NOVA SCHOOL OF BUSINESS & ECONOMICS

Nova School of Business and Economics

Universidade Nova de Lisboa

Dissertation, presented as part of the requirements for the degree of
Doctor of Philosophy in Economics

Three Essays in Applied Microeconomic Theory

Paulo Ignacio Fagandini Ruiz, Student Number 671

A dissertation carried out on the PhD in Economics, under the supervision of
Professor Steffen Hoernig and Professor Guido Maretto.

Submitted: July 30th, 2018
Defended: November 30th, 2018

Abstract

In this thesis, I write about the theory of contracts and auctions. In the first two chapters, I study the principal-agent relationship with moral hazard. The first chapter focuses on the matching between principals and agents, while the second analyzes the implications of the agent having a chance to fix a potential bad outcome before it is revealed to the principal. Finally, in the third chapter, my coauthor and I model first-price sealed-bid auctions, from an ex-ante perspective, to help bidders prepare bids in practice. We study how asymmetries in valuation, information, and sophistication of rivals affect the optimal strategy.

Keywords: Asymmetric Information, Moral Hazard, Matching, Auctions.

Acknowledgments

Without any doubt individual effort and dedication are necessary conditions to succeed in completing a doctoral dissertation, but, by no means a sufficient one. Behind me, I have been lucky to have many people that always supported me, encouraged me to continue, and finally guided me.

First of all, I must thank my supervisors, Steffen and Guido. The confidence and unconditional support I had from Guido allowed me to think freely and to come up with the ideas that laid the foundations of this dissertation. Steffen's rigor and dedication allowed me to turn those ideas and projects into academic papers. I could not have asked for better supervisors. I am in debt to both of you.

The rest of the Faculty who supported me first on my funding search, and later on my job market quests. Here I cannot skip mentioning Fernando Anjos, Miguel Ferreira, Iliyan Georgiev, Susana Peralta, and Vasco Santos.

I am also grateful to Ingemar Dierickx; I realize now how lucky I was for not getting my grant since the first year of the Ph.D.

To the people in the European Center for Advanced Research in Economics and Statistics of the Université libre de Bruxelles. Prof. Bram De Rock and Nancy De Munk who provided me with everything I could need to make the most of my stance at ECARES. Of course, I am forever grateful to Professor Patrick Legros who kindly helped me to move forward with my dissertation significantly.

Thanks to everyone in the Research Office, current and past, as you also played a crucial role in my progress in the PhD. You always were available and did whatever you could to support me with the funding, conferences, submissions, etc.

To Professor Juan Pablo Torres-Martínez for his support from even before applying to the PhD. He continuously offered his assistance and support until the very late stages of the Doctorate.

Finally, nothing could have been possible without my family: Andreia, with her love, patience, and encouragement managed to keep me on track in pursuing my goal of obtaining a doctorate. Without you, I would have never found the strength and motivation to complete such a project. My parents, who were always there from the beginning, always believing in me, and pushing me to achieve more. All I have accomplished in my life I owe to you, to your support when I needed it and to your advice. My sister and brother, who were always on my side, my godparents and all the rest. Knowing that I have the support of my whole family means the world to me.

This work had the support of the Fundação para a Ciência e a Tecnologia FCT under the grant SFRH/BD/105669/2015.

Contents

Chapter 1

Wealth and the Principal-Agent Matching[1]

1.1 Introduction

Wealth can play a major role in the design of optimal incentives, and therefore in the matching between principals and agents. In order to study the compensation of these agents, it is fundamental to understand how they will endogenously match with the firms they end up working for,[2] and the specific role that the characteristics of the contracting parties play in it. I propose a model in which risk neutral agents, characterized by their wealth and talent, match with firms (principals), characterized by their size, to perform a task and in the presence of moral hazard. This model allows to study the contracts being signed by the parties, and the implications that wealth brings to the matching for the traditional results of positive assortative matching between talented agents and

[1]I have to thank comments from Steffen Hoernig, Guido Maretto, Cláudia Custódio, Patrick Rey, Fernando Anjos, Alper Nakkas, Patrick Legros, Takashi Akahoshi, Daniel Cardona, and Antoni Riera. I also appreciate comments and suggestions from participants at the following seminars/conferences: Nova SBE RG, Nova SBE IRW, UECE Lisbon Meetings 2017, DEA-UIB.

[2]Ackerberg and Botticini (2002) shows the pervasive effects of neglecting the endogenous matching when studying incentive contracts empirically.

bigger firms.

Wealth can affect the agent's behavior, and the optimal contract, in two different ways: through the agent's risk aversion or through limited liability. While the scarce literature studying the effects in compensation focused on the former (Chade and de Serio, 2014, Thiele and Wambach, 1999), this article concentrates on the latter. Limited liability prevents the principal from selling a high participation in the company to the agent in order to achieve an output closer to the one without information asymmetries.[3] Having this in mind, I raise the question: How do principals and agents match? In particular: Do wealthier agents match with larger firms or the opposite? Can be wealth and talent be assigned to agents in such a way that the positive assortative matching of talented agents working in bigger companies does not hold?

Even though the empirical literature studying the matching between principals and agents has focused on the agent's talent as the main driver, wealth, through limited liability, can play a major role.[4] It is a standard result in the moral hazard literature that under asymmetric information the principal can achieve the first best level of effort with a risk neutral agent that is not cash constrained, allowing the principal to define negative transfers to the agent in case of bad outcomes, or as the traditional interpretation indicates, an up-front cash payment from the agent to the principal. This can be observed for example in the franchise model, or in the cab business. This scenario is not common in the corporate world (see Baker, Jensen, and Murphy (1988)). This can be due to the lack of wealth from the agents to make the transfers, legal or social limitations, or it can be simply interpreted in a different manner, for example as

[3]Shetty (1988) studied the effect that limited liability has on the contracts between tenants and landlords.

[4]Dam and Perez-Castrillo (2006), to my knowledge, is the only that explicitly considered the role of wealth in a traditional principal-agent matching setup. However, they consider homogeneous principals, while I allow for heterogeneous principals. I also allow for agents to differ in talent. Legros and Newman (1996) considers a problem of matching agents of the same type with different wealth levels, that get together to form firms between them. Wealth plays an important role in this.

an obligation for the executives to buy shares from the company. Other professions do face these kind of contracts more explicitly. For example, lawyers, when promoted to partners, must buy their partnership with money from their own pockets. Physicians, in some countries, must buy a participation in the clinics where they receive their patients. In general any kind of partnership will involve transfers from the agents to the principal.

As mentioned, in the absence of limited liability the principal can set up a contract that achieves a first best outcome, by selling the firm to the agent for the whole expected surplus leaving the agent without any information rents. This is no longer true once the agent is cash constrained. I focus my study on how this channel affects the matching between agents and principals. In a first stage I study the isolated version of the principal-agent model, finding that the principal's utility increases in matches with wealthier agents, whereas the agent's utility increases by working in bigger firms. This suggests that a positive assortative matching is to be expected, that is wealthier agents should work in bigger firms.

Because utility is not perfectly transferable in the principal-agent model, as the sharing of the surplus affects the strength of the incentives, it is necessary to use the concept of generalized increasing differences, introduced by Legros and Newman (2007), in order to analyze the matching. Doing so, requires to study carefully the utility possibility frontier (UPF) that arises from the model. After doing that, I find that there is positive assortative matching (PAM) between principals and agents where the types for principal is the firm size, and the agents' is their wealth. The same PAM is valid when the agents' type is their talent: more talented agents work for bigger firms.

Then I describe the contracts that each pair will sign and the expected output they will produce. I also consider gains in efficiency (by generating contracts closer to the first best for some matches) that are generated by the market pressure introduced by competition, that is, the fact that other principals could offer more convenient contracts to my own agent. In particular, I provide conditions under which high-type principals give larger incentives to their agents than what they give in the isolated version of the model. For this market pressure to be effective, though, firms should have similar sizes.

Finally I give an example in which a wealthy agent has lower skills than a poor one, a situation in which generalized increasing differences may not hold anymore, and therefore, neither needs to hold the positive assortative matching between principals and agents when considering talent as the agent's type. This situation can be less rare as expected, as if talent is similar, the correlation between talent and wealth can be less clear. This situation can also apply to agents in the beginning of their careers. As a consequence, the assumptions made on the correlation between wealth and talent are critical to claim that, for example, there is positive assortative matching between bigger companies and more talented agents.

In Section 1.2 I describe and solve the baseline principal-agent model I use in this article, to proceed with the analysis of the utility possibility frontier and the matching analysis in Section 1.3. In Section 1.4 I provide an example of a distribution of wealth and talent in which the positive assortative matching does not hold. Finally, in Section 1.5 I present a discussion on how this model fits the literature, to conclude in Section 1.6.

1.2 The Principal-Agent Framework

In this section I develop a slightly modified version of the classical model of moral hazard with risk neutral agents.[5] The modifications I introduce allow to capture the effects of different levels of cash constraint for the agent and the size of the firm in the

[5]See Bolton and Dewatripont (2005), Laffont and Martimort (2002), and Salanié (2005) for examples.

optimal compensation scheme and welfare allocation, by parameterizing the model in the firm's size, agent's wealth, and the agent's talent.

Both principal and agent are assumed to be risk neutral. I assume that the firm's output x can have two states, $x \in \{0, \xi\}$, where $\xi \in [0,1]$ reflects the firm's size.[6] The agent's effort e can be chosen between 0 and 1, and affects positively the probability of success — having $x = \xi$ as the outcome — in the following way: $Pr(x = \xi|e) = p(e) = e$.[7] Let a and b be the base wage and the bonus respectively. Although a is paid independently of the outcome, b is paid by the principal only if $x = \xi$ is observed. Effort is costly for the agent, and is represented by the cost function $c(e) = e^2/(2\tau)$. Here $\tau \in (0,1)$ measures the ability - or talent - of the agent: a higher τ implies lower effort cost. Assume that the reservation utility for principal and agent are 0 and $\overline{u}$ respectively.[8] Let $\overline{u}$ be the maximum between 0 and whatever he can get by working somewhere else. Also assume that the agent has personal wealth $\omega \geq 0$. The agent's cash constraint is therefore determined by $-\omega$, meaning that the principal can never set wages such that the agent transfers more than ω to her.

First, the principal makes a take-it-or-leave-it offer to the agent, specifying the base wage and the bonus for success; the agent accepts or rejects the offer; conditional on accepting it, he decides how much effort to exert. Finally the outcome is realized and the principal makes the transfer to the agent.

The maximization problem of the principal is given by:

[6] Considering ξ as a parameter allows to use it as firm size. For example in Gabaix and Landier (2008), the authors mention that variables such as earnings or capitalization can be considered as firm size, and Bandiera, Guiso, Prat, and Sadun (2015) consider the number of employees as firm size. Given the characteristics of this single-period model, defining size as the earnings in the good state avoids the inclusion of more variables that would needlessly complicate the model.

[7] The parameterization of the model in the other variables makes the constraint for effort, to be lower than 1, not binding.

[8] By the nature of the problem, I consider the reservation utility for the agent to be the minimum change in utility he is willing to accept. Note that if the reservation utility $\overline{u}$ is included in the participation constraint, for the risk neutrality, it can be canceled out.

$$\max_{e,a,b} \quad -a + p(e)[\xi - b] \tag{1.1}$$

$$s.t. \quad a + p(e)b - c(e) \geq \overline{u} \tag{PC}$$

$$e \in \arg\max_{\hat{e}} \{a + p(\hat{e})b - c(\hat{e})\} \tag{IC}$$

$$a \geq -\omega \tag{CC}$$

Equation (PC) is the participation constraint that ensures that the agent will accept the contract proposed by the principal.[9] Equation (IC) represents the incentive compatibility constraint that ensures that the agent will choose endogenously what the principal has chosen as optimal effort, and finally (CC) represents the cash constraint. Usually in the literature (CC) is represented as $a \geq 0$ to model a cash-constrained agent, or in other words, a situation with limited liability. In this article, instead, we allow for different levels of wealth.

The solution to this problem depends on the value that ω takes. In fact, it is well known in the moral hazard literature that if ω is high enough (and this will be shown in the model as well), the principal can achieve first best effort with a contract that is equivalent to selling up-front the outcome to the agent.[10] Doing so, he can extract the whole expected surplus, minus $\overline{u}$.

When solving the problem not all the constraints are going to be binding. If the (CC) is too severe, then the (PC) is not binding. Conversely if the agent has large wealth, and/or a high reservation utility, one finds that only the (PC) is binding. There are intermediate cases in which both constraints are binding, and the solution is given by

[9] The participation constraint would be indeed $\omega + a + p(e)b - c(e) \geq \overline{u} + \omega$, as the agent's utility depends on his own wealth and on his outside option, but ω cancels out on both sides.

[10] First best refers to the situation in which there is no asymmetry of information between the agent and the principal.

the system of equations provided by the constraints of the problem. Because this model is quite standard in the literature I leave its detailed derivation for Appendix A.1 and directly present the main results in Table 1.1. The utility of the principal is represented by v, whereas the agent's utility is represented by u.

	Binding Constraint		
Variable	(CC) $\omega + \overline{u} < \frac{\xi^2\tau}{8}$	(CC) and (PC) $\frac{\xi^2\tau}{8} \leq \omega + \overline{u} \leq \frac{\xi^2\tau}{2}$	(PC) $\frac{\xi^2\tau}{2} < \omega + \overline{u}$
e	$\frac{\xi\tau}{2}$	$\sqrt{(\overline{u}+\omega)2\tau}$	$\xi\tau$
a	$-\omega$	$-\omega$	$\overline{u} - \frac{\xi^2\tau}{2}$
b	$\frac{\xi}{2}$	$\frac{\sqrt{(\overline{u}+\omega)2\tau}}{\tau}$	ξ
$E[u]$	$\frac{\xi^2\tau}{8}$	$\overline{u}+\omega$	$\overline{u}+\omega$
$E[\Delta u]$	$-\omega + \frac{\xi^2\tau}{8}$	$\overline{u}$	$\overline{u}$
$E[v]$	$\omega + \frac{\xi^2\tau}{4}$	$\xi\sqrt{(\overline{u}+\omega)2\tau} - 2\overline{u} - \omega$	$\frac{\xi^2\tau}{2} - \overline{u}$
Δ Surplus	$\frac{3\xi^2\tau}{8}$	$\xi\sqrt{(\overline{u}+\omega)2\tau} - \overline{u} - \omega$	$\frac{\xi^2\tau}{2}$

Table 1.1: Solution to problem in (1.1). Δu and Δ Surplus represent the induced changes in both variables.

It can be immediately observed that when only (PC) is binding the first best can be implemented, by selling the firm to the agent at his reservation utility minus the expected surplus, and letting him keep the whole outcome ($b = \xi$). The agent receives utility $\overline{u}$ and the implemented optimal level of effort is $\xi\tau$. When only the (CC) is binding, the agent obtains information rents, defined as whatever the agent obtains above his reservation utility. The optimal effort is reduced to half when compared to the first best and the surplus has dropped by 25%. What happens in between? The agent gets utility $\overline{u}$ as the (PC) is binding, and also $a = -\omega$, as the (CC) is binding. Note that a lower value of ω implies an increase in the fixed wage that would lead the agent to increase his utility; nevertheless, a decrease in ω leads also to a decrease in the bonus and the implemented effort, keeping the agent at his reservation utility level. The principal, though, is strictly worse off, as less is being produced and she keeps a share of a smaller surplus.

The conclusions of this simple principal-agent model can be summarized as:

1. The higher the agent's wealth, the larger has to be the firm for him to be considered cash-constrained.

2. If cash-constrained, the agent's information rents decrease with his personal wealth.

3. The bigger the firm, the larger the set of agents that are to be considered cash-constrained.

4. The larger the firm the (weakly) higher the bonus for the agent and the expected utility for principal and agent, however the bonus is (weakly) lower as a share of the firm's profits.

5. If the agent is not cash constrained the firm is indifferent as to the level of agent's wealth and the outcome is first best.

Baker and Hall (2004) finds, in an empirical work, that the size of the bonus for the agent is smaller, as a percentage of the firm, the bigger the firm. This is consistent with point 4. The model predicts that when the agent is cash constrained, the bonus b is linear or constant on the firm's size, and therefore there is a decreasing relationship between the share of profits the agent is keeping and the size of the firm.

We haven't focused on the comparative statics of τ. However it is worth mentioning that the results are in line with the findings in the literature (Gabaix and Landier, 2008) in the sense that more talent leads to greater compensation.

Two results are to be learned from this. First an agent would always like to work for a firm in which he is cash constrained, as he can extract more surplus. A good way to do that is to work in a big company, as this makes him more cash-constrained, for a given ω, and also increases his expected utility, as he is getting a share of a bigger surplus. On the other hand, a firm would like to hire wealthy agents (or less cash-constrained or even better an agent that is not cash-constrained at all), as then the principal can keep most or all of the surplus. For the principal, how wealthy the agents needs to be for her to be able to keep the whole surplus, depends on the size of the firm: the smaller the firm, the less wealthy the agent needs to be, and therefore smaller firms would be indifferent between less and more wealthy agents, as long as both could buy the firm. The agent, however, will always look for bigger firms to work for. This result is consistent with the findings of Dam and Perez-Castrillo (2006).

1.3 The Utility Possibility Frontier and the Matching

In this section we will study assortative matching between principals and agents. In a broad sense we have assortative matching when the matching between economic agents is determined by their types. For example, in our case, a result of positive assortative matching (PAM) implies that wealthier agents match with bigger firms. Negative assortative matching (NAM) means exactly the opposite, i.e., that wealthier agents work in smaller firms. It is useful to remind the reader at this point that I have defined $\overline{u}$ as the maximum between 0 and the agent's outside option. Now the outside market option is going to be an input in the utility possibility frontier (UPF), and describes how much utility the agent is obtaining with a particular contract.

In the previous section, I provide arguments suggesting that PAM is a reasonable matching outcome to expect when the agent's type is wealth. When the agent's type is talent, it is considered standard in the literature that more talented agents match with higher types of principals. Consistently with the assumption in the previous section, that the principal makes a take-it-or-leave-it offer to the agent, I assume here that there are more agents than principals.

It can be shown that the surplus of a match is supermodular, in the above model, in the principal and agent's types, which quite often is enough to have PAM in a matching problem with transferable utility. However, in the principal-agent setup, utility is not perfectly transferable. In a contract, the principal can find many ways to transfer utility to the agent, and the case of perfectly transferable utility would be simply a money transfer to the agent (a higher fixed wage a). This is not an optimal alternative for the principal. She will instead set a higher bonus b, incentivizing the agent to exert a higher level of effort, and thus increasing the surplus. The transfer of utility to the agent is financed in part by giving up a share of the surplus, but also by the increase of it, so the utility loss for the principal is lower than the gains in utility for the agent.

Legros and Newman (2007) [LN henceforward] introduced a methodology to address the matching problem when utility is not perfectly transferable. They consider a match between individuals type R and S. Let $s > s'$ and $r > r'$ be two different types within each category, and the outcome be a match between an individual from R with another individual from S. They argue that r is going to match with s if r can outbid r' for s. In particular, they introduce the concept of generalized increasing differences (GID) which relies on the utility possibility frontier (UPF) in order to determine the matching between individuals. The UPF describes the combinations of utilities u and v that are Pareto efficient.[11] For the sake of consistency I will use the same notation for the UPF, outlined in the following definition:

Definition 1. *Let $\theta = (\omega, \tau)$ be the agent's type, and ξ be the principal's type. The utility possibility frontier is described by the following functions:*

- *Let $\phi(\xi, \theta, u)$ be the maximum utility that a principal type ξ can obtain when matching with an agent of type θ, and this agent gets utility u.*

[11]That is, neither principal nor agent can have higher utility without the other being worse off.

- *Let $\psi(\theta, \xi, v)$ be the maximum utility that an agent type θ can obtain when matching with a principal of type ξ, and this principal gets utility v.*

Thus, ϕ and ψ represent the UPF from the points of view of the principal and agent respectively, and for a bijective UPF and given types, one is the inverse of the other, with respect to utility levels.

The main result in LN - that GID implies PAM - intuitively states that the low type principal, for any level of utility he can get by matching with the high type agent, will provide a certain utility level for each type of agent (high or low type). In order for the high type principal to keep the high type agent, she should be able to outbid the other principal, that is, to provide at least as much utility to the high type agent as he would receive from the low type principal. This will only occur if the principal gets more utility by providing that utility to the high type agent than what she would get by giving the low type agent what he gets from the low type principal. By doing this the high type principal can outbid the low type principal and PAM will arise as the market outcome.

In the model with only one principal and one agent, developed in section 3, it has been highlighted that, for the optimal contract, there are 3 possible situations: only (CC) is binding, only (PC) is binding, or both (CC) and (PC) are binding. The (IC) is always binding. From table 1.1, we can write down an analogous table for the UPF (Table 1.2).

	CC	PC & CC	PC
$E[\Delta u] = \psi(\omega, \xi, v)$	$\frac{\xi^2\tau}{8} - \omega$	u	u
$E[v] = \phi(\xi, \omega, u)$	$\omega + \frac{\xi^2\tau}{4}$	$\xi\sqrt{(u+\omega)2\tau} - 2u - \omega$	$\frac{\xi^2\tau}{2} - u$

Table 1.2: Utility Possibility Frontier.

The situation when only the (CC) is binding represents a single point on the UPF. This happens because the agent is getting a fixed rent, and the (PC) is not binding.

If we want to move to the right along the UPF we need to give higher utility to the agent, and then the (PC) starts binding. The (CC) is binding until the unconditional transfer a implied by when only (PC) is binding exceeds $-\omega$ (in other words, (CC) stops binding), leading to:

$$\underbrace{\overline{u} - \frac{\xi^2\tau}{2}}_{a_{pc}} \geq \underbrace{-\omega}_{a_{cc}} \tag{1.2}$$

or when looking at the UPF, when $u \geq \xi^2\tau/2 - \omega$. For all the values of $u \in [\xi^2\tau/8 - \omega, \xi^2\tau/2 - \omega]$, both constraints are binding. Of course the case can appear in which $\xi^2\tau/8 \leq \omega$, or even $\xi^2\tau/2 \leq \omega$, nevertheless, as mentioned earlier, I assume that the agent cannot obtain negative expected utility, and therefore $\overline{u} \geq 0$. In particular for the UPF I assume that $\overline{u} = 0$.

What is different to the usual UPFs is that this UPF's domain varies with ω, τ and ξ. For example, when ω is small, it considers values of u that are strictly positive. This represents a small variation with respect to LN's UPF, as they assume that its domain is between 0 and some upper bound for both individuals. On the other hand, a bigger company, or a more talented agent increases the expected value of the company, making an agent cash constrained, and therefore creating information rents. To address this fact, I define the following correspondences:

Definition 2. *Given the UPF, define:*

- $\underline{u}_\theta(\xi) := \max\{0, \frac{\xi^2\tau}{8} - \omega\}$ *as the minimum level of utility that an agent of type θ can get from a match with a principal of type ξ.*

- $V_\xi(\theta) := [0, \phi(\xi, \theta, \underline{u}_\theta(\xi))]$ *as the feasible utility levels that an optimal contract can give to a principal whose firm has size ξ contracting with an agent of type θ. By* *definition, $V_\xi(\theta)$ is also the domain of $\psi(\theta, \xi, \cdot)$.*

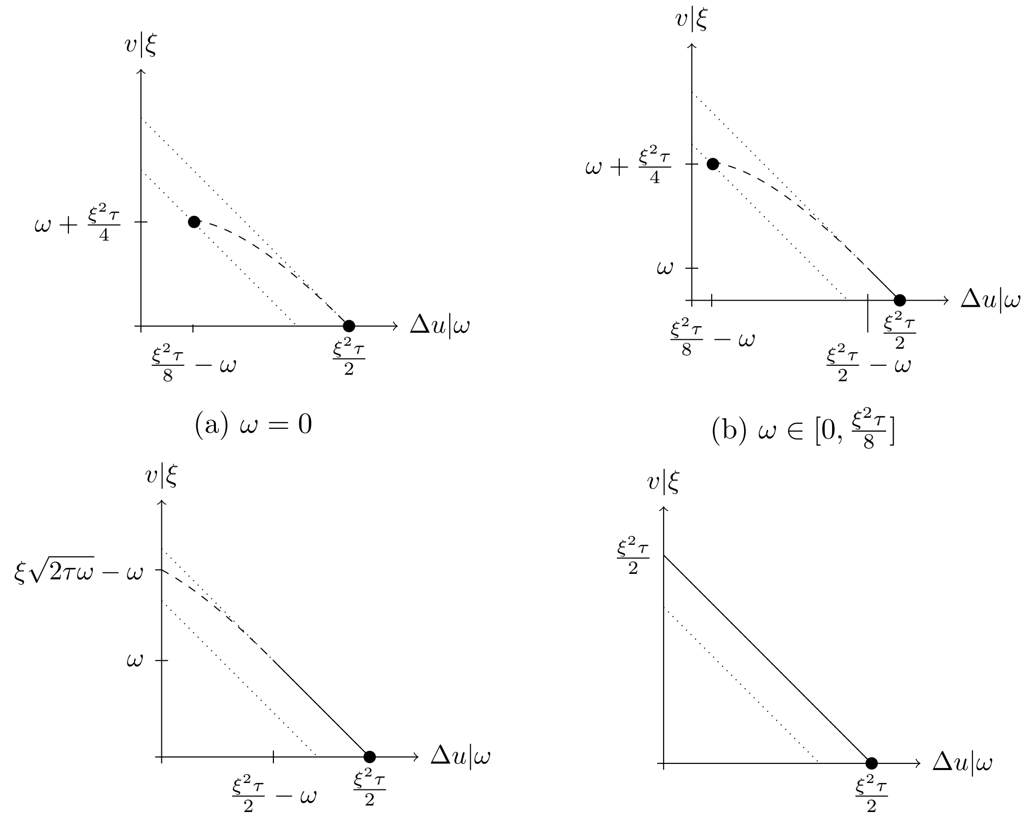
- $U_\xi(\theta) := [\underline{u}_\theta(\xi), \frac{\xi^2\tau}{2}]$ *as the set of feasible utility levels that an optimal contract can give to an agent of type θ working for a firm of size ξ. By definition, $U_\xi(\theta)$ is also the domain of $\phi(\xi, \theta, \cdot)$.*

In Figure 1.1 I draw the UPF for different levels of ω. The upper dotted line represents the surplus in the first best, whereas the lower dotted line represents the surplus when only the (CC) binds. Both lines have a slope of -1. In Figure 1.2 I repeat the same exercise but changing the value of τ instead.

The left and right solid points rest over the second and first best surplus respectively. Specifically the left dot represents the contract in a second best when only the CC is binding, whereas the right dot represents the principal giving away the firm for free to the agent. The solid line represents sections of the UPF when the principal has sold the firm to the agent in exchange of some fixed fee. The dashed line represents the UPF when both constraints, CC and PC, are binding.

By increasing the agent's wealth, the UPF expands upward until it reaches the first best surplus. As the utility for the agent increases, the principal implements the contract indicated by the system of the three constraints CC, PC, and IC. At the same time, the higher the utility for the agent, the more relevant the PC becomes compared to the CC, and therefore the solution gets closer to the first best outcome. Here, the principal is selling cheaper a share in the outcome and therefore the agent, for the same price, is obtaining more of the outcome, hence exerting more effort. If the agent has some wealth, the principal will decide, once the agent is receiving a high amount of utility, just to sell the whole firm for whatever wealth the agent has. By moving along the UPF to the right, the principal will sell the firm cheaper, increasing the utility received by the agent.

As the agent gets wealthier, the principal selling the firm to the agent happens



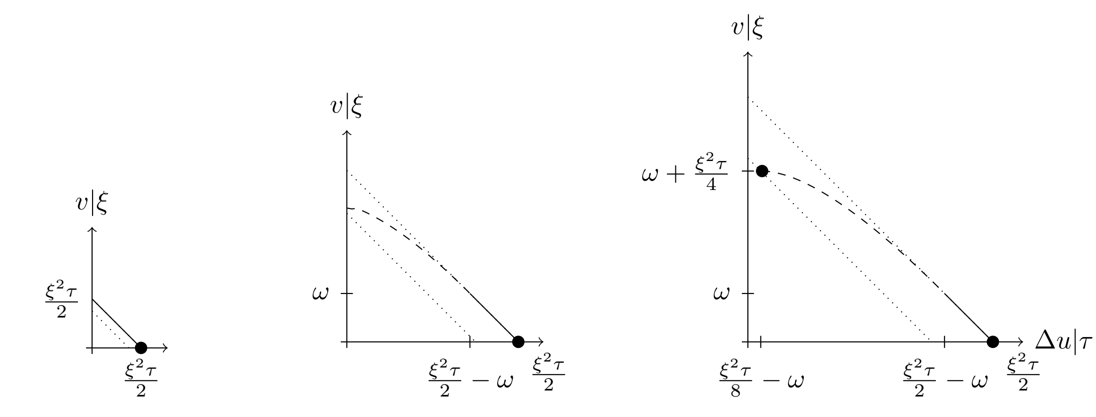
(a) $\omega = 0$

(b) $\omega \in [0, \frac{\xi^2\tau}{8}]$

(c) $\omega \in [\frac{\xi^2\tau}{8}, \frac{\xi^2\tau}{2}]$

(d) $\omega \in [\frac{\xi^2\tau}{2}, \infty)$

Figure 1.1: UPF for different levels of ω.



(a) $\tau \in \left[0, \frac{2\omega}{\xi^2}\right]$

(b) $\tau \in \left[\frac{2\omega}{\xi^2}, \frac{8\omega}{\xi^2}\right]$

(c) $\tau \in \left[\frac{8\omega}{\xi^2}, \infty\right]$

Figure 1.2: UPF for different levels of τ.

sooner in the UPF, that is, for lower values of Δu. In the extreme case when the agent has enough wealth to pay the whole surplus, the principal will be able to write the first best contract and extract the whole surplus.

On the other hand, when the agent has a fixed amount of wealth, increasing his talent will cause the value of the firm, in first and second best outcomes, to increase. This of course is good for both (the share he is receiving increases, and the share of the principal increases as well), however there is a caveat. The more valuable the firm is, the relative wealth, that is the amount of wealth the agent has compared to the value of the firm (ξ), is decreasing, and therefore the agent with more talent is relatively more affected by the CC than a less talented agent. We can observe in Figure 1.2 how by increasing the level of τ the agent moves from a situation in which he is unaffected by the cash constraint, and therefore a first best outcome is always achieved, to another in which he is considerably cash constrained, up to a point where information rents are created. Let us start with a benchmark result.

Fact 1. *If talent is homogeneous among the agents, and they are wealthy enough such that they are not cash constrained for any firm, then nothing can be said about how firms and agents are going to match.*

Proof. If the poorest agent is rich enough to buy the biggest firm, the principal will always set up a first best contract, and therefore will sell the firm exactly at its surplus. For the agents, then, all the firms represent the same utility, that is zero, and therefore are indifferent between them. For this reason, any kind of matching can arise. □

Fact 1 represents the simplest case, in which always first best contracts are written and the principals are able to extract the whole surplus of their firms. Putting this case out of the way, we can study more interesting situations.

In order to look for GID or PAM in this model, I will refer to Corollary 1 in LN, where they use the assumption that ϕ is twice differentiable to obtain PAM, by looking at the signs of its second derivatives. There is a caveat though: in our model ϕ is not twice differentiable. In Appendix A.2 I discuss that for LN's corollary, it is enough for PAM that ϕ is differentiable in ξ, and that this derivative is increasing in u and the agent's type. These conditions establish that the gains of a principal by matching with a higher type counterpart are greater (a supermodularity condition) than when matching with lower type counterparts, plus that for higher values of utility given to the agent along the UPF, these gains also increase. This will ensure that the high type principal can outbid the low type principal for the high type agent.

Lemma 1. *The UPF described by $\phi(\xi, \theta, u)$ is:*

- *Continuous and strictly decreasing in u for $u \in U_\xi(\theta)$.*
- *Differentiable in u for $u \in int(U_\xi(\theta))$.*
- *Differentiable in ξ.*

Proof. To prove continuity it is enough to verify that $\phi(\xi, \theta, \xi^2\tau/8 - \omega) = \omega + \xi^2\tau/4$ and that $\phi(\xi, \theta, \xi^2\tau/2 - \omega) = \omega$. Both are verified using simple algebra.

Every piece of the UPF is differentiable in u. Therefore, it is sufficient to verify that the derivatives coincide when the function changes its functional form. At the first best surplus (and therefore in that section of the UPF) the UPF has slope -1. The derivative of the UPF when PC and CC are binding is $(\xi\sqrt{2\tau(u+\omega)})/(2(u+\omega)) - 2$. After replacing $u = \xi^2\tau/2 - \omega$ we obtain -1, and then the UPF is differentiable in the interior of $U_\xi(\theta)$. The part of the UPF in which only the CC is binding is a point, and lies in the minimum value of utility the agent can get from the matching, and therefore does not belong to the interior of $U_\xi(\theta)$.

Finally, as the function is continuous and every piece of it is strictly decreasing in u, the UPF is decreasing in u.

Now analyzing $\partial\phi/\partial\xi$, we need to study again its intervals. ϕ written as a function of ξ is:

$$\phi\left(\xi,\theta,\overline{u}\right)=\begin{cases}\frac{\xi^2\tau}{2}-\overline{u} & \text{if } \xi\leq\sqrt{\frac{2(\overline{u}+\omega)}{\tau}}\\ \xi\sqrt{(\overline{u}+\omega)2\tau}-2\overline{u}-\omega & \text{if } \sqrt{\frac{2(\overline{u}+\omega)}{\tau}}<\xi\leq\sqrt{\frac{8(\overline{u}+\omega)}{\tau}}\\ \frac{\xi^2\tau}{4}+\omega & \text{if } \xi>\sqrt{\frac{8(\overline{u}+\omega)}{\tau}}\end{cases}$$

Using simple algebra it can be shown that this function is continuous in ξ. Moreover, taking the derivatives in each piece, and replacing the boundaries of each piece of the function, it can also be verified to be differentiable in ξ. □

Lemma 1 is instrumental in the construction of the proof of the final proposition. It says that the UPF is well behaved and follows a standard principle: The more utility is given to the agent, the less utility the principal will obtain.

Lemma 2. *$\partial\phi/\partial\xi$ is continuous and increasing in τ, ω, and u.*

For the proof of Lemma 2 please refer to Appendix A.2. This is the first stone to build up the supermodularity type of characteristics for ϕ in order to obtain PAM. Further, it is necessary that ϕ does not decrease in the agent's type, nor ψ decrease in the principal's type. This is formalized in the following lemma,

Lemma 3. *$\phi\left(\xi,\theta,u\right)$ and $\psi\left(\theta,\xi,v\right)$ are type increasing, that is:*

- *ϕ is non decreasing in ω and τ.*
- *ψ is non decreasing in ξ.*

Proof. From Table 1.1 it can be seen that, at least piece-wise, both ϕ and ψ are non decreasing in ω or τ the former, and ξ the latter. Continuity of ϕ and ψ in each of the relevant variables, which is easily shown, is sufficient then to obtain the type increasing property.

□

Lemma 3 shows that the surplus increases when increasing ω, τ, or ξ, formalizing what can be observed in figures 1.1, 1.2, and 1.3 in which the UPF shifts upward when increasing any of those parameters. This implies that for any level of u that is feasible, the principal obtains at least as much utility with a richer, or more talented agent.

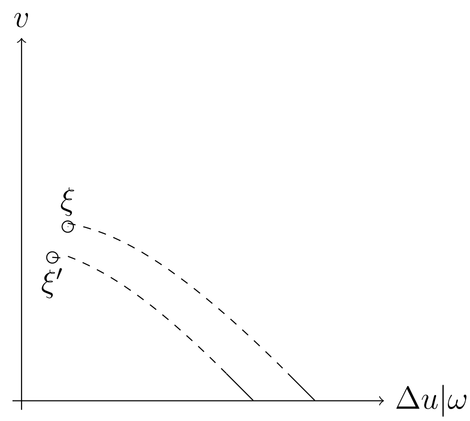


Figure 1.3: $\xi' < \xi$, $\omega \in [0, \frac{\tilde{\xi}^2\tau}{8}]$ with $\tilde{\xi} \in \{\xi', \xi\}$

Proposition 1. *The economy with principals and agents with moral hazard satisfies generalized increasing differences in (ξ, ω) and (ξ, τ), which implies that:*

- *For equally talented agents, larger firms will match with wealthier agents.*
- *For equally wealthy agents, larger firms will match with more talented agents.*

Proof. Given that the function ϕ is differentiable in ξ (Lemma 1), the fact that Lemma 2 ensures that $\partial\phi/\partial\xi$ is non decreasing in τ, ω, and u, and that Lemma 3 shows that ϕ and ψ are type increasing, Corollary 1 in (Legros and Newman, 2007, p.1097) can be applied, as the requirements for its proof are satisfied, and therefore PAM is obtained for both matches. For a detailed proof refer to Appendix A.2.

□

Additional results with discrete types

Wealthier agents will work in bigger firms, and more talented agents will also work in bigger firms. However, what is the equilibrium going to look like? Finding how principals and agents will construct their contracts is not simple, as an equilibrium would imply that everyone is maximizing their expected utility by matching with their partner, and no other principal *steals* the agent of another one. This means not only that the match is stable, which is implied by GID, but that the contracts signed by each party can be clearly identified. The conclusions of GID can be applied to a continuum of firms and agents, as well as for a discrete set of them. For simplicity, I will focus in what is left of this article on working with the discrete case.

Proposition 2. *Let $\theta = (\omega, \tau)$ and $\theta' = (\omega', \tau')$ be the types of two consecutive agents such that $\omega' \leq \omega$ and $\tau' \leq \tau$. One coordinate (τ or ω) is equal among all the agents, whereas the other (ω or τ) is strictly larger. Let $\xi' < \xi$ bet the sizes of two consecutive firms. The equilibrium outcome must satisfy:*

- *If (θ', ξ') represents the match of the lowest matching types, agent and principal will obtain:*

$$u^* = \underline{u}_{\theta'}(\xi')$$
$$v^* = \phi(\xi', \theta', u^*)$$

- *Otherwise, let $\tilde{v}$ be the utility of the low type principal. The high type match will* *obtain:*

$$u^* = \max\{\psi\left(\theta, \xi', \tilde{v}\right), \underline{u}_\theta(\xi)\}$$
$$v^* = \phi\left(\xi, \theta, u^*\right)$$

Proof. The first part of the proposition is trivial, as by GID no principal would want to outbid the lowest type principal for the lowest type agent, and therefore there are no incentives for the lowest type principal to provide more utility than the minimum possible, that is $\underline{u}_{\theta'}(\xi')$.

For the second part, assume that $u^* = \psi\left(\theta, \xi', \tilde{v}\right)$. Note that if the high type principal were to offer less than u^*, say $u^* - 2\epsilon$, then the low type principal could offer $u^* - \epsilon$ and be strictly better off because the UFP is strictly increasing in θ. By offering higher utility to the high type agent, she outbids the high type principal for the high type agent. If $u^* = \underline{u}_\theta(\xi)$, the same reasoning applies, as it is sufficient that the high type principal offers enough utility to the agent to avoid the outbidding from the low type principal. □

In a 2 by 2 world, that is two principals and two or more agents, the low type match will write down a second best contract as if they were in an isolated situation, whereas the high type match will write down a contract that provides the agent at least as much utility as the high type agent would obtain with the low type principal, when this principal is getting his second best utility when matching with the low type agent, that is, when the low type principal is unable to outbid the high type principal for the high type agent.

Proposition 3. *Consider consecutive matches of firms of size $\xi' < \xi$, and agents with* *types $\theta' < \theta$, such that only ω or τ is equal for all the agents, and for every match, the agents are cash constrained for both firms.*

If difference in wealth or talent is large enough, i.e. if:

- $\omega - \omega' > \tau \left(2\xi'^4 + \frac{(\xi^2-\xi'^2)}{4} - \xi'^2\sqrt{\xi^2 - \xi'^2 + 4\xi'^4}\right)$*, when the agents' type is wealth.*
- $\frac{\tau'}{\tau} < 2 - \frac{\xi^2}{2\xi'^2}$*, when the agents' type is talent.*

Then the high type match will write a contract with stronger incentives than the contract it would write out of the market, and therefore closer to the first best output.

Proof. Let $\xi' < \xi$, $\omega' \leq \omega$ and $\tau' \leq \tau$ with only one of these two last inequalities being strict. Let v' represent the utility the low type principal obtains by matching with the low type agent, and assume both agents are cash constrained for both firms. From Proposition 2 we can write the maximum utility the high type agent could get from the low type principal:

$$u = \frac{1}{4}\left(-2(\omega' + \omega) + \xi'^2\frac{\tau + \Delta\tau}{2} + \xi'\sqrt{\tau(\xi'^2\Delta\tau + 4\Delta\omega)}\right), \tag{1.3}$$

where $\Delta\omega = \omega - \omega'$ and $\Delta\tau = \tau - \tau'$. If the contract of the high type match is equivalent to the one they would write out of the market, then the high type agent would receive $\xi^2\tau/8 - \omega$. The conditions in the proposition come from comparing the utility level in (1.3) against this last expression. □

Proposition 3 states the conditions such that the competitive pressure imposed by the market makes the high type match write a contract that provides incentives in which the outcome is closer to the first best, compared to the expected outcome in the one-to-one version of the model, reflecting the efficiency brought by competition into this economy. The first part of Proposition 3 says that when the agents' type is wealth, then the more talented they are, and the more different in size are the firms, how much different wealth between high and low type agents needs to be, in order for the competitive pressure to be enough to motivate the high type principal to strengthen the incentives for the high type agent. Note that a bigger firm makes the agent more cash constrained, increasing the information rents, and therefore decreasing the need of more compensation to avoid outbidding from the low type principal (as part of this cost is already covered by the higher information rents). This explains, then, the necessary increase in the wealth of the high type agent compared to the low type to have a contract that provides him more than the information rents.

The second part of Proposition 3 says that when the agents' type is talent, and the size of the smaller firm is less than half of the size of the big one, then for any relationship of talent between the agents, there is no competition enough to make the high type principal to provide extra incentives to the high type agent. However if $\xi' \geq (\sqrt{2}/2)\xi$, then for any pair τ' and τ the contract written by the high types is going to give the high type agent more information rents than what he would obtain with the high type principal in the absence of market pressure.

The conclusions of Propositions 3 and 4 go in line with the literature, in the sense that market pressure (competitive factors) can increase the difference in expected compensation among the agents more than the difference in the information rents created by the firm size.[12] This would explain why concentrated distributions in talent, for example, can lead to dispersed distributions of compensation as found by Terviö (2008).

[12]Because the agents are risk neutral, this is equivalent to gains in expected utility.

1.4 Wealth, Talent and the Matching

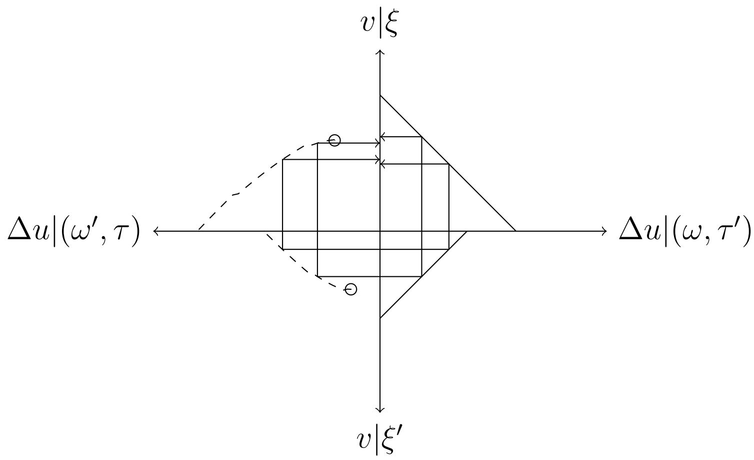
I have shown that the economy described in this work satisfies positive assortative matching when the agent's type is either talent or wealth. This suggests that if talent and wealth are positively correlated, that is, more talented agents have higher amounts of wealth, indeed positive assortative matching will arise. However, there are situations in which this does not necessarily happen. In particular, for young agents, their wealth is not correlated with their talent but maybe their cash constraint is influenced by their networks or their family wealth, so the question that remains is: Does positive assortative matching with respect to talent holds for all joint distributions of the agents' talent and wealth?

The corporate finance literature (Terviö, 2008) considered positive assortative matching between agents and principals when considering talent and size as their types. I exploit the model introduced in this article to analyze the impact that wealth can have on the matching between principals and agents. In other words, are there distributions of wealth and talent among agents that could compromise the PAM with respect to talent? Can wealthier but poorly talented agents match with big firms at the same time that talented and poorer agents end up working in small firms?

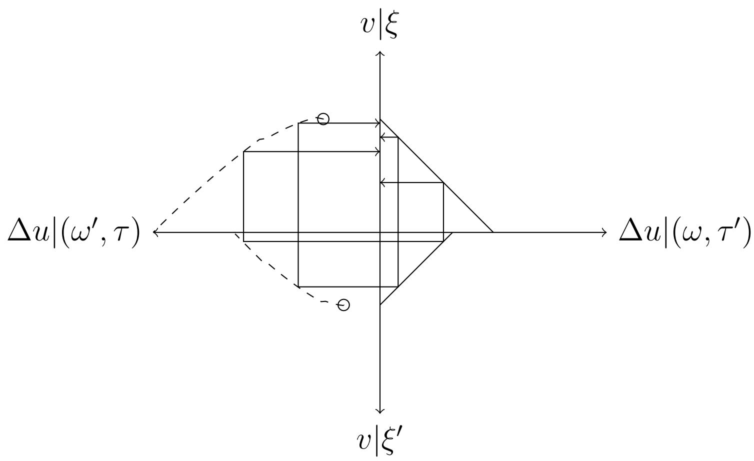
In Figure 1.4 we observe a graphical representation of the UPF to look for GID or GDD,[13] following Legros and Newman (2007). In detail, we compare four possible matches: In the upper half of each vertical axis we consider the big firm matching with the agents, whereas in the bottom half we consider the small firm matching with the agents. The left half considers the poor but talented agent matching with the firms, whereas the right half considers the wealthier but not so skilled agent matching also with the firms as well. Each axis represents the utilities of each actor, and the curves are

[13]LN define generalized decreasing differences (GDD) as a sufficient condition for negative assortative matching (NAM), where the lower type principal can outbid the high type principal for the high type agent.

the UPF derived from each matching. For simplicity I assume that the poor agent has $\omega = 0$, whereas the rich agent has $\omega > 1$, which ensures that he is cash unconstrained with any firm.



(a) $\tau - \tau' < \underline{\tau}$.



(b) $\tau - \tau' \geq \underline{\tau}$.

Figure 1.4: Assessment of the bidimensional matching between agents and principals.

We observe (Figure 1.4a) that when having a poor but skilled agent against a rich but less capable agent we might no longer have PAM with respect to talent and firm size. This happens when the difference in their talent is below some threshold ($\tau - \tau' < \underline{\tau}$). Furthermore, if we start assuming a level $\overline{v}$ as reservation profits for the firms, then we can even end up with negative assortative matching when considering firm size and agent talent, if rich agents have poor skills and poor agents are talented. Conversely, if the difference in talent between the two agents is sufficiently high (Figure 1.4b), then the effect of the limited liability becomes irrelevant, and positive assortative matching between talent and firm size arises.

Note that where the UPF intersects the horizontal axis (the agent's change in utility) only changes with τ and ξ, whereas ω is crucial to determine where the UPF reaches the maximum value for ϕ (the principal's utility). Because of this, if we analyze the matching between both of the agents, (ω,τ') and (ω',τ), with a firm of a given size, then the poorer agent will always have the bottom of his UPF to the right of the UPF generated by the matching of the wealthier one. However, if the difference in talent is small, the poorer agent's UPF will reach a lower maximum (the isolated second best contract outcome) on the vertical axis than the UPF of the wealthier agent (the first best outcome, of a slightly smaller outcome because of the lower talent). This crossing of the UPFs is critical, as the firms prefer one agent or the other, depending on how much utility they will have to provide to the agent in order to make the matching stable. The UPFs will cross if and only if the difference in talent is low enough.

1.5 Discussion

The model studied in this paper provides a simple framework to study the implications of the matching between principals and agents, when we consider, besides talent, the wealth owned by the agents. Doing that, on the one side we obtain predictions for future empirical work, with a model that considers the effects of endogenous matching. On the other, the model replicates some empirical findings in the literature of executive compensation. In this section I will expand on these two points.

Dam and Perez-Castrillo (2006) takes the principal-agent model, with moral hazard, into the matching framework. They conclude that market conditions, such as competition for jobs or competition for workers, are critical to determine the market outcome. A key similarity with this work is the consideration of wealth as the agent's type, the consideration of non-perfectly-transferable utility between principal and agents, and the introduction of the effects of wealth through the limited liability channel. A key difference, though, is that they consider homogeneous principals while I let principals to differ in the size of their projects. I also let the agents have their talent as another type. These differences allow to consider a much wider range of scenarios, and instead of answering who is working (or who is hiring), they allow to answer the question of *who* is working *where*. Letting talent be part of the model, allows to contrast the conclusions with a wider set of empirical literature, and check if there are crossed implications, as the agent's could have a multidimensional type.

Earlier, Legros and Newman (1996) set up a general equilibrium model in which agents with different levels of wealth get together to form a firm. They find that wealth plays a crucial role in defining the type of firms that are going to be formed, in particular in the presence of moral hazard (when monitoring costs are nontrivial).[14] In this article, instead of studying firm formation, I try to model the environment in which empirical compensation studies are based on, such as, firms hiring executives. I do so, by letting the market to be two-sided, differentiating firms from agents. Furthermore, in the model presented in this article, agents are endowed with talent as well. Both articles share, though, characteristics such as risk neutrality and wealth playing a crucial role in the matching of agents through the limited liability channel.

The empirical literature on compensation, has started to incorporate some degrees

[14]This market is closer to the one presented in Lucas (1978), in which the market is one-sided, and economic agents might assume different roles in the resulting organizations. In a more recent work Eeckhout and Kircher (2018) expands this matching concept by allowing the firm to choose endogenously how many and how skilled are those workers.

of matching and moral hazard. Of particular relevance are the works of Gabaix and Landier (2008) and Terviö (2008), which provide interesting conclusions relating the variables that might drive the matching between executives and firms. Gabaix and Landier (2008) present a model of CEO and firms matching, based on the distribution of the CEOs' talent. They assume that compensation is based on talent, while firms and CEOs are differentiated between them in size and talent respectively. Their main theoretical contribution is how compensation on talent reacts to the talent distribution, reaching the conclusion (supported by their data) that even highly similar talented CEOs can show big differences in their wages. They also conclude that bigger firms lead to higher wages in equilibrium. Terviö (2008) on the other hand, develops a matching model that tries to obtain the distribution of CEO's ability from the known distribution of pay and firm's market value. The author concludes, among other important insights, that the wide differences in compensation for a small distribution of CEO's ability is given by firms characteristics. It happens as well that competitive factors are crucial to explain the huge differences in compensation levels among managers. Again, Terviö (2008) neglects moral hazard problems, and neglect any impact wealth might have on the the design of incentives. An important lesson though is the effects competition has in the level of wages. The model introduced in this article provides a theoretical ground for these findings. The more similar the agents and firms, the latter needs to pay more to keep the high type agents agents, otherwise, smaller firms could outbid the bigger ones for them. This competitive effect generates efficiency gains, which reflects in higher expected surplus.

Edmans, Gabaix, and Landier (2009) expands the framework of Gabaix and Landier (2008), by incorporating the agency problems within a competitive assignment model. They conclude that the compensation, as a share of the firm, is decreasing in firm size, and they evaluate the effectiveness of incentive compensation, a result that can also be replicated in the model introduced in this article. Moreover, in their conclusions they raise a question that relates directly with this article: Are CEO incentives increasing in wealth? They refer to this problem and the impossibility to solve, at least empirically, given that there is no information on absolute wealth for the CEO, and it is only possible to obtain the wealth inside the firm, in terms of stocks and options. This is one of the things that Baker and Hall (2004) tried to address. They were trying to investigate the relationship between CEO's compensation with the firm's size. To do this they developed a single and multitask agency problem, and find the optimal compensation scheme for the CEO. They later estimate their model focusing on understanding how the marginal effect of a manager depends on the firm's size. An interesting finding is that the CEO's bonus decreases, as a share of the company, with the firm's size. They do not use a matching model as the previous article, nevertheless this result is confirmed by the model presented in here. The authors use wealth to determine the risk aversion only, given a firm size, and later make three assumptions that allow them to proxy wealth in three different ways: first they assume that wealth is proportional to total annual compensation, later they assume that wealth is the CEO's holding in the firm (options plus stock), and finally they assume that CEOs of big firms aren't richer nor poorer than CEOs of smaller firms. For us, this is not the case, as the agent's wealth is key to determine 1) if the agent is cash constrained, and 2) if it is cash constrained, how much information rents can the agent extract.

1.6 Conclusions

The model developed here allows to understand some implications of the traditional moral hazard framework on the matching between principals and agents. In particular I focus on the effect of the agent's wealth on his relationship with firms of different size. I show that, with risk neutral principals and agents, wealth makes the agent cash constrained for a lower amount of firms (the smaller ones), and further, if he is not cash constrained, his information rents decrease with his wealth, and increasingly so in the firm's size. Another result is that the size of the firm, measured in earnings, increases both: agent and principal's utility and compensation. Therefore a first conclusion is that the agent would prefer to work in big firms, for which he would be cash constrained and therefore able to extract information rents. On the other hand, firms prefer to hire wealthier agents, as this would allow them to reduce the information rents the agent can extract from the surplus. This is true only as long as the agent is cash constrained. This will happen for higher levels of wealth, the higher size of the firm, or conversely, for a given size of the firm, for lower levels of wealth. If the firm is small, the principal can do just fine with poorer agents, as less wealth from their side is required to keep them from being cash constrained.

In order to tackle the question of how principals and agents match, I adapt the techniques developed in Legros and Newman (2007) and use generalized increasing differences to obtain endogenous positive assortative matching between principals and agents when considering the firm's size and the agent's wealth or talent. I also describe conditions on the parameter space to describe the contract associated to each match, and its efficiency. In particular I find that when types are closer, the market pressure is higher and makes the high type match to set up a contract that creates a higher surplus than that which could be obtained in a 1 to 1 situation. I also find that the lowest type will always write a contract that is equivalent to an outside the market outcome, when principals retain the bargaining power.

Finally I provide an example in which a poor agent with high talent, and a wealthy agent with poor talent, match with firms of different size. I find that if the difference in talent is small, then there will be no positive assortative matching with respect to talent and firm size. However, if the difference between the agents' talent is sufficiently high, PAM is maintained. Considering that wealth is not necessarily perfectly correlated with talent in a sample with similar agents, this can be an issue for empirical considerations. In particular, given that Ackerberg and Botticini (2002) has already shown the negative impact that neglecting the matching considerations can have on empirical results.

Chapter 2

Hunting with two Bullets: Moral Hazard with a Second Chance[1]

2.1 Introduction

In real life we can think of many situations in which principals face agents who, while unable to hide the outcome, can hide the way it was achieved (for example, when and how much effort was exerted). The agent may have ways to achieve the outcome that are completely unknown to the principal when offering the contract.

In this paper, I introduce a moral hazard model in which the agent takes sequential decisions. In the beginning, the principal offers a contract, which is contingent on the final outcome. Later the agent exerts effort, observes the outcome of his effort, and

[1] I have to thank comments from Steffen Hoernig and Guido Maretto, Fernando Anjos, Igor Cunha, Pedro Vicente, Armin Schmutzler, Patrick Rey, Tobias Kretschmer, Stéphane Gauthier, Patrick Legros, and Bernhard Eckwert. I am also grateful for the useful discussions with the participants in the Nova SBE RG, the QED Jamboree 2016, the ASSET Annual Meeting 2016, and the PEJ 2018 Annual Meeting.

in case the outcome is bad, can decide to exert further effort in an attempt to obtain a better outcome. A key piece of the model is that effort is cumulative, that is, the probability of success is higher the more effort has been exerted in total.

This setup allows us to consider different situations. For example, the agent might at first exert low levels of effort, with the hope of having a good outcome, and knowing that in case this *bet* goes wrong, he will have another chance to work hard and increase the chances of delivering. Consider the example of a honey dealer buying from cheap low reputation suppliers. If he has enough time until delivery, he might gamble with the cheap suppliers, and buy from more expensive and reliable suppliers only if he received low-quality products from the cheap ones. This decreases the chances for the principal observing a good outcome, when compared to going to the good suppliers from the beginning, as the agent has a higher probability of success, and higher chances to fix a unlikely bad outcome.

The extra chance is, however, not necessarily a bad thing for the principal. Indeed, it might even reduce the cost of effort, by introducing the option of gambling at first. In this article, I show that the principal, under certain conditions, designs contracts that make the agent gamble even in scenarios without moral hazard. However, moral hazard increases the agency costs significantly compared to the standard agency model (it is easier for the agent to hide what his actions are because of his larger set of options). I show that these facts create non-convexities and non-monotonicities in the implementation of effort as a function of its cost.

Finally, I study the case where the extra chance represents an undesired activity, which can trigger later a punishment to the principal if caught. I show that if the penalty is big enough, the principal will never contract a strategy involving gambling (no effort and later trying to fix an adverse outcome). As the agent does not suffer the consequences other than the cost of the extra chance,[2] the agency costs increase substantially in the regions where gambling is the agent's best response, and therefore no effort is contracted in a broader set of parameters than in the case without the punishment.

In Section 2.2, I review the literature that relates to the problem presented in this paper. In Section 2.3, I present the baseline model with its efficiency implications. Then, in Section 2.4, I describe how a cost for the principal, of having used the second chance, affects the implementation of strategies involving the use of that extra chance with and without moral hazard. Finally, I conclude in Section 2.5.

2.2 Related Literature

The traditional framework used in moral hazard consists on an agent that has to perform a task for a principal, and the principal cannot observe the effort exerted by the agent. For that reason, the principal sets a payment schedule contingent on the outcome. Examples of this can be found for example in Bolton and Dewatripont (2005), Laffont and Martimort (2002), and Salanié (2005). The model introduced in this paper, incorporates a second chance to exert effort by the agent, before the principal observes the outcome.

The literature so far has a set of different branches to which this model can relate. In particular, Holmstrom and Milgrom (1987) introduced a pure moral hazard model in which the agent has to exert effort a number of times before the principal compensates him. In their work the authors' model has the following timeline: in each of the N periods of the game the agent has to exert effort, and after each of those periods there

[2] "The Justice Department has lost the will and ability to prosecute top corporate executives. They focus on settlements with corporations for money...", Jessie Eisinger in an interview with Knowledge@Wharton on August 2017 - http://knowledge.wharton.upenn.edu/article/why-wrongdoing-executives-are-rarely-prosecuted/.

will be an instant realization that cannot be obscured by the agent. At the end of the N periods, the principal will be able to observe the whole history of realizations and then proceed to compensate the agent. They conclude that the optimal contract involves aggregation of realizations and linear compensation on this aggregated performance along with constant effort from the agent. While they were pointing to show that not always it is necessary to use all the information to reach optimal compensation schemes and that sometimes simple functions, as observed in the real world, turn out to be optimal solutions for the principal, I focus more on the behavior of the agent trying to exploit the fact of having more than one period to achieve a final output, and I do not provide more information to the principal than a single outcome. Besides, in my framework effort is cumulative, and the agent can stop working after the first period if desired.

Another one is multitasking (Holmstrom and Milgrom, 1991), as the two actions are different (as effort is cumulative) and both go in the direction of increasing the principal's utility. Even though the complementarity of both actions is a similarity with the model presented in this article, the timing is quite different. As effort is cumulative in my model, the first time the agent exerts effort impacts on the productivity of effort in the extra chance; however, the effort exerted during this second chance does not affect the productivity of the first one.

Zhao (2008) provides a model in which two contracting parties might be unaware about their own or the counterpart's strategy set. This was extended by articles like Auster (2013) and von Thadden and Zhao (2014), although in these situations, usually the agent is unaware of some of his options, and the principal decides to reveal — or not — information about those to the agent through the contract. The model of this paper considers a similar framework, however it departs from the main stream by considering that: the principal might be unaware of the agent's strategies, the principal makes a take-it-or-leave-it contract, and therefore the agent has no way to pass information to the principal, and finally, I explore the possibility of the principal removing some of the agent's strategies by using deadlines.

Varas (2017) provides a model to explain why contracts exhibit low turnover rates and deferred compensation. In his setup, managers can shorten the time they take to carry out a project by sacrificing quality. The principal, therefore, delays compensation to the future, so the quality of the project is revealed. This framework is related to the one introduced in this paper in the sense that the agent can exert actions that can create costs for the principal in the future. However, Varas' setup is intended for longer principal-agent relationships, in which termination is fundamental in the stationary contracts. In this article, I focus on a short-term relationship between a principal and an agent. This can be applied to suppliers, contractors, etc.

There is another branch of the literature that studies fraud using counterfeit signals, but with a very important difference: the literature considers the problem as pure adverse selection or instead as moral hazard followed by adverse selection, in which the agent can choose the signal to present to the principal about a previous realization which in turn can depend on some effort measure. In this literature, we find for example Maggi and Rodríguez-Clare (1995) which propose a model in which the agent must reveal to the principal its true type. Crocker and Morgan (1998), and Crocker and Slemrod (2007) incorporate a first stage in which the agent must indeed exert an effort level, and later he can decide to reveal or not the outcome (that is, reveal the true outcome or a false one).

Clausen (2013) introduces a novel concept that extends the previous models incorporating the fact that the agent can decide to obscure the true outcome from the principal, but he cannot control what it is going to be the signal that indeed the principal observes. Clausen considers a model in which the agent exerts effort once, and later there are successive accurate signals about the outcome realizations that are privately observed by the agent each time before the principal. The agent can then decide to counterfeit each signal for a better one deceiving the principal. The type of situations that can be represented by this model are different from the ones I try to describe. Indeed, Clausen mentions internet advertising click fraud, in which companies exaggerate the reported clicks for internet advertising, or security companies that hide breaches, to obscure the fact of having failed in their mission. A crucial difference with respect to the model I present in this work is that the agent, instead of sending a counterfeit signal to the principal, can exert effort and *de facto* improve the outcome. He cannot deceive the principal by obscuring the outcome.

Finally, the *gambling for resurrection* literature (Calveras, Ganuza, and Hauk, 2004, Thaler and Johnson, 1990) might also seem to be close to this work. Nevertheless, there is a key difference. While the model presented in this paper focuses on the effects of having the possibility of improving a poor outcome after the first effort was exerted, the gambling for resurrection focuses on the risk attitudes of agents given a previous event. The classical example is what Thaler and Johnson (1990) call the break-even effect, as when agents with previous losses will take risky opportunities to recover even if those carry even more risk. I am more interested in how the possibility of additional effort can change the agent's behavior from the beginning and how this affects the incentive scheme design.

2.3 The Model

The classical moral hazard models in the literature considers a principal that makes a take-it-or-leave-it offer of a contract to an agent. This contract establishes a payment schedule, from the principal to the agent, conditional on publicly observed outcomes.

The outcomes are stochastic, but their distribution is influenced by the amount of effort the agent has exerted. This effort is not observable, and that is precisely what creates the moral hazard problem. I extend this classical model by giving the agent the opportunity to improve a bad outcome after he has already exerted some effort, but before it is observed by the principal.

Principal and agent are both assumed to be risk neutral. I assume further that the agent is cash constrained. First principal and agent sign the contract, establishing payments $w \in \{w_l, w_h\}$ contingent on the observed output $y \in \{y_l, y_h\}$. The effort the agent can exert is denoted by $e \in \{0, 1\}$, which impacts the probability of obtaining a high output y_h. This probability depends on the amount of effort exerted in the present period and the past. Let (e_1, e_2) be the agent's strategy that works in the following way: the agent in the first period will exert e_1 and later will observe an interim accurate signal $\hat{y} \in \{y_l, y_h\}$. If $\hat{y} = y_h$, the agent will not exert more effort whatsoever (as the outcome cannot be improved) and the publicly observed outcome is $y = \hat{y} = y_h$. However, if $\hat{y} = y_l$ then what happens next depends on the agent's choice of e_2. If $e_2 = 0$, there is no second lottery and the previous outcome is maintained. The principal observes $y = y_l$ and pays w_l.[3] If $e_2 = 1$ the outcome y is drawn from a lottery that assigns a higher probability of occurrence to y_h than in the previous period, because the total amount of exerted effort has increased, and then capturing this model's feature that effort is considered cumulative. A contract under this setup is a wage schedule that induces the agent to choose a particular effort strategy. Without any loss of generality and for the sake of simplicity I will assume from now on that $y_l = 0$.

The utility function of the principal is $u_p = y_i - w_i$, $i = h, l$, while the utility function of the agent is $u_a = w_i - e_1 - \beta e_2$, $i = h, l$, where $\beta > 0$ represents the cost of exerting

[3] I assume a second lottery only if effort is exerted in the second period, while for the first one, no effort still has a positive probability of success. This assumption brings tremendous gains in simplicity and parsimony to the model. I have verified that giving the agent a free draw when $e_2 = 0$ does not change the main conclusions of the model.

effort *to improve upon the already realized outcome*. Let the agent's reservation utility be $\overline{u} = 0$. Furthermore, $w_0 \geq 0$, as the agent was assumed to be cash constrained.

Effort influences the probability of having a good outcome. This probability may take the values p_0, p_1, and p_2 defined as:

- $p_0 > 0$ is the probability of success when no effort was exerted in the past, nor the present.

- $p_1 > p_0$ is the probability of success when effort was exerted only once.

- $p_2 = 1$ is the probability of success when effort is exerted now and in the past.

Note that the subscripts indicate how many times the agent has exerted effort at that time. It is also worth noting that the probability of having a good outcome after the first period can only be p_0 or p_1, while in the second lottery it can be p_1 or p_2. Letting $p_2 = 1$ allows focusing on the significance of p_0 and p_1. Finally, p_0 is the probability of success without exerting any effort, and as that it helps to measure how much the effort can add to the outcome.

The complexity of the task is captured by p_1. If p_1 is high, for a given p_0, then exerting effort once is enough to have the goal achieved, and therefore it represents a simple task, but if p_1 is very low, then exerting effort once is most likely not to be enough to achieve the desired goal, representing a complex task.

The model gives the agent the alternative to delay effort if convenient. If the agent observes a bad interim outcome, exerting effort to fix that outcome implies the same probability of success as having exerted effort at the beginning (p_1). Moreover, when choosing the strategy, the probability of success of $(0, 1)$ is strictly higher when compared to $(1, 0)$. Note that the agent would delay effort, not because of impatience (which is not modeled), but because of the value of the option of exerting the effort in the future. In this model, delaying effort increases the chances of success compared to

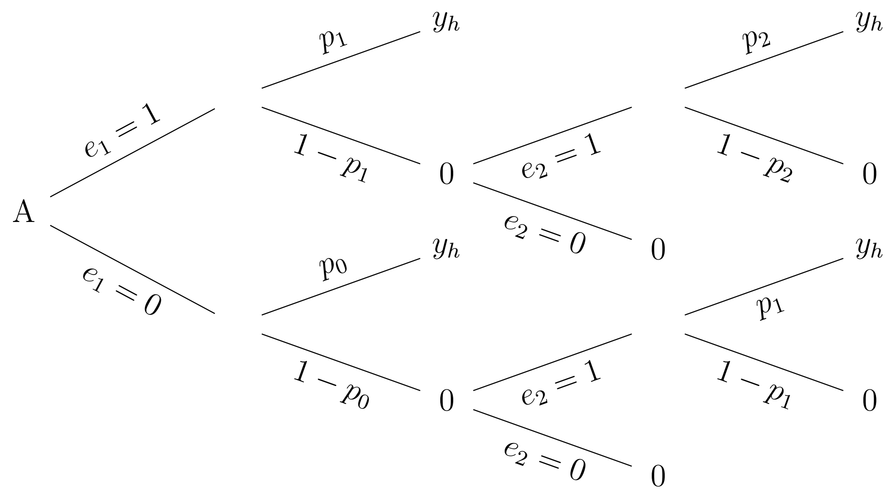


Figure 2.1: Agent's decision tree.

exerting effort only in the first period. Of course this option is valuable as long as its cost (β) is not too high. An illustration of the decisions the agent can make and their impact on the outcome can be observed in Figure 2.1.

In the following subsections, I find the contracts under full information, with an unaware principal with moral hazard, and with an aware principal also in the presence of moral hazard. Later I present a discussion on the inefficiencies created by the information asymmetry.

Full Information

Without any information asymmetry, the maximization problem is equivalent to the one the agent would solve if owning the project. Note that removing all information asymmetry implies the principal is aware of the existence of both chances the agent can use to exert effort. As such, she considers the direct trade-off between the expected outcome and the cost of each strategy, allowing us to disregard the wage schedule. Following this setup has many advantages. Firstly it allows identifying the welfare maximizing strategies for the parameters of the problem. Secondly, it allows to set up very explicitly the best response for the agent in the problem with moral hazard. The maximization problem under full information is:

$$\max_{(e_1,e_2)} p_{e_1} y_h - e_1 + e_2(1 - p_{e_1})(p_{e_1+1} y_h - \beta) \tag{2.1}$$

Lemma 4. *There exist $\underline{\beta}_1 < \overline{\beta}_1$ such that with full information:*

1. *$\forall\ \beta > \underline{\beta}_1$, $(0, 1)$ is never contracted.*
2. *$\forall\ \beta < \overline{\beta}_1$, $(1, 0)$ is never contracted.*

Proof. In Appendix B.1, I find the strategies that are implemented for different levels of y_h, given β. It can be established that $\underline{\beta}_1 = p_1/[(1 - p_0) - p_1(1 - p_1)]$ and $\overline{\beta}_1 = 1/(p_1 - p_0)$ are such that the lemma is satisfied. Note that $\underline{\beta}_1$ can be higher than 1, if $p_1 > (3 - \sqrt{5})/2$ and $p_0 > (1 - p_1)^2$. □

Lemma 4 shows that, if the cost of exerting effort in the extra chance is too high, it is never optimal to use it. In the same direction, if the cost of this extra chance is low, and if it is worth to exert some effort, this chance is to be used. Moreover, if this extra chance is very cheap, effort is going to be implemented in the form of $(0, 1)$ even for very low values of y_h. Another interesting implication of Lemma 4 is that as $p_1 - p_0 \to 0$, or in words, either the task is so hard that needs effort twice, or the task is easy enough that effort (once or twice) adds little, $\underline{\beta}_1 \to p_1/(1 - p_1)^2$ while $\overline{\beta}_1 \to \infty$, so while $(0, 1)$ is still going to be implemented, $(1, 0)$ is going to be implemented only for an infinitely large cost of creating the second chance.

Lemma 4 considerably facilitates the computation of the contract with full information, as it reduces the strategies to consider within intervals of β. The solution to the problem stated in (2.1) is represented in Proposition 4.

Proposition 4. *The contract under perfect information is given by:*

1. *For* $\beta < \underline{\beta}_1$,

   (a) $(0,0)$ *for* $y_h \leq \frac{\beta}{p_1}$.

   (b) $(0,1)$ *for* $\frac{\beta}{p_1} \leq y_h \leq \frac{1-\beta(p_1-p_0)}{(1-p_0)(1-p_1)}$.

   (c) $(1,1)$ *otherwise.*

2. *For* $\underline{\beta}_1 < \beta < \overline{\beta}_1$,

   (a) $(0,0)$ *for* $y_h \leq \frac{1+(1-p_1)\beta}{1-p_0}$.

   (b) $(1,1)$ *otherwise.*

3. *For* $\overline{\beta}_1 < \beta$.

   (a) $(0,0)$ *for* $y_h \leq \frac{1}{p_1-p_0}$.

   (b) $(1,0)$ *for* $\frac{1}{p_1-p_0} \leq y_h \leq \beta$.

   (c) $(1,1)$ *otherwise.*

The contract under full information is represented in Figure 2.2. Note that it results from the possibility of delaying effort that at least some effort is contracted even for very low levels of y_h. This is not only true because close to the origin the future effort is cheap, as the fact that even for values of β greater than 1 this is still the case, but because of the value of the option of delaying effort.

At higher levels of y_h we observe that for low levels of β the strategy involving exploiting the option of exerting future effort is preferred, and as β increases then $(1,1)$ becomes optimal. This is because even though for those levels of β exerting the option is still profitable, as β is now higher, the agent tries to diminish the probability of having to use that option, and he achieves that by exerting effort in the first period. If

y_h is high enough, by increasing β, we only disregard the use of the option (as it will never give positive expected profits), while if y_h is lower, then from $(1,1)$ by increasing β, the optimal strategy will be to exert no effort at all. This is another very graphical way to observe the value of the option of delaying effort, as for a fixed y_h the strategy $(1,1)$ changes to $(0,0)$ instead of moving to $(1,0)$ as β increases.

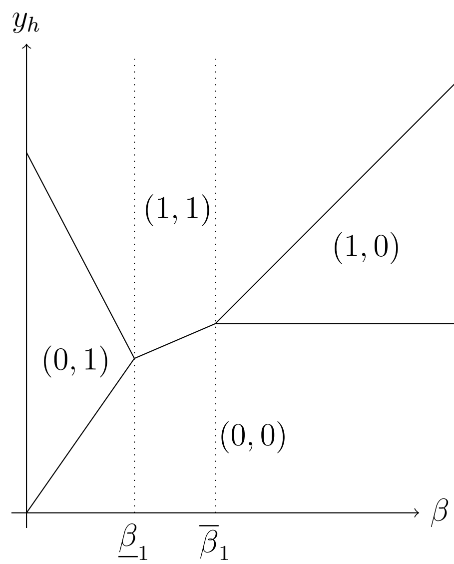


Figure 2.2: Contract under perfect information.

Finally, we can observe that while e_2 does not increase in the cost of the second effort (β), we also observe that, for some levels of y_h, e_1 is not monotonic in β. It is surprising that e_1 is exerted optimally at all for levels below $y_h = 1/(p_1 - p_0)$. To pin down the cause of this, recall the previous discussion: setting $e_1 = 1$ affects e_2 in two ways. First, it makes it less likely that the second chance is going to be used, and second, it increases the expected revenue for doing so, by increasing the chance of success of $e_2 = 1$ from $p_1 < 1$ to $p_2 = 1$. When increasing β, and y_h is such that the strategy moves from $(1,1)$ to $(0,0)$, the increasing cost of the second chance has destroyed the benefits of this complementarity, leading to the optimality of no effort at all.

Moral Hazard and Unaware Principal

In this subsection, I portray the situation in which the principal does not know about the possibility of the extra chance.

This unaware principal will propose the classical textbook moral hazard contract, with $w_h = 1/(p_1 - p_0)$ for a good observed outcome, and $w_l = 0$ otherwise.

We can use the optimal strategies under full information to describe the best response of the agent, by considering $w_h = 1/(p_1 - p_0)$. The principal will offer this wage level if and only if $y_h \geq p_1/(p_1 - p_0)^2$. Replacing w_h in the agent's best response, we obtain the strategies followed by the agent. These depend on the value of β. Note that while $(1, 0)$ and $(0, 0)$ are the strategies the principal expects, $(0, 1)$ and $(1, 1)$ strategies the principal does not expect, as she is unaware of them.

It can be shown (Appendix B.1) that, for $w_h = 1/(p_1 - p_0)$, the agent will never choose $(0, 0)$. Moreover, it can be shown that there exists thresholds on β such that the agent will choose $(0, 1)$, $(1, 1)$ or $(1, 0)$ within different intervals for β. Each of these strategies will lead to different levels of outcome for the principal:

Strategy	Exp. outcome for the principal	Exp. outcome by unaware principal
$(0, 1)$	$[p_0 + p_1(1 - p_0)]\left(y_h - \frac{1}{p_1 - p_0}\right)$	$p_1\left(y_h - \frac{1}{p_1 - p_0}\right)$
$(1, 1)$	$y_h - \frac{1}{p_1 - p_0}$	$p_1\left(y_h - \frac{1}{p_1 - p_0}\right)$
$(1, 0)$	$p_1\left(y_h - \frac{1}{p_1 - p_0}\right)$	$p_1\left(y_h - \frac{1}{p_1 - p_0}\right)$

Table 2.1: Principal's expected outcome.

In Table 2.1, we observe that for each of these scenarios, the principal is at least as good as she expect to be in the traditional model without the extra chance. However, that is not all. The agent also benefits from this, as he obtains larger rents which are created by the extra chance. More specifically, the principal could have induced the same strategies with a much lower wage, keeping a higher share of the outcome.

Strategy	w_h - unaware principal	Optimal w_h	Agent's gains
$(0,1)$	$\frac{1}{p_1-p_0}$	$\frac{\beta}{p_1}$	$\frac{1}{p_1-p_0}-\frac{\beta}{p_1}$
$(1,0)$	$\frac{1}{p_1-p_0}$	$\frac{1}{p_1-p_0}$	0
$(1,1)^{*}$	$\frac{1}{p_1-p_0}$	$\frac{1-\beta(p_1-p_0)}{(1-p_0)(1-p_1)}$	$\frac{1}{p_1-p_0}-\frac{1-\beta(p_1-p_0)}{(1-p_0)(1-p_1)}$
$(1,1)^{**}$	$\frac{1}{p_1-p_0}$	$\frac{1+(1-p_1)\beta}{1-p_0}$	$\frac{1}{p_1-p_0}-\frac{1+(1-p_1)\beta}{1-p_0}$

Table 2.2: Rents distribution with unaware principal. $*$ when $\beta < \overline{\beta}_1$, and $**$ when $\beta \geq \overline{\beta}_1$.

It can be shown that the agent's gains are always positive, given his optimal chosen strategy for the parameters involved (probabilities of success and future cost of effort, β).

Moral Hazard

Under moral hazard, I consider the traditional participation constraint and the incentive compatibility constraint ensuring that the agent accepts to sign the contract and chooses the desired strategy. The cost for the principal is represented by the wages w_h and w_l for the good and bad outcome respectively. Recall that the agent is assumed to be cash constrained, and therefore I set immediately $w_l = 0$. The principal's maximization problem when facing asymmetric information is represented by:

$$\begin{aligned}
\max_{w_h,e_1,e_2} \quad & p_{e_1}(y_h - w_h) + e_2(1-p_{e_1})p_{e_1+1}(y_h - w_h) \qquad (2.2)\\
\text{s.t.} \quad & p_{e_1}w_h - e_1 + e_2(1-p_{e_1})(p_{e_1+1}w_h - \beta) \geq 0 \\
& (e_1,e_2) \in \arg\max_{(\hat{e}_1,\hat{e}_2)} p_{\hat{e}_1}w_h - \hat{e}_1 + \hat{e}_2(1-p_{\hat{e}_1})(p_{\hat{e}_1+1}w_h - \beta)
\end{aligned}$$

The incentive compatibility is given by the solution to the problem with full information. The only difference is that it is necessary to replace y_h with w_h, as the agent is now getting only w_h instead of the whole outcome. This gives immediately the optimal incentive compatible wages for a given β and outcome distribution. As such $\underline{\beta}_1$ and $\overline{\beta}_1$ from Lemma 4 are also critical points for the case with moral hazard, as for the principal, it is impossible to implement a contract involving $(1,0)$ or $(0,1)$ between those parameters, no matter the wage. Having the incentive compatible wage for each strategy, and having the implementable contracts as a function of β, it is just a matter of comparing expected profits of implementing each strategy, leading to the first result in the presence of moral hazard, expressed in Lemma 5.

Lemma 5. *There exist $\underline{\beta}_2$ ($< \underline{\beta}_1$) and $\overline{\beta}_2$ ($> \overline{\beta}_1$) such that under asymmetric information:*

1. *$\forall\ \beta > \underline{\beta}_2$, $(0,1)$ is never implemented.*

2. *$\forall\ \beta < \overline{\beta}_2$, $(1,0)$ is never implemented.*

Proof. Comparing the expected profits of implementing each strategy, given the wages are incentive compatible, it happens that there exist y_h such that $(0,1)$ is optimal only if

$$\beta < \underline{\beta}_2 = \frac{p_1^2}{p_0 + p_1 - (1-p_1)[p_1^2 + p_0^2 - p_0^2 p_1] - p_0 p_1 (3 - p_1)},$$

while there exists y_h such that $(1,0)$ is optimal only if

$$\beta > \overline{\beta}_2 = \frac{p_1(1-p_0)}{(p_1-p_0)^2}.$$

It can be shown after some algebra that $\underline{\beta}_2 < \underline{\beta}_1$ since $p_1 p_0 (1-p_0)(1-p_1)^2 > 0$ and $\overline{\beta}_1 < \overline{\beta}_2$ since $0 < p_0 < p_1 < 1$. □

An important implication of Lemma 5 is that the contracts implementing $(0,1)$ and $(1,0)$ have more restrictive requirements over β than when compared to the case with perfect information. An important remark is that, even though the β_2s found in Lemma 5 play the same role that the β_1s played in the case without moral hazard, the β_1 are still very important for the case with asymmetric information. They now affect directly the incentive compatibility constraint, as they define the regions for which the agent will never play $(1,0)$ or $(0,1)$, no matter w_h. The contract under moral hazard is described for five different intervals for β as stated in Proposition 5.

Proposition 5. *Under asymmetric information, the contract with a cash constrained agent is given by:*

1. *For* $\beta < \underline{\beta}_2$,
   (a) $(0,0)$ *for* $y_h \leq \beta\left(\frac{p_0+p_1(1-p_0)}{(1-p_0)p_1^2}\right)$.
   (b) $(0,1)$ *for* $\beta\left(\frac{p_0+p_1(1-p_0)}{(1-p_0)p_1^2}\right) \leq y_h$, *and* $y_h \leq \frac{1}{[(1-p_0)(1-p_1)]^2} - \beta\left[\frac{p_1-p_0}{[(1-p_0)(1-p_1)]^2} + \frac{p_0+p_1(1-p_0)}{p_1(1-p_0)(1-p_1)}\right]$
   (c) $(1,1)$ *otherwise.*

2. *For* $\underline{\beta}_2 < \beta < \underline{\beta}_1$,
   (a) $(0,0)$ *for* $y_h \leq \frac{1-\beta(p_1-p_0)}{(1-p_0)^2(1-p_1)}$.
   (b) $(1,1)$ *otherwise.*

3. *For* $\underline{\beta}_1 < \beta < \overline{\beta}_1$,
   (a) $(0,0)$ *for* $y_h \leq \frac{1+(1-p_1)\beta}{(1-p_0)^2}$.
   (b) $(1,1)$ *otherwise.*

4. *For* $\overline{\beta}_1 < \beta < \overline{\beta}_2$,
   (a) $(0,0)$ *for* $y_h \leq \frac{\beta}{1-p_0}$.
   (b) $(1,1)$ *otherwise.*

5. *For* $\overline{\beta}_2 < \beta$.

   (a) $(0,0)$ *for* $y_h \leq \frac{p_1}{(p_1-p_0)^2}$.

   (b) $(1,0)$ *for* $\frac{p_1}{(p_1-p_0)^2} \leq y_h \leq \frac{\beta(p_1-p_0)-p_1}{(1-p_1)(p_1-p_0)}$.

   (c) $(1,1)$ *otherwise.*

Focusing the attention on the first interval for β, we observe how $(0,1)$ becomes optimal for a broader set of parameter values than compared to the case with full information. Note that the denominator in the interval is quite small, leading to a very big intercept on the frontier between $(0,1)$ and $(1,1)$ as optimal contracts. Looking at Figure 2.3 there is a very interesting fringe of y_h about the middle. There are some values of y_h for which, by increasing β, we have the following transition: starts with $(0,1)$, then moves to $(0,0)$, followed by $(1,1)$ to finally come back to $(0,0)$.

Proposition 6. *With information asymmetries there exist y_h such that e_2 is not monotonically decreasing in β.*

Proof. The proof follows directly Proposition 5 and Lemma 5. Let

$$y_h \in \Upsilon := \left(\frac{1-\underline{\beta}_2(p_1-p_0)}{(1-p_0)^2(1-p_1)}, \frac{1-\underline{\beta}_1(p_1-p_0)}{(1-p_0)^2(1-p_1)} \right)$$

The set Υ is nonempty since $\underline{\beta}_2 < \underline{\beta}_1$. From Proposition 5, it can be seen that between $\underline{\beta}_2$ and $\underline{\beta}_1$ the optimal contract will change from $(0,0)$ to $(1,1)$, as the slope of the frontier between both is decreasing in β. It can be observed as well, that for β increasing above $\underline{\beta}_1$ the optimal contract will move from $(1,1)$ to a contract for which e_2 is zero, as the frontier between both has always a positive slope.

$\square$

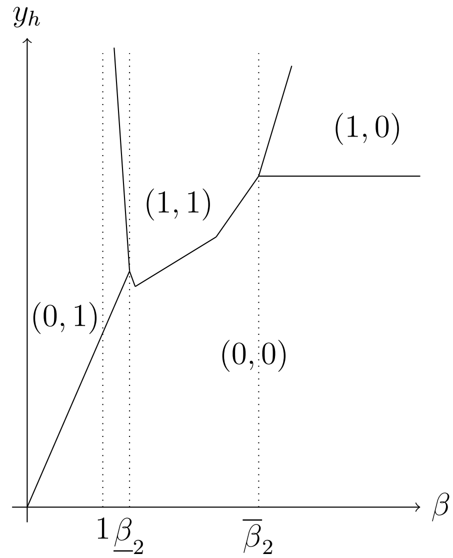


Figure 2.3: Contract with moral hazard.

The set of $(0, 0)$ is not convex, which follows directly from Proposition 5 where the slope of the frontier of $(0, 0)$ changes from positive to negative and later to positive again. At very low levels of β, $(0, 1)$ is implemented following the same logic it had in the case with full information. The second chance, being so cheap, makes it preferable to bet all in taking the risk and later try to fix any bad outcome. As β gets closer to $\underline{\beta}_2$ (for the same level of y_h) the extra chance becomes more costly, so it would be better to avoid it. The logic behind the solution with full information to achieve that was to implement $(1, 1)$ instead, in order to decrease the probability of using the extra chance. However, as β is still low, the agent would prefer to deviate to gamble with $(0, 1)$, and therefore the incentives to make him stick to $(1, 1)$ become too high. In this situation, the principal decides to implement $(0, 0)$. As β increases further, approaching now $\underline{\beta}_1$, the principal knows that the incentives for the agent to deviate from $(1, 1)$ decrease, diminishing the agency costs, so she will start implementing $(1, 1)$ again. The lower y_h inside the fringe, the closer the β needs to be to $\underline{\beta}_1$. In fact in the limit, it will coincide and will implement $(1, 1)$ just in $\beta = \underline{\beta}_1$. When β is above $\underline{\beta}_1$, the logic of the model follows the case with full information. $(1,1)$ is implemented instead $(0,1)$ as β becomes larger, to decrease the possibility of paying the cost of the extra chance, and as β increases even further, the principal will implement strategies with $e_2 = 0$, and setting $e_1 = 1$ or $e_1 = 0$ depending on the level of y_h. While, as expected, introducing agency costs in the model changes the optimal solution for the principal, the information rents around $\underline{\beta}_1$ create the non-convexity in the model. Note that to the left of $\underline{\beta}_1$ the agent has the option to deviate between three contracts, while to the right of $\underline{\beta}_1$ he endogenously will never choose $(0,1)$, so the principal does not require to provide incentives to prevent that deviation.

As it was the case under full information, we observe that e_1 is not monotonic on β for some values of y_h either. This follows the same rationale previously discussed in the sense that, increasing β, the benefits of the complementarity between e_1 and e_2 are offset by the higher cost, requiring higher values of y_h to justify this strategy.

The other comparisons that can be made between the cases with full or imperfect information are in line with what would be expected in the traditional models of moral hazard and are illustrated in Figure 2.4. First, we observe that $(0,0)$ expands against all the other contracts. Besides the change in convexity of the set of $(0,0)$ discussed previously for the contract with imperfect information, the change in the extension of the sets implementing $(0,1)$ and $(1,0)$ is clear. In particular, the change in the slope of the frontier between $(0,1)$ and $(1,1)$ makes it very hard (require very high values of y_h) to implement $(1,1)$ when β falls below $\underline{\beta}_2$. It is remarkable that $\underline{\beta}_2$ is not required to be below 1 for this to happen, so this case happens even when the creation of this extra chance is relatively more costly than exerting effort on the first chance.

The conclusions of Lemma 5 can also be observed in Figure 2.4. Note how larger the interval of $[\underline{\beta}, \overline{\beta}]$, for which only $(0,0)$ or $(1,1)$ contracts are implemented, is in when compared to the case with full information. This happens because without full information, it is easier for the agent to deviate from one of those contracts to the other, or from $(1, 1)$ to $(1, 0)$, and therefore the principal will prefer to implement contracts that give him better information about what was done by the agent. When β is very high (above $\overline{\beta}_2$), the principal can incentivate the agent to not deviate from his strategy, starting to implement $(1, 0)$ again. Basically to the left of $\underline{\beta}_2$ the principal is resigned not to implement $(1, 1)$ when it would be optimal.

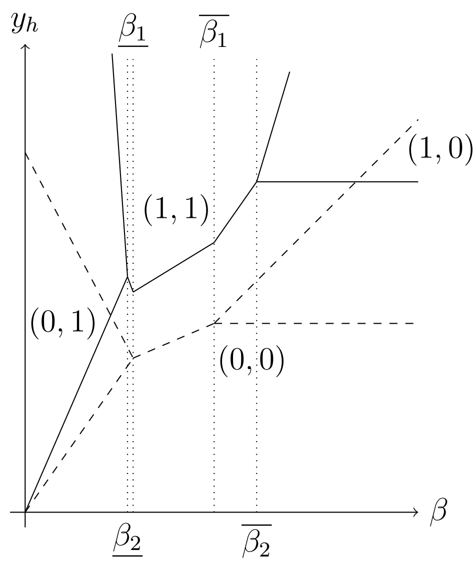


Figure 2.4: Contracts with full (dashed) and imperfect (solid) information.

A final conclusion that can be made from the comparison between the contracts with and without moral hazard is that the space in which e_2 is used decreases considerably because of the expansion of $(0, 1)$ to regions for which $(1, 1)$ was optimal. Also by the change in the slope of the frontier between $(1, 0)$ and $(1, 1)$ from the case with full to imperfect information, it is clear that not only $(0, 0)$ is to be implemented for a broader set of parameters, but $(1, 0)$ as well, although for higher values of y_h and β.

The study of the scenario with moral hazard shows that, when the agent has a low cost for creating the extra instance, it is very costly for the principal to dissuade him from shirking in the beginning with the hope of fixing a bad outcome in future scenarios.

2.4 Externalities

In this section, I assume that the principal cares about how the outcome was achieved. Some of the analogies previously made with the strategies that the agent can follow imply risky financial activities or dubious process manipulation. Even further, creating the extra chance can be sometimes illegal (forcing workers to work extra hours when it is not allowed), or incurring in practices that might be censored by a regulator or the industry.

These costs might be the expected punishment the principal faces if caught by the regulator. Consider the recent case in the automobile industry, in which many brands modified the computers in their motors to pass the emission regulation tests. If the people in charge knew that software like this could be produced, that has a high probability of succeeding to passing the test, at a very cheap cost (in the model this would be very high p_1 with a very low β) they have strong incentives to take their chances with the software. If the expected cost of the transgression is low, then the principal might contract strategies in which gambling is encouraged, obtaining higher profits: if β is low, the required compensation is also low, and therefore the principal obtains a bigger surplus.

In what follows, I will assume that exerting e_2, if detected by the regulator, triggers a punishment to the principal, such that its expected cost, if e_2 was exerted, is ξ. The immediate effect we can depict is that the strategies involving the second effort will see their expected return decreased, and therefore the strategies without e_2 will be optimal for a wider set of parameters.

The introduction of an externality cost, under perfect information, that is if the principal could indeed choose the optimal level of effort to maximize aggregated surplus, has a similar effect than facing a massive cost of e_2, or β. The detailed derivation of this version of the model is left for Appendix B.2.1. Figure 2.5 contains the main implications of introducing an externality in the case without moral hazard.

To start with, the strategy involving shirking and exerting effort when facing a bad outcome becomes less optimal than when there is no externality. Exerting effort in both periods becomes preferred over $(0, 1)$, but it is nevertheless implemented less than in the original model, losing against $(0, 0)$ and $(1, 0)$. As the shape of the plot remains more or less equal, we obtain the first result of introducing the externality. The variables $\underline{\beta}$ and $\overline{\beta}$ are shifted to the left along with the plot, exactly in the value of ξ. This implies that the strategy $(0, 1)$ might end up being completely wiped off in the absence of moral hazard for some value of ξ high enough, this threshold is stated in the following proposition:

Proposition 7. *With full information, if the cost* ξ *exceeds* $p_1/[(1-p_0)-p_1(1-p_1)]$ *then* $(0, 1)$ *is never optimal.*

Proof. In Appendix B.2.1 can be observed that $(0, 1)$ is implemented only for values of β between 0 and $p_1/[(1-p_0)-p_1(1-p_1)]-\xi$. As $\beta \geq 0$ then for $\xi > p_1/[(1-p_0)-p_1(1-p_1)]$ the set for which $(0, 1)$ results optimal is empty. $\square$

Now when considering moral hazard, I assume as usual that the principal cannot observe the agent's actions. Again, the introduction of the externality should carry similar consequences of those that higher values of β would carry.

The detailed solution of this problem can be found in Appendix B.2.2. As expected, now the strategies $(0, 1)$ and $(1, 1)$ have higher potential cost for the principal, and therefore they turn out to be optimal in a smaller parameter space when compared to the problem without the cost of ξ. The effect on the model with moral hazard is not

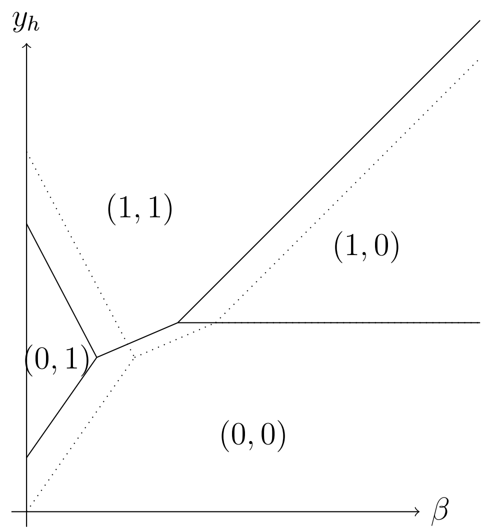


Figure 2.5: Contract with full information and externality.

identical to the one that happens in the version with perfect information though. Now, because of the agency problem, the plot does not move to the left as it happened before. Note that the values of $\underline{\beta}_1$ and $\overline{\beta}_1$ are not affected at all.

A similarity, though, with the full information case is that there is a value for ξ that makes $(0,1)$ never optimal.

Proposition 8. *If the cost ξ is higher or equal than*

$$\frac{p_1}{(1-p_0)(1-p_1)[(1-p_0)-p_1(1-p_1)]}$$

then $(0,1)$ is never optimal.

Proof. In Appendix B.2.2 it can be observed that $(0,1)$ is implemented only for values of beta such that

$$\beta \quad \in \quad \left[0, \frac{p_1^2+[p_1(1-p_0)(1-p_1)\{(1-p_1)p_1-(1-p_0)\}]\xi}{[p_0+(1-p_0)p_1](1-p_0)(1-p_1)+(p_1-p_0)p_1^2}\right]$$

As $\beta \geq 0$ then for $\xi \geq p_1/\{(1-p_0)(1-p_1)[(1-p_0)-p_1(1-p_1)]\}$ the set for which

(0, 1) is optimal is empty. □

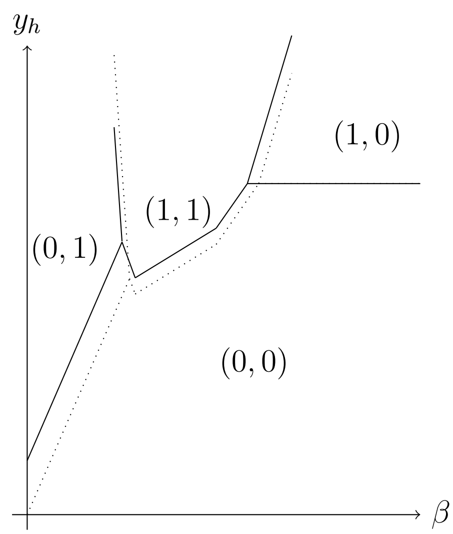


Figure 2.6: Contract under moral hazard and externality.

The strategy (1, 1) can never be ruled out for a finite externality cost, because even if $\xi > 1/(p_1 - p_0)$, which would make $\overline{\beta}_1$ negative, there can always be found a y_h big enough such that it will be optimal to exert effort in both opportunities. As a consequence, if the regulator would like to prevent second effort in any scenario, it would need, as expected, to relate the punishment (ξ) to y_h, such that the principal avoids implementing strategies with $(-, 1)$ at all. Furthermore, this would make the principal to implement shorter deadlines more often.

The agency costs are aggravated, which is reflected in Figure 2.6. In the region where the non-convexity of (0, 0) occurs, with the externality these costs are higher. As the principal is bearing the costs ξ and not the agent, the principal contracts strategies involving (0, 0) more aggressively to the left of $\underline{\beta}_1$.

2.5 Conclusions

I propose a model of moral hazard in which the agent has more than one chance to exert effort before revealing the output to the principal. The agent, after exerting some effort, observes an interim signal and then decides whether to make more effort to improve an adverse outcome or to deliver the outcome immediately.

Both agent and principal value this extra chance. I show how this situation creates extra rents that are split between principal and agent, when the principal offers the classical contract without the second chance, to an agent that can actually create this extra chance. Moreover, I find that when the principal knows about the extra chance, and there is no information asymmetry, both agent and principal value the possibility of fixing a bad outcome as an option, by being able to shirk at the beginning, if convenient, and later trying to fix a potentially adverse outcome.

However, once the information asymmetries are introduced, implementing a never shirking contract, and when the cost of creating the additional instance is low, happens for very high values of positive outcomes only. One of the main findings of the model is that with moral hazard the effort in the extra chance is not necessarily decreasing on its cost, existing values of output for which an increase in the cost of the second chance might increase the effort contracted in that second chance.

Finally, I study what happens if the extra chance can bring later consequences, and therefore costs, to the principal. This is done in order to adapt the model to situations in which this extra chance represents the use of illegal techniques to deliver the outcome as promised. I find that this possible cost for the principal causes that the strategies that do not involve the use of this extra chance become optimal for a wider set of parameters. Moreover, the principal will write contracts that incentive effort in the first period where with full information would have contracted effort only in the second chance, with the hopes of decreasing the probability of the agent using the second chance.

Chapter 3

Bidding in First Price Sealed Bid Common Value Auctions: A Computational Approach[1]

This chapter is a joint work with Ingemar Dierickx.

3.1 Introduction

In this paper we address the problem that bidders face in first price sealed bid (FPSB) auctions for common value goods: How much to shade their signals about the value of the object being sold. Countless firms are grappling with this problem every day. Yet, the auction literature offers little practical advice on how to determine an optimal bid. Two different disciplines, economics and decision sciences, have taken totally different approaches to the auction problem. Excellent reviews of the different approaches to auctions can be found in Rothkopf (2007) and Lorentziadis (2016).

[1]This chapter profited specially from the comments of Steffen Hoernig. We also appreciate valuable comments from Guido Maretto, Susana Peralta, Fernando Anjos, Pedro Vicente, Patrick Rey, Robert Wilson, and the participants at the Nova SBE RG Seminar.

The economics literature models auction outcomes as Bayes Nash equilibria. While it offers valuable advice for auctioneers on auction design, it provides little practical guidance to bidders. Some papers, most notably Wilson (1967, 1969, 1977, 1998), Milgrom and Weber (1982a), and more recently Hoernig and Fagandini (2018) actually try to pin down a bidding function for the bidders, but elegant analytical solutions are derived at the cost of a good amount of simplifying assumptions that are rarely satisfied in real life.

Another limitation of this literature is that theoretical results are usually derived in models with uniformly additive noise, an assumption that is required to solve those models analytically.[2] In the applied literature, normal and log-normal distributions are usually thought to be best suited to model real life problems. For example, electricity markets are usually modeled with additive normal noise, while hydrocarbon reservoirs in the petroleum industry are treated with a log-normal distribution.[3]

Although some of the economic theory work includes asymmetries on the parameters of the distributions of the signal noise — e.g. Wilson (1998) — most of the literature assumes symmetry at least in one dimension (for example, rationality of all the bidders). This is an important limitation. While symmetry assumptions are often necessary to ensure tractability of the models, they are, as Armantier and Sbai (2006) have pointed out, not often found in the real world. Furthermore, as we show in this paper, even small deviations from the classic symmetry assumptions result in sharply different optimal bids.

Recently the economic theory literature has tried to deal with more degrees of asymmetries among the population of bidders. In particular, Eyster and Rabin (2005), with their concept of *Cursed Equilibrium*, provide a new explanation for the Winner's

[2]Klemperer (2004), Krishna (2010), Milgrom (2004), Salant (2014).
[3]Crawford (1970), Smiley (1979).

Curse.[4] In their model, some agents fail to take into account how information impacts the other players' strategies. The cursedness hypothesis indeed helps to improve the fitting of laboratory data to the models, however it (and the data) also suffers from simplifying assumptions common to the rest of the theoretical literature: in general uniform distributions (and bounded domains), symmetry assumptions; and is hard to put into practice.

Crawford and Iriberri (2007) also study out-of-equilibrium models to explain bidders' behavior that is inconsistent with the traditional Bayes Nash solution. In particular, they apply the concept of *Level-K Thinking* introduced by Stahl and Wilson (1994, 1995). That means that bidders are assumed to reach a certain level of best responding. Say, if there are two bidders, both bidders assume a particular strategy by the other bidder. They best respond to that strategy. Later, they realize that the other player could have done the same reasoning, and so decide to best respond to their previous strategy. They can go in this reasoning forever, reaching a Nash Equilibrium. How many times they follow the reasoning of best responding to the previous strategy is the value of "k." The more times they are allowed to do that, the closer is the equilibrium to a Nash Equilibrium (which would require "∞-level"). They manage to explain data of overbidding in first price common value auctions — as well as in independent-private-value auctions — when compared to the Bayesian Equilibrium. There is no doubt about their contribution to explaining laboratory data; however, it does not look to offer practical advice on how a bidder should prepare a bid, suffering again from the classic simplifying assumptions of the economic theory literature we mentioned previously.

We believe that our model, with its simplifying assumptions, makes it more tangible for the practitioner while respecting other authors' appreciations. For example Dyer,

[4]We consider slightly above a decade as fairly recent in the auction literature.

Kagel, and Levin (1989) consider that some experienced bidders might follow simple strategies when facing an auction: *"We believe that in the field the executives have learned a set of situation specific rules of thumb..."* which would work only in invariant environments.

Decision theory has focused more on the bidders' problem. Since it is not, in general, feasible to analytically derive the Bayes Nash equilibrium bidding strategies in realistic scenarios, bidding strategies are optimized against a given distribution of competitors' bids relying mostly on Monte Carlo simulations.[5] There are some early decision theoretic and experimental contributions from the 50's and the 60's,[6] there are papers that address the bidders' problem within the context of a specific industry,[7] and there is a body of papers that propose computational methods to determine optimal bids, this time not in the sense of the Bayes Nash equilibrium.[8] One computational approach uses data from previous auctions and runs Monte Carlo simulations to determine optimal bids. Key papers include David (1993), Wen and David (2001), Ma, Wen, Ni, and Liu (2005), among others. In a second computational approach, data from earlier auctions are used to estimate the moments of the distribution from which bidders' signals are drawn, assuming that Bayes Nash equilibrium strategies were played by the other bidders, and optimal bids are then computed using classical symmetric models. Key papers include Bajari (1998), Bajari and Hortaçsu (2005), Campo, Perrigne, and Vuong (2003), among others.

In this paper, we build upon the pioneering work of Rothkopf (1969) and Wilson (1984), who use non-Bayes Nash equilibrium models, and propose a practical computational method that enables firms to submit bids that maximize ex-ante expected profits

[5] R. Wilson and M. Rothkopf are probably among the most prolific authors that actually looked at both sides, economic theory and decision sciences.

[6] For example Friedman (1956), Ortega Reichert (1968), among others.

[7] *E.g.* Capen, Clapp, Campbell, et al. (1971).

[8] From now on we will refer to equilibrium or optimal shading factor as the solution from the ex-ante signal point of view. When we refer to a Bayesian equilibrium, we will do so explicitly.

in a broad range of realistic valuation and information scenarios.

Specifically, we find a constant shading factor (SF) that is computed ex-ante of receiving the signal and:[9] (i) allows for a common value component as well as a firm-specific component in valuations, (ii) allows for differences in the accuracy of bidder signals, and (iii) allows for the introduction of non-rational bidders. Rothkopf (1980), and Compte and Postlewaite (2012) provide discussions on why constant strategies, and shading before observing the signal, should approximate Bayesian equilibrium strategies when the prior is diffuse.

To test whether the SF results in Bayes Nash equilibrium bids, we use as benchmark the solution to the symmetric problem found in Hoernig and Fagandini (2018),[10] and find that the SF exactly replicates those results when the prior is diffuse. In addition, we checked the programming code and the derived first order conditions with results from brute force Monte Carlo simulations in a broader range of scenarios.

We also generalize Robert Wilson's bias factor (BF) to obtain a measure of the Winner's Curse. In short, the BF shades the bidder's signal by the expected error of the signal conditional on winning. Therefore, this correction allows the bidder to obtain zero expected winning profits avoiding the Winner's Curse. If the BF allows the bidder to obtain zero expected winning profits, it is natural that an optimal constant shading should be larger. An optimal constant shading should take into account two effects, the Winner's Curse and the competitive effect. The more bidders are playing, the more aggressive a bidder must be in order to win the auction, however, the more bidders the more severe the Winner's Curse, and therefore the bidders should be more conservative. To disentangle these two opposite effects we consider the BF as the portion of the

[9] These types of strategies are to be expected in real life. For example Shachat and Wei (2012) find that while Bayesian equilibrium fits better English auctions data, constant shading strategies are more often found in first price sealed bid auctions.

[10] Wilson (1969) was the first one to find an explicit bidding function, however it was limited to only two symmetric bidders.

SF that takes care of avoiding the Winner's Curse, while the other accounts for the competitive effect.

Finally, we allow for a subset of "naive" bidders. These naive bidders follow a simple rule of thumb, and shade their signals by an arbitrary fixed amount. While the model allows to set any fixed shading for these naive bidders we assume they only neglect the Winner's Curse: they shade their bids by shading factor minus the bias factor. By doing so, we let these players to properly account for the competitive effect, but not to account for the Winner's Curse,[11] and analyze their impact on the optimal bids of sophisticated bidders. The presence of naive bidders in real life bidding problems cannot be denied. Occurrence of the Winner's Curse in common value auctions has been acknowledged since more than a half a century (Kagel and Levin, 2002). Dyer et al. (1989) document that in laboratory experiments even experienced executives in the construction industry, who are successful in their jobs, suffer from the Winner's Curse. They suggest that industry specific learning and situation-specific rules of thumb, which could not be applied in a laboratory setting, may help them avoid overbidding in the field. Furthermore, experienced contractors do suffer unanticipated losses when bidding on a type of project they are not familiar with.[12] Without going further, the study that brought the Winner's Curse into light by Capen et al. (1971) is based on field data. All these findings suggest that in real life auctions it is not unlikely that a subset of bidders may be "naive". We find that failing to account for the presence of naive bidders results in underbidding when there is one naive competitor and severe overbidding when the population of naive competitors is large.

The paper is organized as follows: In Section 3.2 we present the model and introduce the shading factor. In Section 3.3, we present the bias factor. In Section 3.4 we show

[11] The term "naive" is used in the same sense by Kagel and Levin (2002), and in a similar way in Lorentziadis (2012).

[12] Thaler (1988) provides a very good revision of laboratory and field studies that put in evidence the existence of the Winner's Curse.

how shading factors react to specific asymmetries in the bidding population. Section 3.5 summarizes key conclusions and implications of our work and suggests avenues for future research.

3.2 The Model

Assume that there are n bidders, competing in a first price sealed bid (FPSB) auction for an item with unknown value $\mu_i = \Delta_i + \mu$, where μ is the unknown common value part for everyone, and Δ_i represents a constant private value component. All bidders receive an unbiased signal of their valuation for the auctioned object $s_i = \mu_i + \epsilon_i$, with ϵ_i independently distributed $N(0, \sigma_i^2)$. This setup allow us to consider three sources of heterogeneity in the bidder population: bidder-specific valuation differences (Δ_i), differences in the accuracy of bidders' signals (σ_i^2), and the presence of naive as well as rational bidders. The heterogeneity that might exist among bidders and the number of bidders is common knowledge.

Specifically, we allow for two groups of bidders,[13] whose members receive a signal drawn from the same family of distributions but with different moments.[14] These may stem from a variety of factors such as different economies of scale, experience, technological factors, and so on. If $\Delta_i = 0$ for every bidder we have a pure common value scenario.

Finally, we allow for the presence of "naive" bidders. These bidders will shade some amount that does not maximize expected profits in any sense (nor ex-ante as we do, nor ex-post as in the Bayes Nash equilibrium). In particular, we will set a shading assuming

[13]The analysis can be extended to three or more groups. However, 2 groups are sufficient to show the main comparative statistics.

[14]Note, however that while the moments may differ the nature (Normal or Log Normal) of the underlying distribution is the same for all bidders. This is consistent with the auction literature and reflects the fact that the distribution of the error term is related to the characteristics of the auctioned item and not to the characteristics of the bidders.

these bidders fail to account for the Winner's Curse.[15] Rational bidders choose ex-ante equilibrium bids against the population of - rational and/or naive - opponents they are facing. What naive bidders do is assumed to be known by the rational ones.

Our model applies to two main cases: one with additive noise $s_i = \mu + \Delta_i + \epsilon_i$ with $\epsilon_i \sim N(0, \sigma_i^2)$, the other with multiplicative noise $s_i = \mu \Delta_i \eta_i$ with $\eta_i \sim LN(0, \sigma_i^2)$. We focus on the additive model, as the multiplicative model can be transformed into the additive model by taking the natural logarithm.

We use the fact that Monte Carlo simulations are often used in the bidding process, by simulating competitors' bids, but instead of actually assuming bids for other rational players, we look for ex-ante equilibrium bids among all the rational bidders. Also, instead of simulating many scenarios that could consume a huge amount of computing power, deliver not very precise estimates and be very inefficient to optimize with, we obtain the equilibrium constant shadings from the first order conditions. This helps to obtain accurate estimates quickly.

Let c_i be the shading applied by bidder i, for $i = 1, 2, ..., n$. That is, she submits a bid $b_i = s_i - c_i$, and consequently, conditional on μ, her bids are distributed $N(\mu + \Delta_i - c_i, \sigma_i^2)$.

The ex-ante expected profits for bidder i are:

$$\begin{aligned} E_{\mu,\epsilon_i,\epsilon_{-i}}[\pi_i] &= E_{\mu,\epsilon_i,\epsilon_{-i}}\left[(\mu_i - b_i)\prod_{j\neq i} \mathbb{1}_{b_i > b_j}\right] = \\ &= E_{\mu,\epsilon_i,\epsilon_{-i}}\left[(\mu + \Delta_i - (\mu + \Delta_i + \epsilon_i - c_i))\prod_{j\neq i} \mathbb{1}_{\mu+\Delta_i+\epsilon_i-c_i > \mu+\Delta_j+\epsilon_j-c_j}\right] = \\ &= E_{\epsilon_i}\left[(c_i - \epsilon_i)\prod_{j\neq i} F_j(\Delta_i - c_i - (\Delta_j - c_j) + \epsilon_i)\right] \end{aligned} \tag{3.1}$$

[15] While the model allows for any constant shading for these naive bidders, we consider they shade the SF minus the BF, which accounts only for the Winner's Curse and is further explained in Section 3.3.

where $f_k(\cdot)$ and $F_k(\cdot)$ are the density and cumulative distribution functions of ϵ_k for $k = 1, ..., n$ respectively. It is clear how from the ex-ante perspective, the common component μ cancels out, and therefore the ex-ante expected profits do not depend on μ nor on its prior distribution. The decision variable of bidder i is c_i, the amount she will shade her bid.

The shading factor for bidder i then is the c_i that maximizes (3.1). A Nash Equilibrium in pure strategies in shading factors is defined as:

Definition 3. *Let there be M bidders, from which $N \leq M$ are rational bidders. Let c_i be the shading applied by each bidder. Fix the naive bidders' shading at $(c_{N+1}, ..., c_M)$. The vector $(c_1, c_2, ..., c_N)$ is an equilibrium if, for a given $(c_{N+1}, ..., c_M)$ and for all i from 1 to N, it holds that:*

$$c_i \in \arg\max_{\hat{c}} E_{\epsilon_i}\left[(\hat{c} - \epsilon_i)\prod_{j \neq i} F_j(\Delta_i - \hat{c} - (\Delta_j - c_j) + \epsilon_i)\right] \tag{3.2}$$

The shadings applied by the rational bidders in equilibrium $\{c_i\}_1^N$ are called the Shading Factors.

When considering the optimization problem of bidder i, we obtain the following first order condition:

$$E_{\epsilon_i}\left[\left(\prod_{j \neq i} F_j(z_{ij} + \epsilon_i)\right)\left(1 - (c_i - \epsilon_i)\sum_{j \neq i}\frac{f_j(z_{ij} + \epsilon_i)}{F_j(z_{ij} + \epsilon_i)}\right)\right] = 0 \tag{3.3}$$

where $z_{ij} = \Delta_i - c_i - (\Delta_j - c_j)$. In Figure 3.1 it can be seen that the second order condition is satisfied as well. We generate that plot for each bidder, in every simulation, to verify that we are in the presence of a local maximum. Computational methods are used to solve this problem. The steps for the algorithm are as follows: (i) assume

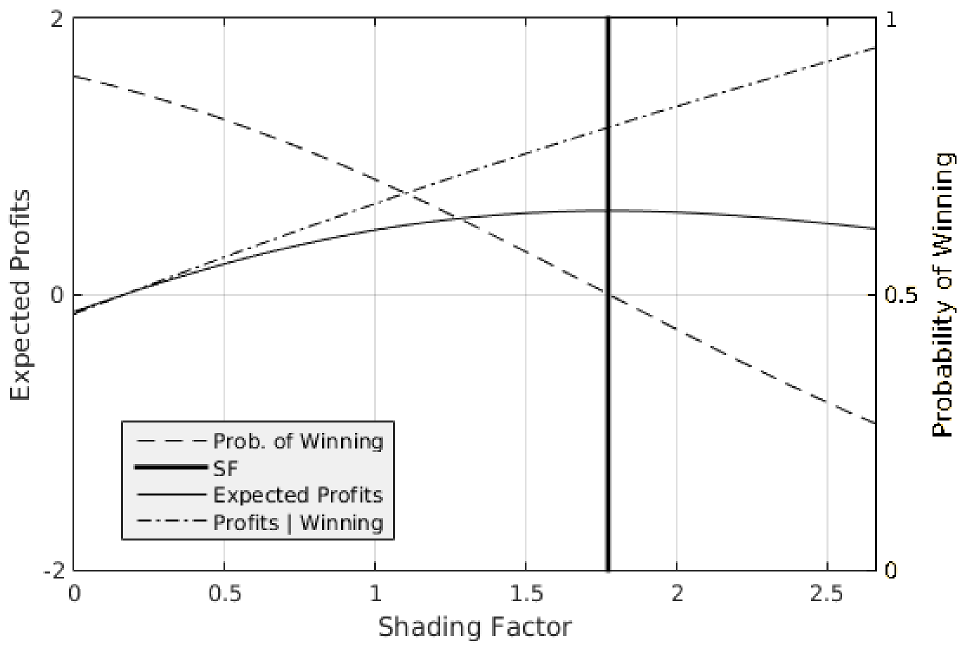


Figure 3.1: Expected profits, expected winning profits, and probability of winning for different values of the shading factor. Simulation for a bidder facing a symmetric competitor playing the equilibrium $SF \approx 1.77$.

that the initial shading for all the rational bidders is zero, (ii) find the optimal shading from the first order condition (3.3) for all the rational bidders, assuming all the other rational bidders' are shading according to the previously assumed shading, (iii) update the optimal shading for every rational bidder, (iv) assess how much has the SF changed compared to the previous value. If some bidder's SF has changed more than 10^{-4}, we iterate from (ii), using the last shadings found for the rational bidders. For cases where the standard deviation or the number of bidders is not too large, the algorithm converges relatively fast (in the order of a couple of seconds).

In Figure 3.1 we plot the ex-ante expected profits for a bidder facing a symmetric competitor, who is already applying the SF, as a function of her shading. We also plot the probability of winning and the ex-ante expected winning profits. Quick inspection of the results for larger numbers of bidders also shows that as the number of bidders increases, expected profits converge to zero — as they should.

We checked our results with Monte Carlo simulations in the following sense. We generated signals for the bidders, applied a shading, and averaged profits. Doing that we looked for the shading that gave bidders the highest expected profits, when considering the others' shadings. The estimates confirmed our results obtained by using the first order condition in (3.3).

Finally, we verified that our model replicates the optimal shading found in Hoernig and Fagandini (2018), who extended the original 2-bidders result from Wilson (1969) to multiple bidders. They derive the Bayes Nash equilibrium bidding function for the symmetric problem with n bidders and a diffuse prior. Their solution corresponds to:

$$b(s) = s - \sigma \times \frac{(n(n-1))^{-1} + \int_{\mathbb{R}} w f(w)^2 F(w)^{n-2} dw}{\int_{\mathbb{R}} f(w)^2 F(w)^{n-2} dw} \tag{3.4}$$

where σ is the standard deviation of each bidder's signal error. The fact that our model coincides with the Bayes Nash equilibrium for the symmetric case makes us believe that it should produce optimal bids for asymmetric auctions as well.

3.3 The Bias Factor

In any first price auction the bidder should bid below his valuation to have any profits. In a common value auction this is particularly true, as the expected winner's signal is above the true valuation, giving birth to the Winner's Curse.

More specifically, a bidder bidding in a first price sealed bid common value auction should take into account two considerations: the Winner's Curse and competition. Accounting for the Winner's Curse requires bidders to bid more cautiously as the number of bidders increases. On the other hand, increased competition decreases the probability of winning, leading the bidders to bid more aggressively.

The bias factor (BF) introduced by Wilson (1984) — and revisited by Cramton (1995) — enables us to disentangle these two effects.

The BF indicates by how many standard deviations the signal received by the winner exceeds her true value of the auctioned item.

$$BF_i = \frac{E[\varepsilon_i|win]}{\sigma_i} \tag{3.5}$$

For the symmetric case considered by Wilson, the BF corresponds to the expected signal error conditional on winning divided by the standard deviation.[16] Moreover, independently of the strategy chosen by the bidders, as long it is symmetric and increasing in the signal, the BF will tell exactly how the expected signal received by the winner will over estimate the true value of the auctioned object. Applying the (additive) bias factor to the signal ensures that the adjusted signal $\hat{s}_i$ is unbiased conditional on winning:

$$\hat{s}_i = s_i - BF_i \times \sigma_i$$
$$E[\hat{s}_i|win] = \mu + \Delta_i$$

Shading signals by the bias factor, i.e. $b_i = \hat{s}_i$, results in zero expected winning profits. Thus the BF given c_{-i} can be computed by solving

[16]Wilson divided by σ just to have a result that was scale-independent.

$$E_{\epsilon_i}\left[(c_i - \epsilon_i)\prod_{j\neq i} F_j(\Delta_i - c_i - (\Delta_j - c_j) + \epsilon_i)\right] = 0 \tag{3.6}$$

With heterogeneous bidders, the bias factor is no longer the highest order statistic of the signal error. Naturally, it is different for every bidder.

Bidders	BF	Prob. Win
2	0.56	0.500
3	0.85	0.333
4	1.03	0.250
5	1.16	0.200
6	1.27	0.167
7	1.35	0.143
8	1.42	0.125

Table 3.1: Bias factors as found by Wilson (1984) for symmetric bidders. $\epsilon_i \sim N(0,1)$.

We find that the classical results in the literature for the Winner's Curse (Capen et al., 1971, Kagel and Levin, 2002, Rothkopf, 1969, Wilson, 1967) hold. The Winner's Curse effect is stronger — and the bias factor is correspondingly larger — when the number of rivals is greater (Table 3.1), when rivals have an intrinsic firm-specific valuation advantage, and when the bidder signals are less accurate. We also find that the bias factor increases sharply when some rivals are naive. Intuitively, it is clear that winning against a naive bidder, who does not account for the Winner's Curse, would be "worse news" than winning against a rational bidder, and therefore that a larger adjustment is required to avoid the Winner's curse. However, the magnitude of this effect — particularly when there are two or more naive rivals — is unexpected.

While the bias factor is by no means a tool to generate optimal bids, it is useful in three very important ways. First, it provides a lower bound for an optimal shading. Second, it enables us to isolate the Winner's Curse effect, which helps to explain why bids do not change monotonically in the number of bidders, a fact that might at first seem surprising. We do that by computing the BF for a bidder facing competitors who use the correct SF. The SF, as we mentioned earlier, encompasses two effects, the Winner's Curse and competition. By isolating the Winner's Curse with the BF, we can obtain the competitive effect, which describes the trade-off between the gains in case of winning, and the probability of beating the other bidders. This is shown later in Table 3.2. Third, the bias factor enables us to suggest a plausible candidate for bids submitted by naive bidders; in turn this makes it possible to assess the impact of naive bidders on the bids — and expected profits — of rational bidders. These issues are taken up in Section 3.4 below. Although in our model — and in general — naive bidders could take any shading behavior we decide — for example, a constant markup of 20% of a standard deviation — we decided to give our naive bidders the benefit of the doubt, and let them account at least for the competitive effect, only ignoring the Winner's Curse. If the competitive factor is the difference between the SF and the BF, we let naive bidders shade their signals by $CF = SF - BF$.

3.4 Simulations and Predictions

In this section we first examine the shading factor in a few simple cases to verify whether our methodology produces results that are in accord with the standard results in the literature. Subsequently, we analyze the impact of valuation asymmetries (firm-specific valuation differences), information asymmetries (some bidders receiving a noisier signal than others), and the presence of naive bidders on equilibrium bidding strategies.

3.4.1 Rational Symmetric Bidders

In this subsection, we consider the symmetric all-rational model to verify consistency with predictions in the literature. Examining the SF, winning profits and the probability of winning for different numbers of bidders (N) and different levels of noise in their signals, confirms that our model reproduces standard results for a population of rational symmetric bidders.

For example, general results corroborate the intuition that the BF is a lower bound to the SF, and that the SF converges to the BF as the number of bidders increases, implying zero expected profits when bidders approach infinity, in accordance with the literature.

Table 3.2 shows an interesting pattern in the SF for a symmetric all-rational auction when signals are drawn from a standard normal distribution.

Bidders	BF	CF	SF	Exp. Winning Profits	Prob. Win
2	0.17	1.60	1.77	1.21	0.50
3	0.43	1.08	1.51	0.66	0.33
4	0.62	0.89	1.51	0.48	0.25
5	0.76	0.79	1.55	0.38	0.20
6	0.87	0.73	1.60	0.33	0.17
7	0.96	0.68	1.64	0.29	0.14
8	1.04	0.65	1.66	0.26	0.13

Table 3.2: Correction for symmetric bidders with $\epsilon_i \sim N(0, 1)$.

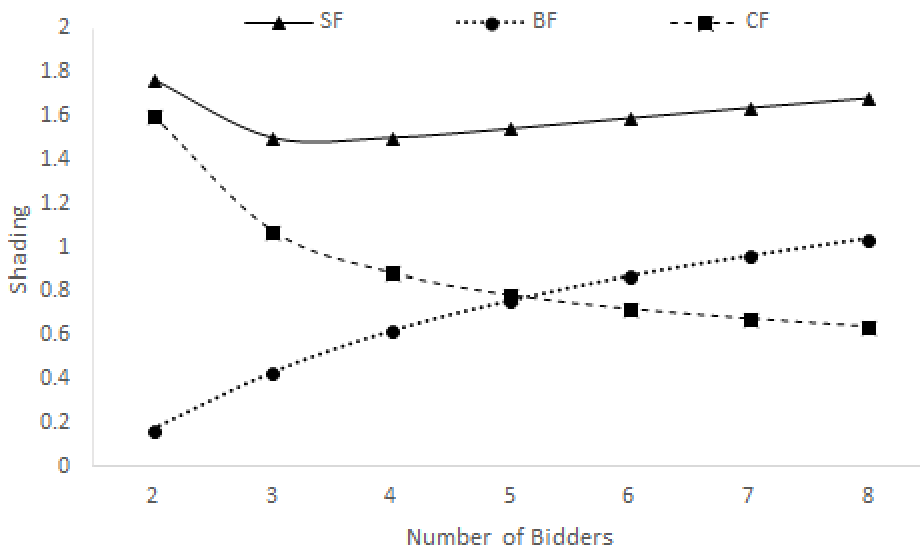


Figure 3.2: Decomposition of the shading factor in the Winner's Curse and the competitive effect. Symmetric bidders with $\epsilon_i \sim N(0,1)$.

As the number of bidders increases, the probability of winning and expected winning profits monotonically decline as they should. However, there appears to be an anomaly in the SF: the amount by which bidders shade their bids first decreases when N goes from 2 to 4, and then steadily increases from there on.[17] As suggested earlier in this paper, this non-monotonic pattern is the result of two effects working in opposite directions. As a bidder faces more rivals, (i) the Winner's Curse looms larger, which requires a larger adjustment of her signal, and (ii) competition increases, which requires her to bid closer to her signal, *i.e.* to sacrifice winning margin for a higher probability of winning. Initially, the competition effect dominates. With 2 bidders, low rivalry results in bidders shading their bids substantially; as the number of bidders gets larger, increasing rivalry leads them to bid closer to their valuation. With more than 4 bidders, however, the need to compensate for the Winner's Curse becomes the dominant effect, driving more conservative bids as the number of bidders increases.

Table 3.3 shows the impact of all bidders having less accurate signals, for a fixed

[17]This finding was confirmed also with Monte Carlo simulations to rule out potential coding issues.

number of bidders. As expected, the adjustment implied by the shading factor increases. It is interesting to note that in equilibrium profits increase when bidders have less accurate signals.

σ	SF	Winning Profits
0.5	0.75	0.24
0.7	1.05	0.33
0.9	1.36	0.43
1.1	1.66	0.53
1.3	1.96	0.62
1.5	2.26	0.72

Table 3.3: SF for 4 rational bidders. $\epsilon_i \sim N(0, \sigma)$.

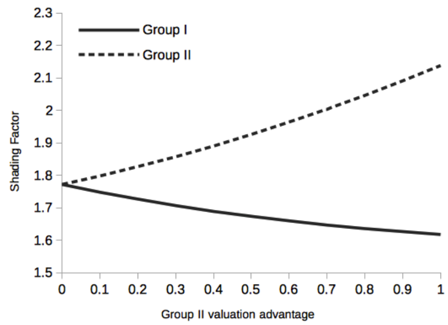
This result is consistent with a general finding that the efficiency of auctions increases as bidders' information (symmetrically) improves. For example, in a very different model Milgrom and Weber (1982b) analyze various information policies and show that auctioneer revenues increase when bidders are provided better information.

3.4.2 Valuation Asymmetries

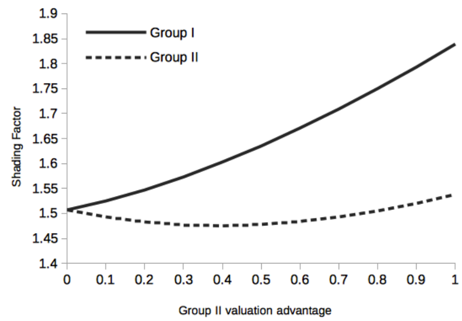
We consider two groups of bidders. The valuation advantage for Group I is held constant (Δ_I), whereas this parameter is varied for bidders in Group II (Δ_{II}). Throughout this subsection, both groups receive signals with errors generated $N(0, 1)$.

First we examine the case with two sophisticated bidders. Starting from the symmetric case where $\Delta_I = \Delta_{II}$, we progressively increase the mean of the distribution for Group II (i.e. Δ_2). Remember that Δ_I and Δ_{II} are assumed to be common knowledge. Without loss of generality, we assume that $\Delta_I = 0$.

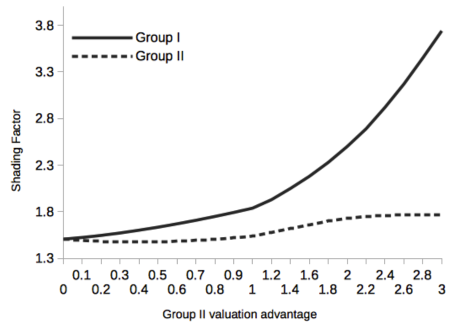
When there is no valuation asymmetry, both bidders obviously shade the same. As the valuation asymmetry increases we observe that the bidder with the lower valuation is shading less while the bidder with the higher valuation is shading more. The valuation advantage provides space for the high-valuation bidder to increase profits by shading



(a) 1 vs. 1



(b) 2 vs. 2



(c) 2 vs. 2

Figure 3.3: Valuation asymmetry between two groups of sophisticated bidders.

more without decreasing her probability of winning too much. The low-valuation bidder knows that his rival is bidding more conservatively, alleviating the Winner's Curse he would suffer by beating a bidder with a higher valuation. This motivates more aggressive behavior, as can be seen from the decreasing SF.

The picture changes dramatically when we have multiple bidders and thus competition within each group. In Figure 3.3b we observe that Group I always shades more (*i.e.* bids less) than Group II, and increasingly so as the valuation asymmetry increases. This makes sense: competition amongst high-valuation bidders in Group II requires low valuation bidders in Group I to bid conservatively to avoid the Winner's Curse. However, we also note a very interesting phenomenon: the pattern for Group II is non-monotonic on its valuation advantage. Initially, for small valuation asymmetries the shading factor declines (reaching a minimum at about $0.4 \times \sigma$) and then increases.

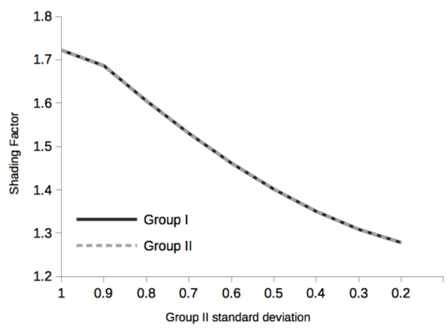
As can be seen in figure 3.3c, when valuation differences become very large the Shading Factor converges to the value for an auction with only two bidders (about 1.77). The increasing valuation gap combined with vigorous competition between Group II bidders requires Group I bidders to bid more cautiously to avoid the Winner's Curse. Put simply, when valuation differences become large, bidders in Group I become largely irrelevant, approaching a symmetric two bidder scenario.

3.4.3 Signal Quality

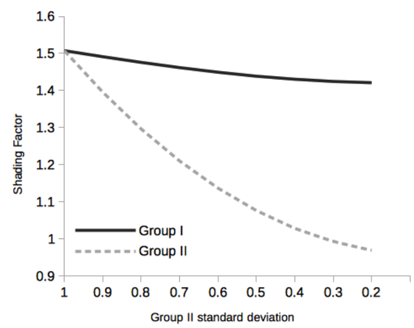
Information asymmetries have been widely studied in the context of bidding for oil and gas leases on the Outer Continental Shelf. As Hendricks and Porter (2007, p. 45) put it:

> *Oil and gas leases are classified into two categories. Wildcat tracts are located in previously unexplored areas. Prior to a wildcat auction, firms are allowed to conduct seismic studies, but they are not permitted to drill any exploratory wells. The seismic studies provide noisy, but roughly equally informative signals about the amount of oil and gas on a lease. We argue that wildcat auctions are likely to satisfy the symmetry assumption on the signal distribution. Drainage leases are adjacent to wildcat tracts where oil and gas deposits have been discovered previously. Firms that own adjacent tracts possess drilling information that makes them better informed about the value of the drainage tract than other firms, who are likely to have access only to seismic information. We argue that these auctions can be modeled by assuming one bidder has a private, informative signal and all other bidders have no private information.*

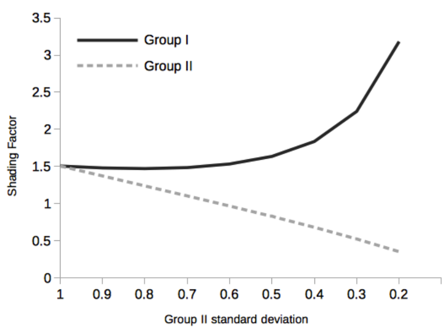
In this section, we again consider two groups, I and II. To focus on the effect of the quality of the signal only, we assume that $\Delta_I = \Delta_{II} = 0$. We assume that while bidders in Group I receive estimates with error $N(0,1)$, bidders on Group II receive signals with error $N(0,\sigma)$. Starting with $\sigma = 1$, we progressively improve Group II's signal quality — decreasing the standard deviation on their signal's error — to study the impact of an increasing information asymmetry. Note in the horizontal axis, that a decreasing value implies a better signal quality for Group II.



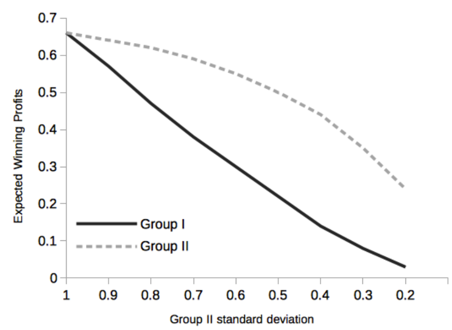
(a) 1 bidder in G I vs. 1 bidder in G II



(b) 2 bidders in G I vs. 1 bidder in G II



(c) 1 bidder in G I vs. 2 bidders in G II



(d) 1 bidder in G I vs. 2 bidders in G II

Figure 3.4: The impact of an information asymmetry on optimal bids.

Figure 3.4a shows that for the case of one less informed bidder competing against one better-informed rival, both bidding schedules roughly coincide. In equilibrium, the better-informed bidder shades her bid just enough to get close to her less-informed rival. Hendricks, Porter, and Wilson (1994) obtain a similar result in a somewhat different common value auction where one uninformed bidder, who only observes a public signal, competes with one informed bidder who also observes a private signal of the true value.[18]

However, the bidding schedules no longer coincide when there are more than two bidders, as can be seen in Figures 3.4b and 3.4c. Hendricks et al. (1994) argue that in the case studied — bidding for OCS oil and gas leases — informed bidders collude so that they can effectively be thought of as a single informed bidder. As for uninformed bidders, Porter (1995, p. 18-19) points out that:

> *The assumption that there is only one uninformed bidder is not important, provided that all uninformed bidders observe only public information. Then the equilibrium distribution function $G(b)$ characterizes the highest bid submitted by the uninformed bidders. If uninformed bidders have access to informative private signals of V then winning is an informative event for the informed firm (if it has less than perfect information concerning V), and the strategy of the informed bidder depends on the potential number of (relatively) uninformed bidders.*

The number of uninformed bidders is unimportant only *if they have no private information at all*, which is a strong assumption. In our model, less-informed bidders do have a private signal of the true value, albeit a less accurate one, and therefore the number of less-informed bidders does matter, as can be seen from the different bidding schedules in Figure 3.4b, 3.4c, and 3.4d.

In Figure 3.4d we see that expected profits of a less-informed bidder who competes with better-informed rivals rapidly vanish as the latter's information advantage increases. This is entirely as expected: as their information gets better, rivalry between the better-informed bidders becomes more intense, forcing the less-informed bidder to bid more conservatively in order to avoid the Winner's Curse. Note that in our model,

[18] See also Porter (1995), and Hendricks and Porter (2007).

the less-informed bidder does have a private signal of the true value, albeit a less accurate one. Therefore, while his expected profits rapidly vanish, they are not zero.

3.4.4 Naive Bidders

In this section we illustrate the effects of naive bidders on equilibrium bidding and expected profits of a sophisticated bidder. We consider in this section no valuation or information advantages for any bidder, so we isolate the effect of the sophistication asymmetry.

As defined in this paper, "naive" bidders, are aware that they need to shade somehow their bids in order to obtain positive profits. Even though the evidence (Kagel and Levin, 2002) suggests even experienced bidders use rules of thumb, and the model allows us to use any shading for naive bidders, we decided to give some degree of sophistication to the naive bidders: we allow them to understand that, the more bidders, the more aggressive they need to bid — although always below their signal. Therefore, we consider the competitive factor $CF = SF - BF$ as a sensible shading for the naive bidders.

		Number of Naive Rivals				
		1	2	3	4	5
Total Number of Rivals	1	0.52				
	2	0.15	0.10			
	3	0.08	0.04	0.02		
	4	0.05	0.02	0.01	0.00	
	5	0.03	0.02	0.01	0.00	0.00

Table 3.4: Expected profits for sophisticated bidders. $\epsilon_i \sim N(0,1)$.

For a sophisticated bidder, facing naive rivals is bad news. Table 3.4 shows the different dimensions in which naive rivals can impact the profits of sophisticated bidders, and the naive rivals' margin.

Total Number of Rivals	CF
1	1.60
2	1.08
3	0.89
4	0.79
5	0.73

Table 3.5: Shading applied by the naive bidders, CF, in the scenarios of Table 3.4.

In each column, the number of naive rivals is kept constant while the total number of competitors — hence the number of sophisticated bidders — varies. Moving down in any particular column shows the competitive pressure of rational competitors.

In each row moving to the right shows, for a given number of rivals, the increased competitive pressure as a result of replacing rational by naive competitors.

Finally, the diagonal shows what happens when a rational bidder faces a pool of naive rivals only: profits vanish *extremely* fast when their number increases. The devastating impact of naive rivals is due to the fact that they ignore the Winner's Curse and therefore shade *less* when they are facing more competitors, as can be seen in table 3.5.

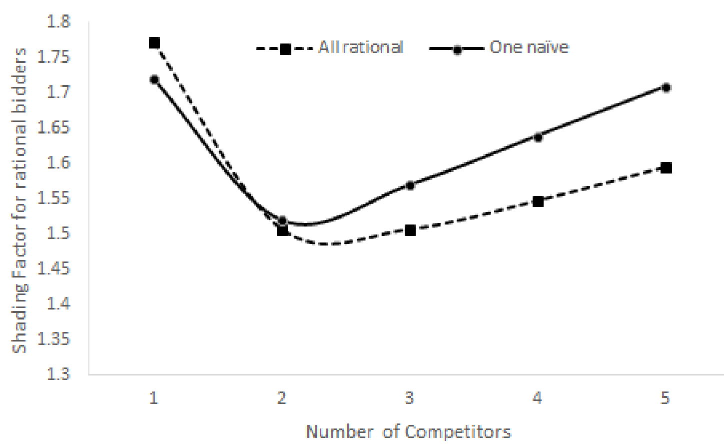


Figure 3.5: Optimal shading when all competitors are rational *vs.* one is naive.

Figure 3.5 shows Shading Factors for two cases: when all rivals are rational (square marks) and when one of the rivals is naive (round marks). When facing only one rival, a rational bidder bids more aggressively (*i.e.* shades less) when facing a more aggressive bid from a naive rival. This is because concerns about the Winner's Curse — which in this case is relatively weak — are dominated by competition. Facing more rivals, the opposite holds: concerns about the Winner's Curse dominate the impact of increased competition. This is because a naive bidder, as shown in Table 3.5, bids more aggressively when facing more competitors; in turn, this aggressive bidding behavior aggravates the Winner's Curse, requiring a more cautious bid to guard against it. In Figure 3.5, one can clearly see what happens when a bidder fails to account for the fact that one of her opponents is naive rather than rational: underbidding in the case of one opponent, and overbidding in all other cases.

3.5 Conclusions and Implications

In this paper, we compute the shading factor (SF) to obtain optimal bids in first price sealed bid common value auctions. The SF is computed ex-ante of receiving a signal, does not require a bounded support of either signals or bids, allow for differences in the accuracy of bidders' estimates, as well as firm-specific valuation differences. Furthermore, the SF also allow for (a subset of) "naive" bidders, who fail to account for the Winner's Curse.

We find that our model reproduces the solution to the Bayes Nash equilibrium with n symmetric bidders and a diffuse prior found in Hoernig and Fagandini (2018).

We also generalize Wilson's (1984) Bias Factor (BF) to obtain a measure of the Winner's Curse effect, allowing us to disentangle the SF's shading in two parts: the Winner's Curse Effect, and the competitive effect. Even though the interplay of different dimensions of bidder heterogeneity may lead to surprising shadings that are not monotonic in the number of bidders, all results can be understood by analyzing how those two effects, the competition effect and the Winner's Curse effect, work in opposite directions.

The BF also enables us to suggest a plausible candidate for bids submitted by naive bidders and to assess their impact on the expected profits of sophisticated bidders. Our results show that this impact is devastating. Hence, a critical task in real life bidding problems is to correctly gauge the level of sophistication of one's competitors. Absent specific information, it may be better to underestimate their capacities — and shade one's bid accordingly — rather than overestimate them and fall victim to the Winner's Curse. This, of course, opens a Pandora's box of tactical opportunities, since it would be advantageous for a sophisticated bidder to be *perceived* as naive. Given that bidders in real life do not know each other's type (rational or naive) with certainty, this opens up an interesting line of inquiry into the incentives for sophisticated bidders to pose as naive. Savvy rational bidders may wish to cultivate an image of being unsophisticated!

Bibliography

Ackerberg, D. A. and M. Botticini (2002). Endogenous matching and the empirical determinants of contract form. *Journal of Political Economy 110*(3), 564–591.

Armantier, O. and E. Sbai (2006). Estimation and comparison of treasury auction formats when bidders are asymmetric. *Journal of Applied Econometrics 21*(6), 745–779.

Auster, S. (2013). Asymmetric awareness and moral hazard. *Games and Economic Behavior 82*, 503–521.

Bajari, P. (1998). Econometrics of sealed bid auctions. In *Proceedings of the Business and Economic statistics section of the American Statistical Association*, pp. 41–49.

Bajari, P. and A. Hortaçsu (2005). Are structural estimates of auctions models reasonable? evidence from experimental data. *Journal of Political Economy 113*(4), 703–741.

Baker, G. P. and B. J. Hall (2004). Ceo incentives and firm size. *Journal of Labor Economics 22*(4), 767–798.

Baker, G. P., M. C. Jensen, and K. J. Murphy (1988). Compensation and incentives: Practice vs. theory. *The Journal of Finance 43*(3), 593–616.

Bandiera, O., L. Guiso, A. Prat, and R. Sadun (2015). Matching firms, managers, and incentives. *Journal of Labor Economics 33*(3), 623–681.

Bolton, P. and M. Dewatripont (2005). *Contract Theory*. The MIT Press.

Calveras, A., J. J. Ganuza, and E. Hauk (2004). Wild bids. Gambling for resurrection in procurement contracts. *Journal of Regulatory Economics 26*(1), 41–68.

Campo, S., I. Perrigne, and Q. Vuong (2003). Asymmetry in first-price auctions with affiliated private values. *Journal of Applied Econometrics 18*(2), 179–207.

Capen, E. C., R. V. Clapp, W. M. Campbell, et al. (1971). Competitive bidding in high-risk situations. *Journal of Petroleum Technology 23*(6), 641–653.

Chade, H. and V. N. V. de Serio (2014). Wealth effects and agency costs. *Games and Economic Behavior 86*, 1–11.

Clausen, A. (2013). Moral hazard with counterfeit signals. SIRE Discussion Papers 2013-13, Scottish Institute for Research in Economics (SIRE).

Compte, O. and A. Postlewaite (2012). Simple auctions. Working paper.

Cramton, P. C. (1995). Money out of thin air: The nationwide narrowband pcs auction. *Journal of Economics & Management Strategy 4*, 267–343.

Crawford, P. B. (1970). Texas offshore bidding patterns. *Journal of Petroleum Technology 22*(3), 283–289.

Crawford, V. P. and N. Iriberri (2007). Level-k auctions: Can a nonequilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions? *Econometrica 75*, 1721–1770.

Crocker, K. J. and J. Morgan (1998). Is honesty the best policy? curtailing insurance fraud through optimal incentive contracts. *Journal of Political Economy 106*(2), 355–375.

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

Dam, K. and D. Perez-Castrillo (2006). The principal-agent matching market. *Frontiers of Theoretical Economics 2*(1), 1–32.

David, A. K. (1993). A bidding model of perfect competition. In *IEE proceedings-C*, pp. 421–426.

Dyer, D., J. H. Kagel, and D. Levin (1989). A comparison of naive and experienced bidders in common value offer auctions: A laboratory analysys. *The Economic Journal 99*(394), 108–115.

Edmans, A., X. Gabaix, and A. Landier (2009). A multiplicative model of optimal ceo incentives in market equilibrium. *The Review of Financial Studies 22*(12), 4881–4917.

Eeckhout, J. and P. Kircher (2018). Assortative matching with large firms. *Econometrica 86*(1), 85–132.

Eyster, E. and M. Rabin (2005). Cursed equilibrium. *Econometrica 73*, 1623–1672.

Friedman, L. (1956). A competitive-bidding strategy. *Operations Research 4*(1), 104–112.

Gabaix, X. and A. Landier (2008). Why has ceo pay increased so much? *The Quarterly Journal of Economics 123*(1), 49–100.

Grossman, S. J. and O. D. Hart (1983). An analysis of the principal-agent problem. *Econometrica 51*(1), 7–45.

Hendricks, K. and R. H. Porter (2007). Chapter 32 an empirical perspective on auctions. In M. Armstrong and R. Porter (Eds.), *Handbook of Industrial Organization*, Volume 3 of *Handbook of Industrial Organization*, pp. 2073 – 2143. Elsevier.

Hendricks, K., R. H. Porter, and C. A. Wilson (1994). Auctions for Oil and Gas Leases with an Informed Bidder and a Random Reservation Price. *Econometrica 62*(6), 1415–1444.

Hoernig, S. and P. Fagandini (2018). An exact solution to the multiple-bidder pure common value first-price sealed bid auction with gaussian noise and diffuse prior. Nova SBE mimeo.

Holmstrom, B. and P. Milgrom (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica 55*(2), 303–328.

Holmstrom, B. and P. Milgrom (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization 7*, 24–52. Special Issue: Papers from the Conference on the New Science of Organization, January 1991.

Kagel, J. H. and D. Levin (2002). *Common value auctions and the winner's curse.* Princeton University Press.

Klemperer, P. (2004). *Auctions: Theory and Practice*. Princeton University Press.

Krishna, V. (2010). *Auction Theory* (Second Edition ed.). San Diego: Academic Press.

Laffont, J.-J. and D. Martimort (2002). *The Theory of Incentives The Principal-Agent Model*. Princeton University Press.

Legros, P. and A. Newman (1996). Wealth effects, distribution, and the theory of the organization. *Journal of Economic Theory 70*(2), 312–341.

Legros, P. and A. Newman (2007). Beauty is a beast, frog is a prince: Assortative matching with nontransferabilities. *Econometrica 75*(4), 1073–1102.

Lorentziadis, P. (2012). Optimal bidding in auctions of mixed populations of bidders. *European Journal of Operational Research 217*(3), 653–663.

Lorentziadis, P. (2016). Optimal bidding in auctions from a game theory perspective. *European Journal of Operational Research 248*, 347–371.

Lucas, R. J. (1978). On the Size Distribution of Business Firms. *The Bell Journal of Economics 9*(2), 508–523.

Ma, X., F. Wen, Y. Ni, and J. Liu (2005). Towards the development of risk-constrained optimal bidding strategies for generation companies in electricity markets. *Electric Power Systems Research 73*(3), 305–312.

Maggi, G. and A. Rodríguez-Clare (1995). On countervailing incentives. *Journal of Economic Theory 66*, 238–263.

Milgrom, P. and R. J. Weber (1982a). The value of information in a sealed-bid auction. *Journal of Mathematical Economics 10*(1), 105–114.

Milgrom, P. R. (2004). *Putting auction theory to work*. Cambridge University Press.

Milgrom, P. R. and R. J. Weber (1982b). A theory of auctions and competitive bidding. *Econometrica 50*(5), 1089–1122.

Ortega Reichert, A. (1968). *Models for Competitive Bidding Under Uncertainty*. Ph. D. thesis, Department of Operations Research and Department of Statistics, Standford University.

Porter, R. H. (1995). The Role of Information in U.S. Offshore Oil and Gas Lease Auction. *Econometrica 63*(1), 1–27.

Rothkopf, M. H. (1969). A model of rational competitive bidding. *Management Science 15*(7), 362–373.

Rothkopf, M. H. (1980). Equilibrium linear bidding strategies. *Operations Research 28*(3), 576–583.

Rothkopf, M. H. (2007). Decision analysis: The right tool for auctions. *Decision Analysis 4*, 167–172.

Salanié, B. (2005). *The Economics of Contracts, A Primer.* The MIT Press.

Salant, D. J. (2014). *A Primer on Auction Design, Management, and Strategy.* MIT Press.

Shachat, J. and L. Wei (2012). Procuring commodities: First-price sealed-bid or english auctions. *Marketing Science 31*(2), 317–333.

Shetty, S. (1988). Limited liability, wealth differences and tenancy contracts in agrarian economies. *Journal of Development Economics 29*(1), 1–22.

Smiley, A. K. (1979). *Competitive Bidding Under Uncertainty: The Case of Offshore Oil.* Ballinger Pub. Co.

Stahl, D. O. and P. W. Wilson (1994). Experimental evidence on players' models of other players. *Journal of Economic Behavior & Organization 25*, 309–327.

Stahl, D. O. and P. W. Wilson (1995). On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior 10*, 218–254.

Terviö, M. (2008). The difference that ceos make: An assignment model approach. *American Economic Review 98*(3), 642–668.

Thaler, R. H. (1988). Anomalies: The winner's curse. *Journal of Economic Perspectives 2*, 191–202.

Thaler, R. H. and E. J. Johnson (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Science 36*(6), 643–660.

Thiele, H. and A. Wambach (1999). Wealth effects in the principal agent model. *Journal of Economic Theory 89*(2), 247–260.

Varas, F. (2017). Managerial short-termism, turnover policy, and the dynamics of incentives. *The Review of Financial Studies*. Forthcoming.

von Thadden, E.-L. and X. Zhao (2014). Multi-task agency with unawareness. *Theory and Decision 77*(2), 197–222.

Wen, F. and A. K. David (2001). Optimal bidding strategies and modeling of imperfect information among competitive generators. *Power Systems, IEEE Transactions on 16*(1), 15–21.

Wilson, R. B. (1967). Competitive bidding with asymmetric information. *Management Science 13*(11), 816–820.

Wilson, R. B. (1969). Competitive bidding with disparate information. *Management Science 15*(7), 446–448.

Wilson, R. B. (1977). A bidding model of perfect competition. *The Review of Economic Studies 44*(3), 511–518.

Wilson, R. B. (1984). Lecture notes for course 310b analytical foundations of pricing strategy, stanford business school.

Wilson, R. B. (1998). Sequential equilibria of asymmetric ascending auctions: The case of log-normal distributions. *Economic Theory 12*(2), 433–440.

Zhao, X. (2008). Moral hazard with unawareness. *Rationality and Society 20*(4), 471–496.

Appendix A

Appendix Wealth and the Principal-Agent Matching

A.1 Derivation of the Partial Model

In this Appendix I provide the detailed steps to arrive from the maximization problem of the principal in (1.1) to the results shown in Table 1.1.

Given the simplifying assumptions in the model, it is possible to solve the problem by using the first order approach (as the agent's problem has a unique solution),[1] the optimal e for the agent, given a pair of a and b is given by the following condition:

$$\begin{aligned} p'(e)b &= c'(e) \quad \text{or,} \\ b &= \frac{c'(e)}{p'(e)} \end{aligned} \tag{A.1}$$

Equation (A.1) implies that $e = b\tau$. As expected, the amount of effort exerted by

[1] Grossman and Hart (1983) show how, under certain conditions, satisfied in this model, it can be solved in two stages, first the agent's problem, given the wages, and later the principal's problem.

the agent is increasing in the distance between the wage in the good and bad state (in other words, the size of the bonus), as well as in the level of his ability. We can replace then the first order condition (A.1) into the other equations of the principal's problem eliminating the variable b from it. The new problem for the principal is:

$$\max_{e,a} \quad -a + e[\xi - c'(e)] \tag{A.2}$$

$$s.t. \quad a + ec'(e) - c(e) \geq \overline{u} \tag{PC.2}$$

$$a \geq -\omega \tag{CC.2}$$

This reduced problem has a particular advantage. Besides having less variables to consider, it allow us to put the attention in the cash constraint for the agent. In particular we are interested on how affected are the incentives when (CC.2) is relevant for the principal's optimization. This will be the case as long as (CC.2) is binding. We will proceed assuming that the agent is not cash constrained, and solve the optimal compensation scheme. This is equivalent to think that ω is big enough such that the a obtained by maximizing subject to (PC.2) is higher than $-\omega$. Later we will solve the problem considering the opposite case, to finalize with the case in which both are binding.

Let ω be such that (CC.2) is not binding. From the new participation constraint we obtain the minimum a that would make the agent sign the contract. This a is given by:

$$a = \overline{u} + c(e) - ec'(e) = \overline{u} - c(e)$$

The optimal a depends positively on the agent's reservation utility but negatively on effort. This is the most direct way in which we can observe how the cash constraint (that forbids at some point decreasing a no matter the level of e) impedes the principal to achieve an efficient outcome. The optimal level for a can be replaced in the objective function for the principal's maximization problem described in (A.2) to end up with the following problem:

$$\max_{e} -\overline{u} + e\xi - c(e)$$

Whose first order condition with respect to e is $\xi = c'(e)$. This implies that the marginal benefit of e should be equal to its marginal cost, which is exactly the optimality condition in a first best situation. The solution should satisfy $e^* = \tau\xi$. It should not be surprising to find that for higher values of the output in the good state, the higher the contracted effort, as it happens with the agent's talent.

The wages set by the principal are $a = \overline{u} - \frac{\xi^2\tau}{2}$ and $b = \xi$. As stated previously, the principal charges the expected surplus of the operation whereas letting the agent keep the whole good outcome of the firm.

The expected utility for agent and principal are $\overline{u}+\omega$ and $\frac{\xi^2\tau}{2} - \overline{u}$ $(= -a)$ respectively. This shows also that the agent, when not cash constrained, is unable to extract any information rents from the principal. This result is widely known in the moral hazard literature. The principal therefore keeps the whole surplus of the project, and she only needs to satisfy the agent's reservation utility. This is equivalent to a first best outcome, when the principal has full knowledge of the agent's actions.

To obtain these results, it is important to recall that we have assumed that equation (CC.2) is not binding. Having an expression for a from the participation constraint

(CC.2), we can find a condition that indicates if either the participation constraint (PC.2) or the cash constraint (CC.2) is binding. For equation (CC.2) not to be binding it must hold that:

$$\overline{u} - \frac{\xi^2\tau}{2} > -\omega \tag{A.3}$$

From equation (A.3) we obtain already some conclusions. The higher the agent's reservation utility ($\overline{u}$) and the agent's wealth (ω), the less likely the cash constraint (CC) is going to be binding. In the opposite direction, the larger the size of the firm (ξ), the more likely the cash constraint is to be binding.

Assume now that equation (CC.2) is binding. That implies immediately that $a = -\omega$, and given that equation (IC) hasn't changed, $b = c'(e)$. Replacing that in the objective function, now the principal optimizes:

$$\max_{e} -(-\omega) + e[\xi - c'(e)]$$

That yields as solution $e^* = \frac{\tau\xi}{2}$. It is direct to see how the optimal effort has diminished. The optimal compensation scheme is now given by $a = -\omega$ and $b = \frac{\xi}{2}$. The expected utility for the agent and the principal are $\frac{\xi^2\tau}{8}$ and $\omega + \frac{\xi^2\tau}{4}$ respectively.

An interesting result is that the agent's expected utility is decreasing in ω and increasing in ξ. Remember that the higher ω, the less possibilities has the agent to extract information rents from the principal. Looking to the principal, $\tilde{u}_p$ is increasing in both ω and ξ.

Finally there is the situation in which both constraints PC and CC are binding (IC must always be binding with asymmetric information). In this case, we determine the optimal level of effort without having to maximize the principal's utility, as this level is determined by the system of equations given by the PC, CC and IC. We obtain that the implemented level of effort is:

$$e = \sqrt{(\overline{u} + \omega)2\tau} \tag{A.4}$$

Equation (A.4) implies that e is increasing in ω, τ, and $\overline{u}$. As the participation constraint is binding, the agent is still getting the expected utility $\overline{u}$. However, as $a = -\omega$, a is decreasing in the agent's wealth, and therefore the effectively paid bonus should increase. As it can be observed, to implement higher levels of effort it is necessary to have higher values of b as well. As consequence the expected bonus increases in ω, compensating the agent for the decrease in his fixed pay a and keeping his utility at $\overline{u}$. Conversely, decreasing the agent's wealth implies a higher fixed wage a, and therefore for the agent to obtain his reservation utility, a lower bonus is required implementing a lower level of effort. The principal suffers by having to implement lower levels of effort (compared to first best) and paying a higher fixed compensation, decreasing her expected utility (as she is getting lower probability of receiving ξ). The principal's expected utility can be written as:

$$E[u_p|\text{CC and PC binding}] = \xi\sqrt{(\overline{u} + \omega)2\tau} - 2\overline{u} - \omega$$

It is very important to recall that when the CC is binding for the agent, the PC must be satisfied, so the agent is always getting $\overline{u}$ or more. This implies that the agent is always at least as good (always weakly better) when he is cash constrained. On the opposite side, the principal is getting all the surplus when the agent is not cash constrained.

A.2 Proof of Proposition 1

In this Appendix I explore if the conditions which are sufficient to have generalized increasing differences (GID) are satisfied by the principal-agent model developed in the main text. The utility possibility frontier (UPF) generated by the original model is:

	CC	PC & CC	PC
$E[\Delta u]$	$\frac{\xi^2\tau}{8} - \omega$	u	u
$E[v]$	$\omega + \frac{\xi^2\tau}{4}$	$\xi\sqrt{(u+\omega)2\tau} - 2u - \omega$	$\frac{\xi^2\tau}{2} - u$

Table A.1: Utility Possibility Frontier.

In this UPF $E[v]$ represents the utility obtained from the principal after signing the contract with the agent. Each column represents the situation in which the agent is cash constrained, but the participation constraint is not binding, when the cash constraint and the participation constraint are both binding, and finally when only the participation constraint is binding. The relation between the variables that define in what situation we are, is given by the base wage determined when the agent is not cash constrained, against his wealth. If the fixed part of the optimal wage, assume an cash unconstrained agent, is lower than his minus wealth, then the agent is not cash constrained. Then, the agent is cash constrained if:

$$\overline{u} - \frac{\xi^2\tau}{2} < -\omega$$

This implies that the agent is not cash constrained if $\xi \geq \sqrt{\frac{2(\overline{u}+\omega)}{\tau}}$, so $\phi(\xi, \theta, \overline{u}) = \frac{\xi^2\tau}{2} - \overline{u}$ if $\xi \geq \sqrt{\frac{2(\overline{u}+\omega)}{\tau}}$. Note that this condition is equivalent to $\overline{u} \leq \frac{\xi^2\tau}{2} - \omega$ when looked from the point of view of $\overline{u}$.

With that, $\phi()$ can be written as:

$$\phi(\xi,\theta,\overline{u}) = \begin{cases} \frac{\xi^2\tau}{2} - \overline{u} & \text{if } \xi \leq \sqrt{\frac{2(\overline{u}+\omega)}{\tau}} \\ \xi\sqrt{(\overline{u}+\omega)2\tau} - 2\overline{u} - \omega & \text{if } \sqrt{\frac{2(\overline{u}+\omega)}{\tau}} < \xi \leq \sqrt{\frac{8(\overline{u}+\omega)}{\tau}} \\ \frac{\xi^2\tau}{4} + \omega & \text{if } \xi > \sqrt{\frac{8(\overline{u}+\omega)}{\tau}} \end{cases}$$

Which is above written as a function of ξ. This function is continuous in ξ. In order to be differentiable, let's look at its derivatives:

$$\frac{\partial\phi(\xi,\theta,\overline{u})}{\partial\xi} = \begin{cases} \xi\tau & \text{if } \xi \leq \sqrt{\frac{2(\overline{u}+\omega)}{\tau}} \\ \sqrt{(\overline{u}+\omega)2\tau} & \text{if } \sqrt{\frac{2(\overline{u}+\omega)}{\tau}} < \xi \leq \sqrt{\frac{8(\overline{u}+\omega)}{\tau}} \\ \frac{\xi\tau}{2} & \text{if } \xi > \sqrt{\frac{8(\overline{u}+\omega)}{\tau}} \end{cases}$$

The derivatives coincide in each interval, and therefore ϕ is differentiable in ξ. It remains to answer if $\partial\phi/\partial\xi$ is differentiable in $\overline{u}$, ω, and τ.

A.2.1 $\partial\phi/\partial\xi$ for $(\overline{u}+\omega)$

$$\frac{\partial\phi(\xi,\theta,\overline{u})}{\partial\xi} = \begin{cases} \frac{\xi\tau}{2} & \text{if } \overline{u}+\omega < \frac{\xi^2\tau}{8} \\ \sqrt{(\overline{u}+\omega)2\tau} & \text{if } \frac{\xi^2\tau}{8} \leq \overline{u}+\omega < \frac{\xi^2\tau}{2} \\ \xi\tau & \text{if } \overline{u}+\omega \geq \frac{\xi^2\tau}{2} \end{cases}$$

Which is continuous in $\overline{u}+\omega$. Derivatives with respect to $\overline{u}+\omega$ in each interval are 0, $\frac{\tau}{\sqrt{2}\sqrt{\tau(u+\omega)}}$, and 0 respectively. Evaluating in the extremes of the interval it gives $\frac{2}{\xi}$ and $\frac{1}{\xi}$ respectively. Therefore it is not differentiable in $\overline{u}+\omega$.

However, it is necessary that $\partial\phi/\partial\xi$ is increasing in $\overline{u}$ which it is. As the function is continuous, and increasing in $\overline{u}$ inside each interval, then $\partial\phi/\partial\xi$ **is increasing in** $\overline{u} + \omega$, which implies that it is increasing in $\overline{u}$ and ω.

A.2.2 $\partial\phi/\partial\xi$ for τ

$$\frac{\partial\phi\left(\xi,\theta,\overline{u}\right)}{\partial\xi} = \begin{cases} \xi\tau & \text{if } \tau \leq \frac{2(\overline{u}+\omega)}{\xi^2} \\ \sqrt{(\overline{u}+\omega)2\tau} & \text{if } \frac{2(\overline{u}+\omega)}{\xi^2} < \tau \leq \frac{8(\overline{u}+\omega)}{\xi^2} \\ \frac{\xi\tau}{2} & \text{if } \tau > \frac{8(\overline{u}+\omega)}{\xi^2} \end{cases}$$

Which is continuous in τ. The derivatives with respect to τ are: ξ, $\frac{\overline{u}+\omega}{\sqrt{2\tau(\overline{u}+\omega)}}$, and $\frac{\xi}{\tau}$ respectively. All of them positive, and therefore $\partial\phi/\partial\xi$ is increasing in τ in each interval. This added to continuity gives that $\partial\phi/\partial\xi$ **is increasing in τ.**

A.2.3 GID

The sufficiency conditions expressed in corollary 1 in Legros and Newman (2007, p. 1083) are:

- For GID in ξ and ω,

$$\frac{\partial^2\phi(\xi,\theta,u)}{\partial\xi\partial\omega} \geq 0 \qquad \text{and} \qquad \frac{\partial^2\phi(\xi,\theta,u)}{\partial\xi\partial u} \geq 0$$

- For GID in ξ and τ,

$$\frac{\partial^2\phi(\xi,\theta,u)}{\partial\xi\partial\tau} \geq 0 \qquad \text{and} \qquad \frac{\partial^2\phi(\xi,\theta,u)}{\partial\xi\partial u} \geq 0$$

In their proof, they use $\partial^2\phi/\partial\xi\partial\omega$ or $\partial^2\phi/\partial\xi\partial\tau$ to obtain that the derivative of ϕ with respect to ξ is increasing in the agent's type and utility (Legros and Newman, 2007, p.1097). Even though this function is not twice differentiable in the model presented here, we have shown that it is increasing in all the necessary variables and therefore the model in this economy satisfies GID in ξ, ω and ξ, τ.

Appendix B

Appendix Hunting with two Bullets

B.1 Baseline Model

B.1.1 Perfect Information

In this appendix I solve the problem with perfect information. For simplicity and without loss of generality I assume $e_h = 1$. For the same reason, I solve the problem assuming the agent owns the firm, or conversely, the principal exerts the effort. Because of this assumption, the maximization problem lacks considerations about the wage schedule. The maximization problem is:

$$\max_{(e_1,e_2)} p(e_1)y_h - e_1 + e_2[1 - p(e_1)]\{p(e_1, 1)y_h - \beta\}$$

There is no need to compare all the strategies for a given pair of β and y_h. For example, if $e_1 = 1$, we know that if $y_h \geq \beta$, then the agent will always exert effort for e_2, and therefore the strategy $(1, 0)$ will never be followed.[1] Conversely, if $y_h < \beta$ the

[1]Note that if $y_h \geq \beta$, given $\hat{y} = 0$ it is always optimal to exert $e_2 = 1$, as the revenues (y_h) are higher that the cost of that effort.

agent will never exert effort in his second chance, and therefore the strategy $(1, 1)$ should never be considered. We can say even more, as $y_h < \beta \Rightarrow y_h < \frac{\beta}{p_1}$, we can also discard $(0, 1)$ from the possibilities.[2] With this information, we can define a correspondence, relating values of β to strategies to be implemented for some value of y_h.

The agent will stick with a strategy involving $e_2 \neq 0$ if, after observing $\hat{y} = 0$, the return of exerting effort again is positive. In practical terms $(0, 1)$ will be optimal after observing $\hat{y} = 0$ if and only if $p_1 y_h > \beta 1$. In the same way $(1, 1)$ will be optimal after observing $\hat{y} = 0$ if and only if $y_h > \beta 1$.

From these two constraints, we obtain two critical values of y_h: $\frac{\beta}{p_1}$ and β. These critical points can be described in the diagram in Figure B.1.

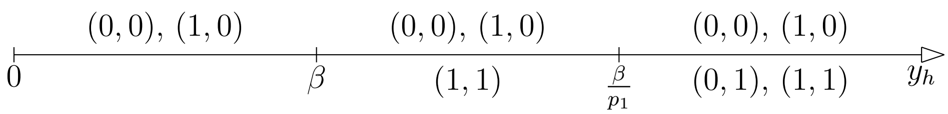


Figure B.1: Strategies to be considered, for some level of y_h, given β.

Figure B.1 shows which strategies are profitable as a function of y_h given β. We can say more though. If it is rational to exert effort in the second period, given effort in the first period (which implies that the probability of success is now 1), then not exerting effort in the second period given effort in the first period will never be played. In the same way it can be argued that, if it is rational to exert effort in the second period, given no effort in the first period (which implies that the probability of success is now p_1), then it is not rational to exert no effort in the first period without exerting effort in the second period. The new graphical description is in Figure B.2.

Comparing the profits between each pair, we can find the intervals that are relevant.

1. For the first interval $(0, 0)$ versus $(\geq)$ $(1, 0)$,

[2]Note that if $y_h < \beta$ then, given e_1, the expected revenues of $e_2 = 1$ are $p_1 y_h$, which is lower than β.

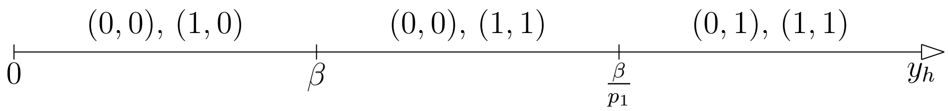


Figure B.2: Strategies that can be implemented for different levels of y_h.

$$\begin{aligned} p_0 y_h &\geq p_1 y_h - 1 \\ y_h &\leq \frac{1}{p_1 - p_0} \end{aligned}$$

2. For the second interval $(0,0)$ versus $(\geq)$ $(1,1)$,

$$\begin{aligned} p_0 y_h &\geq y_h - 1 - (1-p_1)\beta \\ \frac{1+(1-p_1)\beta}{1-p_0} &\geq y_h \end{aligned}$$

3. For the final interval $(0,1)$ versus $(\geq)$ $(1,1)$,

$$\begin{aligned} p_0 y_h + (1-p_0)[p_1 y_h - \beta] &\geq y_h - 1 - (1-p_1)\beta \\ 1 - (p_1 - p_0)\beta &\geq y_h (1-p_0)(1-p_1) \\ \frac{1-\beta(p_1-p_0)}{(1-p_0)(1-p_1)} &\geq y_h \end{aligned}$$

The next step is to check which of the strategy sets is empty.

Name the intervals from left to right as A, B, C, etc.

It is clear that set A is nonempty. For B to be nonempty, it is necessary that:

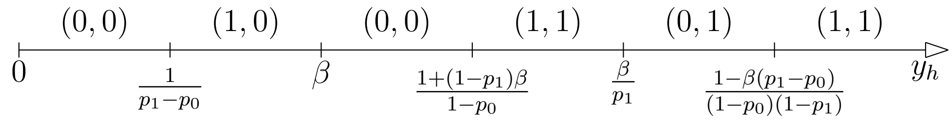


Figure B.3: Strategies that can be implemented for different levels of y_h.

$$\frac{1}{p_1 - p_0} \leq \beta$$

For C to be nonempty, it is necessary:

$$\begin{aligned} \beta &\leq \frac{1 + (1 - p_1)\beta}{1 - p_0} \\ (1 - p_0)\beta &\leq 1 + (1 - p_1)\beta \\ \beta &\leq \frac{1}{p_1 - p_0} \end{aligned}$$

From here we get that B and C are mutually exclusive sets. Now for D to be nonempty,

$$\begin{aligned} \frac{1 + (1 - p_1)\beta}{1 - p_0} &\leq \frac{\beta}{p_1} \\ p_1 + p_1(1 - p_1)\beta &\leq (1 - p_0)\beta \\ p_1 &\leq \beta[(1 - p_0) - p_1(1 - p_1)] \end{aligned}$$

The only way this could hold is if $(1 - p_0) - p_1(1 - p_1) > 0$, which happens when $p_0 < 1 - p_1 + p_1^2$. As $p_0 < p_1$ by assumption, the last inequality always holds, and therefore the condition for D being non empty is:

$$\frac{p_1}{(1-p_0)-p_1(1-p_1)} \leq \beta$$

For E to be nonempty:

$$\begin{aligned}\frac{\beta}{p_1} &\leq \frac{1-\beta(p_1-p_0)}{(1-p_0)(1-p_1)}\\ \beta &\leq \frac{p_1}{(1-p_0)-p_1(1-p_1)}\end{aligned}$$

So again, E and D are mutually exclusive. One more comparison I want to show is that indeed or B or C is always empty.

$$\begin{aligned}\frac{1}{p_1-p_0} &\leq \frac{1+(1-p_1)\beta}{1-p_0}\\ p_0(p_1-1) &\leq (1-p_1)^2\end{aligned}$$

Which always holds. Finally, analyzing the relationship between these two important thresholds for β,

$$\begin{aligned}\frac{1}{p_1-p_0} &> \frac{p_1}{(1-p_0)-p_1(1-p_1)}\\ 1-p_0-p_1+p_1^2 &> p_1^2-p_1p_0\\ 1 &> p_0(1-p_1)+p_1\end{aligned}$$

Which always holds, for being the right-hand side a convex combination of something strictly lower than 1, and 1.

For a given effort level e_h, the optimal contracts with moral hazard would exclude the combination A-C-E-F but consider only A-B-E-F, A-B-D-F, and A-C-D-F.

In Figure 2.2 it is represented the contract with full information. As can be seen in the derivation, this was done as if the principal and the agent were the same single person. Also the two key betas have been labeled as $\underline{\beta_1} = \frac{p_1}{(1-p_0)-p_1(1-p_1)}$ and $\overline{\beta_1} = \frac{1}{p_1-p_0}$.

B.1.2 Moral Hazard and Unaware Principal

When the principal is unaware of the possibilities of the agent, she will offer a contract with $w_h = \frac{1}{p_1-p_0}$ as the wage for a successful project, and $w_0 = 0$ otherwise.

In the agent's best response, this is exactly the division between $(0,0)$ and $(1,0)$, and as the contour of the $(0,0)$ region is increasing, it happens that, for $w_h = 1/(p_1-p_0)$ the agent never chooses $(0,0)$, independently of the value of β.

Moreover, by equalizing the frontier between $(0,1)$ and $(1,1)$ when $\beta < \underline{\beta}_1$, we obtain the intervals for β for which the agent will exert $(0,1)$, $(1,1)$ and $(1,0)$ respectively.

$$\frac{1}{p_1-p_0} = \frac{1-\beta(p_1-p_0)}{(1-p_0)(1-p_1)}$$
$$\tilde{\beta} = \frac{(p_1-p_0)-(1-p_0)(1-p_1)}{(p_1-p_0)^2}$$

If $\beta < \tilde{\beta}$, the agent's strategy is $(0,1)$, if $\beta \geq \tilde{\beta}$, the agent's strategy is $(1,1)$, and if $\beta > \underline{\beta_1}$ the agent chooses $(1,0)$.

B.1.3 Moral Hazard

For the contract under moral hazard what was considered interim rationality constraint in the case with perfect information, is now an additional incentive compatibility. First, it is necessary that the agent accepts a wage that will make him choose the strategy at $t = 0$ as decided the principal, but later it is further necessary that after the first realization he sticks with that strategy. Again this would imply that $w_h \geq \beta e_h$ for (e_h, e_h) and $w_h \geq \frac{\beta e_h}{p_1}$ for $(0, e_h)$, while the opposite should be true for $(e_h, 0)$ and $(0, 0)$ respectively.

The maximization problem for the principal is now:

$$\begin{aligned}
\max_{w_h,(e_1,e_2)} \quad & p(e_1)\{y_h - w_h\} + e_2[1 - p(e_1)]p(e_1, 1)\{y_h - w_h\} \\
\text{s.t.} \quad & p(e_1)w_h - e_1 + e_2[1 - p(e_1)]\{p(e_1, 1)w_h - \beta\} \geq 0 \\
& (e_1, e_2) \in \arg\max_{(\hat{e}_1,\hat{e}_2)} p(\hat{e}_1)w_h - \hat{e}_1 + \hat{e}_2[1 - p(\hat{e}_1)]\{p(\hat{e}_1, 1)w_h - \beta\}
\end{aligned}$$

The first constraint is the participation constraint that will make the agent to sign the contract at $t = 0$. The second constraint is going to make the agent to choose the strategy (e_1, e_2) that the principal desires.

The minimum necessary wage for each strategy, which satisfies all the constraints is given by the solution to the problem under perfect information, recalling though that the agent keeps only w_h and not y_h. The expected profits for each strategy for the principal are:

	$\beta \leq \frac{p_1}{(1-p_0)-p_1(1-p_1)}$	$\frac{p_1}{(1-p_0)-p_1(1-p_1)} \leq \beta \leq \frac{1}{p_1-p_0}$	$\frac{1}{p_1-p_0} \leq \beta$
$(0,0)$	$p_0 y_h$	$p_0 y_h$	$p_0 y_h$
$(1,0)$	-	-	$p_1\left(y_h - \frac{1}{p_1-p_0}\right)$
$(0,1)$	$\left(y_h - \frac{\beta}{p_1}\right)[p_0 + (1-p_0)p_1]$	-	-
$(1,1)$	$y_h - \frac{1-\beta(p_1-p_0)}{(1-p_0)(1-p_1)}$	$y_h - \frac{1+(1-p_1)\beta}{1-p_0}$	$y_h - \beta$

Table B.1: Expected profits for principal for each feasible strategy according to β, with moral hazard.

We then can separate the three cases, each one of them represented in a column.

The most complex is the first column while the simplest is the second column.

1. $\beta \leq \frac{p_1}{(1-p_0)-p_1(1-p_1)}$

   (a) $(0,0)$ versus $(0,1)$,

$$
\begin{aligned}
p_0 y_h &\geq \left(y_h - \frac{\beta}{p_1}\right)(p_0 + (1-p_0)p_1) \\
\beta\left(\frac{p_0 + p_1(1-p_0)}{(1-p_0)p_1^2}\right) &\geq y_h
\end{aligned}
$$

   (b) $(0,1)$ versus $(1,1)$,

$$
\left(y_h - \frac{\beta}{p_1}\right)[p_0 + (1-p_0)p_1] \geq y_h - \frac{1-\beta(p_1-p_0)}{(1-p_0)(1-p_1)}
$$

$$
\begin{aligned}
y_h \leq & \frac{1}{[(1-p_0)(1-p_1)]^2} - \ldots \\
& \ldots - \frac{\beta}{(1-p_0)(1-p_1)}\left[\frac{p_1-p_0}{(1-p_0)(1-p_1)} + \frac{p_0 + p_1(1-p_0)}{p_1}\right]
\end{aligned}
$$

   (c) $(0,0)$ versus $(1,1)$,

$$
\begin{aligned}
p_0 y_h &\geq y_h - \frac{1-\beta(p_1-p_0)}{(1-p_0)(1-p_1)} \\
y_h &\leq \frac{1}{(1-p_0)^2(1-p_1)} - \frac{\beta(p_1-p_0)}{(1-p_0)^2(1-p_1)}
\end{aligned}
$$

   The intersection of the constraints is at

$$\underline{\beta} = \frac{p_1^2}{-p_0^2(p_1-1)^2 + p_0(p_1^2-3p_1+1) + p_1(p_1^2-p_1+1)}$$

2. $\frac{p_1}{(1-p_0)-p_1(1-p_1)} \leq \beta \leq \frac{1}{p_1-p_0}$

   This is achieved by comparing the two implementable contracts that are at hand:

$$\begin{aligned} p_0 y_h &\geq y_h - \frac{1+(1-p_1)\beta}{1-p_0} \\ y_h &\leq \frac{1+(1-p_1)\beta}{(1-p_0)^2} \end{aligned}$$

3. $\frac{1}{p_1-p_0} \leq \beta$

   (a) $(0,0)$ versus $(1,0)$,

$$\begin{aligned} p_0 y_h &\geq p_1\left(y_h - \frac{1}{p_1-p_0}\right) \\ y_h &\leq \frac{p_1}{(p_1-p_0)^2} \end{aligned}$$

   (b) $(1,0)$ versus $(1,1)$,

$$\begin{aligned} p_1\left(y_h - \frac{1}{p_1-p_0}\right) &\geq y_h - \beta \\ y_h &\leq \frac{\beta}{1-p_1} - \frac{p_1}{(1-p_1)(p_1-p_0)} \end{aligned}$$

(c) $(0,0)$ versus $(1,1)$,

$$\begin{aligned} p_0 y_h &\geq y_h - \beta \\ \frac{\beta}{1-p_0} &\geq y_h \end{aligned}$$

The intersection of the constraints is at,

$$\overline{\beta} = \frac{p_1(1-p_0)}{(p_1-p_0)^2}$$

With a Cash Unconstrained Agent

In this appendix, I will develop the model with moral hazard, but assuming that the agent is no longer cash constrained.

The maximization problem for the principal is now:

$$\begin{aligned} \max_{(w_h,w_0),(e_1,e_2)} \quad & -w_0 + p(e_1)\{y_h - (w_h - w_0)\} + e_2[1-p(e_1)]\{p(e_1,1)[y_h - (w_h - w_0)]\} \\ \text{s.t.} \quad & w_0 - e_1 + p(e_1)(w_h - w_0) + e_2[1-p(e_1)]\{p(e_1,1)(w_h - w_0) - \beta\} \geq 0 \\ & (e_1,e_2) \in \arg\max_{(\hat{e}_1,\hat{e}_2)} w_0 - \hat{e}_1 + p(\hat{e}_1)(w_h - w_0) + \hat{e}_2[1-p(\hat{e}_1)]\{p(\hat{e}_1,1)(w_h - w_0) - \beta\} \end{aligned}$$

From where we can replace the participation constraint in the objective function and the incentive compatibility constraint. We obtain the following maximization problem:

$$\max_{(e_1,e_2)} -e_1 + p(e_1)y_h + e_2(1-p(e_1))[p(e_1,1)y_h - \beta]$$

Replacing the participation constraint in the first incentive compatibility cancels all the terms, once the optimal must be in the arg max and therefore they should coincide. As w_h and w_0, the principal will choose the effort levels such that the objective function is satisfied, and therefore the effort choice will coincide with the one in the case with full information. He then can adjust the w_h and w_0 to satisfy all the constraints, in particular, the second incentive compatibility.

B.2 Externality

B.2.1 Full information

In this appendix we repeat the process followed for the baseline contract with full information, but including the externality effect ξ. It will be clear that including this externality will imply a shift to the left on the frontiers between strategies. For space-saving reasons, I will omit several steps that are already clarified in Appendix B.1.

The agent/principal will stick with a strategy involving $e_2 \neq 0$ if after observing $\hat{y} = 0$ the return of exerting effort (again) is positive. In practical terms, the strategy $(0, e_h)$ will be optimal after observing $\hat{y} = 0$, if and only if $p_1 y_h > \beta e_h + \xi$. In the same way (e_h, e_h) will be optimal after observing $\hat{y} = 0$ if and only if $y_h > \beta e_h + \xi$.

From these two constraints, we obtain two critical values of y_h: $\frac{\beta+\xi}{p_1}$ and $\beta+\xi$. These critical points can be described in the diagram in figure B.4.

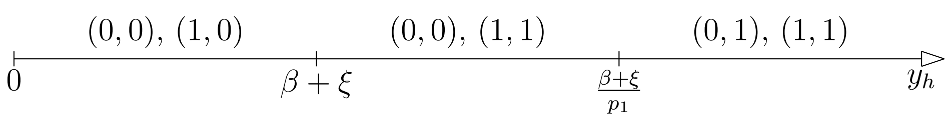


Figure B.4: Strategies that can be implemented for different levels of y_h.

Comparing the profits between each pair, we can find the intervals that are relevant.

1. For the first interval $(0,0)$ versus $(\geq)$ $(1,0)$,

$$y_h \leq \frac{1}{p_1 - p_0}$$

2. For the second interval $(0,0)$ versus $(\geq)$ $(1,1)$,

$$\frac{1+(1-p_1)(\beta+\xi)}{1-p_0} \geq y_h$$

3. For the final interval $(0,1)$ versus $(\geq)$ $(1,1)$,

$$\frac{1-(\beta+\xi)(p_1-p_0)}{(1-p_0)(1-p_1)} \geq y_h$$

The next step is to check if any of the strategy sets is empty.

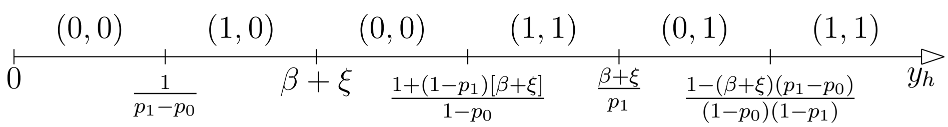


Figure B.5: Strategies that can be implemented for different levels of y_h.

Name the intervals from left to right as A, B, C, etc.

It is clear that set A is nonempty. For B to be nonempty, it is necessary that:

$$\frac{1}{p_1-p_0} \leq \beta+\xi$$

For C to be nonempty it is necessary:

$$\beta+\xi \leq \frac{1}{p_1-p_0}$$

From here we get that B and C are mutually exclusive sets. Now for D to be nonempty,

$$\frac{p_1}{(1-p_0)-p_1(1-p_1)} \leq (\beta+\xi)$$

For E to be non-empty:

$$\beta+\xi \leq \frac{p_1}{(1-p_0)-p_1(1-p_1)}$$

So again, E and D are mutually exclusive.

For a given effort level e_h, the optimal contracts with moral hazard would exclude the combination A-C-E-F, but consider only A-B-E-F, A-B-D-F, and A-C-D-F.

In Figure 2.5 it is represented the contract with full information and the externality. As can be seen in the derivation, this was done as if the principal and the agent were the same single person. The two key betas have been labeled as $\underline{\beta_1^E} = \frac{p_1}{(1-p_0)-p_1(1-p_1)} - \xi$ and $\overline{\beta_1^E} = \frac{1}{p_1-p_0} - \xi$. The superscript E means variables considering the externality.

B.2.2 Moral Hazard

The maximization problem for the principal is:

$$\begin{aligned}
\max_{w_h,(e_1,e_2)} \quad & p(e_1)\{y_h - w_h\} + e_2[1-p(e_1)]\{p(e_1,1)(y_h-w_h) - \xi\} \\
\text{s.t.} \quad & p(e_1)w_h - e_1 + e_2[1-p(e_1)]\{p(e_1,1)w_h - \beta\} \geq 0 \\
& (e_1,e_2) \in \arg\max_{(\hat{e}_1,\hat{e}_2)} p(\hat{e}_1)w_h - \hat{e}_1 + \hat{e}_2[1-p(\hat{e}_1)]\{p(\hat{e}_1,1)w_h - \beta\}
\end{aligned}$$

The first constraint is the participation constraint that will make the agent to sign the contract at $t = 0$. The second constraint is going to make the agent to choose the strategy (e_1, e_2) that the principal desires. Finally, we will also require interim rationality.

The minimum necessary wage for each strategy, which satisfies all the constraints is given by the solution to the problem under full information, recalling though that the agent keeps only w_h and not y_h. The expected profits for each strategy for the principal are:

	$\beta \leq \frac{p_1}{(1-p_0)-p_1(1-p_1)}$	$\frac{p_1}{(1-p_0)-p_1(1-p_1)} \leq \beta \leq \frac{1}{p_1-p_0}$	$\frac{1}{p_1-p_0} \leq \beta$
$(0,0)$	$p_0 y_h$	$p_0 y_h$	$p_0 y_h$
$(1,0)$	-	-	$p_1\left(y_h - \frac{1}{p_1-p_0}\right)$
$(0,1)$	$\left(y_h - \frac{\beta}{p_1}\right)[p_0 + (1-p_0)p_1] - (1-p_0)\xi$	-	-
$(1,1)$	$y_h - \frac{1-\beta(p_1-p_0)}{(1-p_0)(1-p_1)} - (1-p_1)\xi$	$y_h - \frac{1+(1-p_1)\beta}{1-p_0} - (1-p_1)\xi$	$y_h - \beta - (1-p_1)\xi$

Table B.2: Expected profits for principal for each feasible strategy according to β, with moral hazard.

We then can separate the three cases, each one of them represented in a column. The most complex is the first column while the simplest is the second column.

1. $\beta \leq \frac{p_1}{(1-p_0)-p_1(1-p_1)}$

   (a) $(0,0)$ versus $(0,1)$,

$$p_0 y_h \geq \left(y_h - \frac{\beta}{p_1}\right)(p_0 + (1-p_0)p_1) - (1-p_0)\xi$$
$$y_h \leq \beta\left(\frac{p_0 + p_1(1-p_0)}{(1-p_0)p_1^2}\right) + \frac{\xi}{p_1}$$

   (b) $(0,1)$ versus $(1,1)$,

$$\left(y_h - \frac{\beta}{p_1}\right)[p_0 + (1-p_0)p_1] - (1-p_0)\xi \geq y_h - \frac{1-\beta(p_1-p_0)}{(1-p_0)(1-p_1)} - (1-p_1)\xi$$

$$\begin{aligned}
y_h \leq & \frac{1}{[(1-p_0)(1-p_1)]^2} - ... \\
& - \frac{\beta}{(1-p_0)(1-p_1)} \left[\frac{p_1 - p_0}{(1-p_0)(1-p_1)} + \frac{p_0 + p_1(1-p_0)}{p_1} \right] - ... \\
& - \frac{p_1 - p_0}{(1-p_0)(1-p_1)} \xi
\end{aligned}$$

(c) $(0,0)$ versus $(1,1)$,

$$\begin{aligned}
p_0 y_h &\geq y_h - \frac{1 - \beta(p_1 - p_0)}{(1-p_0)(1-p_1)} - (1-p_1)\xi \\
y_h &\leq \frac{1}{(1-p_0)^2(1-p_1)} - \frac{\beta(p_1 - p_0)}{(1-p_0)^2(1-p_1)} + \frac{1-p_1}{1-p_0}\xi
\end{aligned}$$

The intersection of the constraints is at

$$\underset{\sim}{\beta}_2 = \frac{p_1^2 + [p_1(1-p_0)(1-p_1)\{(1-p_1)p_1 - (1-p_0)\}]\xi}{[p_0 + (1-p_0)p_1](1-p_0)(1-p_1) + (p_1 - p_0)p_1^2}$$

2. $\frac{p_1}{(1-p_0)-p_1(1-p_1)} \leq \beta \leq \frac{1}{p_1 - p_0}$

This is achieved by comparing the two implementable contracts that are at hand:

$$\begin{aligned}
p_0 y_h &\geq y_h - \frac{1 + (1-p_1)\beta}{1-p_0} - (1-p_1)\xi \\
y_h &\leq \frac{1 + (1-p_1)\beta}{(1-p_0)^2} + \frac{1-p_1}{1-p_0}\xi
\end{aligned}$$

3. $\frac{1}{p_1 - p_0} \leq \beta$

(a) $(0,0)$ versus $(1,0)$,

$$
\begin{aligned}
p_0 y_h &\geq p_1 \left(y_h - \frac{1}{p_1 - p_0} \right) \\
y_h &\leq \frac{p_1}{(p_1 - p_0)^2}
\end{aligned}
$$

(b) $(1,0)$ versus $(1,1)$,

$$
\begin{aligned}
p_1 \left(y_h - \frac{1}{p_1 - p_0} \right) &\geq y_h - \beta - (1 - p_1)\xi \\
y_h &\leq \frac{\beta}{1 - p_1} - \frac{p_1}{(1 - p_1)(p_1 - p_0)} + \xi
\end{aligned}
$$

(c) $(0,0)$ versus $(1,1)$,

$$
\begin{aligned}
p_0 y_h &\geq y_h - \beta - (1 - p_1)\xi \\
\frac{\beta}{1 - p_0} + \frac{1 - p_1}{1 - p_0}\xi &\geq y_h
\end{aligned}
$$

The intersection of the constraints is at,

$$
\tilde{\beta}_2 = \frac{p_1(1 - p_0)}{(p_1 - p_0)^2} - (1 - p_1)\xi
$$