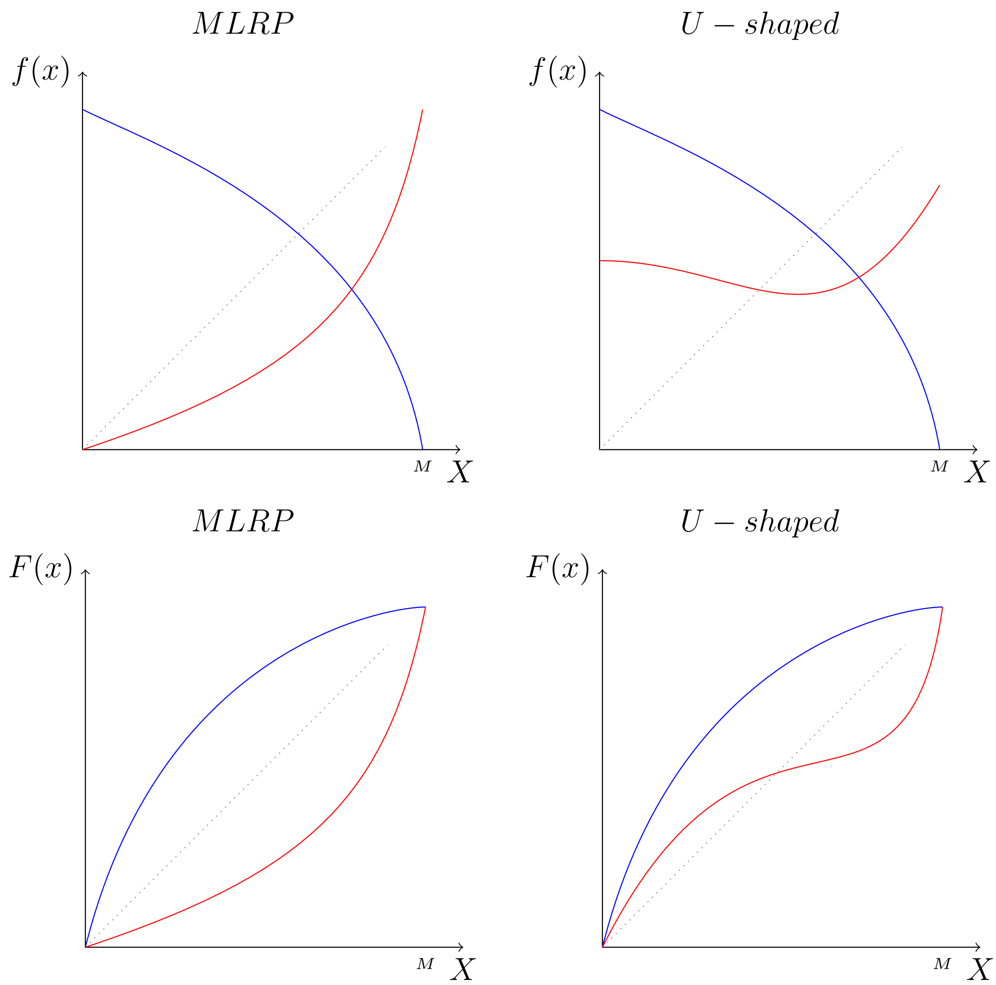
$$\left.\frac{\partial g(z)}{\partial z}\right|_{z=0} = 0 = \inf_{z>0} \left\{ \frac{\partial g(z)}{\partial z} \right\}$$

and *any* upward sloping contract induces manipulation. With such manipulation technology it is impossible to incentivize working hard with a manipulation-proof contract. That is, the optimal contract *always* entails manipulation in equilibrium when the cost of small lies is low and incentivizing hard work is valuable.

Which brings us to the second key moving part of the proof. As mentioned, the bonus contract can be thought of as a local variation of the best manipulation-proof contract, the debt contract. This local change to the contract implements a higher level of effort if the effort "moves enough probability weight from the left to the right" of the distribution. By definition the MLRP does precisely that, and thus is essentially sufficient to show that this local change improves on the original contract.

However, the MLRP condition is not necessary for such perturbation argument to be *globally* true. Indeed, many distributions which do not satisfy the MLRP still exhibit the property that a well-chosen bonus contract implements a higher level of effort than a debt contract. For instance, many "U-shaped" distributions have this property, which is precisely the intuition that helped us build the examples in section 1.

Figure 3.3: MLRP & U-shaped distributions



We would like to conclude by explaining our own interpretation of our results, an interpretation with which the reader may very-well disagree. Piecing together the observations made above we consider that manipulation is often a necessary evil. In our model's restricted world there are many situations where the optimal contract induces certain acceptable manipulations in equilibrium. As mentioned, the model of Crocker and Slemrod (2007) theoretically links the growth of performance-based executive compensation to the explosion of accounting scandals of the early twenty-first century. Ours suggest that such theoretical link is not restricted to high-level executives, as contracts with high-powered incentives are the staples of our modern economy.

Assessing the scale of these acceptable manipulations remains an empirical question. One to which we, the authors, are skeptical can ever be answered precisely. The literature on earnings management consistently observes a positive correlation between CEOs' incentive pays and earnings management. Papers like Sun (2014) use simplified manipulation models to argue that this correlation is likely to be driven by optimal contracting and does not reflect inefficiencies in the market, further evidence that acceptable manipulations exist.

We do not believe that more could be done. Our fundamental objection is one of logical consistency, as assessing precisely the losses due to acceptable manipulations would require that the econometrician observe both hidden actions of exerting effort and profit manipulation. However, we postulate that the Principal cannot observe these actions, as such observation would preclude moral hazard. In other words, evaluating empirically such phenomena with any precision would require for the econometrician to be better informed than the contracting parties, an assumption which is hardly tenable in any situation we can think of.

3.5 Conclusion

The literature on ex-post moral hazard is well established, dating at least to the Costly state verification model of Townsend (1979). However, the subsequent literature considers many different definitions for a manipulation, and the conclusions for each particular model is highly sensitive to the assumptions made about the manipulation technologies.

The recent literature highlights a trade-off between the provision of incentives to work hard and the prevention of manipulation. The importance of this trade-off is supported by the empirical literature on earnings management, which consistently observes a positive correlation between CEOs' incentive pay and earnings management.

Despite many theoretical models and empirical evidence pointing to the existence of such trade-offs, no previous article provides a set of general conditions under which the optimal contract entails acceptable manipulations in equilibrium. This state of knowledge is unfortunate given the strong normative implications of some models, which implicitly imply that unethical behaviours such as fraud are a normal part of a well-functioning economy.

This article sheds light on the interplay between the manipulation technology and the productivity of effort. The optimal contract is non-decreasing whenever the agent can burn the business' profit. This is because burning money is unambiguously wasteful and no gains in incentives can be made by allowing it.

The optimal contract is always manipulation-proof when the manipulation technology is linear. This feature is entirely driven by the linearity assumption, which guarantees that any contract can be replaced by the value function of the optimisation problem it defines in the manipulation stage of the game without changing incentives. In other words, the reason why the optimal contract prevents manipulation is not because a manipulation is "bad" *per see*, but because there are no losses in doing so when the technology is linear. This is not true with convex manipulation technologies.

When the manipulation technology is convex then the optimal contract sometimes entails acceptable manipulations in equilibrium. This feature depends on the specificity of the interplay between the manipulation technology and the "productivity of effort". Intuitively, when hard work is productive enough to be worth rewarding, then upward manipulations are justified, provided they are not so frequent that the expected losses to manipulation stays low. Bonus contracts have the desirable property of incentivizing hard work while maintaining the expected losses to acceptable manipulations low.

A mathematical definition of "productive enough" effort is elusive. We have shown that the monotone likelihood ratio is somewhat of a sufficient condition for the optimal contract to entail manipulations in equilibrium. It is not, however, necessary, as bonus contracts can incentivize a high level of effort for many other types of distributions. However, our results still suggest that the monotone likelihood ratio is a stronger assumption than previously thought, as it is essentially the type of assumption which justifies acceptable manipulations. That is, it is precisely the type of assumption for which it is true that unethical behaviours such as fraud are a normal part of a well-functioning economy.

3.6 Appendix A: Omitted definitions

We collect standard results on stochastic orders. We mainly follow the treatment of Shaked and Shanthikumar (2007) but we also incorporate some results known in the literature. We assume throughout that every random variable has support $\chi \subset [0, M]$ for $0 < M < \infty$.

Let X, Y be two random variables. We say that X **is smaller than** Y **in the usual stochastic order**, denoted by $X \leq_{FOSD} Y$, if

$$\mathbb{P}[X > x] \leq \mathbb{P}[Y > x] \quad \text{for all } x \in \chi. \tag{FOSD}$$

Condition (FOSD) is often called **first-order stochastic dominance**. Let F and G denote the cumulative distribution function of X and Y respectively. It holds that $X \leq_{FOSD} Y$ if and only if

$$G(x) \leq F(x) \text{ for all } x \in \chi \text{ with strict inequality for some } x.$$

Accordingly, we write $F \leq_{FOSD} G$ to denote $X \leq_{FOSD} Y$ when it is not ambiguous. Let $(X(\theta))_{\theta \in \Theta}$ be a family of random variable with parameters $\theta \in \Theta \subset \mathbb{R}$. Let $(F(x|\theta))_{\theta \in \Theta}$ be their corresponding conditional cumulative distribution functions and assume that $F(\cdot|\theta)$ is differentiable in θ. Then $(F(x|\theta))_{\theta \in \Theta}$ satisfy FOSD in θ, $\theta \leq \theta' \Rightarrow F(x|\theta) \leq_{FOSD} F(x|\theta')$, if

$$F_\theta(x|\theta) \leq 0 \text{ for all } x \in \chi \text{ with strict inequality for some } x.$$

Let X, Y be random variables and let f, g be their corresponding density. Let

$$L(x) := \frac{g(x)}{f(x)}$$

be their likelihood ratio. We say that **X is smaller than Y in the likelihood ratio order**, denoted by $X \leq_{LR} Y$, if

$$\frac{\partial L(x)}{\partial x} \geq 0 \text{ for all } x \in \chi \tag{MLRP}$$

where $a/0 := \infty$ whenever $a > 0$. Condition (MLRP) is sometimes called the *monotone likelihood ratio property*, and is equivalent to the condition that

$$f(x)g(y) \geq f(y)g(x) \text{ for all } x \leq y.$$

Integrating the previous equation over $x \in A$ and $y \in B$ for A, B measurable subsets of χ we obtain the following equivalent condition:

$$\mathbb{P}[X \in A]\mathbb{P}[Y \in B] \geq \mathbb{P}[X \in B]\mathbb{P}[Y \in A] \text{ for all measurable sets } A, B \text{ such that } A \leq B$$

where $A \leq B$ means $x \in A$ and $y \in B$ implies $x \leq y$. This last condition is interesting because it does not involve densities and applies uniformly to continuous, discrete or mixed distributions.

Let X and Y have full support and denote by F and G their respective cumulative distribution functions. Then

$$X \leq_{LR} Y \iff G/F \text{ is convex.}$$

Let $(X(\theta))_{\theta\in\Theta}$ be a family of random variables with full support and let $(f(x|\theta))_{\theta\in\Theta}$ be their corresponding conditional densities. Assume that $f(\cdot|\theta)$ is differentiable in θ. Then $(f(x|\theta))_{\theta\in\Theta}$ satisfies the MLRP in θ, $\theta \leq \theta' \Rightarrow X(\theta) \leq_{LR} X(\theta')$, if

$$f(x|\theta)f(y|\theta') \geq f(x|\theta')f(y|\theta) \text{ whenever } x > y \text{ and } \theta' > \theta.$$

The previous condition is equivalent to

$$\frac{\partial}{\partial x}\left[\frac{f_\theta(x|\theta')}{f(x|\theta')}\right] \geq 0 \text{ for all } \theta' \in \Theta \text{ and for all } x \in \chi.$$

We say that $(f(x|\theta))_{\theta\in\Theta}$ satisfies the strict MLRP in θ if the previous inequality is strict. Alternatively, the strict MLRP states that for every $\theta < \theta'$ it is

$$f(x|\theta)f(y|\theta') > f(x|\theta')f(y|\theta) \text{ whenever } x > y \text{ and } \theta' > \theta.$$

Of course the strict MLRP implies that $F(\theta')/F(\theta)$ is strictly convex.

Finally, note that

$$X \leq_{LR} Y \implies X \leq_{FOSD} Y$$

but the converse is not generally true unless $|\chi| = 2$.

3.7 Appendix B: Ommitted proofs

Proof of theorem 6: By theorem 12 the optimal contract is non-decreasing. Suppose by contraposition that the contract Y induces manipulation in equilibrium and let e_Y be a level of effort implemented by Y. Since Y is monotonic the contraposition assumption states that there exists some $x_{s,e_Y} \in [0, M]$ for which every $z \in \sigma(Y, e, s)$ are such that

$$z > 0.$$

Let $V(s; Y, e_Y)$ be the value function of the manipulation stage of the game and let $U \subset [0, M]$ be the set of $x_{s,e_Y} \in [0, M]$ for which the contract Y induces manipulation in equilibrium. By lemma 5, $V(s, Y, e_Y)$ is continuous and has slope $\leq 1 + r$, with equality on U. Consider the alternative contract Y defined by the value function $V(s, Y, e_Y)$, i.e.

$$\overline{Y} = V(s, Y, e_Y).$$

Since g is linear it suffice to prove that $\overline{Y}$ is manipulation-proof to obtain that $\overline{Y}$ implements e_Y and dominates Y, similarly to the proof of theorem 12.

Redefine the optimal manipulation correspondence as

$$\sigma(Y, x) = \arg\max_{z \in [0,M]} \{Y(x + z) - g(z)\} \quad \text{and}$$
$$\sigma(\overline{Y}, x) = \arg\max_{z \in [0,M]} \{\overline{Y}(x + z) - g(z)\}.$$

Notice that since Y is manipulation-proof on $[0, M] \setminus U$ the two contracts are equal on this set and thus $\overline{Y}$ is also manipulation-poof on $[0, M] \setminus U$.

Let $x \in U$ be given. We want to show that $0 \in \sigma(\overline{Y}, x)$. By construction to every $z \in \sigma(Y, x)$ it is $\overline{Y}(x) = Y(x + z) - (1 + r)z$. Suppose by contradiction that $0 \notin \sigma(\overline{Y}, x)$. Since $\overline{Y}$ is monotonic this implies that every $z' \in \sigma(\overline{Y}, x)$ are such that $z' > 0$. However, this assumption implies that

$$\begin{aligned} \overline{Y}(x + z') - (1 + r)z' > \overline{Y}(x) \quad &\Longleftrightarrow \\ [Y(x + z') - (1 + r)z'] - (1 + r)z' > \overline{Y}(x) \quad &\Longleftrightarrow \\ Y(x + z') - 2(1 + r)z' > \overline{Y}(x). & \end{aligned}$$

If $z' \in \sigma(Y, x)$ then the previous equality becomes

$$Y(x + z') - 2(1 + r)z' > \overline{Y}(x) = Y(x + z') - (1 + r)z',$$

an absurd given that $z' > 0$ and $r > 0$. If $z' \notin \sigma(Y,x)$ then there exists a $\tilde{z} \in \sigma(Y,x)$, $\tilde{z} \neq z'$, such that simultaneously

$$Y(x+\tilde{z}) - (1+r)\tilde{z} > Y(x+z') - (1+r)z'$$

and

$$Y(x+\tilde{z}) - (1+r)\tilde{z} = \overline{Y}(x) < \overline{Y}(x+z') - (1+r)z',$$

another absurd. Thus, $0 \in \sigma(\overline{Y},x)$ and we are done. ■

Proof of theorem 8: We begin with a few preliminary claims.

Claim 19 *The Individual Rationality constraint IR must be binding at any solution of problem Problem Entrepreneur.*

Proof of the claim: Let e_Y be the level of effort implemented by contract Y. The claim follows immediately by contraposition observing that if

$$\int X(e_Y) - Y(X(e_Y))d\mathbb{P} > (1+r)Q$$

then there exists an alternative contract $\tilde{Y} \neq Y$ which implements effort $e_{\tilde{Y}} \geq e_Y$, which is feasible because

$$\int X(e_Y) - Y(X(e_Y))d\mathbb{P} > \int X(e_{\tilde{Y}}) - \tilde{Y}(X(e_{\tilde{Y}})d\mathbb{P} \geq (1+r)Q$$

and which strictly dominates Y because

$$\int \tilde{Y}(X(e_{\tilde{Y}})d\mathbb{P} > \int Y(X(e_Y))d\mathbb{P}. \qquad \square$$

Claim 20 *The best manipulation-proof contract is a debt contract, the function*

$$Y^{MP}(x) = \max\{0, x-d\}$$

for $d \in (0,m)$ satisfying

$$\int X(s,e_{MP}) - \max\{0, X(s,e_{MP}) - d\}d\mathbb{P} = (1+r)Q,$$

where e_{MP} is the level of effort implemented by Y^{MP}.

Proof of the claim: The feasibility constraint states that $0 \leq Y \leq X$ and thus it is $Y^{MP}(0) = 0$. By theorem 12 the contract $Y^{MP}(x)$ is a non-decreasing function and since

$e_{MP} < e_{2nd}$ the best manipulation-proof contract implements the highest possible effort level. By Assumption 4 it is

$$\inf_{z>0}\left\{\frac{\partial g(z)}{\partial z}\right\} = 1$$

and Y^{MP} must be continuous and with a slope ≤ 1 in order to prevent manipulations.

Set

$$Y^{MP}(x) = \max\{0, x-d\}$$

for $d > 0$ making constraint IR an equality. Taking derivative we have

$$\left.\frac{\partial Y(x)}{\partial x}\right|_{x\in(0,M)\setminus\{d\}} = \begin{cases} 0 & \text{if } x < d \\ 1 & \text{if } x > d \end{cases}$$

and the monotone likelihood ratio property guarantees that Y^{MP} is the manipulation-proof contract that implements the highest possible effort level. □

We want to show that there exists a manipulation technology for which there is a pair (Y, e_Y) such that Y implements e_Y and Y strictly dominates Y^{MP}, i.e

$$\int Y(X(s, e_Y) + z(s) - g(z(s))d\mathbb{P} - c(e_Y) > \int Y^{MP}(X(s, e_{MP})d\mathbb{P} - c(e_{MP}) \tag{domination}$$

for

$$z(s) \in \sigma(Y, e_Y, s)$$

a selection. Rearranging we obtain

$$\mathbb{E}[Y(X(e_Y)] - \mathbb{E}[Y^{MP}(X(e_{MP})] - [c(e_Y) - c(e_{MP})] > \mathbb{E}[g(z(s))].$$

The assumption that $e_{MP} < e_{2nd}$ guarantees that there exists effort levels $e_{\tilde{Y}} \in (e_{MP}, e_{2nd}]$ for which we can find a contract $\tilde{Y}$ which implements $e_{\tilde{Y}}$ and is such that

$$\mathbb{E}[\tilde{Y}(X(s, e_{\tilde{Y}})] - \mathbb{E}[Y^{MP}(X(s, e_{MP})] - [c(e_{\tilde{Y}}) - c(e_{MP})] > 0.$$

Thus, we will be done if we can find a pair (Y, e_Y) such that $e_Y \in (e_{MP}, e_{2nd}]$ and

$$\mathbb{E}[g(z(s))] \to 0,$$

a convergence which we will define precisely below.

Consider the following bonus contract:

$$Y^{Bonus}(x) = \begin{cases} 0 & \text{if } x < d' \\ x - \beta & \text{if } x \geq d' \end{cases}$$

for $0 < \beta < d'$ and $d < d' < M$. By theorem 2 of Lauzier (2020d) for every given g satisfying Assumption 4 there exists a partition

$$\mathcal{M}_g = \{[0, d' - \tilde{g}^{-1}(b)), [d' - \tilde{g}^{-1}(b), d'), [d', M]\}$$

for which the function

$$z(x) = \begin{cases} 0 & \text{if } x \in \mathcal{M}_g \setminus [d' - \tilde{g}^{-1}(b), d') \\ d' - x & \text{if } x \in [d' - \tilde{g}^{-1}(b), d') \end{cases}$$

is a selection of the optimal choice correspondence

$$\sigma(Y^{Bonus}, x) = \arg \max_{z \in [-M, M]} \{Y^{Bonus}(x + z) - g(z)\}.$$

There exists a net of functions $(g_\gamma)_{\gamma \in \Gamma}$ such that

1. for every $\gamma < +\infty$ the function g_γ satisfies assumption 4;
2. the net $(g_\gamma)_{\gamma \in \Gamma}$ consists of increasingly steeper functions: if $\gamma' > \gamma$ then for every z it is $g_{\gamma'}(z) \geq g_\gamma(z)$, with strict inequality for some z;
3. it is

$$\lim_{\gamma \to +\infty} g_\gamma(z) = g_\infty(z) = \begin{cases} 0 & \text{if } z \leq 0 \\ +\infty & \text{if } z > 0 \end{cases}.$$

The middle interval $[d' - \tilde{g}_\gamma^{-1}(b), d')$ converges to the singleton $\{d'\}$, i.e

$$[d' - \tilde{g}_\gamma^{-1}(b), d') \stackrel{\gamma \to +\infty}{\longrightarrow} \{d'\}.$$

Since the random variables $X(e)$ are continuous it holds for every $e \in E$ that

$$\mathbb{P}[X(e) \in [d' - \tilde{g}_\gamma^{-1}(b), d') | e] \stackrel{\gamma \to +\infty}{\longrightarrow} 0$$

and thus

$$\mathbb{E}[g_\gamma(z(s))] \stackrel{\gamma \to +\infty}{\longrightarrow} 0.$$

As $\gamma \to +\infty$ the Incentive Compatibility constraint IC converges to the Incentive Compatiblity constraint of the problem with ex-ante moral hazard only, i.e. to the Incentive Compatibility constraint of the problem

$$\sup_{Y \in B_+(\mathcal{B}([0,M])), \overline{e} \in E} \int Y(X(s,\overline{e}))d\mathbb{P} - c(\overline{e})$$
$$s.t.\ 0 \leq Y \leq X$$
$$\int X(s,\overline{e}) - Y(X(s,\overline{e})))d\mathbb{P} \geq (1+r)Q$$
$$\overline{e} \in \arg\max_e \left\{ \int Y(X(s,e))d\mathbb{P} - c(e) \right\}.$$

Since the distribution satisfies the monotone likelihood ratio property we can define the bonus contrat Y^{Bonus} such that, in the limit when $\gamma \to +\infty$, the contract implements the effort level $e \in (e_{MP}, e_{2nd}]$ and satisfies the constraint IR at equality. Thus, the inequality (domination) is satisfied in the limit as $\gamma \to +\infty$ and we are done. ■

3.8 Appendix C: Solution concept

This appendix expands on the solution concept we used in the main text and explains further some of the assumptions we made. We focus on the case where the financier makes the take-it or leave-it offer at the initial stage of the game.

Recall that the entrepreneur have to make two sequential choices after being presented with an offer; he first chooses a level of effort $e \in E$, then he observes the realisation $X(s,e)$ ("Nature moves") and finally he chooses a manipulation $z \in \mathbb{R}$. We would like to emphasis that this wording already implies that we are restricting our attention to equilibria in pure strategies.

Now recall that the family $(X(e))_{e \in E}$ consists of random variables which all have common and full support $[0,M]$. We can thereby slightly abuse notation and write the optimal manipulation correspondence $\sigma(Y,e,s)$ as $\sigma(Y,x)$. In other words, our assumption guarantees that the optimal manipulation correspondence is a mapping $[0,M] \to \mathcal{B}([0,M])$.

The optimal manipulation correspondence is rarely single-valued, and not every actions which are payoff equivalent to the entrepreneur are equal. This is better seen by considering a flat part of a contract and noticing that the entrepreneur might be indifferent between burning the money or telling the truth, but that the former action hurts the financier while the latter does not. That is, the latter manipulation Pareto dominates the former. This motivate our focus on equilibria for which the entrepreneur takes the manipulation most favoured by the financier whenever indifferent, which is tantamount to focusing on equilibria for which the manipulation is Pareto efficient.

A similar problem arises for the choice of effort $e \in E$. Let the contract Y be given and denote by $E^*(Y)$ the set of effort that maximize the expected payoff for the entrepreneur. The set $E^*(Y)$ we considered in the text does not need to be a singleton. In other words, a given contract does not necessarily implement only one level effort. This is because we worked with the weakest assumptions on the probability distribution as we can manage. Thereby we considered pure strategy equilibria where the entrepreneur chooses the highest level of effort whenever indifferent. This choice is motivated by our interpretation that working hard is a "good thing" and not by any mathematical properties of the model.

This discussion incidentally sheds lights on our notion of sub-game perfectness. Our assumptions about the entrepreneur's actions at indifference guarantee that the financier's conjecture about the behaviour of the entrepreneur is correct in equilibrium. Without these assumptions there might exist equilibria where this conjecture is incorrect.

Formally let Y be an offer and let $E^*(Y) = \{e_l, e_h\} \subset E$ for $e_l < e_h$. Suppose that the financier's believe that the entrepreneur plays e_h with probability 1, i.e. that his conjecture is

$$\mu^F(e^*(Y)) = \delta_{e_h},$$

where $\mu^F(e^*(Y))$ is the financier's conjecture about the entrepreneur's choice of effort given offer Y and where δ is the Dirac measure. Since $e_h \in E^*(Y)$ there might exists equilibria where the financier offer Y and wrongly believes that the entrepreneur takes the action e_h while the entrepreneur truly takes the action e_l. In other words, without our assumptions on the entrepreneur's behaviour, we would need to be very careful about our definition of belief and sub-game perfectness.

Chapter 4

Insurance design and arson-type risks

Abstract

We design the optimal insurance contract when the insurer faces arson-type risks using a novel proof technique developed in previous work. The optimal contract must be manipulation-proof. The optimal contract is therefore continuous and has a bounded slope. Ipso facto, any contract which mixes deductible and coinsurance is robust to these types of risks.

Keywords

Insurance design, ex-post moral hazard, arson-type risks, positioning, discontinuous optimisation, positioning choice problems

4.1 Introduction

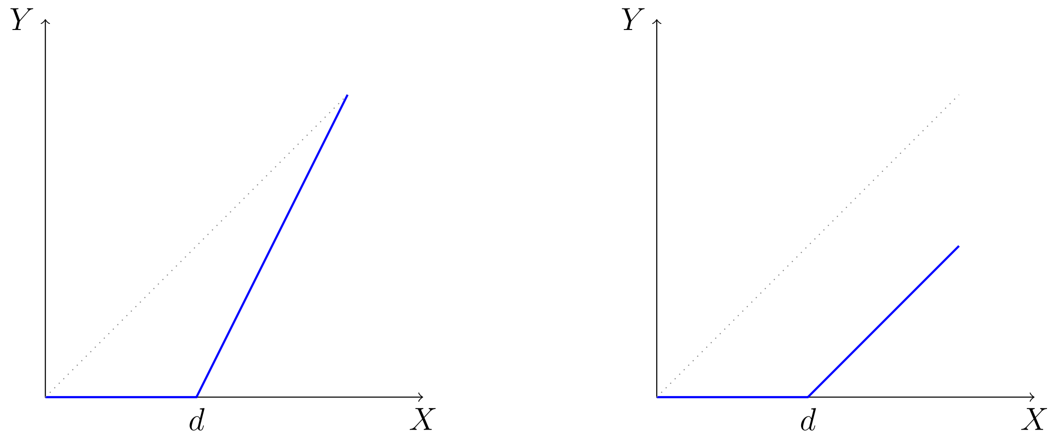
Suppose Bob was just involved in a bicycle accident. After the fact, an officer of the law provided Bob with a certificate indicating that the automobile driver was responsible for the accident. The certificate does not, however, specify how bad the damage inflicted to Bob's steel steed was. Which types of insurance contracts will incentivize Bob to take a sledgehammer to his bicycle before taking a picture and filling his insurance claim?

Huberman et al. (1983) was first to point out that the possibility of inflating an insurance claim by physically destroying an object imposes strong structure on the type of contract which can be offered by an insurer. They show that (completely) vanishing deductibles strongly incentivize the insured to augment the damage. In their model, the next best contract is a simple deductible.

Picard (2000) uses a similar argument to guarantee that his model's optimal contract is continuous. The author also defines the risk of destroying the object as an "arson-type" risk to distinguish it from other ex-post risks of defrauding the contract[1].

We use results from Lauzier (2020d) and Lauzier (2020a) to show that the arguments made in Picard (2000) and Huberman et al. (1983) are two sides of the same coin. We show that when there are arson-type risks the optimal contract must be continuous and have a bounded slope. Ipso facto, any contract mixing coinsurance and deductible will be robust to these risks.

After a brief literature review, we introduce and solve the model then conclude by providing a few observations about our findings' robustness.



(a) completely disappearing (b) simple

Figure 4.1: deductibles

Literature review

Huberman et al. (1983) contains the earliest mention of arson-type risks which we are aware of. The authors analyze the optimal insurance contract when respecting the contract involves non-actuarial costs such as administrative cost. Their model's optimal contract is a completely disappearing deductible, and the authors observe that this type of contract is infrequently observed in real life, if at all. They show that if the insured can cause extra damage then their model's next best contract is a simple deductible.

Picard (2000)'s introduction of arson-type risks is similar to Huberman et al. (1983)'s. Picard (2000) designs the optimal insurance contract when the insured can defraud the

[1]See Picard's chapter in Dionne et al. (2000) for a review of the different notions of fraud risks.

contract and manipulate the audit costs. The model's optimal contract is discontinuous, another oddity infrequently observed in real life. The author then shows that this discontinuity disappears when there are arson-type risks.

Both the contract's continuity and its bounded slope obtain from the same mathematical result. Formally, the manipulation stage of the game defines an optimisation problem. As these contracts can be discontinuous as stated in Picard (2000) the first-order conditions cannot be used. This is where the results of Lauzier (2020d) become necessary. These results also inform that the value function of the manipulation stage's optimisation problem is continuous and has a bounded slope.

Using an argument similar to one made in Lauzier (2020a) we determine that the shape of the contract is continuous and has a bounded slope. This is because replacing a contract by the value function of the optimisation problem it defines does not change the amount of insurance provided. The new contract does not induce manipulation, it is therefore cheaper as the expected waste due to manipulation is priced by the insurer at nil.

While in Lauzier (2020a) we obtain acceptable manipulations we cannot phantom them in the case of insurance contracts. In insurance contracts, the possibility of manipulations simply hurts the insured by lowering the protection by the insurer in equilibrium. Since all contracts mixing coinsurance and deductibles are robust to arson-type risks, there does not seem to be a trade-off between the prevention of arson-type manipulations and the provision of incentives to prevent ex-ante moral hazard. This is exactly the opposite result as that obtained in Lauzier (2020a).

We will discuss Spaeter and Roger (1997) in the main text and we refer to Dionne et al. (2000) for an introduction to the Arrow-Borch-Raviv model.

4.2 Model

The three stage proof will structure the presentation. We start by showing that optimal contracts accounting for arson-type risks must be continuous and have a bounded sloped. This implies that we can substitute our optimisation problem by a simpler one. This new optimisation problem is formally almost identical to the one found in Spaeter and Roger (1997), modulo an extra constraint which is not always binding. The latter allows us to characterize many cases, though the final derivations will be left to the reader.

Notation and problem

We use standard notation throughout. Let S be a set of states of the world and let $\mathbb{P}$ be the probability of state $s \in S$. Let the function $X : S \to [0, M]$ be a continuous random variable representing the risk to be insured against, with X having full support $[0, M]$ and some mass at 0, said mass representing the possibility that no accident ever occurs. Let the function $c : [0, M] \to \mathbb{R}_+$ be the cost of respecting the insurance contract and let $Y : [0, M] \to [0, M]$ be the state-contingent, variable part of the contract[2]. The amount $H \geq 0$ denotes the price of the contract while $W_0 > 0$ is the initial wealth of the insured and $\rho \geq 0$ is the loading factor of the insurer. As usual, the function $u : \mathbb{R}_+ \to \mathbb{R}_+$ is a twice differentiable and strictly concave Bernoulli utility function satisfying Inada conditions.

The game proceeds as follows:

Stage 1 the insured buys the insurance contract Y at price H;

Stage 2 the state s realizes and loss $X(s)$ occurs (Nature moves);

Stage 3 the insured observes the loss and decides to take hidden action $z \in [0, M]$ to augment the damages;

Stage 4 the contract is implemented without renegotiation.

The solution concept is a weak Perfect Bayesian equilibrium where we assume that the insured takes the insurer's most favoured action whenever indifferent[3].

By backward induction the optimal contract (H, Y) solves the following optimization program:

$$\sup_{H \geq 0, Y \in B_+(\mathcal{B}([0,M]))} \int u(W_0 - H - X(s) - z(s) - g(z(s)) + Y(X(s) + z(s)))d\mathbb{P} \qquad \text{(Problem I)}$$

$$s.t.\ 0 \leq Y \qquad \text{(LLI)}$$

$$Y \leq X \qquad \text{(BI)}$$

$$\int Y(X(s) + z(s)) + c(Y(X(s) + z(s)))d\mathbb{P} \leq (1 + \rho)H \qquad \text{(PCI)}$$

$$\forall s,\ z(s) \in argmax_z\{X(s) - z - g(z) + Y(X(s) + z)\} \qquad \text{(IIC)}$$

where (LLI) is the insured's limited liability constraint, (BI) is the "boundedness constraint" stating that the insurer will never pay more than the observed loss, (PCI) is the

[2] So $Y \in B_+(\mathcal{B}([0, M]))$, the space of non-negative and bounded functions (sup-norm) which are measurable with regard to the Borel σ-algebra of $[0, M]$.

[3] We refer to Mas-Colell et al. (1995) for a definition of our equilibrium concept.

insurer's participation constraint, (IIC) is the insured's interim incentive compatibility constraint and the function

$$g(z) = \begin{cases} +\infty & \text{if } z < 0 \\ \beta z & \text{if } z \geq 0 \end{cases}$$

represents an extra cost of inflicting damage (Bob buying a sledgehammer).

We now aim to prove that Y is Lipschitz continuous with constant $\leq 1 + \beta$.

Assumption 1 *We assume throughout that Y is a non-decreasing function.*

The next statement is standard and will not be proved:

Lemma 21 *The insurer's participation constraint* (PCI) *must be binding.*

The interim incentive compatibility constraint defines an optimisation problem that defined in Lauzier (2020d) as a positioning choice problem. It immediately follows from Theorem 3 (p.12) that the value function $V(s)$ is Lipschitz continuous with constant $\leq 1 + \beta$. Paralleling the argumentation of Theorem 6 (p.19) in Lauzier (2020a) we conclude:

Proposition 22 *Any optimal contract Y^* is manipulation-proof: for every $s \in S$ it is $0 \in argmax_z\{X(s) - z - g(z) + Y(X(s) + z)\}$.*

Proof. Suppose, by the way of contradiction, that Y^* is optimal but that there exists a $s \in S$ such that such that for every $z(s) \in argmax_z\{X(s) - z - g(z) + Y(X(s) + z)\}$ it is $z(s) > 0$. Consider the alternative contract $(\tilde{H}, \tilde{Y})$ for which we set $\tilde{Y}(s) = V(s)$, where $V(s)$ is the value function of the optimisation problem (IIC) defined by Y^*. Since $\tilde{Y}(s)$ is Lipschitz with constant $1 + \beta$ it is manipulation-proof so $V(s) = \tilde{V}(s)$, where $\tilde{V}(s)$ is the value function of the optimisation problem (IIC) defined by $\tilde{Y}$. In words, this means that the insured receives state-by-state the same final payoff under both contracts. By lemma 1, it holds that $\tilde{H} < H$ and $(\tilde{H}, \tilde{Y})$ dominates (H, Y^*), a contradiction. ■

We can intuitively understand proposition 2 as stating that arson-type risks are fully priced by the insurer. The insured thus will prefer the cheapest contract as he receives state-by-state the same final amount under both contracts. Or, from Bob's perspective, he was offered two contracts offering the same protection. An expensive one which would allow him to take a sledgehammer to his bike and a cheaper one which did not, so Bob, being a rational person, chose the cheaper option.

Corollary 23 *Any contract Y must be Lipschitz and with slope $\leq 1 + \beta$.*

Furthermore, together, corollary 3 and assumption 1 imply that, by Theorem 3 of Lauzier (2020d), the family of contracts

$$\{Y \in B_+(\mathcal{B}([0,M])) : Y(s) = V(s)\}$$

consist of functions which are almost everywhere differentiable and everywhere directionnally differentiable.

We can thus rewrite (Problem I) as

$$\max_{H\geq 0, Y\in B_+(\mathcal{B}([0,M]))} \int u(W_0 - H - X(s) + Y(X(s))d\mathbb{P} \qquad \text{(Problem S)}$$

$$s.t.\ 0 \leq Y \leq X \qquad (4.1)$$

$$slope(Y) \leq 1+\beta \qquad (4.2)$$

$$\int Y(X(s)) + c(Y(X(s)))d\mathbb{P} = (1+\rho)H \qquad (4.3)$$

under the implicit assumption that $Y \in C^0[0,M]$ [4] is almost everywhere differentiable and everywhere directionnally differentiable. Notice how this problem, as rewritten, is almost identical to the problem studied in Spaeter and Roger (1997) except for the extra constraint $slope(Y) \leq 1+\beta$.

A deduction from the author's work informs when the constraint is binding. While the constraint is not binding the authors' solution is also a solution to Problem S. While the constraint is binding we must resolve Problem S.

Lemma 24 *If $c = 0$ then the optimal contract entails full insurance, i.e. $Y = X$.*

Proof. Observe that for $Y = X$ constraint (4.2) is never binding and for $c = 0$ constraint (4.3) collapses to

$$\int Y(X(s))d\mathbb{P} = (1+\rho)H$$

so problem (Problem S) is the standard Arrow-Borch-Raviv problem. ■

This lemma tells us that (Problem S) becomes interesting only when $c > 0$ somewhere. This is because when $c = 0$ Bob is completely insured this precludes any incentive to take a sledgehammer to his bike.

[4] The space of continuous functions on the close interval $[0, M]$.

Fact 25 *If Y^* solves the reduced problem*

$$\max_{H\geq 0, Y\in B_+(\mathcal{B}([0,M]))} \int u(W_0 - H - X(s) + Y(X(s))d\mathbb{P} \qquad \text{(Reduced problem)}$$
$$s.t.\ 0 \leq Y \leq X$$
$$\int Y(X(s)) + c(Y(X(s)))d\mathbb{P} = (1+\rho)H$$

of Spaeter and Roger (1997) and Y^* satisfies constraint (4.2) *then Y^* solves* (Problem S).

This fact informs us that the only problematic case which must be handled is when the contract Y^* found in Spaeter and Roger (1997) does not satisfy constraint (4.2) somewhere. Intuitively it seems natural to attempt fattening Y^* sufficiently to satisfy $slope(Y) \leq 1 + \beta$ thus solving Problem S. This approach is sometimes fruitful but does not work in certain cases as we explain later. While intermediate cases will be left to the reader we will characterise a simple case when $\beta = 0$.

Lemma 26 *If $\beta = 0$, Y_R^* solves* (Reduced problem) *and Y_R^* is a completely disappearing deductible then the unique solution Y_I^* to problem* (Problem I) *is a simple deductible, i.e. a function of the form*

$$Y_I^*(s) = \max\{0, X(s) - d\}$$

for $d \geq 0$.

Proof. If Y_R^* is a completely disappearing deductible then there exists a $\tilde{s} \in S$ such that for every $X(s) \geq X(\tilde{s})$ it is $Y_R^*(s) = X(s)$. This implies that there is a state $\tilde{\tilde{s}} \in S$ such that for every $X(\overline{s}) \geq X(\tilde{\tilde{s}})$ the constraint (4.2) of problem (Problem S) will be binding and so $slope(Y_I^*(s)) = 1$ on the set $[X(\tilde{\tilde{s}}), M]$. It is easy to check that[5] $slope(Y_I^*(s)) = 0$ on $[0, X(\tilde{\tilde{s}}))$. Setting $d = X(\tilde{\tilde{s}})$ we obtain that Y_I^* can be written as $Y_I^*(s) = \max\{0, X(s) - d\}$. ■

The possibility that people have to take a sledgehammer to their bicycles is priced into insurance contracts is the reason why Bob will never be offered a completely disappearing deductible contract. Simply, though Bob would honestly like to purchase such contract, he cannot. This is because no Bob can commit to being honest.

The intuition that the contract solving problem (Problem S) is simply a "flattening" of the solution to (Reduced problem) is misleading. By additivity of the Lebesgue integral, the insurer's participation constraint states that the insurer should recoup the cost *on average* and *not state-by-state*. This, fortunately, gives us some leeway in solving

[5] Either by "Guessing & Verifying" in (Problem S) or working by contradiction assuming $Y_I^*(s)$ is increasing somewhere on $[0, X(\tilde{\tilde{s}}))$.

(Problem S). However, this also means that there are some cases where we have to roll-up our sleeves and directly attack the problem.

4.3 Concluding remarks

We conclude with a few observations. First, notice that in proving Proposition 2 we never used the underlying probability properties except for the fact that the two integrals in (Problem I) are finite. This means that we could have considered different decision criterion. For instance, by replacing $\mathbb{P}$ with a capacity ν in one integral and integrating in the sense of Choquet instead of Lebesgue. This would have changed nothing to the veracity of Proposition 2, provided that the set of solutions to (Problem I) remains non-empty.

Which brings us to the two problems of existence and uniqueness. Existence was implicitly guaranteed by the formal association with the problem of Spaeter and Roger (1997). Uniqueness is not. While the contracts found in lemma 4, fact 1 and lemma 5 are unique, we were unable to prove uniqueness in the general case. The difficulty comes from the observation that the cost c must be recouped *on average* and not state-by-state. The problem remains open.

Finally, notice the following: (A) we worked by backward induction from the last action taken in the game and (B) the class of contracts robust to arson-type risks contains all contracts mixing co-insurance and deductible. This means that there are no reasons to think *a priori* that complexifying the model would fundamentally change the observations made in the text. In particular, enriching the model with notions of ex-ante hidden action or information would most likely change nothing about the fact that the optimal contract must be robust to arson-type risks.

Bibliography

Attar, Andrea and Eloisa Campioni, "Costly state verification and debt contracts: a critical resume," *Research in Economics*, 2003, *57* (4), 315–343.

Bernanke, Ben S, Mark Gertler, and Simon Gilchrist, "The financial accelerator in a quantitative business cycle framework," *Handbook of macroeconomics*, 1999, *1*, 1341–1393.

Beyer, Anne, Ilan Guttman, and Ivan Marinovic, "Optimal contracts with performance manipulation," *Journal of Accounting Research*, 2014, *52* (4), 817–847.

Borwein, Jonathan and Adrian S Lewis, *Convex analysis and nonlinear optimization: theory and examples*, Springer Science & Business Media, 2010.

Clarke, Francis, *Functional analysis, calculus of variations and optimal control*, Vol. 264, Springer Science & Business Media, 2013.

Clarke, Frank H, "Generalized gradients and applications," *Transactions of the American Mathematical Society*, 1975, *205*, 247–262.

_ , "Generalized gradients of Lipschitz functionals," *Advances in Mathematics*, 1981, *40* (1), 52–67.

Crocker, Keith J and Joel Slemrod, "The economics of earnings manipulation and managerial compensation," *The RAND Journal of Economics*, 2007, *38* (3), 698–713.

Danskin, John M, *The theory of max-min and its application to weapons allocation problems*, Vol. 5, Springer Science & Business Media, 1967.

Diamond, Douglas W, "Financial intermediation and delegated monitoring," *The Review of Economic Studies*, 1984, *51* (3), 393–414.

Dionne, Georges et al., *Handbook of insurance*, Springer, 2000.

Huberman, Gur, David Mayers, and Clifford W Smith Jr, "Optimal insurance policy indemnity schedules," *The Bell Journal of Economics*, 1983, pp. 415–426.

Hébert, Benjamin, "Moral hazard and the optimality of debt," *The Review of Economic Studies*, 2017, *85* (4), 2214–2252.

Innes, Robert D, "Limited liability and incentive contracting with ex-ante action choices," *Journal of economic theory*, 1990, *52* (1), 45–67.

Ke, Rongzhu and Christopher Thomas Ryan, "A general solution method for moral hazard problems," *Theoretical Economics*, 2018, *13* (3), 1425–1481.

_ **and** _ , "Monotonicity of optimal contracts without the first-order approach," *Operations Research*, 2018, *66* (4), 1101–1118.

Kiyotaki, Nobuhiro and John Moore, "Credit cycles," *Journal of political economy*, 1997, *105* (2), 211–248.

Koufopoulos, Kostas, Roman Kozhan, and Giulio Trigilia, "Optimal Security Design under Asymmetric Information and Profit Manipulation," *The Review of Corporate Finance Studies*, 11 2018, *8* (1), 146–173.

Lacker, Jeffrey M, "Collateralized debt as the optimal contract," *Review of Economic Dynamics*, 2001, *4* (4), 842–859.

_ **and John A Weinberg**, "Optimal contracts under costly state falsification," *Journal of Political Economy*, 1989, *97* (6), 1345–1363.

Lauzier, Jean-Gabriel, "Ex-Post moral hazard and manipulation-proof contracts," *Mimeo*, 2020.

_ , "Insurance design and arson-type risks," *Mimeo*, 2020.

_ , "Positioning choice problems: economic applications," *Mimeo*, 2020.

_ , "Positioning choice problems: the mathematics," *Mimeo*, 2020.

Maggi, Giovanni and Andrés Rodriguez-Clare, "Costly distortion of information in agency problems," *The RAND Journal of Economics*, 1995, pp. 675–689.

Mas-Colell, Andreu, Michael Dennis Whinston, Jerry R Green et al., *Microeconomic theory*, Vol. 1, Oxford university press New York, 1995.

Milgrom, Paul and Ilya Segal, "Envelope theorems for arbitrary choice sets," *Econometrica*, 2002, *70* (2), 583–601.

Morand, Olivier, Kevin Reffett, and Suchismita Tarafdar, "A nonsmooth approach to envelope theorems," *Journal of Mathematical Economics*, 2015, *61*, 157–165.

__ , __ , **and** __ , "Generalized Envelope Theorems: Applications to Dynamic Programming," *Journal of Optimization Theory and Applications*, Mar 2018, *176* (3), 650–687.

Picard, Pierre, "On the design of optimal insurance policies under manipulation of audit cost," *International Economic Review*, 2000, *41* (4), 1049–1071.

Shaked, Moshe and George Shanthikumar, *Stochastic orders*, Springer Science & Business Media, 2007.

Spaeter, Sandrine and Patrick Roger, "The design of optimal insurance contracts: a topological approach," *The GENEVA Papers on Risk and Insurance-Theory*, 1997, *22* (1), 5–19.

Sun, Bo, "Executive compensation and earnings management under moral hazard," *Journal of Economic Dynamics and Control*, 2014, *41*, 276–290.

Townsend, Robert M, "Optimal contracts and competitive markets with costly state verification," *Journal of Economic theory*, 1979, *21* (2), 265–293.