Essays in Mechanism Design

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Mannheim

vorgelegt von

Frank Rosar

Mannheim 2007

Referent: Prof. Konrad Stahl, Ph.D.
Koreferent: Prof. Dr. Hans Peter Grüner
Abteilungssprecher: Prof. Dr. Hans Peter Grüner

Tag der mündlichen Prüfung: 18. Dezember 2007

Meinen Eltern

Acknowledgements

First of all, I wish to thank my supervisor, Konrad Stahl, for his encouraging support, his patience and his valuable advice. I learned very much from his numerous suggestions and comments. Moreover, I am indebted to Hans Peter Grüner for the discussion of my ideas, the encouraging support and the motivation to work on the problem in chapter 2 of this dissertation.

In addition, I would like to thank my colleagues at the Chair for Applied Microeconomics and my fellows at the CDSEM and in the Department of Economics for providing a lively atmosphere and a stimulating research environment. Chapter 4 benefitted a lot from the collaboration with Florian Müller. Furthermore, I thank Stefanie Brilon, Thomas Gall, Verena Glutting, Johannes Koenen and Elisabeth Schulte for their comments on the papers in this dissertation, and Jens Lehmann as well as Melanie Schienle for their support while I was working on this thesis. I am particularly grateful for endless discussions with Elisabeth.

Moreover, I thank the University of Toulouse I, where large parts of this thesis were written, for their hospitality. Financial support from the Deutsche Forschungsgemeinschaft and the University of Mannheim is gratefully acknowledged.

Contents

Chapter 1

General Introduction

This dissertation consists of three self–contained papers. In chapter 2 we analyze a collective decision problem, while in chapters 3 and 4 we study a theoretical and a more applied auction problem, respectively. Chapter 4 originated from joint work with Florian Müller.

1.1 Group Decisions with Dispersed Information and Interdependent Preferences

In chapter 2 we study a decision problem in which a group has to take a collective scalar decision, decision–relevant information is dispersed among the members of the group, and the individual group members' preferences are interdependent. Such a problem occurs, for instance, in a company whose department managers have to take collective decisions, for example about the level of the company's R&D expenditures. Although each manager is interested in the well–being of the company as a whole, it is likely that he is biased towards the well–being of his own department. Moreover, each manager may be endowed with superior information about his own department, for instance about the effect of the decision on it. Thus, the department managers face two conflicting goals: On the one hand, they want to aggregate information of the different departments in order to achieve decisions which are good for the company, but on the other hand, they want to use their private information strategically to push the decision into the direction which is best for their own departments. The question arises which decision making processes are capable of making use of as much information as possible without being too prone to opportunistic behavior.

Grüner and Kiel (2004) compare two distinct decision making processes within this environment. In the first one, each group member has to secretely vote for a real–valued decision and the average vote is implemented as collective decision. In the second one, the median vote is implemented. The two decision making processes differ in their ability to aggregate information and in their vulnerability to opportunistic behavior. Suppose a department manager believes that an extreme decision is optimal for his own department. Then, in case the collective decision is the average vote, this manager has the ability to push the decision far into his preferred direction by voting for a decision which is even more extreme than the one he considers best. For instance, referring to the R&D expenditure example, a manager who wants the company to spend one million may vote for an expenditure of three millions because he expects other managers to vote for less than one million. Since each group member can use extreme positions to exercise a large influence on the collective decision, this process is particularly vulnerable to opportunistic behavior. By contrast, if the median vote is implemented, extreme positions are disregarded such that the effects of opportunistic behavior are less severe. However, this comes at the cost that less of the agents' information is reflected in the decision. The main result of Grüner and Kiel (2004) is that, from a welfare point of view, either decision rule may be preferable depending on the degree of interdependence in preferences.

The reason why it is sometimes better to implement the median instead of the average vote is to confine opportunistic behavior which has more severe consequences when the average vote is implemented. The idea behind this paper is that there are often better means available to cope with opportunistic behavior than changing the decision rule. For instance, when the collective decision is the average vote, the problem of opportunistic behavior can be mitigated by forbidding extreme votes. Referring again to the example from above, one may forbid the managers to vote for expenditures exceeding one million. This reduces the ability of mangers who prefer extreme positions to manipulate the collective decision, while it maintains a larger responsiveness of the decision to the managers' information than when the median vote is implemented. Our main result is that if it is possible to restrict the group members' discretion in voting, then implementing the average vote is under general conditions preferable to implementing the median vote.

1.2 First–Price Auctions, Seller Information and Commitment to Sell

In chapter 3 we study the problem of a seller who wants to conduct a first–price sealed bid auction to sell an indivisible object. In such an auction each buyer is allowed to submit a sealed bid. The buyer who submits the highest bid obtains the object and has to pay an amount corresponding to his bid. The seller's degree of freedom when he designs the auction is the set of admissible bids. For instance, he may only allow bids above a certain reserve price, bids below a certain bid cap, or he may only allow a finite number of bids.

While the optimal auction problem is well understood under standard assumptions, some of the commonly made assumptions are often not met in practice. For instance, it is commonly assumed that the seller's reservation value does not change between the point in time at which he designs the auction and at which the auction is conducted. However, it often happens that the seller has to leave the buyers a reasonable amount of time to think about their bidding strategies. An example may be a procurement contract for sale that leaves bidders to prepare construction plans or prototypes. In the meantime the seller's information may improve and his estimate about his reservation value may become more accurate. Another commonly made assumption concerns the seller's commitment power. Normally it is assumed that he can commit not to withdraw the object after obtaining only unsatisfactory bids. However, even if he is not legally allowed to withdraw the object, he can often affect the outcome of the auction by submitting himself bids via a third party. Although this is usually also forbidden, it is hard to detect and there is empirical evidence that it does indeed happen in practice. Therefore we are interested in a setting which differs from the standard independent private values setting with risk–neutral buyers in two respects: firstly, the seller has to fix the rules of the auction before he learns his valuation of the object and secondly, he has the opportunity to keep the object after observing the bids.

In the first part of the chapter we consider the case with a risk–neutral seller to explain the effects which drive our results: If the seller was able to postpone the announcement of the reserve price until he becomes better informed, he would wait and choose different reserve prices for different realizations of his reservation value. But since he is not allowed to do this in our setting, he can only choose one (real) reserve price. Yet to a certain degree, he can mimic further reserve prices by choosing a non–connected set of admissible bids. A gap in this set forces each buyer to decide between bidding below the gap which means bearing a higher risk that the seller decides not to sell the object at the end, and bidding above the gap which increases the seller's eagerness to sell. The lowest bid above the gap can be interpreted as the reserve price relevant for buyers who want the seller to sell with a high probability. We show that for general distributions of buyer valuations it may be optimal for the seller to choose a set of admissible bids which exhibits gaps.

Crucial for the analysis in the first part of the chapter is that the seller cannot postpone the announcement of the auction rules until he is better informed. However, in many applications it is not clear why the seller should not be able to wait, at least with the announcement of some specifics of the auction such as the reserve price. In the second part of the chapter we relax this assumption. We show that a risk–averse seller might have a strict incentive not to wait. If he fixes a first–price auction with a non–connected set of admissible bids before he gets informed rather than waiting until he gets informed and announcing the rules of the auction at that point, he can sometimes obtain less variation in the buyers' bidding behavior without sacrificing expected revenue. Due to his risk–aversion, this makes him strictly better off. Thus, with a risk–averse seller who cannot commit to sell, the information structure and timing we assumed in the first part of the chapter may arise endogenously.

1.3 Asymmetric Procurement Systems

In chapter 4 we study the problem of a procurer, e.g. a carmaker, who needs to procure a specific part on a regular basis. In each period he has to decide between buying the part from the previous period's supplier, the incumbent, or from one of several entrants. Since incumbent and entrants differ in nature, the problem is inherently asymmetric. Due to his collaboration with the procurer in the previous period, the incumbent possesses process–specific knowledge which cannot be easily transferred. By switching to one of the entrants, the procurer has to incur switching costs. The level of these switching costs can largely be influenced by the incumbent. For instance, if the incumbent incurs relationship–specific investments to improve just–in–time production, it becomes more valuable for the procurer to continue the relationship with him. The procurer faces two conflicting goals: On the one hand, he wants to protect the incumbent in order to endow him with the right incentives to invest in the relationship, but on the other hand, he wants to exercise competitive pressure on him such that procurement prices are bid down competitively.

How the procurer's trade–off is resolved differs largely in different parts of the world. In the automotive industry, American carmakers try to exercise as much competitive pressure on their suppliers as possible, while Japanese carmakers try to build deep relationships with their incumbent supplier to obtain benefits from collaboration. Although the differences in the two systems are mainly rooted in differences in business culture and a different evolution of industry in the two countries, the question arises whether one system has structural advantages over the other one, in particular since in the recent past Western car producers tried to imitate their Asian counterparts.

One of the main differences between the two systems lies in the way the procurer treats the incumbent. In the "American system" the procurer does not constrain himself in the way he exercises competitive pressure on the incumbent, whereas in the "Japanese system" he grants the incumbent a distinct standing: He first tries to come to an agreement with him and only approaches the entrants in case they are not able to reach an agreement. Furthermore, to capture in our model that in practice there is often much more opportunistic behavior than assumed in large parts of the procurement literature (e.g. in Laffont and Tirole (1988) and in Bag (1997)), we assume that the procurer can commit himself to one of the two systems, but within the systems he behaves opportunistically. In particular, he cannot credibly promise the incumbent to reward his relationship–specific investments in a pre–specified way, although he can observe it. As consequence, the procurer does not design his procurement process to induce a certain investment behavior by the incumbent, but the incumbent invests to have the procurer choose a procurement process which gives him some preferential treatment.

With this framework we find a theoretical foundation for some empirical findings in the literature on the two procurement systems such as that the characteristics of the part procured determine which system performs better (see, e.g., Hahn, Kim, and Kim (1986) and Dyer (1996)). In particular, we find that the procurer's preferences over the two systems depend in a non–monotonic way on the importance of the investment associated with the part procured. While for low and for high investment costs, i.e. for important and relatively unimportant investment, the "American system" is superior, the “Japanese system” is better for intermediate costs, i.e. for intermediate important investment. Furthermore, we find that the incumbent is not inevitably better protected if the procurer does not use the entrants to exercise direct pressure on him.

Chapter 2

Group Decisions with Dispersed Information and Interdependent Preferences

2.1 Introduction

There are two main reasons for delegating a decision to a group. First, although the preferences of group members may be interdependent because of altruism or due to spill–over effects, it is likely that they are not completely aligned. By involving several individuals in the decision making process, divergent interests can better be taken into account. Second, information relevant to evaluate the possible decisions might be distributed among the members of the group. Hence, information aggregation may be improved.

We are interested in situations in which a group is already installed to take a collective scalar decision, information relevant for the decision is dispersed among the members of the group, and the group members' preferences are interdependent. Such a problem occurs, for instance, in a company whose department managers have to take collective decisions. Each head of department is endowed with superior information about his own department and, although he is also interested in the well–being of the company as a whole, he is biased towards the well–being of his own department. Other examples include national parliaments, in which the representatives of different regions meet; or decisions taken by the members of a family.[1]

A large number of decision making processes is conceivable and may be classified according to whether cheap talk is possible between the group members before the decision is taken and according to whether monetary incentives are feasible. If monetary incentives are feasible, the problem of finding the welfare maximizing decision making process turns into a standard mechanism design problem. In this case the first–best process can be implemented as ex post Nash equilibrium using an expected externality mechanism.[2] However, it is often not possible[3] or not desirable[4] to set monetary incentives. Under absence of monetary incentives the problem is non–trivial and implementability is affected by whether cheap talk is possible before the collective decision is taken or not.

In this paper we are concerned with the class of decision making processes in which monetary incentives are not feasible and cheap talk is excluded a priori. In particular, we are interested in processes which can be implemented by first asking the group members to secretely vote for some decision, and then applying a decision rule mapping their votes into a collective decision. Thereby we allow for the possibility that group members are restricted in choosing a vote. For instance, they may only be allowed to choose from a discrete set of possible votes or from votes belonging to a certain interval. Two particular decision rules are of special interest: In the first, the decision is the average of all votes (henceforth *mean decision rule*) and in the second, it is the median vote (henceforth *median decision rule*).

If the preferences of all group members are completely aligned (henceforth *common preferences*), the only concern of each group member is that the decision mechanism can appropriately make use of his information. In contrast, if a member's preferences depend only on his own information (henceforth *private preferences*), his sole aim is to push the collective decision into his preferred direction. For the intermediate cases in which preferences are interdependent but not completely aligned, a group member faces a trade-off between these goals.

To what extent the collective decision reflects the group members' information depends on the decision making process employed. On the one hand, a process which

[1]For further examples see section 1.2 in Grüner and Kiel (2004).

[2]See Bergemann and Morris (2006).

[3]E.g., for legal reasons or because group members do not respond to monetary incentives.

[4]Think of decisions taken in a parliament or by a jury in a court. Furthermore, monetary incentives may not be desirable if it is too costly to design them.

endows each group member with more influence on the decision can result in more of the members' information being utilized for the decision. But on the other hand, such a process improves also the group members' opportunities to behave opportunistically. A welfare maximizing planner designing the decision making process therefore faces a trade–off between avoiding opportunistic behavior and the responsiveness of the decision to the information. How severe this trade–off is depends on how eager the members are to behave opportunistically and thus on the degree of interdependence in preferences.

We are interested in comparing the performance of the two classes of decision making processes in which either the mean or the median decision rule is used and the set of admissible votes is part of the mechanism design problem. The questions we are going to address are 'Which decision rule should be used for which degrees of interdependence?' and 'How large should each individual group member's influence on the decision be?'[5]

This Contribution

A broader categorization of the related literature can follow the same lines as in section 1.3 of Grüner and Kiel (2004) since our paper builds directly upon theirs. Grüner and Kiel (2004) analyze the mean and the median decision rule when voting is not restricted, i.e. when group members are allowed to vote for any real–valued decision. Their main finding is that, depending on the degree of interdependence, either decision rule may attain a higher level of welfare. For preferences close to private the median decision rule is preferable, while for preferences close to common the mean decision rule is.

We also consider the mean and the median decision rule, but we allow restrictions on admissible votes. Since for either decision rule and for almost all degrees of interdependence group members prefer decisions that are more extreme than is socially optimal, welfare can be enhanced by forbidding at least the most extreme votes.

Our main results describe how the findings in Grüner and Kiel (2004) change when we compare the mean and the median decision rule for the case with optimally restricted instead of unrestricted voting. Asymptotically, i.e. for large groups, we find that the mean decision rule is not only preferable for common preferences, but also for private preferences. Thus, if it is possible to restrict the group members' discretion in voting, the median decision rule is no longer needed. The heuristics behind this result are the

[5] A group member's influence on the collective decision is jointly determined by his discretion in choosing a vote and the decision rule.

following: If preferences are private, group members are very prone to opportunistic behavior. This makes the decision process that relies on an unrestricted mean decision rule perform badly, because for this process the influence of each individual group member's vote on the group decision is large, and thus also his ability to behave opportunistically. By consequence, welfare can be improved by reducing the group members' influence on the decision. Two different ways of doing this are choosing the median decision rule (then only the median vote affects the decision) and sticking to the mean decision rule but restricting the group members' discretion in choosing a vote (for instance by forbidding extreme votes). Since the second way is also capable of reducing opportunistic behavior, but allows for a larger responsiveness of the decision to the group members' information, it outperforms the first way.

For the non–asymptotic case the intuition behind the comparison result remains valid, but is technically harder to prove. However, for uniformly distributed signals we have enough structure to prove that for any number of group members and for any degree of interdependence in preferences the optimally restricted mean mechanism is preferable to the optimally restricted median mechanism. Finally, we address the robustness of this result. We show numerically that it stays valid for small numbers of group members for distributions with linear densities as well as for symmetric distributions with quadratic densities.

Structure of the Paper

We present the model in the subsequent section. In section 2.3 we discuss mechanisms relying on the mean and the median decision rule separately before we compare them in section 2.4. Finally, we conclude in section 2.5. All proofs can be found in the Appendix.

2.2 The Model

There is a group consisting of n symmetric agents which has to take a collective decision $x \in \mathbf{R}$. For convenience we assume n to be an odd number larger than three such that $m := \frac{n+1}{2}$ is an integer. Throughout the paper we will denote a generic agent by i.

The signal θ_i is private information of agent i. All signals are independently drawn from a cumulative distribution function $\Phi(\cdot)$ with a connected support $\Theta = [\underline{\theta}, \overline{\theta}]$. We assume that a probability density function $\phi = \Phi'$ exists and is strictly positive on the support. Furthermore, we assume the distribution to be normalized such that the expected signal is zero, i.e. $\mathbf{E}[\theta_i] = 0$. We denote the variance of the distribution by $\sigma^2 := \mathbf{E}[\theta_i^2]$ and the median of the distribution by $\theta_{\text{Med}} := \Phi^{-1}(\frac{1}{2})$. Later on we will be interested in the distribution of the median signal, i.e. the mth highest of n independently drawn signals. We denote this signal by $\theta_{m:n}$, the probability density function according to which $\theta_{m:n}$ is distributed by $\phi_{m:n}(\theta) := \frac{n!}{(m-1)!^2}\Phi(\theta)^{m-1}\phi(\theta)(1 - \Phi(\theta))^{m-1}$, and the respective cumulative distribution function by $\Phi_{m:n}(\theta)$.

Agent i's payoff is $u_i = -(x-\theta_i^*)^2$ such that he is interested in minimizing the distance between the collective decision x and his preferred decision θ_i^*. His preferred decision is a convex combination of his own private information and the private information of all other agents: $\theta_i^* := (1-\alpha)\theta_i + \frac{\alpha}{n-1}\sum_{j\neq i}\theta_j$ with $\alpha \in [0, \overline{\alpha}_n] := [0, \frac{n-1}{n}]$. The polar cases $\alpha = 0$ and $\alpha = \overline{\alpha}_n$ describe *private* and *common preferences*, respectively. If preferences are private, an agent's preferred decision depends only on his own private information. If preferences are common, it depends symmetrically on each agent's private information. In this case the preferred decisions of all agents coincide. The parameter α measures the degree of interdependence in preferences. The higher α is, the stronger is interdependence.

A decision making process is characterized by a *mechanism* $\Gamma = (V, x)$ consisting of a set of *admissible votes* $V \subset \mathbf{R}$ and a *decision rule* $x : V^n \to \mathbf{R}$ mapping votes into decisions. The performance of a mechanism depends on how well it aggregates the agents' true signals. Therefore we introduce a specific notation for how the decisions made in equilibrium depend on the agents' signals instead of their votes. We call a function $d : \Theta^n \to \mathbf{R}$ mapping types into decisions a *decision function* and we call a decision function d *implementable* if there exists a mechanism $\Gamma = (V, x)$ and an equilibrium (described by the voting rules $v_i : \Theta \to V$) inducing it, i.e. if $d(\theta_1, \dots, \theta_n) = x(v_1(\theta_1), \dots, v_n(\theta_n))$.

We are interested in mechanisms relying on the *mean decision rule*, defined by $x_1(v_1, \dots, v_n) := \frac{1}{n}\sum_i v_i$, and the *median decision rule*, defined by $x_2(v_1, \dots, v_n) := \text{median}\{v_1, \dots, v_n\}$. To any mechanism relying on the mean decision rule, $\Gamma = (V, x_1)$, we refer as *mean mechanism* and to any mechanism relying on the median decision rule, $\Gamma = (V, x_2)$, we refer as *median mechanism*.

As equilibrium concept we adopt the notion of Bayesian Nash equilibrium.

2.2.1 Welfare

The sum of the agents' utilities is maximized by decision $d_1^*(\theta_1, \dots, \theta_n) := \frac{1}{n}\sum_i \theta_i$.[6] We define welfare attained by a decision function $d(\cdot)$ as the negative of the expected squared difference between the decisions taken and the first–best decisions multiplied by n:

$$\mathcal{W}(d) = -n\mathbf{E}[(d(\theta_1, \dots, \theta_n) - d_1^*(\theta_1, \dots, \theta_n))^2]. \tag{2.1}$$

This welfare functional is a positive linear transformation of the sum of expected utilities.[7] We adopt this particular transformation as our welfare functional because it has some nice properties concerning interpretability (e.g., welfare only depends on the degree of interdependence via the agents' strategic behavior) and computability (e.g., asymptotic welfare converges for the most important decision functions).

2.2.2 Optimal Decisions and Upper Bounds on Welfare

The following Lemma describes the welfare maximizing decisions when different information about the agents' signals is available.

Lemma 2.1 (Optimal decisions)

(i) The first–best decision is $d_1^*(\theta_1, \dots, \theta_n) = \frac{1}{n}\sum_i \delta_1^*(\theta_i)$ *with* $\delta_1^*(\theta_i) := \theta_i$.

(ii) If only the median signal is known, the best decision is $d_2^*(\theta_1, \dots, \theta_n) = \delta_2^*(\theta_{m:n})$ *with*

$$\delta_2^*(\theta) := \mathbf{E}[d_1^*(\cdot)|\theta_{m:n} = \theta] = \frac{1}{n}\theta + \frac{1}{2}\frac{n-1}{n}\left(\mathbf{E}[\theta_i|\theta_i < \theta] + \mathbf{E}[\theta_i|\theta_i > \theta]\right).$$

(iii) The best uninformed decision is $d_3^*(\theta_1, \dots, \theta_n) = 0$.

The functions δ_1^* and δ_2^* describe the agents' voting behavior that would be best for welfare when the mean and the median decision rule are used, respectively. When the mean decision rule is used, it is socially optimal that each agent votes for his signal. When the median decision rule is used, it is optimal that each agent votes for his expectation of the first–best decision conditional on his own signal being the median signal. This expectation normally differs from his signal.[8]

---

[6] This follows directly from the FOC $\sum_i -2(d - \theta_i^*) = -2(nd - \sum_i \theta_i) \stackrel{!}{=} 0$.

[7] $\mathcal{W}(d) = \sum_i \mathbf{E}[-(d(\cdot) - \theta_i^*(\cdot))^2 + \theta_i^*(\cdot)^2 - d_1^*(\cdot)^2] = \sum_i \mathbf{E}[u_i(\cdot)] + \sum_i \mathbf{E}[\theta_i^*(\cdot)^2 - d_1^*(\cdot)^2]$.

[8] Whether it is more or less extreme, i.e. whether $|\theta| < |\delta_2^*(\theta)|$ or $|\delta_2^*(\theta)| < |\theta|$ is true, is ambiguous.

The performance of a mechanism is subject to informational constraints (Which agents' information is used in equilibrium?) and to incentive compatibility constraints (How is the information used in equilibrium?). For instance, if some mechanism implements a decision function which depends only on the median signal, this mechanism can never do better than always choosing the best decision conditional on knowing only the median signal. From the informational constraints we obtain the following upper bounds on welfare:

Lemma 2.2 (Upper bounds on welfare)

(i) For any decision function $d : \Theta^n \to \mathbf{R}$ *we have* $\mathcal{W}(d) \leq \mathcal{B}_1 := \mathcal{W}(d_1^*) = 0$.

(ii) For any decision function $d : \Theta^n \to \mathbf{R}$ *depending only on the median signal we have* $\mathcal{W}(d) \leq \mathcal{B}_2 := \mathcal{W}(d_2^*) = -\sigma^2 + n\mathbf{E}[\delta_2^*(\theta_{m:n})^2]$.

(iii) The highest welfare level that can be achieved by a decision function which does not depend on the signals is $\mathcal{B}_3 := \mathcal{W}(d_3^*) = -\sigma^2$.

The optimal level of welfare that can be achieved when only a certain information is used is increasing in the amount of information. Therefore we have the following ordering of bounds:

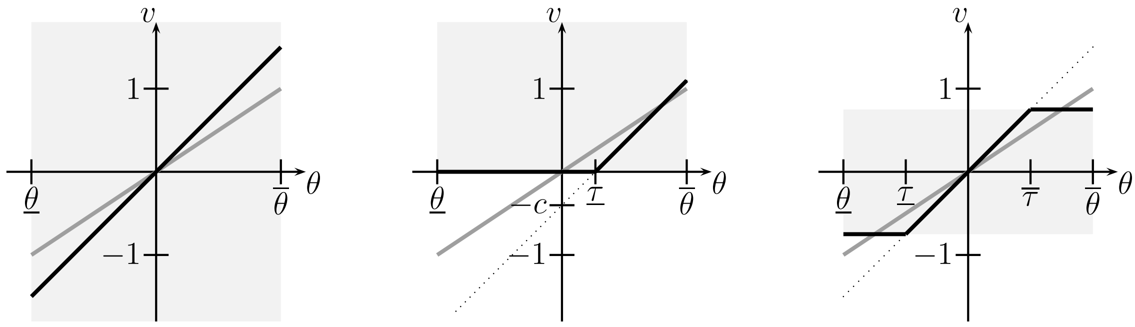
$$0 = \mathcal{B}_1 > \mathcal{B}_2 > \mathcal{B}_3 = -\sigma^2.$$

Since the best uninformed decision $d_3^*(\cdot) = 0$ is always implementable, a welfare level of at least $\mathcal{B}_3 = -\sigma^2$ can be achieved for any distribution, any number of agents and any degree of interdependence. Henceforth we will refer to a mechanism which always implements the best uniformed decision as *best uninformed mechanism.*

2.3 Mechanisms

In this section we derive the equilibria of mechanisms relying on the mean and the median decision rule. Moreover, we show that agents exaggerate (relative to the welfare maximizing voting behavior) for either rule. As a consequence, in both cases exaggeration can be mitigated and welfare can be enhanced by forbidding extreme votes.

Appendices A.4 and A.5 supplement the analysis in this section. In Appendix A.4 we derive necessary and sufficient conditions for decision functions to be implementable



(a) $V = \mathbf{R}$ (b) $V = \mathbf{R}_+$ (c) $V = [-\frac{3}{4}, \frac{3}{4}]$

Figure 2.1: Voting under the mean decision rule; actual voting $v(\theta)$: black solid curve; welfare maximizing voting $\delta_1^*(\theta)$: dark gray curve; admissible votes: light gray region ($\theta \sim U[-1,1]$, $n = 3$, $\alpha = \frac{1}{2}$)

by general mechanisms, i.e. by mechanisms which do not necessarily rely on the mean or the median decision rule. In Appendix A.5 we briefly discuss stochastic mechanisms.

2.3.1 Mean Mechanisms

If the mean mechanism is used and agents' votes are unrestricted, there is a unique symmetric equilibrium in which each agent votes according to

$$\delta_1(\theta) := n(1-\alpha)\theta$$

(Grüner and Kiel, 2004, Proposition 2). If preferences are common (i.e. if $n(1-\alpha) = 1$), there is no conflict of interest between the agents. Each agent's vote is in accordance with welfare maximizing behavior: $\delta_1(\theta) = \delta_1^*(\theta) = \theta$. However, if preferences are not common (i.e. if $n(1-\alpha) > 1$), then an agent who votes for his signal believes that the collective decision will be less extreme than the decision he prefers.[9] Therefore he does not vote for his signal but tries to push the decision into his preferred direction by exaggerating relative to the welfare maximizing voting behavior, i.e. $|\delta_1^*(\theta)| < |\delta_1(\theta)|$. This is displayed in figure 2.1(a). The gray curve depicts the welfare maximizing voting behavior, the black curve shows the agents' equilibrium voting behavior.

If voting is restricted, two things may change. First, an agent may not be allowed to choose his preferred vote. Then he chooses the closest vote which is still allowed.

[9] I.e. he believes the decision to be smaller (larger) than the decision he prefers if his signal is positive (negative).

Second, an agent might expect the votes of others to be biased. He then takes this bias into account and adjusts his vote accordingly. Figure 2.1(b) displays an example for this: Suppose only positive votes are allowed. Then each agent believes that the other agents' expected votes are strictly positive, which means that they tend to be too high. As a consequence, each agent adjusts his voting behavior downwards by c. Therefore some agents with positive signals still prefer negative votes. However, since negative votes are not allowed, there is pooling at the lower end of the distribution. Agents with signals from the interval $[-\underline{\theta}, \underline{\tau}]$ all vote for decision zero. We obtain the following Proposition.

Proposition 2.1 (Mean mechanism, equilibrium)

(i) The mean mechanism $\Gamma = ([\underline{v}, \overline{v}], x_1)$ *possesses a unique symmetric equilibrium in which each agent votes according to voting rule*

$$v(\theta) := \begin{cases} \delta_1(\overline{\tau}) - c & \text{if } \overline{\tau} < \theta \\ \delta_1(\theta) - c & \text{if } \underline{\tau} \leq \theta \leq \overline{\tau} \\ \delta_1(\underline{\tau}) - c & \text{if } \theta < \underline{\tau} \end{cases}$$

with $\underline{\tau} := \delta_1^{-1}(\underline{v} + c)$, $\overline{\tau} := \delta_1^{-1}(\overline{v} + c)$ *and* $c = (n-1)\mathbf{E}[v(\theta)]$. *The implemented decision function is* $d_1(\theta_1, \ldots, \theta_n) := \frac{1}{n} \sum_i v(\theta_i)$.

(ii) The welfare level attained by the unique symmetric equilibrium is

$$\mathcal{W}([\underline{v}, \overline{v}], x_1) := \mathcal{W}(d_1) = -\mathbf{E}[(v(\theta_i) - \theta_i)^2].$$

Besides the symmetric equilibrium in which each agent votes according to voting rule $v(\theta)$, also equilibria with asymmetric voting strategies exist. For example, in a game with three players strategies $v_1(\theta) = v(\theta) + k/2$, $v_2(\theta) = v(\theta) + k/2$ and $v_3(\theta) = v(\theta) - k$ constitute an equilibrium in case voting is not restricted. Agents 1 and 2 increase their votes by $k/2$ relative to the symmetric equilibrium and agent 3 offsets these increases by decreasing his vote by k. Since the adjustments just cancel each other out, the implemented decision function and the attained welfare level stay unaffected.

If voting is restricted, asymmetric equilibria do not necessarily implement the same decision function that is implemented in the unique symmetric equilibrium.[10] In the

[10] Reconsider the example from above and suppose now that votes above $\overline{v}$ and below $-\overline{v}$ are forbidden. Since agent 3 is supposed to decrease his vote more strongly than agents 1 and 2 increase theirs, it is

proof of Proposition 2.1 we derive all equilibria, but in the remainder of this paper we restrict attention to the unique symmetric equilibrium stated in Proposition 2.1.

In part (ii) of Proposition 2.1 a simple expression for welfare is stated. From this expression it can be directly observed that welfare is higher the closer an agent's vote lies to his signal. When only votes from a certain interval are admissible, voting strategies consist of flat parts for high and low signals, and of an increasing part connecting the two flat parts. Figure 2.1(c) shows the non–degenerated case, figure 2.1(b) displays an example for the degenerated case in which the restriction is only binding on one side. Recall that agents exaggerate when voting is not restricted and preferences are not common. As a consequence, they may vote for decisions which are strictly higher or lower than any signal (see figure 2.1(a) again). Therefore it is always welfare enhancing to forbid at least the most extreme votes. This introduces flat parts in the voting rule which bring some votes strictly closer to the respective signals without bringing any votes farther away.

Proposition 2.2 (Mean mechanism, restriction is optimal)
Suppose a mean mechanism $\Gamma = (V, x_1)$ shall be used and preferences are not common, i.e. $\alpha \in [0, \overline{\alpha}_n)$. Then it is optimal to counteract the agents' exaggeration by restricting the set of admissible votes.

It can also be shown that the optimal set of admissible votes is indeed an interval.[11] Since we will not need this result later on, we will not prove it formally.

General conditions characterizing the optimal interval of admissible votes can be given but are not very tractable because they rely on two–dimensional optimization and the endogenous parameter c. Tractable conditions can be obtained for symmetric distributions, because then attention can be restricted to a one–dimensional optimization

more likely that the restriction detains him from choosing his preferred vote than agents 1 and 2. In contrast, in the symmetric equilibrium, this is equally likely for all agents. Thus, the implemented decision functions must differ.

[11]A proof can make use of a similar reasoning as in section 7 of Alonso and Matouschek (2007). Although they consider a principal–agent model, the implied mathematical problem is similar to the derivation of the optimal set of admissible votes given the equilibrium voting behavior of the mean mechanism in our paper. Furthermore, for the agents' signals being distributed according to a truncated normal distribution, they obtain a similar condition characterizing the optimal set of admissible votes as we do in Proposition 2.3.

problem and c is always zero. If we additionally assume the hazard rate $\phi/(1-\Phi)$ to be increasing, the solution to the optimization problem is unique.

Proposition 2.3 (Mean mechanism, optimal restriction)
Let ϕ be symmetric around zero and let the hazard rate, $\phi/(1-\Phi)$, be increasing. Then the problem $\max_{\overline{v}} \mathcal{W}([-\overline{v}, \overline{v}], x_1)$ possesses a unique solution which is implicitly characterized by $\mathbf{E}[\theta_i|\theta_i > \tau] = \overline{v}$ with $\tau = \delta_1^{-1}(\overline{v})$.

If signals are uniformly distributed on $[-1, 1]$, the optimal interval of admissible votes is $V^* = [-\frac{n(1-\alpha)}{2n(1-\alpha)-1}, \frac{n(1-\alpha)}{2n(1-\alpha)-1}]$.[12] Thus, for any fixed $\alpha < 1$ and for n large it is optimal to allow votes $V^* \approx [-\frac{1}{2}, \frac{1}{2}]$. Later on we will show that using $V = [-\frac{1}{2}, \frac{1}{2}]$ as a rule of thumb often works quite well, even for non–uniform distributions and for small numbers of agents.

Before we go on to the median mechanism, we derive a lower bound on the welfare level that is achieved by the optimally restricted mean mechanism. For reasons of tractability we do this, again, only for distributions which are symmetric around zero. Instead of allowing the agents to choose from an interval, we now allow them only to choose between the two discrete options $-\overline{v}$ and $\overline{v}$. There is an (almost) unique symmetric equilibrium in which it is optimal for agents with positive signal to vote for $\overline{v}$ and for agents with negative signal to vote for $-\overline{v}$.[13] Optimizing welfare over $\overline{v}$ leads to the following Proposition.

Proposition 2.4 (Mean mechanism, lower bound on welfare)
Let ϕ be symmetric about zero. A lower bound on the level of welfare attained by the optimally restricted mean mechanism is

$$\max_{\overline{v}} \mathcal{W}(\{-\overline{v}, \overline{v}\}, x_1) = -\sigma^2 + \mathbf{E}[\theta_i|\theta_i > 0]^2.$$

Note that the bound derived in Proposition 2.4 does not depend on the number of agents and the degree of interdependence. Moreover, it follows directly from the Proposition that the welfare level attained by the optimally restricted mean mechanism

[12] For the uniform distribution on $[-1, 1]$ we have $\mathbf{E}[\theta_i|\theta_i > \tau] = \overline{v} \Leftrightarrow \frac{1}{2}(\tau + 1) = \overline{v}$. Using $\tau = \delta_1^{-1}(\overline{v}) = \frac{\overline{v}}{n(1-\alpha)}$ and solving for $\overline{v}$ yields the result.

[13] The equilibrium is only *almost* unique because there is a degree of freedom regarding the behavior of an agent with type zero.

is at least by the discrete amount $\mathbf{E}[\theta_i|\theta_i > 0]^2$ higher than the welfare level attained by the best uninformed mechanism.

2.3.2 Median Mechanisms

If the median decision rule is used the agents are not restricted in voting, there are multiple equilibria,[14] but only one symmetric equilibrium in which the agents' voting behavior is strictly monotonic. In this equilibrium each agent votes for what he expects to be his preferred decision conditional on having the median signal. I.e. agent i votes according to

$$\begin{aligned} \delta_2(\theta) &:= \mathbf{E}[\theta_i^*(\cdot)|\theta_i = \theta_{m:n} \text{ and } \theta_i = \theta] \\ &= (1-\alpha)\theta + \frac{\alpha}{2}\left(\mathbf{E}[\theta_i|\theta_i < \theta] + \mathbf{E}[\theta_i|\theta_i > \theta]\right) \end{aligned}$$

(Grüner and Kiel, 2004, Proposition 1). If extreme votes are forbidden, an agent is not always allowed to vote according to $\delta_2(\theta)$. He then chooses the vote closest to $\delta_2(\theta)$ which is still allowed.

Proposition 2.5 (Median mechanism, equilibrium)

(i) Median mechanism $\Gamma = ([\underline{v}, \overline{v}], x_2)$ possesses an equilibrium in which each agent votes according to voting rule

$$v(\theta) := \begin{cases} \delta_2(\overline{\tau}) & \text{if } \overline{\tau} < \theta \\ \delta_2(\theta) & \text{if } \underline{\tau} \leq \theta \leq \overline{\tau} \\ \delta_2(\underline{\tau}) & \text{if } \theta < \underline{\tau} \end{cases}$$

with $\underline{\tau} := \delta_2^{-1}(\underline{v})$ and $\overline{\tau} = \delta_2^{-1}(\overline{v})$. This is the only symmetric equilibrium in which the agents' voting behavior is strictly monotonic for signals in $[\underline{\tau}, \overline{\tau}]$. The implemented decision function is $d_2(\theta_1, \ldots, \theta_n) = v(\theta_{m:n})$.

(ii) The welfare level attained by the equilibrium in (i) is

$$\mathcal{W}([\underline{v}, \overline{v}], x_2) := \mathcal{W}(d_2) = \mathcal{B}_2 - n\mathbf{E}[(v(\theta_{m:n}) - \delta_2^*(\theta_{m:n}))^2]. \tag{2.2}$$

[14] For instance, it is an equilibrium if all agents vote for the same decision irrespective of their types. In this case no agent is ever pivotal such that no agent has a strict incentive to deviate from this behavior.

If signals are uniformly distributed, the preferred vote of an agent with signal θ is $\delta_2(\theta) = (1 - \frac{\alpha}{2})\theta$, but it would be best for welfare if he voted according to $\delta_2^*(\theta) = \frac{n+1}{n}\frac{1}{2}\theta$. It follows that

$$|\delta_2^*(\theta)| \leq |\delta_2(\theta)| \leq |\theta| \tag{2.3}$$

for any θ. By the right inequality each agent prefers votes which are less extreme than his signal, by the left inequality he prefers votes which are more extreme than is socially optimal. Crucial for these inequalities is that an agent's bliss point is a weighted sum of his own signal and the average of all other signals. In equilibrium, the fact that an agent is pivotal conveys the information that his signal is the median signal. With uniformly distributed signals an agent can infer from this information that his signal is likely to be more extreme than the average of the other agents' signals. By consequence, his preferred vote is less extreme than his signal, i.e. $|\delta_2(\theta)| \leq |\theta|$. However, since he still puts a stronger weight on his own signal than is socially optimal, his preferred vote is still more extreme than the vote which would be best for welfare. Therefore $|\delta_2^*(\theta)| \leq |\delta_2(\theta)|$.

While condition (2.3) holds for any signal θ when signals are uniformly distributed, this is not necessarily true for other distributions.[15] However, condition (2.3) holds at least for signals close to $\overline{\theta}$ and $\underline{\theta}$ for general distributions: Suppose an agent with signal $\overline{\theta}$ obtains the information that his signal is the median signal. Then he knows that one half of the other agents' signals is also $\overline{\theta}$ and the other half lies somewhere between $\underline{\theta}$ and $\overline{\theta}$. This implies that the signals of agents having smaller signals lie farther away from his own signal than those of agents having larger signals. As a consequence, he adjusts his preferred vote downwards below his own signal. I.e. $\delta_2(\overline{\theta}) < \overline{\theta}$. Since he puts a stronger weight on his own signal than on the other signals, he does not adjust his vote far enough downwards. Thus, his preferred vote is larger than the vote which is best for welfare, i.e. $\delta_2^*(\overline{\theta}) < \delta_2(\overline{\theta})$.

We can conclude that at least agents with very high and very low signals exaggerate relative to the socially optimal use of their information. Therefore it is welfare–enhancing to forbid at least votes outside the interval $[\delta_2^*(\underline{\theta}), \delta_2^*(\overline{\theta})]$. These votes are always more extreme than is socially optimal.

[15]Condition (2.3) requires that $\delta_2(0) = \delta_2^*(0) = 0$. Since this can only hold for distributions with $\theta_{\text{Med}} = 0$, condition (2.3) at least breaks down for any distribution in which the median of the distribution is not zero.

Proposition 2.6 (Median mechanism, restriction is optimal)
Suppose a median mechanism $\Gamma = (V, x_2)$ *shall be used and preferences are not common, i.e.* $\alpha \in [0, \overline{\alpha}_n)$. *Then it is optimal to counteract the agents' exaggeration by restricting the set of admissible votes.*

2.4 Comparison of Mechanisms

In this section we compare mean with median mechanisms. We first explain why the mean decision rule often performs badly when voting is not restricted (subsection 2.4.1). Then we show for some classes of distributions and a finite number of agents (subsection 2.4.2), and for general distributions and a large number of agents (subsection 2.4.3) that the optimally restricted mean mechanism performs better than the optimally restricted median mechanism.

2.4.1 Comparison of the Unrestricted Mechanisms

In this subsection we consider the performance of the mean and the median decision rule when there are no restrictions on voting, i.e. when $V = \mathbf{R}$.

If preferences are common, agents do not behave strategically. As a consequence, the unrestricted mean mechanism implements the first–best, but the unrestricted median mechanism implements only the best conditional on using only the median signal. Since the use of more information is clearly beneficial in absence of strategic behavior, the unrestricted mean mechanism performs better.

Proposition 4 in Grüner and Kiel (2004)
If preferences are common, the unrestricted mean mechanism attains a higher level of welfare than the unrestricted median mechanism:

$$\mathcal{W}(\mathbf{R}, x_1) = \mathcal{B}_1 > \mathcal{B}_2 = \mathcal{W}(\mathbf{R}, x_2) \text{ if } \alpha = \overline{\alpha}_n.$$

Now we consider what happens for other degrees of interdependence. If the unrestricted mean mechanism is used, decision $d_1(\cdot) = \frac{1}{n}\sum_i n(1-\alpha)\theta_i = n(1-\alpha)d_1^*(\cdot)$ is implemented in equilibrium. $n(1-\alpha)$ is the factor by which each agent exaggerates his signal and $d_1^*(\cdot)$ is the first–best decision. The implemented decisions and the first–best decisions are perfectly correlated, but because of the agents' exaggeration they may nevertheless fall far apart. Although it is common knowledge that each agent's preferred decision lies in $[\underline{\theta}, \overline{\theta}]$, the range of decisions is $[n(1-\alpha)\underline{\theta}, n(1-\alpha)\overline{\theta}]$ and becomes arbitrarily large as $n(1-\alpha)$ increases. The reason for this is that the decision is additive in all votes such that individual exaggerations might add up.

Even though there might also be exaggeration when the unrestricted median mechanism is used (see Proposition 2.6), exaggeration is less problematic there because only the median agent's vote affects the implemented decision. The range of the implemented decisions is $[\delta_1(\underline{\theta}), \delta_1(\overline{\theta})]$ which is always a subset of $[\underline{\theta}, \overline{\theta}]$.

If preferences are private, the fact that the consequences of exaggeration are more harmful under the mean decision rule makes the unrestricted median mechanism generally preferable.

Proposition 3 in Grüner and Kiel (2004)
If preferences are private, the unrestricted median mechanism attains a higher level of welfare than the unrestricted mean mechanism:

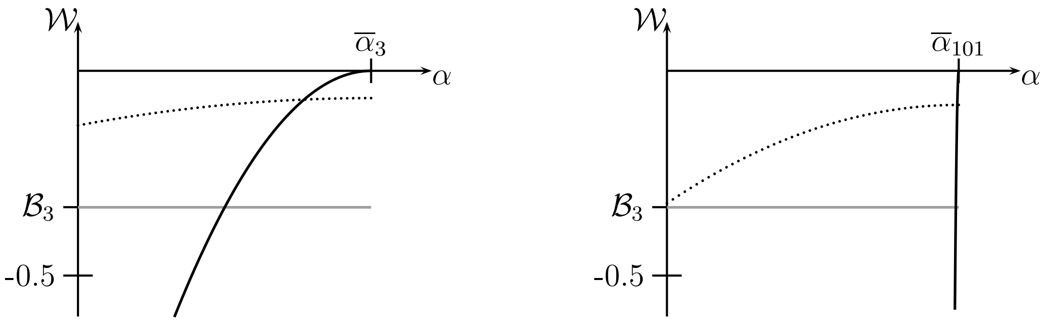
$$\mathcal{W}(\mathbf{R}, x_1) < \mathcal{W}(\mathbf{R}, x_2) \text{ if } \alpha = 0.$$

Since the above inequalities are strict and welfare is continuous in α, Propositions 3 and 4 in Grüner and Kiel (2004) also hold for intermediate degrees of interdependence sufficiently close to private and common preferences, respectively (Proposition 5 in Grüner and Kiel (2004)).

However, asymptotically, the preferability of the mean mechanism is a peculiarity of perfectly common preferences. For any fixed degree of interdependence $\alpha \in [0, 1)$ the factor $n(1-\alpha)$ becomes arbitrarily large as n increases. Since this has, as described above, very harmful consequences when the mean but not when the median decision rule is used, the set of parameters α for which the unrestricted median mechanism is preferable over the unrestricted mean mechanism converges towards the set $[0, 1)$, whereas the set for which the converse holds converges towards the singleton set $\{1\}$.

Proposition 2.7 (Asymptotic comparison, voting unrestricted)
If preferences are not common and the number of agents is sufficiently large, the unrestricted mean mechanism performs worse than the unrestricted median mechanism:

$$\mathcal{W}(\mathbf{R}, x_1) < \mathcal{W}(\mathbf{R}, x_2) \text{ for any fixed } \alpha \in [0, 1) \text{ if } n \text{ is sufficiently large.}$$



(a) $n = 3$ (b) $n = 101$

Figure 2.2: Welfare; mean mechanism $\Gamma = (\mathbf{R}, x_1)$: black solid curve; median mechanism $\Gamma = (\mathbf{R}, x_2)$: black dotted curve; best uninformed mechanism $\Gamma = (V, 0)$: gray curve ($\theta \sim U[-1, 1]$)

In the proof of the Proposition we derive a sufficient condition for the statement which is independent of the distribution of signals. If $n \geq (1 + 2(1 - \alpha))/(1 - \alpha)^2$, the unrestricted mean mechanism is worse than the unrestricted median mechanism. Furthermore, we show also that if $n > 2/(1 - \alpha)$, the unrestricted mean mechanism is worse than the best uninformed mechanism.

Figure 2.2(a) displays how welfare typically looks like. The graph shows the welfare level attained by the unrestricted mean mechanism (black solid curve), by the unrestricted median mechanism (black dotted curve) and by the best uninformed mechanism (gray curve) for uniformly distributed signals and three agents. It can be seen that close to private preferences ($\alpha = 0$) the unrestricted median mechanism performs better (Proposition 3 in Grüner and Kiel (2004)), while close to common preferences ($\alpha = \overline{\alpha}_3$) the unrestricted mean mechanism is better (Proposition 4 in Grüner and Kiel (2004)).

Figure 2.2(b) shows the same functions as figure 2.2(a), but for 101 instead of 3 agents. The graph gives an impression of how bad the unrestricted mean mechanism becomes when the number of agents increases (see Proposition 2.7). While for common preferences the unrestricted mean mechanism still implements the first–best, it is already worse than the best uninformed mechanism for degrees of interdependence very close to common preferences.

2.4.2 Non–Asymptotic Comparison of the Optimally Restricted Mechanisms

The *influence* of an agent's vote on the collective decision depends on two features of the mechanism. First, it depends on his *discretion* in choosing a vote, and second, on the *impact* of his vote on the collective decision. If an agent is allowed to choose any real–valued vote but his vote has no impact on the decision, he has no influence on the decision. Likewise, he has no influence if his vote alone determines the decision but he has to pick his vote from a singleton set.

If the unrestricted mean mechanism is used, an agent's influence is large. He is allowed to pick any real–valued vote and his vote always has an impact on the decision. While the benefit of this large influence is a high correlation between first–best and implemented decisions, it comes at the cost that each agent's vote may affect the decision in a way which might be detrimental for welfare. This happens to be the case if $n(1-\alpha)$ is large, i.e. if interdependence is weak or if the number of agents is large. Then welfare can be significantly improved by reducing an agent's influence on the collective decision.

There are different ways of reducing an agent's influence on the decision. In the preceding subsection we showed that if $n(1-\alpha)$ is large, it is beneficial to choose the unrestricted median mechanism instead of the unrestricted mean mechanism, i.e. to reduce the impact of an agent's vote on the collective decision while leaving his discretion unchanged. In this subsection we show for some classes of distributions that leaving the impact of an agent's vote on the collective decision unchanged but forbidding extreme votes is a better way of reducing his influence on the decision. Heuristically, by forbidding extreme votes it is possible to control for the agents' exaggeration under the mean decision rule while maintaining a correlation between first–best decisions and implemented decisions which is higher than the correlation present under the median decision rule.

We prove that for uniformly distributed signals, any number of agents and any degree of interdependence the optimally restricted mean mechanism performs better than the optimally restricted median mechanism. Although we are not able to prove this result for more general distributions theoretically, numerical examples suggest that it extends at least to the case of distributions with linear densities and symmetric distributions with quadratic densities.

Uniform Distribution

Suppose that signals are uniformly distributed on a normalized support of length 2 and consider first the case in which preferences are private. We can show that the specific mean mechanism $\Gamma = ([-1/2, 1/2], x_1)$ attains a welfare level above $\mathcal{B}_2$ for any number of agents. That is, it attains a welfare level above anything that can be achieved by a median mechanism. If instead of private preferences stronger degrees of interdependence are considered, bound $\mathcal{B}_2$ stays unchanged but the performance of mean mechanism $\Gamma = ([-1/2, 1/2], x_1)$ becomes even better because agents exaggerate less. Hence, we can conclude that the optimally restricted mean mechanism must perform better than *any* median mechanism for *any* degree of interdependence and *any* number of agents.

Proposition 2.8 (Comparison, uniform distribution)
If signals are uniformly distributed, then the optimally restricted mean mechanism attains a higher level of welfare than the optimally restricted median mechanism:

$$\max_{V} \mathcal{W}(V, x_1) > \mathcal{B}_2 \geq \max_{V'} \mathcal{W}(V', x_2) \text{ for any } \alpha \in [0, \overline{\alpha}_n].$$

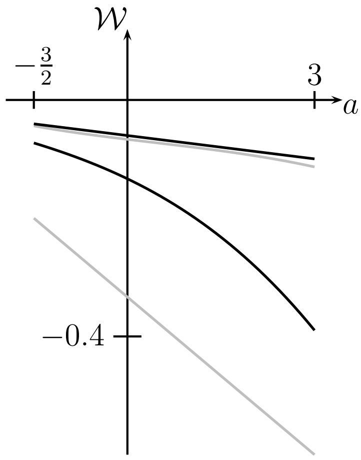
In the proof we make use of the specific structure of the uniform distribution: First, because of the symmetry of distribution and the symmetry of the set of admissible votes we obtain an explicit expression for the equilibrium voting behavior when the mean decision rule is used. Second, the specific structure of the uniform distribution allows us to compute bound $\mathcal{B}_2$ analytically.

Symmetric Distributions with Quadratic Densities

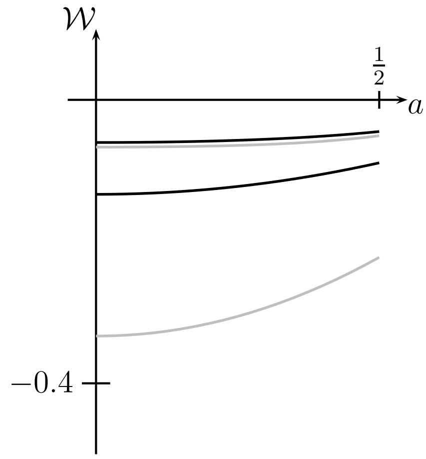
Now we consider symmetric distributions with quadratic densities on a normalized support of length 2.[16] We denote the curvature of the density function by parameter $a \in [-3/2, 3]$. If a is negative, the density function is concave such that most probability mass lies close to the median of the distribution. If a is positive, the density function is convex such that most probability mass lies close to the endpoints of the support.

Again, to show the superiority of the mean decision rule for any degree of interdependence, it suffices to consider private preferences and to find a specific set of ad-

[16] This class of distributions is characterized by $\Theta = [-1, 1]$, $\phi(\theta) = \frac{a}{2}\theta^2 + \frac{1}{2} - \frac{a}{6}$ and $a \in [-3/2, 3]$. The median of any distribution from this class is $\theta_{\text{Med}} = 0$.



(a) ϕ symmetric and quadratic with curvature a



(b) ϕ linear with slope a

Figure 2.3: Welfare for private preferences; mean mechanism $\Gamma = ([-\frac{1}{2}, \frac{1}{2}], x_1)$: upper black curve; median mechanism $\Gamma = (\mathbf{R}, x_2)$: lower black curve; bound $\mathcal{B}_2$: upper gray curve; bound $\mathcal{B}_3$: lower gray curve ($\overline{\theta} - \underline{\theta} = 2$, $n = 3$, $\alpha = 0$)

missible votes V such that the welfare level attained by mean mechanism $\Gamma = (V, x_1)$ is higher than $\mathcal{B}_2$. For $n = 3$ and for any curvature parameter a the same restriction that worked already for the uniform distribution, $V = [-1/2, 1/2]$, works here too. The upper black curve in figure 2.3(a) displays the welfare level attained by mean mechanism $\Gamma = ([-1/2, 1/2], x_1)$. The upper gray curve lying slightly below this curve shows bound $\mathcal{B}_2$. As benchmarks we have drawn the welfare level attained by the unrestricted median mechanism $\Gamma = (\mathbf{R}, x_2)$ (lower black curve), and the welfare level attained by the best uninformed mechanism (lower gray curve).

Distributions with Linear Densities

Finally, we consider distributions with linear densities on a normalized support of length 2.[17] We denote the slope of the density function by parameter $a \in [-1/2, 1/2]$. By a

[17] This class of distributions is characterized by $\Theta = [-\frac{2}{3}a - 1, -\frac{2}{3}a + 1]$, $\phi(\theta) = ax + \frac{2}{3}a^2 + \frac{1}{2}$ and $a \in [-\frac{1}{2}, \frac{1}{2}]$. To obtain $\mathbf{E}[\theta] = 0$ for any parameter a, the support of the distribution varies in a. The median of the distribution characterized by parameter a is $\theta_{\text{Med}} = -(4a^2 + 3 - 3\sqrt{4a^2 + 1})/(6a)$ and differs from zero for $a \neq 0$.

symmetry argument we can restrict attention without loss of generality to positive slope parameters.

Once again, for $n = 3$ and private preferences the specific mean mechanism $\Gamma = ([-1/2, 1/2], x_1)$ performs better than anything that can be achieved using a median mechanism. The upper black curve in figure 2.3(b) displays the welfare level attained by mean mechanism $\Gamma = ([-1/2, 1/2], x_1)$. The upper gray curve lying slightly below the upper black curve shows bound $\mathcal{B}_2$. As benchmarks we have drawn again the welfare level attained by the unrestricted median mechanism $\Gamma = (\mathbf{R}, x_2)$ (lower black curve), and the welfare level attained by the best uninformed mechanism (lower gray curve).

2.4.3 Asymptotic Comparison of the Optimally Restricted Mechanisms

In subsection 2.4.1 we obtained a very strong asymptotic result against the unrestricted mean mechanism: For any degree of interdependence except for common preferences the unrestricted median mechanism is preferable over the unrestricted mean mechanism (Proposition 2.7).

In this subsection we consider the asymptotic performance of the optimally restricted mechanisms. When preferences are common, it is for both, the mean and the median decision rule, optimal not to restrict voting. Therefore the result from section 2.4.1 remains valid: The mean decision rule is preferable. In the following we prove the less obvious result that the mean decision rule is also preferable for private preferences.

We present the argument in three steps. First, we explain why it suffices to compare the optimally restricted mean mechanism with the unrestricted median mechanism, then we derive the comparison result for symmetric and for asymmetric distributions.

Asymptotic Effect of the Restriction of Voting on the Performance of Mean and Median Mechanisms

Restricting the set of admissible votes has different effects on the implemented decision depending on whether the mean or the median decision rule is used.

When the median decision rule is used and agents are allowed to choose from an interval $V = [\underline{v}, \overline{v}]$ instead of from $V = \mathbf{R}$, the implemented decision only changes if the preferred vote of the median agent, $\delta_2(\theta_{m:n})$, does not fall into the interval $[\underline{v}, \overline{v}]$. Since for $n \to \infty$ the median signal $\theta_{m:n}$ lies with probability one close to the median of the distribution, θ_{Med}, the restriction of V effects the implemented decision asymptotically either with probability zero (if $\delta_2(\theta_{\text{Med}}) \in (\underline{v}, \overline{v})$), or with probability one (if $\delta_2(\theta_{\text{Med}}) \notin [\underline{v}, \overline{v}]$).[18] In the former case the optimally restricted median mechanism cannot perform significantly better than the unrestricted median mechanism. In the latter case it cannot perform significantly better than the best uninformed mechanism. For the case in which V is not an interval, an analogous reasoning applies.

By contrast, when the mean decision rule is used, restricting V may significantly improve the performance of the mechanism. Since the restriction affects the implemented decision already if an arbitrary agent's preferred vote does not belong to V and not only if this is true for the median agent's preferred vote, restricting V is a more powerful instrument when the mean decision rule is used. From Proposition 2.4 we know already that the optimally restricted mean mechanism attains a welfare level of at least $-\sigma^2 + \mathbf{E}[\theta_i|\theta_i > 0]^2$ which is by a discrete amount higher than the welfare level $-\sigma^2$ obtained by the best uninformed mechanism. Furthermore, since for all but common preferences the unrestricted mean mechanism becomes arbitrarily bad as the number of agents increases (Proposition 2.7), the optimally restricted mean mechanism also performs significantly better than the unrestricted mean mechanism.

Combining the insights of these two cases we can make the following observation:

Observation 2.1 *To prove the asymptotic superiority of the mean decision rule over the median decision rule, it suffices to show that the optimally restricted mean mechanism performs better than the unrestricted median mechanism.*

Symmetric Distributions

When the mean decision rule is used and agents are only allowed to choose between $-\overline{v}$ and $\overline{v}$, agents with a positive signal vote for $\overline{v}$ and agents with a negative signal vote for $-\overline{v}$. Thus, the information conveyed in equilibrium is the numbers of agents having positive and having negative signals. Since the agents' behavior does not depend on

[18] There is a third case possible. $\delta_2(\theta_{\text{Med}})$ might lie on the boundary of the interval of admissible votes. Then the restriction is binding with probability 1/2. However, using a similar reasoning as for the two other cases, such a mechanism cannot perform significantly better than the unrestricted median mechanism *and* than the best uninformed mechanism at the same time.

the absolute value of $\overline{v}$, there is a degree of freedom regarding what is done with the gathered information. In Proposition 2.4 we have already derived the welfare level that is obtained by choosing $\overline{v}$ optimally.

When the unrestricted median mechanism is used, the information conveyed in equilibrium is the exact position of the median signal. For $n \to \infty$ and private preferences we are able to compute a tractable analytical expression for welfare under the unrestricted median mechanism and we can show that it is worse than the lower bound on welfare achieved by the optimally restricted mean mechanism that we derived in Proposition 2.4. Heuristically, the information conveyed if mean mechanism $\Gamma = (\{-\underline{v}, \overline{v}\}, x_1)$ and if the unrestricted median mechanism $\Gamma = (\mathbf{R}, x_2)$ is used is similar, but only for the mean mechanism there is a degree of freedom regarding what is done with the information obtained. This makes the mean mechanism perform better.

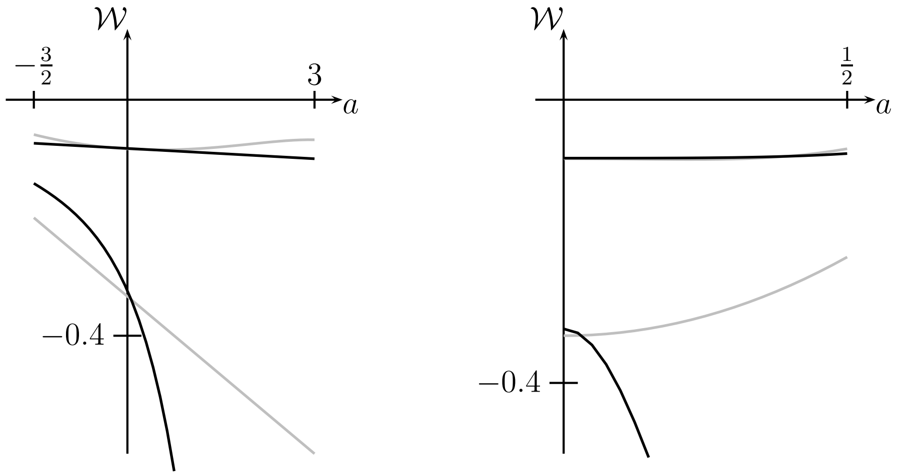
Proposition 2.9 (Asymptotic comparison, symmetric distributions)
Let ϕ be symmetric about zero and let preferences be private, i.e. $\alpha = 0$. If the number of agents is sufficiently large, the optimally restricted mean mechanism attains a higher level of welfare than the unrestricted median mechanism:

$$\lim_{n\to\infty} \max_{V} \mathcal{W}(V, x_1) \geq -\sigma^2 + \mathbf{E}[\theta_i|\theta_i > 0]^2 \geq \lim_{n\to\infty} \mathcal{W}(\mathbf{R}, x_2).$$

Proposition 2.9 together with Observation 2.1 show that for symmetric distributions and a large number of agents the optimally restricted mean mechanism performs better than the optimally restricted median mechanism, while the unrestricted median mechanism is preferable to the unrestricted mean mechanism.

Figure 2.4(a) depicts the same as figure 2.3(a), but for 101 instead of 3 agents. In this case the specific mean mechanism $\Gamma = ([-1/2, 1/2], x_1)$ is only better than bound $\mathcal{B}_2$ for curvature parameters close to zero. Note that it still lies close to it, whereas the unrestricted median mechanism performs much worse. It is notable that for most positive curvature parameters, i.e. when most probability mass lies at the two endpoints of the distribution, the unrestricted median mechanism performs even worse than the best uninformed mechanism.[19]

[19]By restricting the set of admissible votes when the median decision rule is used, it is possible to attain at least a welfare level slightly above what is achieved by the best uninformed mechanism.



(a) ϕ symmetric and quadratic with curvature a

(b) ϕ linear with slope a

Figure 2.4: Welfare for private preferences; mean mechanism $\Gamma = ([-\frac{1}{2}, \frac{1}{2}], x_1)$: upper black curve; median mechanism $\Gamma = (\mathbf{R}, x_2)$: lower black curve; bound $\mathcal{B}_2$: upper gray curve; bound $\mathcal{B}_3$: lower gray curve ($\overline{\theta} - \underline{\theta} = 2$, $n = 101$, $\alpha = 0$)

Asymmetric Distributions

Now we consider distributions for which the mean of the distribution, $\mathbf{E}[\theta_i] = 0$, and the median of the distribution, θ_{Med}, do not coincide. For such distributions and for $n \to \infty$ the first–best decision lies close to zero with probability one while the decision implemented by the median mechanism lies close to $\delta_2(\theta_{\text{Med}}) = (1 - \alpha)\theta_{\text{Med}} \neq 0$ with probability one. Hence, with probability one the unrestricted median mechanism implements decisions which are worse than the best uninformed decision.[20]

[20] By restricting the set of admissible votes the performance of the median mechanism can be improved, but it cannot become much better than the performance of the best uninformed mechanism.

An example for how it may be improved is the following: Suppose that the median of the distribution is positive. Then the vote preferred by the median agent is strictly positive with a probability close to one, but the optimal decision is close to zero with a probability close to one. This problem can be avoided by forbidding positive votes. However, if the unlikely event happens that the median agent prefers a strictly negative vote, it might be optimal to let the decision be strictly negative. Thus, the structure of the optimal set of admissible votes when the median decision rule is used is $V = [\underline{v}, 0]$.

Proposition 2.10 (Asymptotic comparison, asymmetric distributions)
Let $\theta_{Med} \neq 0$ and let preferences be not common, i.e. $\alpha \in [0,1)$. If the number of agents is sufficiently large, the optimally restricted mean mechanism attains a higher level of welfare than the unrestricted median mechanism:

$$\lim_{n\to\infty} \max_{V} \mathcal{W}(V, x_1) > -\sigma^2 > \lim_{n\to\infty} \mathcal{W}(\mathbf{R}, x_2) = -\infty \text{ if } \alpha \neq 1.$$

Figure 2.4(b) depicts the same as figure 2.3(b), but for 101 instead of 3 agents. For $a \neq 0$ the median of the distribution differs from zero. It can be seen that for all values of the slope parameter a but those lying very close to zero the unrestricted median mechanism performs even much worse than the best uninformed mechanism.

2.5 Conclusion

In this paper we considered collective decision problems with interdependent preferences when there is no cheap talk prior to the voting stage and monetary transfers are not feasible. In this setting we compared the mean and the median decision rule when the agents' discretion in voting may be restricted.

Each agent's influence on the decision is large when the mean decision rule is used and voting is not restricted. This gives agents the possibility to behave opportunistically. We showed that it is in many cases optimal to reduce an agent's influence on the decision, but that this is better done by using the mean decision rule and restricting his discretion in choosing a vote than by choosing the median decision rule.

An obvious extension of our model is allowing for cheap talk prior to the voting stage. Since the agents' preferences are interdependent, they might have an incentive to communicate informatively. When cheap talk is possible and preferences are common, the unrestricted median mechanism becomes relatively better,[21] whereas when preferences are private, the unrestricted mean mechanism becomes relatively better.[22] Thus, the

[21]If there is no conflict of interest, any agent truthfully communicates his information prior to the voting stage. Thus, not only the unrestricted mean mechanism but also the unrestricted median mechanism is capable of implementing the first–best.

[22]Without cheap talk the unrestricted mean mechanism is very prone to coordination failures. For instance, if there are 101 agents which all prefer decision 1, the implemented decision is 101. This problem can at least be partially mitigated through communication.

performances of the two mechanisms approach each other and the question arises how this affects their relative performance.

It might also be interesting to look for the optimal mechanism in larger classes of mechanisms. In Appendix A.4 we derive necessary and sufficient conditions for the implementability of decision functions. As a starting point one may therefore look for the optimal mechanism in classes of mechanisms in which these functions are particularly tractable. Furthermore, allowing decisions to be stochastic may improve the tractability of the optimization problem (for this see the discussion in Appendix A.5).

Chapter 3

First–Price Auctions, Seller Information and Commitment to Sell

3.1 Introduction

One of the most crucial assumptions in auction theory addresses the seller's commitment power. Can the seller commit himself not to offer unsold objects again at a later point in time? Can he commit not to affect the auction outcome by placing phantom bids himself? Can he commit not to change the rules of the auction after eliciting some information from the buyers? The seller's ability to commit crucially affects what he can implement and what he optimally should implement. For instance, in the independent private values environment with risk–neutral buyers and a risk–neutral seller any standard auction with an optimally chosen reserve price maximizes the seller's revenue when he *can commit to any behavior* (Myerson, 1981, Riley and Samuelson, 1981). In contrast, when he *cannot commit to anything*, there is basically only a single auction implementable, namely the open English auction without reserve price (Vartiainen, 2007). However, the full commitment and the no commitment assumption are both rather extreme and rarely met in practice. Therefore there is a need to analyze settings in which the seller *can commit to some behavior but not to others*, for example the cases in which he *cannot commit not to sell* and in which he *cannot commit to sell*.

A seller who *cannot commit not to sell* cannot credibly promise not to reauction unsold objects at a later point in time. In this case it is optimal for the seller to conduct a sequence of standard auctions with declining reserve prices (McAfee and Vincent, 1997, Skreta, 2004).

When the seller *cannot commit to sell*, he might decide to keep the object after observing the outcome of the auction. This can either happen directly by withdrawing the object, or indirectly by placing a phantom bid himself.[1] Even when the seller is legally not allowed to refrain from selling the object to the highest bidder or to place own bids, he can often not be prevented from using a third party to place bids on his behalf.

This Contribution

In this paper we are interested in the case in which the seller cannot commit to sell. The first part of the paper addresses the optimal auction with a first–price payment rule in an environment in which the commitment problem is binding with positive probability when the seller sets a reasonable reserve price. This is because we assume that the seller has to design the rules of the auction before he learns his reservation value. Except for this assumption and the seller's lack of commitment power, we stick to the standard independent private values environment with risk–neutral buyers and a risk–neutral seller.

Often some time needs to pass between the announcement of the auction (and its rules) and the beginning of the tendering procedure. For example, this is the case when the seller auctions a procurement contract and the buyers have to prepare prototypes or construction plans before they can reasonably think about their bidding strategies. During the time between the announcement and the start of the bidding procedure the seller may obtain more accurate estimates about his valuation (e.g. due to changed market conditions) or about his outside options (e.g. about his own production possibilities). Moreover, the seller may be uninformed about his valuation at the time he designs the auction if he decides to install general rules for a series of future auctions instead of

[1]The placement of phantom bids can be interpreted as setting a secret reserve price. In Cassady (1967) it is claimed that reserve prices are usually not announced and many examples for this are given. Furthermore, similar observations are made in Hendricks, Porter, and Tan (1993) for auctions for oil and gas leases held by the government, in Elyakime, Laffont, Loisel, and Vuoung (1994) for timber auctions in France and in Ashenfelter and Graddy (2003) for auctions of art.

deciding on the rules on a case–by–case basis.[2]

By the Revenue Equivalence Theorem, any auction format which induces the same allocation and in which the worst buyer type has to make a zero expected payment generates the same expected revenue for the seller.[3] Due to this property the nature of the payment rule plays a more prominent role when the seller cannot commit to sell than when he can: Consider the symmetric equilibria of a first–price and a second–price auction with the same reserve price. Since both auction formats induce monotonic bidding behavior, the resulting allocations coincide when the seller can commit to sell to the highest bidder. Hence, which payment rule is used does not matter for expected revenue. However, the actual payment the highest bidder has to make in the two cases generally differs. As a consequence, the seller's incentive to sell and to keep the object differs too. If the seller cannot commit to sell to the highest bidder, the final allocation of the object and thus also his expected revenue depends on the nature of the payment rule used. We restrict attention to the case in which the seller uses a first–price payment rule. This assumption can be justified for two reasons: First, in practice sellers often want to choose a first–price auction format because of its simplicity and certain other features (e.g., its robustness with respect to collusion). Second, from a theoretical point of view, fixing a specific payment rule allows us to analyze the effects that the commitment problem has on the optimal allocation more thoroughly.

Our main result in the first part of the paper is that the structure of the optimal first–price auction in our setting may differ from the optimal structure in the standard setting. More specifically, while in the standard setting the seller only has an incentive to restrict bidding by setting a reserve price, he may also want to prohibit intermediate bids in our setting. The reason for this is the following: Suppose the seller's reservation value is x_1 with 10 percent probability and $x_2 > x_1$ with 90 percent probability. If the seller learns his reservation value before he has to fix the rules of the auction, he chooses different reserve prices depending on his actual reservation value. If he has to fix rules first, then he can only set one (real) reserve price, yet to a certain degree, he can mimic a second second reserve price by offering a non–connected set of admissible bids

[2] Timber auctions in France are usually conducted as first–price sealed bid auctions without reserve price. Since the seller's value may differ from case to case he has an incentive to install secret reserve prices by placing own bids (Elyakime, Laffont, Loisel, and Vuoung, 1994).

[3] See chapter 3 in Krishna (2002).

$B = [r_1, x_2) \cup [r_2, \infty)$ with $x_1 < r_1$ and $x_2 < r_2$. Then r_1 is the reserve price relevant for buyers who are satisfied with getting the object with at most 10 percent probability, and r_2 is the one relevant for those who want a higher probability. By changing the gap in B, the seller can influence which of the buyers' types aim at getting a high and which ones aim at getting a low probability of obtaining the object, i.e. for which r_1 and for which r_2 is the relevant reserve price. By increasing r_2, fewer buyer types submit a bid above this threshold, but those who do, submit a higher bid than previously. The seller faces a trade–off which is similar to that associated with reserve prices. He can sacrifice efficiency in order to induce more aggressive bids.

We show that for general distributions of buyer types there are distributions of seller types such that the seller wants to prohibit intermediate bids. Furthermore, if the seller's valuation can assume only a finite number of values, this specific kind of first–price auction is sometimes capable of implementing the overall optimal mechanism.

In the first part of the paper we assume that the seller has to design the rules of the auction before he learns his value. In the second part of the paper we relax this assumption by endogenizing the time at which the auction is designed. We show that if the seller is risk–averse instead of risk–neutral, he might have a strict incentive not to wait until he gets informed. If he waits, he chooses a first–price auction with a reserve price only, however the chosen reserve price varies with the realization of his reservation value. As a consequence, also the buyers' bidding behavior depends on the seller's reserve price. In contrast, if the seller designs the auction while he is still uninformed, he might choose a first–price auction in which intermediate bids are prohibited, but the induced bidding behavior does not vary with his type. As this enables him to reduce the variation in bids without reducing expected revenue, the seller may choose the auction before he is fully informed. Thus, the information structure we assumed in the first part of the paper might arise endogenously when the seller is risk–averse.

Structure of the Paper

We present the model in the subsequent section. In section 3.3 we derive which bidding behavior can be implemented by the seller. Then, in section 3.4, we analyze the structure of the optimal first–price auction if the seller is risk–neutral and does not know his value at the time he designs the auction. In section 3.5 we consider the case in which the seller is risk–averse and the time at which he designs the auction is endogenous. The last section concludes. All proofs can be found in the Appendix.

3.2 The Model

We consider the problem between a seller of an indivisible object (player 0) and two symmetric buyers (players 1 and 2). Throughout the paper we will refer to a generic buyer by i and to the other buyer by $-i$.

The valuations that the three players assign to the object are independently drawn and private information. Buyer i's valuation, θ_i, is drawn from a distribution function $F(\cdot)$ with support $[0,1]$. We assume that a density function $f := F'$ exists and is strictly positive on the support. Moreover, we make the assumption, which is commonly adopted in auction theory, that the virtual valuation function $v(\theta) := \theta - (1-F(\theta))/f(\theta)$ is strictly increasing.[4] The seller himself assigns a reservation value of θ_0 to the object, drawn according to a distribution function $G(\cdot)$. To focus on the interesting cases, we assume that for all buyer types $\theta_i > 0$ trade is efficient with positive probability. Technically, this can be stated as $G(\epsilon) > 0$ for any $\epsilon > 0$.

The seller offers the object for sale via a first–price sealed bid auction in which only bids from a set $B \subset \mathbf{R}_+$ are admissible. For instance, $B = [r, \infty)$ specifies a first–price auction with a reserve price, $B = [0, c]$ a first–price auction with a bid cap, and $B = \{b_1, b_2\}$ one in which participating bidders can only choose between a high and a low bid. The timing is the following: The seller specifies the rules of the auction, B, before he and the buyers privately learn their valuations. Then the buyers simultaneously decide between participating in the auction by submitting a bid $b_i \in B$ and not participating. Finally, the seller observes the bids and decides between selling the object to the highest participating buyer[5] and keeping it. A monetary transfer is only made if the seller actually sells the object at the end. In this case the winning bidder pays his bid to the seller. Thus, the seller's payoff is

$$u_0 = \begin{cases} \theta_0 & \text{if the seller keeps the object} \\ b_i & \text{if buyer } i \text{ obtains the object} \end{cases},$$

[4]This assumption is met by the most common distributions and is implied by an increasing hazard rate $f(\theta)/(1-F(\theta))$.

[5]If the highest bid is submitted by several buyers, we assume that the winner is drawn by chance.

and buyer i's payoff is

$$u_i = \begin{cases} \theta_i - b_i & \text{if buyer } i \text{ obtains the object} \\ 0 & \text{if he does not} \end{cases} .$$

For the time being we assume that all players are risk–neutral. Only in section 3.5 we consider a modified version of the model in which the seller is risk–averse.

As equilibrium concept we adopt the notion of Perfect Bayesian equilibrium. A symmetric Perfect Bayesian equilibrium[6] is characterized by a set of admissible bids B, a set of participating buyer types $P \subset [0,1]$, a bid function $b : P \to B$ mapping those types into bids, and a rule describing for which combinations of bids and reservation values the seller sells and keeps the object at the end. We call a bid function implementable by a set B, if the game that is played after the seller has chosen the set B possesses a symmetric Perfect Bayesian equilibrium inducing this bid function. We call a bid function implementable if it is implementable by some set B.

To sum up, we consider a standard auction setting which is modified in two respects: First, the seller has to fix the rules of the auction, the set B, before he learns his value. Second, he cannot commit to actually selling the object after observing the bids and learning his value.

3.3 Implementable Bid Functions

Given a set of admissible bids B we normally obtain a multiplicity of symmetric equilibria which differ only in the behavior of types with mass zero. For instance, there is an equilibrium for which the buyer type that is just indifferent between participating and not participating does participate, and one in which this type does not. This multiplicity complicates stating our results, but it does not affect expected payoffs. We therefore simplify result statements by pinning down the behavior of indifferent players.

Assumption 3.1 *If a buyer is indifferent between participating and not participating, he participates. If he is indifferent between submitting a low and a high bid, he chooses the high bid. If the seller is indifferent between selling and not selling, he sells.*

[6] With symmetric we refer, of course, only to the buyers' behavior.

In this section we first state results for the standard case in which the seller always sells to the highest bidder (subsection 3.3.1), then we derive necessary and sufficient conditions for bid functions to be implementable when the seller cannot commit to sell (subsection 3.3.2).

3.3.1 The Standard Case: The Seller Always Sells to the Highest Bidder

The results in this subsection are standard and can be found in any textbook on auction theory (e.g., in section 2 of Krishna (2002)).

In the standard case the seller chooses a first–price auction and he is committed to selling to the highest bidder. In this case it is optimal for him to set a reserve price only, i.e. to choose a set of admissible bids $B = [r, \infty)$. Such sets of admissible bids implement the following bidding behavior:

Proposition 3.1 (Bid function, standard case)
Let $B = [r, \infty)$ with $r \in [0, 1]$ and suppose the seller has to sell to the highest bidder. Bid function $\beta_r : P \to \mathbf{R}_+$ is implementable by set B if and only if $P = [r, 1]$ and

$$\beta_r(\theta_i) := \theta_i - \int_r^{\theta_i} \frac{F(s)}{F(\theta_i)} ds. \tag{3.1}$$

If no buyer participates in the auction, the seller keeps the object and obtains $\mathbf{E}[\theta_0]$. Thus, when choosing a reserve price, he behaves as if his valuation is $\mathbf{E}[\theta_0]$. By another standard result, the optimal reserve price can be described by the inverse of the virtual valuation function $v(\theta_i) = \theta_i - \frac{1-F(\theta_i)}{f(\theta_i)}$.

Proposition 3.2 (Optimal reserve price, standard case)
If the seller has to sell to the highest bidder and his valuation is y, the revenue maximizing first–price auction can be implemented by the set of admissible bids $B^* = [\rho(y), \infty)$ with

$$\rho(y) := v^{-1}(y). \tag{3.2}$$

3.3.2 The Non–Standard Case: The Seller Cannot Commit to Selling to the Highest Bidder

At the time the seller has to decide between keeping the object and selling it to the highest bidder, he possesses all information which is payoff–relevant for him. He behaves opportunistically by always choosing his best option.

Proposition 3.3 (Selling decision)
The seller keeps the object if and only if either no buyer participates or if the highest bid lies strictly below his reservation value θ_0.

Knowing the seller's selling behavior, we are able to describe how a buyer's bid affects his probability of getting the object. If buyer i submits bid b and buyer $-i$ bids according to bid function $b^*(\cdot)$, buyer i obtains the object with probability

$$\begin{aligned} H(b) \;\; := \;\; & G(b) \cdot \Big[\text{Prob}(-i \text{ does not participate}) \\ & + \text{Prob}(b^*(\theta_{-i}) < b) + \frac{1}{2}\text{Prob}(b^*(\theta_{-i}) = b)\Big]. \end{aligned}$$

Since a buyer's payoff is increasing in his valuation, the set of participating buyer types is always an upper interval, say P. The following Proposition gives implicit necessary and sufficient conditions for bid functions to be implementable for a given set of participating buyer types $P = [r, 1]$.

Proposition 3.4 (Bid function, non–standard case)
Let $b^* : [r, 1] \rightarrow \mathbf{R}_+$ *with* $r \in [0, 1]$. b^* *is implementable if and only if*

$$b^* \text{ is non–decreasing, and} \tag{3.3}$$

$$b^*(\theta_i) := \theta_i - \int_r^{\theta_i} \frac{H(b^*(s))}{H(b^*(\theta_i))} ds. \tag{3.4}$$

Condition (3.4) looks similar to condition (3.1) that characterizes bidding in the standard case, but it is only implicit since the integrand depends on the absolute level of bids and not only on types.

3.4 The Seller Designs the Auction Before He Learns His Value

In this section we show that for general distributions of buyer types the optimal first–price auction might not be implementable by setting a reserve price only. This can be shown by considering only reservation values that are distributed according to a two–point distribution. Furthermore, restricting attention to such distributions makes it easier to explain the relevant effects. Therefore we stick to the following assumption throughout subsections 3.4.1 and 3.4.2:

Assumption 3.2

$$\theta_0 = \begin{cases} 0 & \textit{with probability } p \\ x & \textit{with probability } 1-p \end{cases}$$

with $p \in (0,1)$ *and* $x \in (0,1)$.

One can think of this assumption in the following way: The object has no value to the seller, but with probability $1-p$ he obtains the possibility to sell it at price x to someone not taking part in the auction.

In subsection 3.4.1 we show how the structure imposed by Assumption 3.2 simplifies the conditions for bid functions to be implementable. Then we analyze the seller's revenue maximization problem in subsection 3.4.2. Finally, we show in subsection 3.4.3 that it is not crucial for our results that the seller's value is discretely distributed.

3.4.1 Implementable Bid Functions (Revisited)

In Proposition 3.4 we derived necessary and sufficient conditions for bid functions to be implementable. However, these conditions are only implicit and not very tractable. In this subsection we use the specific structure of Assumption 3.2 to obtain better tractable conditions.

If the seller sets a reserve price which is higher than his highest possible valuation, i.e. $r \geq x$, his commitment problem is never binding so that we are back in the standard case. Standard theory states that in this case the seller might set a reserve price but has no incentive to restrict the set of admissible bids further. For $B = [r, \infty)$ there is a unique symmetric equilibrium in which each buyer bids according to the bid function $\beta_r(\cdot)$ (see Proposition 3.1).

In the remainder of this subsection we consider the case in which the seller sets a reserve price below his highest possible value, i.e. $r < x$. We first explain which bid function is implemented if the seller does not restrict bidding beyond setting a reserve price, then we explain how he can change the implemented bidding behavior by restricting B further.

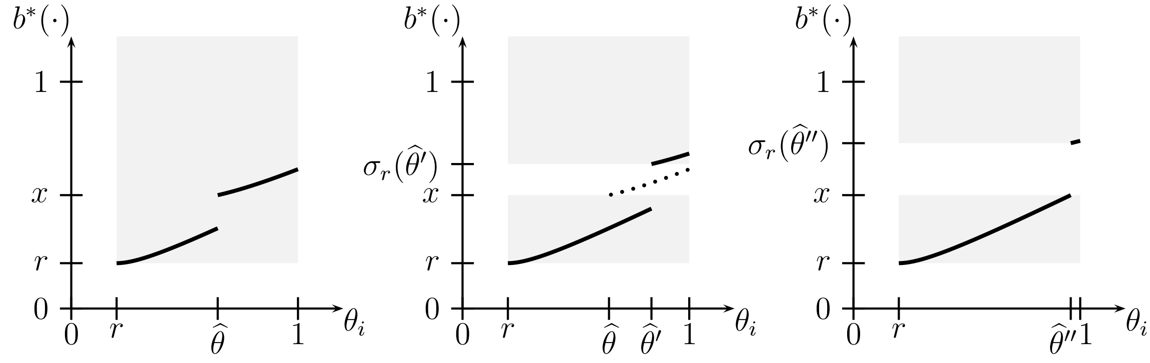
In the standard case a buyer's trade–off is between the probability of being the highest bidder and his payoff conditional on winning. If the seller cannot commit to selling after observing the bids, there is an additional effect. By increasing his bid, a buyer might not only increase the probability of being the highest bidder but also the probability with which the seller is willing to sell to him. For the specific distribution of the seller's reservation value we consider here, a buyer must increase his bid above x to make the seller more eager to sell. But if a buyer's valuation is low, at least if $\theta_i < x$, bidding above x is not attractive for him. If a buyer has such a type, he faces the same trade–off as in the standard case, and, as a consequence, he displays the same bidding behavior as in the standard case, i.e. he bids according to $\beta_r(\cdot)$. However, as the buyer's type increases, some type $\widehat{\theta}$ is eventually reached that is indifferent between bidding $\beta_r(\widehat{\theta})$ and obtaining the object with probability p (conditional on being the highest bidder), and bidding just x in order to obtain it for sure (conditional on being the highest bidder). This type is specified by the condition

$$p(\widehat{\theta} - \beta_r(\widehat{\theta})) = \widehat{\theta} - x \;\Leftrightarrow\; x = (1-p)\widehat{\theta} + p\beta_r(\widehat{\theta}).$$

By defining

$$\sigma_r(\widehat{\theta}) := (1-p)\widehat{\theta} + p\beta_r(\widehat{\theta}) \tag{3.5}$$

we can describe this type by $\widehat{\theta} = \sigma_r^{-1}(x)$. In equilibrium, all types larger than $\sigma_r^{-1}(x)$ increase their bids relative to the standard case and bid above x. An example for such a bid function is depicted in figure 3.1(a). The gray region depicts the set of admissible bids, the black curve is the equilibrium bid function. The bidding behavior is described in the following Proposition:



(a) $B = [r, \infty)$ (b) $B = [r, x) \cup [\sigma_r(\widehat{\theta}'), \infty)$ (c) $B = [r, x) \cup [\sigma_r(\widehat{\theta}''), \infty)$

Figure 3.1: Equilibrium bid functions ($\theta_i \sim U[0,1]$, $x = 0.5$, $p = 0.5$, $r = 0.2$)

Proposition 3.5 (Bid function, two–point distribution, B connected)
Let $B = [r, \infty)$ with $r \in [0, x)$. Bid function $b^* : P \to \mathbf{R}_+$ is implementable by set B if and only if $P = [r, 1]$ and

$$b^*(\theta_i) = \begin{cases} \beta_r(\theta_i) & \text{if } \theta_i \in [r, \widehat{\theta}) \\ \beta_r(\theta_i) + \frac{F(\widehat{\theta})}{F(\theta_i)}(\sigma_r(\widehat{\theta}) - \beta_r(\widehat{\theta})) & \text{if } \theta_i \in [\widehat{\theta}, 1] \end{cases} \tag{3.6}$$

with $\widehat{\theta} = \sigma_r^{-1}(x)$.

Note that if $\sigma_r^{-1}(x) > 1$, all types bid below x. In this case x is so high that bidding above x is not attractive for any buyer type.

We consider now how the seller can affect the implemented bid function by restricting the set of admissible bids. Recall that type $\widehat{\theta} = \sigma_r^{-1}(x)$ is just indifferent between bidding $\beta_r(\widehat{\theta})$ and bidding x when $B = [r, \infty)$. If we do not permit this type to choose bid x, he strictly prefers bidding $\beta_r(\widehat{\theta}) < x$ to any other admissible bid. In the same manner, if we do not only prohibit bid x but all bids in $[x, y)$, all types in $[\widehat{\theta}, \widehat{\theta}') := [\sigma_r^{-1}(x), \sigma_r^{-1}(y))$ bid below instead of above x. An example for the resulting bidding behavior is displayed in figure 3.1(b). The figure shows how the equilibrium bid function in figure 3.1(a) changes if we prohibit bids in $[x, \sigma_r(\widehat{\theta}'))$. The gray region displays, again, the set of admissible bids and the solid curve shows the implemented bid function. To emphasize the change in the bidding behavior, we indicate the bid function from figure 3.1(a) by the dotted curve.

For the subsequent explanations it will be convenient to introduce a specific notation for types in $[0, \widehat{\theta}')$, types in $[\widehat{\theta}', 1]$, and for type $\widehat{\theta}'$. Henceforth we will refer to types bidding below x as *low bidders* and to types bidding above x as *high bidders*. Furthermore, we will refer to the type $\widehat{\theta}'$ separating low bidders from high bidders as *separating type*.

By appropriately restricting the set of admissible bids the seller can induce a higher separating type. In Proposition 3.6 we show that for any $\widehat{\theta}' \in [\sigma_r^{-1}(x), \beta_r^{-1}(x)]$ there is a set of admissible bids B that implements a strictly increasing bid function with separating type $\widehat{\theta}'$. Furthermore, those are the only strictly increasing bid functions that are implementable.

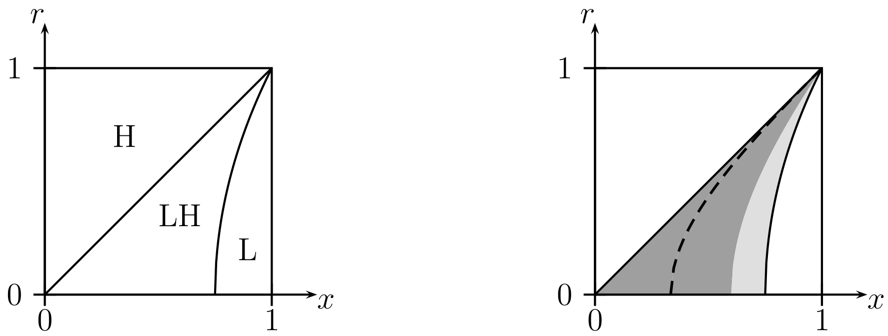
Proposition 3.6 (Bid function, two–point distribution, *B* arbitrary)
Let $r \in [0, x)$. A strictly increasing bid function $b^* : [r, 1] \to \mathbf{R}_+$ is implementable if and only if condition (3.6) is satisfied with $\widehat{\theta}' \in [\sigma_r^{-1}(x), \beta_r^{-1}(x)]$.

The set B by which bid function b^* is implemented must not contain bids from $[x, \sigma_r(\widehat{\theta}'))$. A particular way of implementing b^* is by choosing $B = [r, x) \cup [\sigma_r(\widehat{\theta}'), \infty)$.

We did not explain yet why the seller cannot choose a separating type larger than $\beta_r^{-1}(x)$. For separating type $\widehat{\theta}'' = \beta_r^{-1}(x)$, the implemented bid function is as displayed in figure 3.1(c). The highest low bidders prefer bids lying just below x. If the seller tried to enforce an even higher separating type, some types would want to bid above x, even if this did not increase the seller's eagerness to sell. But since bids just above x are not permitted, these types are effectively constrained by a bid cap. Therefore the implemented bid function would exhibit pooling just below x. Thus, the seller can implement bid functions in which the separating type is larger than $\beta_r^{-1}(x)$, but no strictly increasing ones.

3.4.2 The Revenue Maximizing Auction

In the preceding subsection we derived all implementable bid functions that are strictly increasing. We found that the seller basically has two degrees of freedom, the reserve price r and the separating type $\widehat{\theta}'$. In particular, we found that by prohibiting some intermediate bids the seller can increase the mass of low bidders. In this subsection we first derive the optimal separating type for a given reserve price and explain the rationale for why the seller may actually want to have more low bidders. Then we analyze the optimal choice of the reserve price.



(a) Structure of bidding; H: only high bidders, L: only low bidders, LH: low and high bidders

(b) Structure of B^*; white and light gray regions: B^* connected, dark gray regions: B^* non–connected

Figure 3.2: Structure of bidding and structure of the optimal set of admissible bids ($\theta_i \sim U[0,1]$, $p = 0.5$)

The Optimal Separating Type Given a Reserve Price

If $r \geq x$ (region H in figure 3.2(a)), or if $r < x$ and $\sigma_r^{-1}(x) \geq 1$ (region L), a unique strictly increasing bid function is implementable. In the former case the reserve price forces all participating types to bid above x, in the latter case the reserve price is so much lower than x, that all types prefer bidding below x. Only if $r < x$ and $\sigma_r^{-1}(x) < 1$ (region LH), a multiplicity of strictly increasing bid functions is implementable. While those bid functions have in common that some types bid below and others bid above x, they differ in the type separating low from high bidders.

The separating type does not affect the identity of the highest bidder, but it affects to which types the seller sells the object only if his reservation value is low, and to which types he sells for any realization of his reservation value. A reserve price determines to which types the seller is going to sell with probability zero and with a positive probability. Thus, the separating type and the reserve price have in common that they both determine which buyer types obtain the object with a higher probability and which ones obtain it with a lower probability. By consequence, changes in the reserve price and in the separating type induce similar strategic effects. If the seller increases the reserve price, some types stop participating in the auction, but types that still participate bid higher. If he increases the separating type, some types are forced to bid below x, but types who still bid above x increase their bids (see figure 3.1(b), again). In both cases the seller faces a trade–off between the probability of making an efficient sale and the level of bids (of types who decide to make a high bid).

It turns out that we can describe the optimal separating type using the function $\rho(\cdot)$ which describes the optimal reserve price in standard auctions (see subsection 3.3.1). We obtain the following Proposition:

Proposition 3.7 (Optimal separating type)
Let $r \in [0, x)$ and let $\sigma_r^{-1}(x) < 1$. The optimal strictly increasing and implementable bid function is specified by separating type

$$\widehat{\theta}^* = \begin{cases} \sigma_r^{-1}(x) & \text{if } \rho(x) < \sigma_r^{-1}(x) \\ \rho(x) & \text{if } \sigma_r^{-1}(x) \le \rho(x) \le \beta_r^{-1}(x) \\ \beta_r^{-1}(x) & \text{if } \rho(x) > \beta_r^{-1}(x) \end{cases} . \tag{3.7}$$

The optimum cannot be implemented by a connected set of admissible bids if $\rho(x) > \sigma_r^{-1}(x)$.

By applying a revenue equivalence argument (e.g., by extending Proposition 23.D.3 in Mas-Colell, Whinston, and Green (1995) by a stochastic seller valuation) we obtain that the seller's revenue depends only on the allocation of the object.[7] Furthermore, using standard reasoning (see, e.g., Example 23.F.2 in Mas-Colell, Whinston, and Green (1995)), it is optimal for the seller to sell the object only if the highest bidder's virtual valuation exceeds his reservation value. Hence, if the seller's reservation value turns out to be x, it is optimal for him to sell only to types for which

$$v(\theta_i) \ge x \;\Leftrightarrow\; \theta_i \ge \rho(x).$$

Thus, if the seller can implement a strictly increasing bid function with separating type $\rho(x)$, he achieves the optimal allocation (for a given reserve price). However, as we know from Proposition 3.6, the seller is only able to implement separating types from set $[\sigma_r^{-1}(x), \beta_r^{-1}(x)]$. Although his preferred separating type does not depend on the reserve price, the reserve price determines which separating types are implementable. If $\rho(x) \in$

[7] The seller's revenue depends on (i) for which combinations $(\theta_0, \theta_1, \theta_2)$ he sells the object to the highest bidder and for which combinations he keeps it, and (ii) on the expected payoff of the lowest buyer type. Since we assumed bids to be non–negative and participation in the auction to be voluntary, the lowest buyer type always obtains a payoff of zero. Therefore the seller's revenue depends only on the allocation.

$[\sigma_r^{-1}(x), \beta_r^{-1}(x)]$, the seller can implement his preferred separating type. If $\rho(x)$ lies outside this interval, he chooses the closest separating type which is still implementable.

From Proposition 3.5 we already know that only separating type $\widehat{\theta} = \sigma_r^{-1}(x)$ can be implemented by a connected set of admissible bids. Hence, when $\rho(x) > \sigma_r^{-1}(x)$, the optimum can only be implemented by a non–connected set of admissible bids.

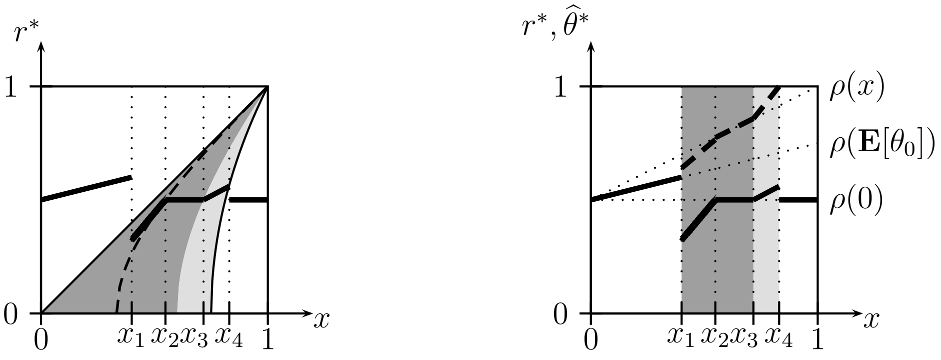
Figure 3.2(b) illustrates the structure of the optimal separating type if the buyers' values are uniformly distributed for different combinations of x and r. Recall that only in the gray regions there is a separating type. In the light gray region the connected set $B^* = [r, \infty)$ implements the optimal separating type (the implemented bid function looks like in figure 3.1(a)). In the dark gray region the optimal separating type can only be implemented by a non–connected set B^*. The dark gray region is divided by a dashed curve into a left and a right part. In the right part the seller can implement his preferred separating type. In the left part his preferred separating type is higher than what he can implement. He therefore chooses the highest separating type possible. This makes the implemented bid function look like in figure 3.1(c).[8]

The Optimal Reserve Price

Up to now we analyzed the optimal choice of a separating type for a given reserve price. Now we address the question how the reserve price should be chosen. First, we show that for general distributions of buyer types there is always an interval of x–values such that the optimal first–price auction can only be implemented by a non–connected set of admissible bids. Then we explain how the optimal reserve price and the optimal separating type depend on the parameter x.

Using a similar reasoning as in the preceding subsection, we easily obtain the allocation which is optimal if the seller has no commitment problem. In this case the seller should sell the object to the highest buyer if this buyer's virtual valuation exceeds his reservation value, otherwise he should keep the object. I.e., if his reservation value is

[8]In this case the seller may have an incentive to implement a bid function which exhibits pooling just below x in order to obtain a higher separating type. Considering also such bid functions may thus further increase the seller's incentive to choose a non–connected set of admissible bids. However, it turns out that restricting attention to strictly increasing bid function suffices to show that for general distributions of buyer types the seller may have an incentive to choose a non–connected set of admissible bids.



(a) Optimal reserve price and regions from figure 3.2

(b) Optimal reserve price (solid curve) and optimal separating type (dashed curve)

Figure 3.3: Optimal reserve price and optimal separating type; dark gray region: B^* non–connected, white and light gray regions: B^* connected ($\theta_i \sim U[0,1]$, $p = 0.5$)

zero, he should sell only to buyers with $v(\theta_i) \geq 0 \Leftrightarrow \theta_i \geq \rho(0)$, and if his reservation value is x, he should sell only to buyers with $v(\theta_i) \geq x \Leftrightarrow \theta_i \geq \rho(x)$. A specific way of implementing this allocation is by choosing a first–price auction with reserve price $r = \rho(0)$ and separating type $\widehat{\theta} = \rho(x)$. Yet, in presence of the commitment problem it might not be possible for the seller to obtain this allocation. While the seller can always choose reserve price $r = \rho(0)$, he might not be able to implement separating type $\widehat{\theta} = \rho(x)$. This is only possible if $\rho(x) \in [\sigma^{-1}_{\rho(0)}(x), \beta^{-1}_{\rho(0)}(x)]$ (Proposition 3.6). However, we can prove that for general distributions of buyer types this is possible at least for some x–values. Moreover, there are always x–values such that the optimal first–price auction can only be implemented by a non-connected set of admissible bids. This happens if $\rho(x) \in (\sigma^{-1}_{\rho(0)}(x), \beta^{-1}_{\rho(0)}(x)]$ (Proposition 3.7). We obtain the following Proposition:

Proposition 3.8 (Optimal reserve price)

(i) For any distribution function $F(\cdot)$ and any probability p there is a non–empty interval of x–values such that $\rho(x) \in (\sigma^{-1}_{\rho(0)}(x), \beta^{-1}_{\rho(0)}(x)]$.

(ii) If $\rho(x) \in (\sigma^{-1}_{\rho(0)}(x), \beta^{-1}_{\rho(0)}(x)]$, the first–price auction with set of admissible bids $B^* = [\rho(0), x) \cup [\sigma_{\rho(0)}(\rho(x)), \infty)$ implements the generally optimal mechanism. Furthermore, this mechanism would be optimal even if the seller knew his reservation value upfront and he had no commitment problem.

We conclude this subsection by describing the dependence of the optimal reserve price and the optimal separating type on the parameter x when the buyers' values are uniformly distributed and the seller's value is high and low with equal probability. Figure 3.3(a) depicts the optimal reserve price for different values of x (thick black curve) and the regions introduced in figure 3.2. Figure 3.3(b) shows the optimal reserve prices (solid curve) together with the optimal separating types (dashed curve). The dark gray regions indicate where the optimal set of admissible bids is non–connected, the white and the light gray regions show where it is connected.

If x is small ($x < x_1$), it is optimal for the seller to choose a reserve price which is higher than both possible realizations of his value. The equilibrium bid function is continuous and lies above x everywhere. Since the seller obtains expected payoff $\mathbf{E}[\theta_0]$ when no buyer is willing to pay the reserve price, the seller chooses the reserve price which is optimal when his value is $\mathbf{E}[\theta_0]$, i.e. $r^* = \rho(\mathbf{E}[\theta_0])$.

If x is large ($x > x_4$), it is too costly for the seller to provide the buyers with incentives to bid above x. Without providing such incentives, he keeps the object if his value is x and sells it through the auction only if his value is zero. Effectively, there is only an auction if the seller's value is zero. Therefore the optimal reserve price is $r^* = \rho(0)$. The equilibrium bid function is continuous and lies everywhere below x.

For intermediate values of x ($x_1 < x < x_4$) the seller chooses a reserve price such that there are low and high bidders. The equilibrium bid function jumps at the separating type $\widehat{\theta}^*$ from below x to above x. If x lies between x_2 and x_3, the generally optimal mechanism can be implemented (see Proposition 3.8). In this case types below $r^* = \rho(0)$ do not participate, types between $r^* = \rho(0)$ and $\widehat{\theta}^* = \rho(x)$ do participate but bid below x, and types above $\widehat{\theta}^* = \rho(x)$ participate and bid above x. This leads to the same allocation that would occur if the seller learned his reservation value first and only then designed the auction. If x lies between x_1 and x_2, the commitment problem prevents the seller from obtaining the optimal allocation. He can only induce separating types below the revenue maximizing one. To implement a higher separating type, he chooses a reserve price below $\rho(0)$. This makes bidding below x more attractive such that the separating type, i.e. the buyer type that is indifferent between bidding above and below x, increases. For x lying between x_3 and x_4 it is the other way around. In this case the seller cannot have the separating type as low as he wants. By increasing the reserve price, he makes bidding above x more attractive such that the separating type decreases.

3.4.3 Robustness: The Seller's Value Is Distributed According to a Continuous Distribution Function

Until now we considered only the case in which the distribution function according to which the seller's reservation value is distributed is discontinuous at x (see Assumption 3.2). We now show that the discontinuity is not crucial for obtaining our results. We can show this by modifying Assumption 3.2 slightly. In this subsection we assume that the seller's value is zero with probability p and is drawn from a uniform distribution on $[x-\epsilon, x]$ with probability $1-p$.

Assumption 3.3

$$\theta_0 \begin{cases} = 0 & \textit{with probability } p \\ \sim U[x-\epsilon, x] & \textit{with probability } 1-p \end{cases}$$

with $p \in (0,1)$, $x \in (0,1)$ *and* $\epsilon \in (0,x)$.

If ϵ is small, the cumulative distribution function is similar to that induced by Assumption 3.2 but it is continuous at x. It turns out that when the seller chooses a first–price auction with $B = [r, \infty)$ and $r < x$, the same bid function that is implemented under Assumption 3.2 is also implemented under Assumption 3.3 if ϵ is sufficiently small.

The reason for this is the following: Suppose that ϵ is small enough such that $r < x - \epsilon$. Then the implement bidding behavior could only differ if some types had an incentive to submit bids from $(x-\epsilon, x)$. Otherwise each type would basically face the same decision problem under both environments such that the same bidding behavior would be induced. Therefore assume to the contrary that it is optimal for some type to choose a bid $b \in (x-\epsilon, x)$. If this type increases his bid marginally, his probability of getting the object increases at least at rate $1/\epsilon$, while his payoff conditional on getting it decreases at rate $1/(\widehat{\theta} - b)$. If $\epsilon < \widehat{\theta} - b$, the type in question has a strict incentive to deviate by increasing his bid. Thus, if ϵ is sufficiently small, nobody wants to submit bids from set $(x-\epsilon, x)$. Hence, we can conclude that the same bidding behavior is implemented under both assumptions.[9]

[9]Note that only types $\theta_i > x - \epsilon$ might have an incentive to submit bids $b \in (x-\epsilon, x)$. Since assumption $r < x - \epsilon$ implies that those types obtain a strictly positive information rent, $\widehat{\theta} - b$ must be bounded away from zero. Therefore there is a threshold $\overline{\epsilon}$ such that for all $\epsilon < \overline{\epsilon}$ the inequality holds.

This proves that if ϵ is sufficiently small, the seller has the same incentives to restrict the set of admissible bids under Assumption 3.3 as he has under Assumption 3.2. As a consequence, the same reasoning as in subsections 3.4.1 and 3.4.2 applies and we obtain the same results.

3.5 The Seller Does Not Want to Learn His Value Before He Designs the Auction

Up to now we assumed that the seller has to fix the rules of the auction before he learns his value. However, it is often not clear why the seller is not able to postpone at least the announcement of some specifics of the auction, e.g. the set B, until he is better informed. In this section we consider a slightly modified version of our model to give a rationale for why the seller might not want to wait. We assume that the seller's information improves at some fixed point in time and that he can choose between fixing rules already before this point in time or only thereafter. From the time at which the auction rules are announced the buyers can infer whether the seller was informed or uninformed when he designed the auction. Furthermore, we now assume that the seller is risk–averse instead of risk–neutral.

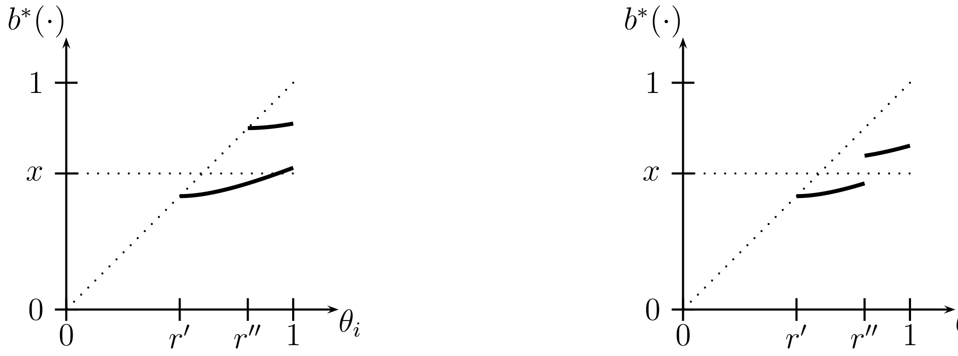
Thus, except for assuming that the seller's expected utility is now $\mathbf{E}[\nu(u_I)]$ with $\nu(\cdot)$ being a strictly increasing and concave function,[10] and except for endogenizing the time at which the auction rules are announced, we stick to the model analyzed in subsections 3.4.1 and 3.4.2. In particular, we assume again that the seller's value is zero with probability p and x with probability $1-p$ (Assumption 3.2).

Case 1: The Seller Designs the Auction After He Learns His Value

Consider first the case in which the seller's reservation value is common information. In this case it is optimal for the seller to set a reserve price without restricting the set of admissible bids further.[11] In contrast to the case in which the seller is risk–neutral, the

[10]See Waehrer, Harstad, and Rothkopf (1998) for a discussion of auctions with a risk–averse seller and risk–neutral buyers.

[11]If the seller restricts bidding further, the implemented bid function exhibits pooling. Since virtual valuations are strictly increasing, it is however better for him to avoid pooling.



(a) Seller learns his value before he chooses B; lower curve: bid function when $\theta_0 = 0$, upper curve: bid function when $\theta_0 = x$

(b) Seller chooses B before he learns his value

Figure 3.4: Equilibrium bid functions ($\theta_i \sim U[0,1]$, $x = 0.6$, $p = 0.5$)

reserve price that is optimal for a risk–averse seller can only be specified implicitly. In the subsequent reasoning we will denote this reserve price by $\widetilde{\rho}(\theta_0)$.

Now consider the case in which the seller's reservation value is private information. In this case multiple equilibria exist, but the equilibrium in which the seller behaves as in the standard case Pareto dominates the other equilibria.[12] I.e., in the Pareto dominant equilibrium he chooses a first–price auction with reserve price $r' := \widetilde{\rho}(0) > 0$ if his reservation value turns out to be zero and one with reserve price $r'' := \widetilde{\rho}(x) > x$ if it turns out to be x. An example of what the implemented bid functions might look like is displayed in figure 3.4(a). Since the seller chooses a reserve price that is higher than the actual realization of his reservation value in both cases, his commitment problem is never binding.

[12] In this footnote we sketch how the Pareto dominant equilibrium can be derived. First note that an equilibrium is characterized by the sets of admissible bids that a seller with valuation zero and one with valuation x choose, the bidding behavior in these cases, and a system of beliefs.

A proof can proceed in three steps: Step 1: The ex ante expected joint surplus of the two seller types is maximized by choosing a first–price auction with reserve price $\widetilde{\rho}(0)$ when the seller's value is zero and one with reserve price $\widetilde{\rho}(x)$ when the seller's value is x. Step 2: For any system of beliefs and for any set of admissible bids that the seller with type zero might choose, it is optimal for the seller with type x to choose a first–price auction with reserve price $\widetilde{\rho}(x)$. Step 3: Given Step 1 and Step 2 there cannot be an equilibrium in which the seller with type zero does better than when he chooses a first–price auction with reserve price $\widetilde{\rho}(0)$. Furthermore, there are beliefs which support this behavior.

Case 2: The Seller Designs the Auction Before He Learns His Value

The only difference between this case and the case analyzed in section 3.4 is that the seller is risk–averse. Since his selling decision is not affected by risk–aversion, the same bid functions that were implementable in section 3.4 are implementable in this case. From Proposition 3.6 we know that not all combinations of reserve price and separating type are implementable, but if the bid function described by reserve price r' and separating type r'' is implementable, the same allocation is implemented as in the Pareto dominant equilibrium of case 1. This bid function is in general not optimal for an uninformed risk–averse seller, but it can serve as a lower bound for the optimum. An example of how a bid function described by a reserve price r' and a separating type r'' looks like is displayed in figure 3.4(b). What is important is that the implemented bidding behavior does not vary with the seller's information.

Comparison of the Two Cases

We now explain why the seller prefers not to wait until he gets informed if the same allocation that is implemented in case 1 can also be implemented in case 2. By a revenue equivalence argument the same expected revenue is generated in both cases. Hence, a risk–neutral seller would be indifferent between the two cases. However, a risk–averse seller also cares about the induced distribution of bids.

A buyer with a type below r'' displays the same bidding behavior in both cases (see Figure 3.4(a) and 3.4(b)). Next consider a buyer whose type is larger than r''. If the seller designs the auction after learning his value, he obtains a high bid when his reservation value is $\theta_0 = x$ and a low bids when it is $\theta_0 = 0$ (see Figure 3.4(a)). On the other hand, if he does not learn his value before he designs the auction, the resulting bid lies between these values (see Figure 3.4(b)). More precisely, the distribution of bids in the informed case is a mean–preserving spread of the bid distribution in the uninformed case. Thus, not waiting and prohibiting intermediate bids induces a smaller dispersion of bids without changing expected bids. This renders not waiting as clearly optimal for a risk-averse seller. We obtain the following Proposition:

Proposition 3.9 (Seller risk–averse)
If the seller is risk–averse, he strictly prefers not waiting and choosing an auction with reserve price r' and separating type r'' over waiting and choosing an auction with reserve

price r' when his value turns out to be zero and an auction with reserve price r'' when it turns out to be x.

3.6 Conclusion

In this paper we studied the problem of an either risk–neutral or risk–averse seller who wants to sell an indivisible object via an auction with a first–price payment rule. The setting we considered differs from standard auction settings because the seller does not know his reservation value at the time he has to design the auction and because he cannot commit to selling the object after observing the buyers' bids. We showed that the structure of the optimal first–price auction in this setting might differ from that in the standard setting. More specifically, the seller might want to prohibit intermediate bids in addition to setting a reserve price. Furthermore, we showed that if the seller is risk–averse, he might have a strict incentive to design the auction before he learns his reservation value. He is sometimes strictly better off fixing an auction format in which he is allowed to set a secret reserve price before he gets informed, than waiting until he gets informed and choosing an auction with an announced reserve price then. Therefore this paper contributes also to the literature which tries to explain the frequent use of secret instead of announced reserve prices.

Chapter 4

Asymmetric Procurement Systems

4.1 Introduction

Although the automotive industries in Europe, North America and Asia are facing similar tasks and incentives in their procurement process, the pattern of procurement has evolved differently. At first glance, the Asian car producers largely engage in a protective long term contract model. As pointed out by Dyer (1996), this system relies on close cooperation between the procurer and his suppliers to build mutual knowledge about each others' production processes and on the sharing of profits achieved through cooperation. Other evidence suggests a downside interpretation with highly demanding car producers exercising high pressure onto the suppliers without using the market. The North American car industry instead relies on frequent competitive auctions between potential suppliers, thus extracting benefits from contracting with the most efficient supplier.

While this difference may largely be rooted in industry history and business culture, one should ask whether there are deeper trade–offs between the two procurement systems. In the recent past, Western car producers tried to imitate their Asian counterparts, as has been observed by McMillan (1990), Dyer (1996) and Liker and Choi (2004). Hence, it seems worth understanding under which circumstances one system is preferable over the other. The main objective of this paper is to analyze and to compare two procurement systems with respect to the expected profit to the procurer, one of which relies more on competitive forces (the "American system"), and the other one relying more on the protection of the incumbent supplier (the "Asian system").

A procurer usually needs to procure goods in regular time intervals. Due to the repetition, the procurement problem becomes inherently asymmetric. At any point in time there are incumbent and entrant suppliers who may differ in several respects in the sense that, for instance, the incumbent may have acquired relationship–specific skills through his collaboration with the procurer.

Thus, the procurer is confronted with the following trade–off: On the one hand, he can realize benefits by continuing the relationship with his incumbent supplier, for instance from a well–functioning just–in–time production and from his engineers working well together with the incumbent's engineers. The level of these benefits can be influenced by the incumbent.[1] Thus, the procurer may want to favor the incumbent in order to provide him with incentives to *invest* in the relationship. On the other hand, he wants procurement prices to be competitively bid down. Therefore he also has an incentive to use the entrants to exercise *competitive pressure* on the incumbent, in order to extract information rents. As Hahn, Kim, and Kim (1986) and McMillan (1990) already point out, these goals are conflicting because higher competitive pressure and thus a higher probability of losing the incumbency status lowers the incumbent's incentive to invest.

The procurement problem is related to three economic problems: (i) a hold–up problem due to the relationship–specific investment that hooks the procurer to the incumbent, (ii) a problem of asymmetric information concerning the suppliers' production costs, and (iii) the relevance of repetition and the influence of future periods on the behavior in the current period.

A *procurement mechanism* describes how the procurer handles the procurement problem in a specific period. It determines to whom the procurement contract is awarded and which monetary transfers have to be made. In contrast, a *procurement system* can be interpreted as a set of general rules concerning the procurement process. The procurer is committed for a longer period of time to the rules imposed by the procurement system,[2] but within the system he can construct mechanisms.

The two procurement systems we are interested in differ only in a procedural aspect.

[1]As Greenstein (1993) and Greenstein (1995) show, investments are sizable and can typically be influenced to a large extent by the incumbent supplier. McMillan (1990) also emphasizes that "there are actions an incumbent can undertake during the course of the initial contract that improve productivity or quality."

[2]This assumption should be interpreted in the light that we model longer term procurement relationships in which the procurer needs to build up reputation with the suppliers for his procurement procedures.

In the *Protective System* (henceforth PS), the procurer is committed to choose a mechanism in which he only negotiates with the entrants after the current relationship with the incumbent irrevocably broke down, while in the *Competitive System* (henceforth CS) he underlies no such restrictions. The two systems differ in the procurer's ability to use the entrants to exercise competitive pressure on the incumbent. This shall reflect the fact that in America the focus is on competition, whereas in Asia incumbent suppliers have a distinct standing. There relationships between procurer and incumbent are deep and the procurer normally tries first to come to an agreement with his incumbent before he looks for possible replacements.[3]

Given either of the procurement systems, we consider the infinite repetition of the procurement problem. Each period begins with the incumbent making an observable investment which generates additional benefits when the relationship is continued. Thereafter, the procurer chooses a procurement mechanism which has to be consistent with the rules imposed by the procurement system he uses. Then, all suppliers learn their production costs as private information and play the chosen mechanism.

Implicit in this way of modeling the strategic interaction between procurer and suppliers is the assumption that the procurer *cannot commit* himself prior to the incumbent's investment decision to a certain course of action afterwards. This commitment assumption is similar to that made by Dasgupta (1990) who analyzes a non–repeated symmetric procurement problem and contrary to Laffont and Tirole (1988) and Bag (1997) who analyze the case in which the procurer *can commit* himself.[4] When the procurer can commit himself, he chooses a mechanism in order to provide the incumbent with the optimal incentives to invest, while when he cannot commit himself and when investment is observable, the causality is reversed: The incumbent invests in order to affect the procurer's mechanism choice.

We are interested in the no commitment case with observable investment for two

[3]Taylor and Wiggins (1997) also compare the American with the Japanese system, but their analysis focuses on different punishment mechanisms within the two systems: In their model the good procured is characterized by an unobservable quality which can be affected by the supplier. In the American system, the procurer can reject delivery and withhold payment if he finds the shipment to be unsatisfactory after a costly inspection. In the Japanese system, there is no inspection, but the procurer punishes unsatisfactory deliveries by cutting the supplier off from his procurement process.

[4]In a more recent paper Arozamena and Cantillon (2004) analyze the incumbent's investment incentives when the procurer is committed to different commonly used (non–optimal) mechanisms.

reasons: First, the kind of investment we have in mind happens gradually over the entire procurement period, for instance through the exchange of technicians or the learning and adjusting to the other's production process. It seems therefore plausible that investment is observable by the procurer but hard to quantify and thus non–verifiable. Second, procurers often seem to behave opportunistically in the short run. For instance, in the automotive industry supply contracts become often binding only when the first part has been delivered, long after sizable investments have been made. It seems therefore not implausible to assume that the procurer is unable to commit himself to a certain procurement mechanism long before he actually needs to procure.

The paper closest to ours is Lewis and Yildirim (2005). They analyze a repeated procurement problem in which a procurer incurs switching costs when he awards the contract to an entrant and they adopt similar commitment assumptions as we do. Their paper complements ours by considering the case in which switching costs, and thus asymmetries, are a strategic choice of the procurer instead of the incumbent.

Results

The procurement mechanism choice in the two systems affects the procurer's revenue in two ways: First, it affects the incumbent's incentives to invest and thus the *investment level* that prevails in equilibrium. Second, it determines the extent to which competitive pressure is exercised on the suppliers and thus the extent to which the procurer can *extract rents* from them. We can distinguish two different kinds of competitive pressure: *Direct competitive pressure* can be interpreted as the extent to which there is competition for the current period's procurement contract between the incumbent and the entrants. The more the procurer favors the incumbent, the lower is direct competitive pressure and the worse is his ability to extract information rents from the suppliers. By contrast, *indirect competitive pressure* arises from the asymmetries between the incumbent and the entrants in the next period. If it is more valuable to a supplier to be incumbent in the next period rather than an entrant, the procurer can bargain for this future advantage and thus extract a future rent today.

Since the procurer behaves opportunistically when he chooses a mechanism, he focusses on direct competitive pressure. He optimizes the extraction of information rents in the current period. Because the procurer is limited in the PS to use the entrance threat to exert pressure on the incumbent, direct pressure is clearly higher in the CS.

However, investment incentives and indirect pressure may be better in the PS.

Regarding the optimal investment level in the two systems we get different results for low, intermediate and high investment costs.

The higher the incumbent's investment, the more to his advantage will the procurer construct the mechanism, i.e. the lower the competition he has to face. If investment costs are low, investment costs are only a minor issue for the incumbent and the optimal investment decision is mainly determined by the effect investment has on competition. In the CS the incumbent competes against the currently best entrant, whereas in the PS, in which the procurer has to negotiate first bilaterally with the incumbent, he only competes against the procurer's expectation thereof. Thus, if the incumbent is very keen on staying in the contract, he must choose an investment level that makes him in the CS preferable to the best conceivable realization of entrants' types, while in the PS it suffices to be preferable to some expected realization. This unambiguously triggers higher investments in the CS when investment costs are low.

In the other polar case with high investment costs, the optimal investment level is negligible in both systems and is thus not crucial for the assessment of the two systems.

If investment costs are intermediate, optimal investment is higher in the PS. The consequence of a higher investment is that the relationship with the incumbent is continued with a higher probability. Since the procurer negotiates in the PS first bilaterally with the incumbent, the continuation probability increases independent of what the entrants' types are. By contrast, since the incumbent is directly confronted with the entrants in the CS, the effect of investment on the continuation probability depends in this system on the entrants' types. In particular, if it turns out that all entrants are very bad, the procurer is willing to continue the relationship with the incumbent independent of whether he increases his investment or not. Thus, a higher investment increases the continuation probability in the CS only if the highest entrant's type is not too bad. This makes the incumbent's marginal revenue from investment structurally higher in the PS and leads to a higher investment there.

The procurer's preferences over the two procurement systems differ for low, intermediate and high investment costs. While for low costs the CS is preferable because investment is higher there, the CS is preferable for high costs because it induces more direct competitive pressure. However, for a region of intermediate cost parameters the procurer prefers the PS due to higher indirect pressure and investment.

Our result that either system may be preferable depending on the characteristics of the part procured is consistent with findings in the empirical literature on procurement systems. Hahn, Kim, and Kim (1986) come to the conclusion that the choice between the systems is not clear–cut and that "a sound purchasing management strategy generally requires a good mix of both approaches for an optimal result." Dyer (1996) presents evidence that for complex products (= products for which investment is important / not so expensive relative to its benefits) the Japanese way of procuring is superior while for simple products (= products for which investment is unimportant / expensive relative to its benefits) the American way is.

Structure of the Paper

In section 4.2, we introduce the model. Then we derive in section 4.3 the optimal procurement mechanism (given investment and procurement system) and in section 4.4 the optimal investment (given a procurement system). In section 4.5 we compare the two systems. Before we conclude in the last section, we qualify our results with respect to case study observations. All proofs can be found in the Appendix.

4.2 The Model

Our model features the infinite repetition of a stage game, which represents one procurement period. We first introduce the stage game, and subsequently we describe the properties of the infinite repetition.

4.2.1 The Stage Game

The model features one procurer who needs to procure an indivisible good from one of $n+1$ suppliers. The suppliers are of two different types: there is one incumbent, I, and n entrants, E1, ..., En. Throughout the paper we will denote a generic entrant by Ej with $j \in \{1, \ldots, n\}$ and a generic supplier by $k \in \{\text{I,E1}, \ldots, \text{E}n\}$.

The only difference between the two types of suppliers is that the incumbent has the opportunity to make an observable relationship–specific investment $i \geq 0$. The investment causes costs $C_\gamma(i) := \gamma C(i)$ to the incumbent and generates a benefit of i in case the relationship between him and the procurer is continued. We assume that the benefit accrues to the incumbent in the first place, but this assumption does not affect our results (as we will discuss below) and is for ease of notation only. Investment costs are the product of a continuously differentiable, strictly increasing and strictly convex cost function $C : \mathbf{R}_+ \to \mathbf{R}_+$ with a parameter $\gamma > 0$. The parameter γ measures the height of investment costs relative to the benefits from investment.[5]

If the procurer comes to an agreement with supplier k, they sign a procurement contract. It is common knowledge that the procurer's value of procuring the good is π, but the costs of providing the good are private information to supplier k. We denote supplier k's costs of providing the good by $\pi - \theta_k$ such that the value generated by the procurement contract (net of potential relationship–specific benefits) is θ_k. As with the relationship–specific benefits, we assume that the entire value generated by the procurement contract accrues to the supplier in the first place. This allows us to describe the contract allocation problem as a standard auction problem instead of a reverse one. As in auction theory, we will refer to θ_k either simply as supplier k's value or as his type.

The suppliers' values are independently and identically distributed according to a cumulative distribution function $\Phi(\cdot)$ with a strictly positive probability density function $\phi(\cdot)$ on the connected support $\Theta = [\underline{\theta}, \overline{\theta}] \subset \mathbf{R}_+$. As common in auction theory, we assume that the hazard rate $\phi/(1-\Phi)$ is increasing in order to prevent bunching in the optimal mechanism. This assumption is satisfied by most of the commonly used distributions and ensures that the virtual valuation function, $v(\theta) := \theta - (1-\Phi(\theta))/\phi(\theta)$, is strictly increasing. Furthermore, we assume the distribution to be such that the virtual valuation function is continuously differentiable.

For the illustration of our results, we will either consider uniformly distributed values or a class of distributions with linear densities on the support $[0, 1]$. Such distributions are completely characterized by the slope of the density function a, $a \in [-2, 2]$. The uniform distribution is a special case in this class, with $a = 0$.

[5] Instead of interpreting γ as a parameter measuring the relative costs of investment, we can also interpret it as a parameter describing the importance of the investment. If we do the monotone transformation $i = C^{-1}(\widetilde{i}/\gamma)$, the costs caused by an investment of level $\widetilde{i}$ become just $\widetilde{i}$ and thus independent of γ. However, the benefits from investment depend after the transformation on γ and are the higher the higher $1/\gamma$ is. Hence, we can reinterpret a cheap investment as an important investment, i.e. an investment generating high benefits, and an expensive investment as an unimportant investment, i.e. an investment generating low benefits.

The outcome of the procurement process is completely characterized by probabilities with which the different suppliers obtain the contract and by transfers that they have to make to the procurer. We denote the vector of probabilities by $y = (y_I, y_{E1}, \dots, y_{En}) \in \Delta_n$ with Δ_n being the set of all vectors $y \in \mathbf{R}^{n+1}$ with $\sum_k y_k = 1$ and $y_k \geq 0$,[6] and we denote the vector of transfers by $t = (t_I, t_{E1}, \dots, t_{En}) \in \mathbf{R}^{n+1}$. Using this notation, we can describe the expected stage game payoffs of the incumbent, an entrant and the procurer: $\widetilde{u}_I := -C_\gamma(i) + y_I \cdot (\theta_I + i) - t_I$, $\widetilde{u}_{Ej} := y_{Ej} \cdot \theta_{Ej} - t_{Ej}$ and $\widetilde{u}_P := t_I + \sum_j t_{Ej}$.

Which outcome is chosen is determined by a *procurement mechanism.* We will denote such a mechanism by $\Gamma = (S_I, S_E; y, t)$. S_I and S_E are strategy sets for the incumbent and the entrants, respectively, $y : S_I \times S_E^n \rightarrow \Delta_n$ is a contract allocation rule and $t : S_I \times S_E^n \rightarrow \mathbf{R}^{n+1}$ is a transfer rule. To allow for the possibility that a supplier decides not to participate in the procurement process, we assume that S_I and S_E contain messages leading to a zero probability of obtaining the object and a zero transfer. We denote the set containing all possible mechanisms by $\mathcal{M}$.

We identify a *procurement system* with a set of mechanisms to which the procurer can commit himself upfront and we are interested in comparing two specific systems: In the *Competitive System* (CS) the procurer is allowed to choose any procurement mechanism, i.e. $\Gamma \in \mathcal{M}_{CS} := \mathcal{M}$. In the *Protective System* (PS) he is restricted in the choice of a mechanism, as he decides about continuing the contract with the incumbent before considering alternative options. Formally, this is $\Gamma \in \mathcal{M}_{PS} := \{\Gamma \in \mathcal{M} | y_I(s_I, s_{E1}, \dots, s_{Ej}) = y_I(s_I)\}$.

The timing of the stage game is such that given either of the procurement systems, the incumbent first chooses an investment. Then the investment is observed and the procurer picks a procurement mechanism out of the relevant procurement system. Finally, the suppliers observe the mechanism chosen, learn their private information θ_k and play the mechanism.

[6] By assuming $\sum_k y_k = 1$, we implicitly assume that the procurer always has to award the contract to one of the suppliers. Think of a situation in which the procurer produces a complex product, e.g. a car, and the part in question is crucial for production. In such situations it is often not an option for the procurer not to procure. Moreover, this property arises endogenously in equilibrium if $\pi \geq v(\underline{\theta})$.

4.2.2 The Repeated Game

We consider the infinite repetition of the stage game and we assume that future payoffs are discounted by a discount factor $\delta \in (0,1)$. Moreover, we assume contract values to be serially independent such that the identity of the incumbent is the only payoff relevant information that carries over from one period to the next one.

Since we are interested in the case in which the procurer is able to commit himself to a procurement system, but that within the system he behaves opportunistically at any point in time,[7] we consider Markov perfect equilibria[8] given either of the procurement systems. Since under this equilibrium notion equilibrium behavior depends only on payoff relevant information, each period is played in the same way and the only part of history that matters for a supplier is whether he enters the current period as incumbent or as one of the entrants. Thus, the strategy of a supplier is completely characterized by his behavior as incumbent and as one of the entrants. Moreover, we restrict attention to equilibria in which the suppliers' strategies are symmetric such that an equilibrium is completely characterized by an equilibrium investment i^* for the current incumbent, a rule mapping investments into mechanism choices for the procurer, and for each possible mechanism choice profiles describing the behavior of the current incumbent, $s_I^*(\cdot)$, and the current entrants, $s_E^*(\cdot)$.

By the one-stage-deviation-principle[9] equilibrium can be computed by considering the non–repeated stage game, however adjusted by continuation values reflecting the payoffs of equilibrium play in all future periods. We denote the continuation values of the next period's incumbent, of one of the next period's entrants, and of the procurer by V_I, V_E and V_P, respectively. Total payoffs are then

$$\begin{aligned} u_I &= -C(i) + y_I \cdot (\theta_I + i + V_I) + (1 - y_I) \cdot V_E - t_I, \\ u_E &= y_{Ej} \cdot (\theta_{Ej} + V_I) + (1 - y_{Ej}) \cdot V_E - t_{Ej}, \text{ and} \\ u_P &= t_I + \textstyle\sum_j t_{Ej} + V_P. \end{aligned}$$

In equilibrium continuation values have to be consistent with the actual equilibrium play, i.e. they are recursively defined by $V_I = \delta \mathbf{E}[u_I]$, $V_E = \delta \mathbf{E}[u_E]$ and $V_P = \delta \mathbf{E}[u_P]$.

[7] This means in particular that he cannot credibly threaten a supplier to penalize him in future periods and that he cannot reward suppliers with multi–period contracts.

[8] A Markov perfect equilibrium is a subgame perfect equilibrium in Markov strategies (see Maskin and Tirole (2001)).

[9] See Fudenberg and Tirole (1991) section 4.2.

4.3 The Optimal Procurement Mechanism for a Given Investment Level in the CS and in the PS

In this section we derive and interpret the procurer's optimal procurement mechanism given a procurement system and an investment. First, using standard mechanism design results,[10] we give conditions for direct mechanisms to be implementable (subsection 4.3.1). Then we derive the optimal mechanisms for both procurement systems (subsection 4.3.2) and describe a convenient way of implementing them indirectly (subsection 4.3.3). Afterwards we compare the optimal mechanisms in both systems regarding protection of the incumbent (subsection 4.3.4), the effect of investment on the implemented contract allocation (subsection 4.3.5) and the efficiency properties of the implemented contract allocation (subsection 4.3.6).

4.3.1 Implementable Direct Mechanisms

By a revelation principle[11] we can without loss of generality restrict attention to direct mechanisms $\Gamma = (\Theta, \Theta; y, t)$ and equilibria in which all suppliers reveal their types truthfully in equilibrium. Because of quasi–linearity of preferences, a supplier's optimal behavior depends only on the contract allocation rule and the transfer rule via the probability with which he expects to obtain the object,

$$\overline{y}_k(\theta) := \mathbf{E}[y_k(\theta_I, \theta_{E1}, \ldots, \theta_{En})|\theta_k = \theta],$$

and the expected transfer he has to make,

$$\bar{t}_k(\theta) := \mathbf{E}[t_k(\theta_I, \theta_{E1}, \ldots, \theta_{En})|\theta_k = \theta].$$

Applying an Envelope Theorem we obtain necessary and sufficient conditions for a direct mechanism to be implementable.

Proposition 4.1 (IC and IR)
The contract allocation rule $y : \Theta^{n+1} \to \Delta_n$ and the transfer rule $t : \Theta^{n+1} \to \mathbf{R}^{n+1}$ specify a direct mechanism for which truth–telling is optimal for all suppliers and individual rationality constraints are binding for the worst types if and only if

[10] See Chapter 23 in Mas-Colell, Whinston, and Green (1995).
[11] See Proposition 23.D.1 in Mas-Colell, Whinston, and Green (1995).

$\overline{y}_k(\theta)$ *is non-decreasing,*

$$\overline{t}_I(\theta) = \overline{y}_I(\theta)(\theta + i + V_I - V_E) - \int_{\underline{\theta}}^{\theta} \overline{y}_I(s)ds \text{ and}$$

$$\overline{t}_{Ej}(\theta) = \overline{y}_{Ej}(\theta)(\theta + V_I - V_E) - \int_{\underline{\theta}}^{\theta} \overline{y}_{Ej}(s)ds.$$

Applying integration by parts we obtain simple expressions for expected transfers.

Lemma 4.1 (Expected transfers)

(i) $\mathbf{E}[t_I(\cdot)] = \mathbf{E}[y_I(\cdot)(v(\theta_I) + i + V_I - V_E)]$

(ii) $\mathbf{E}[t_{Ej}(\cdot)] = \mathbf{E}[y_{Ej}(\cdot)(v(\theta_{Ej}) + V_I - V_E)]$

Lemma 4.1 allows us to describe how the expected payoffs of all players depend on investment, allocation rule and continuation values. The current incumbent's total payoff is

$$\mathbf{E}[u_I] = -C_\gamma(i) + \mathbf{E}[y_I(\cdot)(\theta_I - v(\theta_I))] + V_E. \tag{4.1}$$

The incumbent obtains an information rent, $\mathbf{E}[y_I(\cdot)(\theta_I - v(\theta_I))]$, and the continuation value of an entrant, V_E, even though he might stay incumbent with positive probability. Since values are serially independent there is in the current period no asymmetric information about the expected advantage of being the next period's incumbent. Therefore the procurer can completely extract this future advantage today.

Although the incumbent does not directly obtain any part of the benefits from investment, he has to bear its total costs. However, as we will see in the next subsection, the optimal allocation rule depends on investment. If the incumbent invests more, he obtains the contract more often. This increases his informational rent and provides him with an incentive to invest. Since investment is observable and the incumbent's payoff is additively separable in his private information and investment, the investment does not directly affect the information rent the procurer has to leave to him. As consequence, the procurer can extract the entire benefits from investment. This is the reason why it is not important in our model to whom the relationship–specific benefits accrue in the first place.

An entrant's total rent has the same structure as the incumbent's total rent, except for the costs of investment:

$$\mathbf{E}[u_{Ej}] = \mathbf{E}[y_{Ej}(\cdot)(\theta_{Ej} - v(\theta_{Ej}))] + V_E. \tag{4.2}$$

Finally, we obtain the procurer's per period revenue which is the sum of expected transfers:

$$\mathbf{E}[t_I(\cdot) + \sum_j t_{Ej}(\cdot)] = \mathbf{E}[y_I(\cdot)(v(\theta_I) + i) + \sum_j y_{Ej}(\cdot)v(\theta_{Ej})] + V_I - V_E \tag{4.3}$$

The procurer's revenue consists of two parts: The first part depends on the current period's allocation and the benefits from investment, the second part is a supplier's monetary advantage of being the next period's incumbent instead of one of the entrants.

Since the suppliers have private information about their current period's production costs, the procurer has to leave them an information rent. He can only extract the virtual valuation from the supplier who obtains the contract, which is less than his actual valuation. The virtual valuation of an entrant is given by the function $v(\theta_{Ej})$ known from standard auction theory,[12] the incumbent's virtual valuation has to be adjusted by the benefits from investment and is thus $v(\theta_I) + i$.

Due to the asymmetries in our model, i.e. the investment in both systems and additionally the asymmetries which are inherent in the PS by the construction of the system, the supplier who becomes the incumbent in the next period can expect to get a different information rent than a supplier who starts the next period as an entrant. However, the procurer can bargain for the incumbent's advantage in the next period today and thus extract a future rent today. This suggests that asymmetries between the incumbent and the entrants might be valuable for the procurer.

4.3.2 The Optimal Mechanisms

Knowing how the procurer's revenue depends on the chosen procurement mechanism, we can go on to derive the optimal mechanism in each procurement system. Recall that the two systems differ only in the set of feasible mechanisms. Whereas in the CS the procurer can freely choose a mechanism, he is in the PS committed to make his decision about continuing the relationship with his incumbent before gathering information about the realization of the entrants' values.

For both systems the optimal mechanism is obtained from expression (4.3). Since virtual valuations are monotonic, it is optimal to award the contract to the entrant with the highest value if the relationship with the incumbent is not continued. To abbreviate the description of our results, we introduce notation $\theta_E := \max_j \theta_{Ej}$ for the highest

[12] See, e.g., Myerson (1981) or, for an economic interpretation, Bulow and Roberts (1989).

of the entrants' values. θ_E is a random variable with cumulative distribution function $\Psi := \Phi^n$ and probability density function $\psi := n\phi\Phi^{n-1}$.

Proposition 4.2 (Optimal mechanisms)

(i) The optimal mechanism in the CS awards the contract to the incumbent if $v(\theta_I) + i > v(\theta_E)$ *and to the entrant with the highest value otherwise.*

(ii) The optimal mechanism in the PS awards the contract to the incumbent if $v(\theta_I) + i > \mathbf{E}[v(\theta_E)]$ *and to the entrant with the highest value otherwise.*

Transfers can take any form that is consistent with the interim expected transfers specified in Proposition 4.1.

In the CS it is optimal to allocate the contract to the supplier with the highest virtual valuation. In the PS the procurer has to decide about continuing the relationship with the incumbent before he learns anything about the realization of the entrants' virtual valuations. The best he can do is comparing the incumbent's actual virtual valuation, $v(\theta_I)+i$, with the virtual valuation he can expect to get from the best entrant, $\mathbf{E}[v(\theta_E)]$.

Because maximization in the CS is over a larger set of mechanisms, i.e. $\mathcal{M}_{PS} \subset \mathcal{M}_{CS}$, the procurer would clearly prefer the CS if there were neither differences in investment nor differences in the incumbency advantage. But since these generally differ for the two systems, the commitment to a smaller set of mechanisms may be beneficial for the procurer. Whether and when this is the case is what we analyze in the remainder of the paper.

4.3.3 Indirect Implementation of the Optimal Mechanisms

A simple way of implementing the optimal mechanisms is to use a first–price auction plus a take–it–or–leave–it offer: In the CS the seller holds an auction among the entrants, observes the highest bid and confronts the incumbent with an offer that depends on the outcome of the auction. The entrant who submits the highest bid only obtains the contract if the incumbent does not accept the offer. In the PS the order is reversed. The seller first makes an uninformed offer to the incumbent and holds a first–price auction among the entrants only if the offer is declined.

4.3.4 Protection of the Incumbent

For an outsider observing procurement procedures, procurement tenders and auctions are the main observable characteristics.[13] Looking at the two systems from this perspective, the incumbent seems to be better protected in the PS. In the CS, there are procurement tenders in each and every period suggesting that there is a steady threat of switching to one of the entrants and that the entrants are employed to exercise pressure on the incumbent. By contrast, public tenders are only observed in the PS after a relationship actually broke down. Therefore the incumbent appears to be much less jeopardized there. This notion might be misleading, as we will discuss below. A better indicator for protection of the incumbent might be the probability with which the relationship is continued. In this subsection we analyze whether the incumbent is in our model indeed better protected in the PS, i.e. whether it is justified to call this system *protective*.

A crucial difference between the two systems is that the incumbent competes against the actually best entrant in the CS, whereas he competes only against what the procurer expects to be the best entrant in the PS. Since there are always realizations of the highest entrant's type that are better than some expected realization, the investment necessary to win for sure is structurally higher in the CS. Thus, if the incumbent wants to get full protection against the entrants, he gets it already for a smaller investment in the PS than in the CS. In this respect the incumbent is indeed better protected in the PS.

Proposition 4.3 (Continuation of the relationship for sure)
The incumbent obtains the contract in the CS for sure if $i \geq \bar{i}_{CS} := v(\overline{\theta}) - v(\underline{\theta})$ *and in the PS if* $i \geq \bar{i}_{PS} := \mathbf{E}[v(\theta_E)] - v(\underline{\theta})$. *We have* $\bar{i}_{PS} < \bar{i}_{CS}$.

From Proposition 4.3 we know that for higher investments the incumbent is indeed better protected in the PS. Now we show that this is not necessarily true for lower investments. We show this by considering the case in which the incumbent does not invest. In this case all suppliers are symmetric and thus the optimal mechanism in the CS is also symmetric, as can be seen from Proposition 4.2 (i). As consequence, each supplier is awarded the contract with probability $1/(n+1)$. By contrast, in the PS, in

[13] This is, e.g., observable from newspapers. While there are articles about newly established relationships, this is normally not the case for relationships which are just continued without having engaged in public tender procedures.

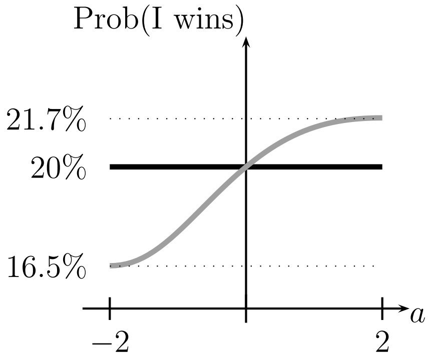


Figure 4.1: Probability that I wins when $i = 0$; gray curve: PS, black curve: CS (ϕ linear with support $[0, 1]$ and slope a, $n = 4$)

which the procurer has to negotiate first bilaterally with the incumbent, mechanisms are necessarily asymmetric. We know from Proposition 4.2 (ii) that the incumbent's virtual valuation, $v(\theta_I)$, is compared with the expected virtual valuation of the highest entrant, $\mathbf{E}[v(\theta_E)]$, instead of with the actual virtual valuation of the highest entrant, $v(\theta_E)$, as in the CS. This has the consequence that the incumbent obtains the contract in the PS but not in the CS if $v(\theta_E)$ turns out to be high (i.e. if $\mathbf{E}[v(\theta_E)] \leq v(\theta_I) < v(\theta_E)$) and he obtains it in the CS but not in the PS if $v(\theta_E)$ turns out to be low (i.e. if $v(\theta_E) < v(\theta_I) \leq \mathbf{E}[v(\theta_E)]$). From this we can infer that there is an advantage for the incumbent in the PS relative to the CS if the probability that $v(\theta_E)$ lies above $\mathbf{E}[v(\theta_E)]$ is high. Whether this is the case depends on the curvature of v and Φ.

Proposition 4.4 (Continuation of the relationship when $i = 0$)
Let $i = 0$. The probability that the contract with the incumbent is continued in the CS is $1/(n + 1)$. The probability that this happens in the PS is strictly larger, equal or strictly smaller than $1/(n+1)$, if $\Phi \circ v^{-1}$ is strictly convex, linear or strictly concave, respectively.

How these probabilities look like for distributions with linear densities is displayed in figure 4.1. The horizontal axis displays the slope parameter of the density function, the gray curve depicts the probability in the PS and the black curve that in the CS. For the uniform distribution, i.e. for $a = 0$, the incumbent is equally well protected in both systems. If the slope is positive, he is better protected in the PS than in the CS. If the slope is negative, the reverse is true. This is consistent with the explanation we gave above. There is an advantage for the incumbent in the PS relative to the CS if the

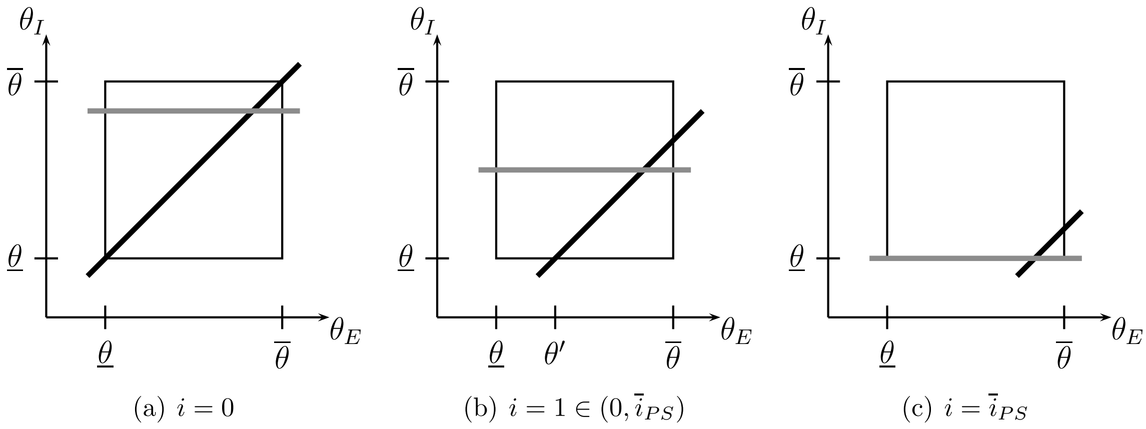


Figure 4.2: Contract allocation; black line: I obtains the contract in the CS above this line; gray line: I obtains the contract in the PS above this line ($\theta_k \sim U[0.5, 2]$, $n = 2$)

probability that $v(\theta_E)$ lies above $\mathbf{E}[v(\theta_E)]$ is high. This is just the case for positive slope parameters.

We can draw two conclusions about the protection of the incumbent. First, observability of procurement tenders might be a bad indicator for protection. For example, when investment is low and the density function is linear with a negative slope, there are much less observable market interactions in the PS, but the probability that the relationship is continued is higher in the CS. Second, for higher levels of investment, i.e. when asymmetries between the entrants and the incumbent are stronger, the incumbent is indeed better protected in the PS.

4.3.5 The Effect of Investment on the Contract Allocation

In this subsection we explain how investment affects the optimal contract allocation in the two systems. Figure 4.2 displays the contract allocations for uniformly distributed values and for three different levels of investment. The square indicates the joint support of θ_E and θ_I. The incumbent obtains the contract in the CS for value combinations (θ_E, θ_I) above the black line and in the PS for combinations above the gray line. The lines describing the contract allocation shift downwards as investment increases such that the incumbent obtains the contract more often.

For the further discussion it will be convenient to introduce a special notation for the lowest of the incumbent's types who obtains the contract for a given realization of the highest entrant's type, θ_E. We will call this type the incumbent's *lowest winning type* and we will denote this type by $\widetilde{\theta}_I(\theta_E)$.

In the PS the lowest winning type is defined as the lowest type θ_I for which

$$v(\theta_I) + i \geq \mathbf{E}[v(\theta_E)].$$

Thus,

$$\widetilde{\theta}_I(\theta_E) = \begin{cases} v^{-1}(\mathbf{E}[v(\theta_E)] - i) & \text{if } i \leq \bar{i}_{PS} \\ \underline{\theta} & \text{if } i > \bar{i}_{PS} \end{cases}.$$

As long as $i < \bar{i}_{PS}$, a marginal increase in investment decreases the lowest winning type at rate $|\mathrm{d}\widetilde{\theta}_I/\mathrm{d}i| = 1/v'(\widetilde{\theta}_I) \in (0,1)$.[14] The marginal effect of investment on the contract allocation does not depend on the actual realization of the highest entrant's type θ_E. If $i > \bar{i}_{PS}$, the incumbent obtains the contract already for sure such that an increase in investment cannot be rewarded by a further decrease in the lowest winning type.

In the CS the lowest winning type is defined as the lowest type θ_I for which

$$v(\theta_I) + i \geq v(\theta_E).$$

Thus,

$$\widetilde{\theta}_I(\theta_E) = \begin{cases} v^{-1}(v(\theta_E) - i) & \text{if } i \leq \bar{i}_{CS} \text{ and } \theta_E > v^{-1}(v(\underline{\theta}) + i) \\ \underline{\theta} & \text{if } i > \bar{i}_{CS} \text{ or } \theta_E \leq v^{-1}(v(\underline{\theta}) + i) \end{cases}.$$

Again, for high levels of investment ($i > \bar{i}_{CS}$) the incumbent obtains the contract already for sure such that a further increase in investment cannot be rewarded by a decrease in the incumbent's lowest winning type. For lower levels of investment ($i \leq \bar{i}_{CS}$) the lowest winning type depends in the CS, in contrast to the PS, on the actual realization of the highest entrant's type θ_E. This can be seen in figure 4.2(b). If the highest entrant's type is bad (i.e. if $\theta_E \leq \theta'$ in figure 4.2(b)), the procurer is very eager to obtain the relationship–specific benefit and awards the contract to the incumbent irrespective of his type. In this case a marginal increase in investment cannot further decrease the incumbent's lowest winning type (i.e. $|\mathrm{d}\widetilde{\theta}_I/\mathrm{d}i| = 0$). Only if the highest entrant's type

[14] When we denote the hazard rate by $h(x) := \phi(x)/(1 - \Phi(x))$, we have $v(x) = x - 1/h(x)$ and $v'(x) = 1 + h'(x)/h(x)^2$. Thus, it follows from the increasing hazard rate property that $1/v'(x) < 1$.

is high (i.e. if $\theta_E > \theta'$ in figure 4.2(b)), an increase in investment decreases the lowest winning type (i.e. $|\mathrm{d}\widetilde{\theta}_I/\mathrm{d}i| = 1/v'(\widetilde{\theta}_I) \in (0,1)$).

We conclude this subsection by emphasizing two properties which we will use repeatedly later on: First, when $i < \bar{i}_{PS}$, investment alters the contract allocation in the PS with certainty, i.e. for any realization of θ_E, whereas this happens in the CS only with some probability, i.e. when θ_E is sufficiently high. When investment increases in the CS from $i = 0$ to $i = \bar{i}_{CS}$, this probability decreases from one to zero. Second, the incumbent's lowest winning type decreases in both systems less than proportional in investment, i.e. $|\mathrm{d}\widetilde{\theta}_I/\mathrm{d}i| < 1$.

4.3.6 Efficiency of the Contract Allocation

In this subsection we assess the optimal contract allocations in the two systems from an efficiency point of view. This will help us deriving the procurer's preferences over the two procurement systems in section 4.5.

While the contract allocation that maximizes the procurer's revenue is determined by virtual valuations and investment, efficiency of the contract allocation depends on real values and investment. It is efficient to award the contract to the incumbent if $\theta_I + i \geq \theta_E$.

To obtain clear effects, we will sometimes make use of the following assumption:

Assumption 4.1 $\mathbf{E}[v(\theta_E)] - v(\underline{\theta}) \geq \overline{\theta} - \underline{\theta}$

This assumption ensures that it is for an investment of $\bar{i}_{PS}$ efficient that the incumbent obtains the contract for sure. The assumption is satisfied for any distribution if there are sufficiently many entrants,[15] and for any number of entrants if the density is at $\underline{\theta}$ smaller than the average density.[16] In the class of distributions with linear densities on $[0, 1]$, the assumption is, e.g., satisfied for slope parameters $a \in [-0.75, 2]$ when $n = 2$ and for any slope parameter when $n \geq 7$.

[15]For $n \to \infty$ the left hand side converges towards $\overline{\theta} - \underline{\theta} + 1/\phi(\underline{\theta})$ which is strictly larger as the right hand side.

[16]Since $\mathbf{E}[v(\theta_E)] > \underline{\theta}$ the left hand side is larger than $1/\phi(\underline{\theta})$. Hence, a sufficient condition for the assumption to be satisfied is $\phi(\underline{\theta}) \leq 1/(\overline{\theta} - \underline{\theta})$.

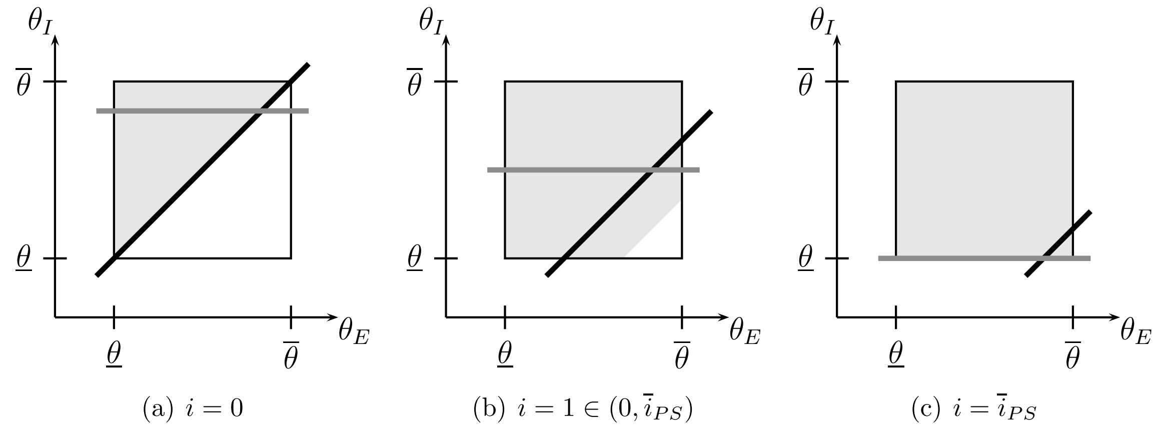


Figure 4.3: Contract allocation and efficiency; gray region: it is efficient that I obtains the contract; black line: I obtains the contract in the CS above this line; gray line: I obtains the contract in the PS above this line ($\theta_k \sim U[0.5, 2]$, $n = 2$)

In the following Proposition we state our efficiency results for the two procurement systems, before we use a graphical example to explain them.

Proposition 4.5 (Efficiency of the contract allocation)

(i) If $i = 0$, the contract allocation is efficient in the CS, but not in the PS.

(ii) If $i \in (0, \bar{i}_{PS})$, neither system allocates efficiently. There are type combinations (θ_E, θ_I) for which the PS allocates efficiently but the CS does not, and vice versa.

Furthermore, in the CS, the highest entrant always obtains the contract when this is efficient, but the incumbent obtains it too seldom.

(iii) If $i \in [\bar{i}_{PS}, \bar{i}_{CS})$ and Assumption 4.1 is satisfied, the contract allocation is efficient in the PS, but not in the CS.

(iv) If $i \in [\bar{i}_{CS}, \infty)$, the contract allocation is efficient in both systems.

Figure 4.3 depicts the same contract allocations that are depicted in figure 4.2 together with the efficient contract allocations. For type combinations in the gray shaded area it is efficient that the incumbent obtains the contract.

If there is no investment, all suppliers are symmetric. In this case it is efficient that the supplier with the highest value obtains the contract and this is what happens in the CS. Since only asymmetric mechanisms are feasible in the PS, efficiency cannot be achieved there. This is depicted in figure 4.3(a).

Consider now $i \in (0, \bar{i}_{PS})$. We know from part (i) of Proposition 4.5 that for $i = 0$ the allocation in the CS is efficient and we know from subsection 4.3.5 that the incumbent's lowest winning type decreases at a rate less than one when investment increases. However, since investment i and type θ_I are perfect substitutes regarding efficiency, the incumbent's lowest winning type has to decrease at rate one to sustain efficiency. Hence, the incumbent obtains the contract too rarely in the CS (see figure 4.3(b)). An analogous reasoning does not go through for the PS since the allocation is for $i = 0$ not efficient in this system.

As for $i \in (0, \bar{i}_{PS})$ the curves describing the allocations implemented in the CS and in the PS intersect in the interior of the type space, there are regions close to the point of intersection in which the incumbent obtains the contract in the CS but not in the PS and vice versa. By the reasoning in the previous paragraph, it is however efficient that the incumbent always obtains the contract in these regions. As a consequence, we cannot rank the two systems regarding efficiency: For some combinations of types the CS allocates the contract efficiently but the PS does not and vice versa.

Thirdly, figure 4.3(c) depicts the case in which $i \in (\bar{i}_{PS}, \bar{i}_{CS})$. For such investments it is by Assumption 4.1 efficient that the incumbent obtains the contract for sure. This happens however only in the PS, rendering this system more efficient.

Finally, for $i \in [\bar{i}_{CS}, \infty)$ it is efficient that the incumbent always obtains the contract and this happens in both systems.

4.4 The Optimal Investment Level in the CS and in the PS

In this section we analyze the incumbent's investment choice for a given procurement system. We first derive the structure of the incumbent's revenue from investment (subsection 4.4.1), then we add the cost function into our considerations and analyze optimal investment (subsection 4.4.2), before we make some statements concerning efficiency of the optimal investment decision (subsection 4.4.3).

4.4.1 The Incumbent's Revenue from Investment

The incumbent invests in order to maximize his expected payoff as specified in (4.1) with the contract allocation rule $y_I(\cdot)$ being as described in Proposition 4.2 (i) and (ii) for the CS and the PS, respectively. To abbreviate notation we will denote the revenue part of his expected payoff by $R(i) := \mathbf{E}[y_I(\cdot) \cdot (\theta_I - v(\theta_I))]$.

In this subsection we proceed in three steps: First, we derive the incumbent's marginal revenue from investment, then we compare the height of marginal revenue in the two procurement systems, before we analyze the curvature of revenue.

Marginal Revenue

Using the notation introduced in subsection 4.3.5, the incumbent's revenue is

$$R(i) = \int_{\underline{\theta}}^{\overline{\theta}} \int_{\widetilde{\theta}_I(\theta_E)}^{\overline{\theta}} (\theta_I - v(\theta_I)) \mathrm{d}\Phi(\theta_I) \mathrm{d}\Psi(\theta_E)$$

such that his marginal revenue is

$$R'(i) = \int_{\underline{\theta}}^{\overline{\theta}} \left| \mathrm{d}\widetilde{\theta}_I(\theta_E)/\mathrm{d}i \right| (1 - \Phi(\widetilde{\theta}_I(\theta_E))) \mathrm{d}\Psi(\theta_E). \tag{4.4}$$

Since the benefits from investment are relationship–specific, the procurer has to continue his relationship with the incumbent to obtain the benefits. Thus, although the procurer does not directly reward investment (see subsection 4.3.1), he rewards it indirectly by awarding the contract more often to the incumbent (see subsection 4.3.5). This increases the incumbent's information rent and provides him with incentives to invest. We can see these effects from expression (4.4): $|\mathrm{d}\widetilde{\theta}_i/\mathrm{d}i|$ describes how fast the incumbent's lowest winning type decreases when investment increases marginally, and $1 - \Phi(\widetilde{\theta}_I(\theta_E))$ describes how fast the incumbent's information rent increases when his lowest winning type decreases marginally.

By imposing the specific structure that the lowest winning type has in the two systems (see subsection 4.3.5), we obtain the following Lemma:

Lemma 4.2 (Marginal revenue)
Let $w := \frac{1-\Phi}{v'}$.

$$\textit{(i)}\ \ R'_{PS}(i) = \begin{cases} w(v^{-1}(\mathbf{E}[v(\theta_E)] - i)) & \textit{if } i < \overline{i}_{PS} \\ 0 & \textit{if } i > \overline{i}_{PS} \end{cases}.$$

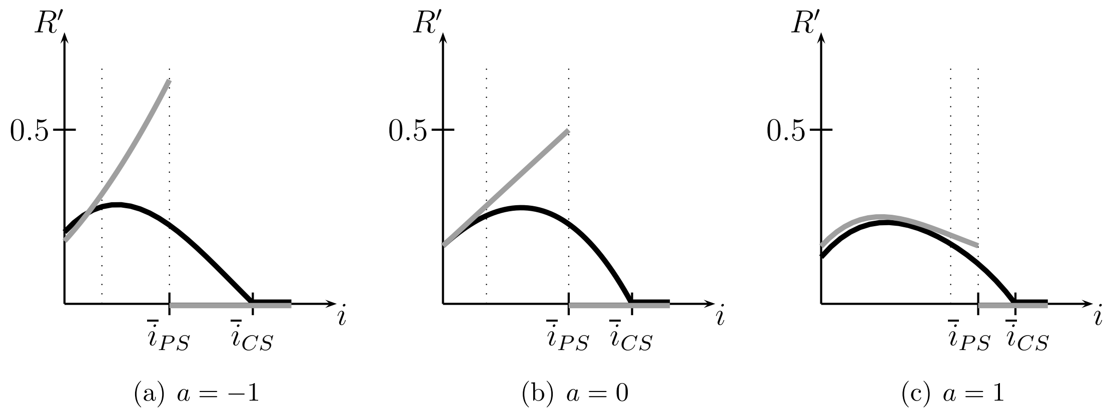


Figure 4.4: Marginal revenue; gray curves: PS, black cuves: CS (ϕ linear with support $[0, 1]$ and slope a, $n = 2$)

On the left of $\bar{i}_{PS}$ marginal revenue is continuous and bounded from below by a strictly positive number. At $\bar{i}_{PS}$ it drops to zero.

$$\textit{(ii)}\quad R'_{CS}(i) = \begin{cases} \int_{v^{-1}(v(\underline{\theta})+i)}^{\overline{\theta}} w(v^{-1}(v(\theta_E) - i))\, d\Psi(\theta_E) & \textit{if } i < \bar{i}_{CS} \\ 0 & \textit{if } i > \bar{i}_{CS} \end{cases}.$$

Marginal revenue is everywhere continuous. In any interval $(0, i')$ with $i' < \bar{i}_{CS}$ it is bounded from below by a strictly positive number.

Relative Height of Marginal Revenue in the Two Systems

Now we derive some structural differences in the height of marginal revenue in the two systems. Figure 4.4 shows how marginal revenue looks like in the two systems for different distributions. Figure 4.4(a) is for a distribution with decreasing density, figure 4.4(b) for the uniform distribution and figure 4.4(c) for a distribution with increasing density. The gray curves depict marginal revenue in the PS, the black ones marginal revenue in the CS.

We can identify three regions as indicated by the dotted vertical lines in figure 4.4: At first, marginal revenue is on a similar level in both systems, then it becomes relatively higher in the PS. At some point, however, it eventually drops to zero in the PS, whereas it stays positive in the CS.

It is easy to see why the third region arises: As soon as the incumbent obtains the contract for sure, marginal revenue drops to zero. From Proposition 4.3 we know that this happens in the PS already for a smaller investment than in the CS. Thus, there exists a region in which marginal revenue is zero in the PS but still positive in the CS. This region is $(\bar{\imath}_{PS}, \bar{\imath}_{CS})$.

Why the first two regions arise is more involved. For the reasoning that we will apply it is helpful to rewrite expression (4.4) in the following way:

$$\begin{aligned} R'(i) &= \text{Prob}\Big(\frac{\mathrm{d}\widetilde{\theta}_I(\theta_E)}{\mathrm{d}i} \neq 0\Big)\mathbf{E}\Big[w(\widetilde{\theta}_I(\theta_E))\Big|\frac{\mathrm{d}\widetilde{\theta}_I(\theta_E)}{\mathrm{d}i} \neq 0\Big]. \\ &= (\text{Probability that } \widetilde{\theta}_I(\theta_E) \text{ decreases}) \\ &\qquad \times (\text{Marginal change in I's information rent} \\ &\qquad\quad \text{conditional on that } \widetilde{\theta}_I(\theta_E) \text{ decreases}) \end{aligned}$$

In subsection 4.3.5 we observed that in the first two regions (i.e. for $i < \bar{\imath}_{PS}$) investment decreases the incumbent's lowest winning type $\widetilde{\theta}_I(\theta_E)$ in the PS for any realization of the entrants' types, but in the CS only if the highest entrant's type is not too bad. Thus, the probability term is always one in the PS but generally smaller than one in the CS. The heuristics for the first two regions are the following: When investment is low, the probability term is only slightly lower in the CS than in the PS and it is ambiguous in which system the marginal change in the incumbent's information rent is higher. This renders the compound effect ambiguous (region 1). However, as investment increases, the probability term stays one in the PS, whereas it falls substantially below one in the CS. Although the information rent effect is still ambiguous, the probability effect dominates and renders marginal revenue higher in the PS (region 2).

That the heuristics for regions 1 and 2 are indeed true for the entire class of distributions with linear densities is confirmed in figure 4.5. Figure 4.5(a) shows that marginal revenue is for $i = 0$ on a similar level in both systems and that it is ambiguous in which system it is higher. Figure 4.5(b) shows that marginal revenue at $\bar{\imath}_{PS}$ is for all slope parameters higher in the PS. If the slope parameter is not too large, the difference is significant.

The following Lemma states sufficient conditions for marginal revenue being higher in the PS for investment levels close to $\bar{\imath}_{PS}$:

Lemma 4.3 (Higher marginal revenue in the PS)

(i) If w is decreasing, i.e. if $R_{PS}(i)$ is convex, then there exists a $\underline{i} \in (0, \bar{\imath}_{PS})$ such that $R'_{PS}(i) > R'_{CS}(i)$ for all $i \in (\underline{i}, \bar{\imath}_{PS})$.

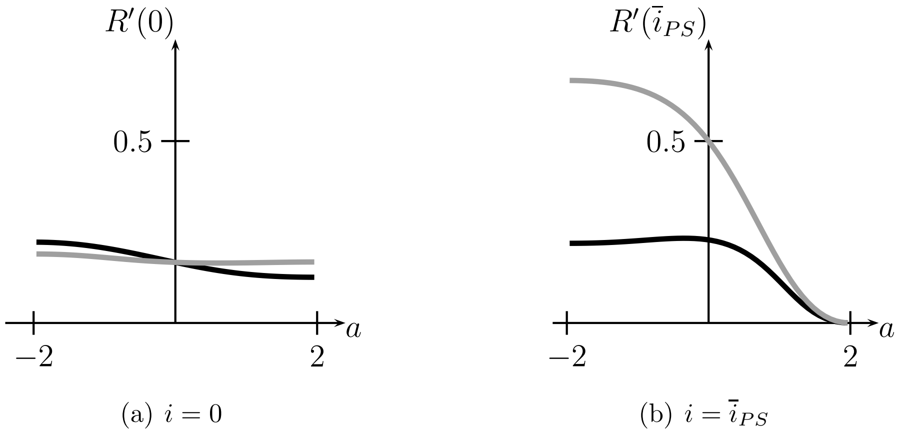


Figure 4.5: Marginal revenue; gray curves: PS; black curves: CS (ϕ linear with support $[0, 1]$ and slope a, $n = 2$)

(ii) If $w \circ v^{-1}$ *is weakly concave on* $[v(\underline{\theta}) - \bar{i}_{PS}, v(\overline{\theta})]$*, then* $R'_{PS}(i) > R'_{CS}(i)$ *for all* $i \in (0, \bar{i}_{PS})$*.*

The condition in (i) is satisfied for many distributions (see the discussion about the curvature of revenue below), but allows no statement about the length of the interval in which marginal revenue is higher in the PS. The result in (ii) is very strong, marginal revenue is everywhere higher in the PS, but the condition that needs to be satisfied to obtain this result is only met by a few distributions, for instance by the uniform distribution.

Curvature of Revenue

As figure 4.4 suggests, marginal revenue is in many cases not monotonous, i.e. the curvature of revenue differs for different levels of investment. In the subsequent paragraphs we discuss the structure of the curvature of revenue in the two systems.

Since revenue depends in both systems only via the incumbent's lowest winning type on investment, we analyze first how this type depends on investment. While we already know from subsection 4.3.5 that the procurer lets more of the incumbent's types win when investment increases, we don't know yet whether his incentive to decrease $\widetilde{\theta}_I(\theta_E)$ becomes larger or smaller when investment increases. I.e., we don't know the second–order effect of investment on $\widetilde{\theta}_I(\theta_E)$. This effect depends on two countervailing effects, a *positive direct effect* and a *negative strategic effect*.

The *positive direct effect* arises as follows: Because the procurer extracts the total benefits from investment when he awards the contract to the incumbent, investment affects his expected payoff via the term Prob(I wins) $\cdot\, i$. Since the probability term and the investment level are complementary, the procurer's incentive to increase the probability by decreasing $\widetilde{\theta}_I(\theta_E)$ becomes stronger for larger levels of investment.

On the other hand, there is a *negative strategic effect*: By awarding the contract more often to the incumbent, the procurer has to leave a higher information rent to all types of the incumbent to which he is already willing to award the contract, i.e. to all types above $\widetilde{\theta}_I(\theta_E)$. As a consequence, the expected information rent he has to leave with the incumbent increases the stronger the more of the incumbent's types already win, i.e. the lower $\widetilde{\theta}_I(\theta_E)$ becomes. Due to this effect, the procurer's incentive to decrease $\widetilde{\theta}_I(\theta_E)$ becomes weaker for higher levels of investment.

Thus, the second–order effect of investment on $\widetilde{\theta}_I(\theta_E)$ is ambiguous and depends on the relative height of the positive direct and the negative strategic effect.

We are now able to discuss the curvature of revenue in the PS. For $i < \bar{\imath}_{PS}$ marginal revenue possesses for this system the structure

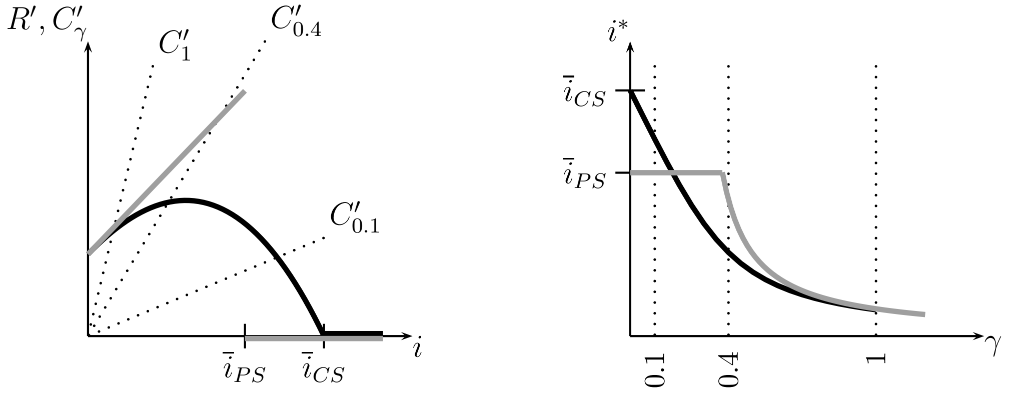
$$R'_{PS}(i) = |\mathrm{d}\widetilde{\theta}_I/\mathrm{d}i|(1-\Phi(\widetilde{\theta}_I)).$$

$|\mathrm{d}\widetilde{\theta}_I/\mathrm{d}i|$ is, as discussed above, increasing or decreasing depending on whether the positive direct effect or the negative indirect effect is stronger. $(1-\Phi(\widetilde{\theta}_I))$ describes the rate at which the incumbent's information rent increases when $\widetilde{\theta}_I$ decreases. Since the information rent increases the stronger the larger investment becomes, we obtain that revenue is in the PS convex if the direct positive effect is either stronger as the indirect negative effect, or at least not too much weaker.

In the class of distributions with linear densities, revenue is only not convex when the slope of the density is sufficiently positive. Examples in which revenue is convex in the PS, i.e. in which marginal revenue is increasing, are depicted in figures 4.4(a) and 4.4(b). An example in which the slope of the density function is large enough to make revenue non–convex is given in figure 4.4(c). In this case revenue is first convex and then concave.

In the CS, marginal revenue can be written as

$$R'(i) = \mathrm{Prob}\Big(\frac{\mathrm{d}\widetilde{\theta}_I(\theta_E)}{\mathrm{d}i} \neq 0\Big)\mathbf{E}\Big[w(\widetilde{\theta}_I(\theta_E))\Big|\frac{\mathrm{d}\widetilde{\theta}_I(\theta_E)}{\mathrm{d}i} \neq 0\Big].$$



(a) Marginal revenue and marginal cost

(b) Optimal investment

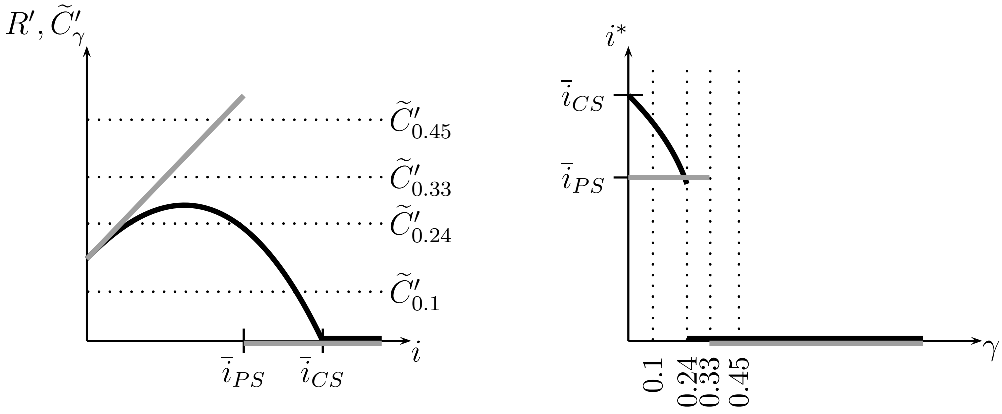
Figure 4.6: Marginal revenue and optimal investment; gray curves: PS, black curves: CS ($\theta_k \sim U[0,1]$, $n = 2$, $C_\gamma(i) = \gamma i^2/2$)

The expectation term exhibits similar effects as marginal revenue in the PS. In particular, this term is increasing in investment when marginal revenue is increasing in the PS. However, the probability term is decreasing and tends towards zero as investment approaches $\bar{i}_{CS}$. This makes revenue inevitably concave for investment levels close to $\bar{i}_{CS}$. All in all, in the CS revenue is likely to be first convex and then concave. Examples for this are the cases depicted in figures 4.4(a), 4.4(b) and 4.4(c).

4.4.2 The Optimal Investment Levels

In this subsection we include the convex cost function $C_\gamma(i) = \gamma C(i)$ into our considerations. Since revenue is generically non–concave in investment, the analysis of optimal investment is involved: second–order effects are relevant, multiple local optima may exist, optimal investment may jump in the parameters of the model, etc. By consequence, optimal investment can look quite differently for different cost functions. Nevertheless, we will derive some properties that hold for general cost functions, $C(i)$, and depend only on the relative height of costs and benefits from investment, i.e. on the parameter γ. We derive results for low, intermediate and high cost parameters resembling cheap investment, intermediate expensive investment and highly expensive investment.

We illustrate our results for uniformly distributed values and for two different cost functions $C(i)$. Figures 4.6 and 4.7 display marginal revenue, marginal costs and optimal



(a) Marginal revenue and marginal cost (b) Optimal investment

Figure 4.7: Marginal revenue and optimal investment; gray curves: PS, black curves: CS ($\theta_k \sim U[0,1]$, $n = 2$, $\widetilde{C}_\gamma(i) = \gamma i$)

investment for quadratic and for linear costs,[17] respectively. The gray curves are for the PS, the black ones for the CS. The dotted lines in figure 4.6(a) and 4.7(a) show marginal costs for different values of the cost parameter. The dotted vertical lines in figures 4.6(b) and 4.7(b) indicate the respective cost parameters.

Cheap Investment

If investment is cheap (consider $C'_{0.1}$ in figure 4.6(a) and $\widetilde{C}'_{0.1}$ in figure 4.7(a)), optimal investment is higher in the CS. This holds generally and arises because marginal revenue is positive over a larger range in the CS. Moreover, optimal investment is for small cost parameters generally flat in the PS. This is because in the PS marginal revenue eventually drops to zero such that corner solutions arise for sufficiently small cost parameters.

Proposition 4.6 (Optimal investment, γ small)
For γ sufficiently close to zero we have $i^*_{PS} = \bar{i}_{PS}$. For $\gamma \to 0$ we have $i^*_{CS} \to \bar{i}_{CS}$. The limit is, however, not reached for positive γ.

It follows from this Proposition that when investment is cheap, there is no conflict between providing the incumbent with incentives to invest and exercising competitive pressure on him. This runs somewhat counter the intuition that the better protection

[17] Since we assumed that the cost function is strictly convex, one might think of an almost linear cost function. For example of $C(i) = i + \epsilon i^2$ with $\epsilon > 0$ very small.

of the incumbent in the PS should increase his incentives to invest. However, protection can be positive as well as negative for incentives. There are two important features of investment: First, the incumbent has to bear the costs of investment even if he does not obtain the contract, and second, investment gives him an advantage in the competition with the entrants. If investment is very cheap, only the second effect is relevant and stronger competition leads to higher investments. Thus, when investment costs are only a minor issue, protection clearly decreases the incumbent's incentive to invest.

Highly Expensive Investment

If investment is expensive (consider C_1' and $\widetilde{C}_{0.45}'$), investment is similarly low in both systems. Which system induces a higher investment depends on the local structure of marginal costs and marginal revenue at $i = 0$. For instance, there is no investment if marginal costs are positive at zero (see figure 4.7(b)) and a small but positive investment if marginal costs are zero (see figure 4.6(b)). In the latter case marginal revenue at $i = 0$ determines in which system investment is higher. Since it is ambiguous in which system marginal revenue is higher (see figure 4.5(a)), either system may induce a higher investment.

Proposition 4.7 (Optimal investment, γ large)

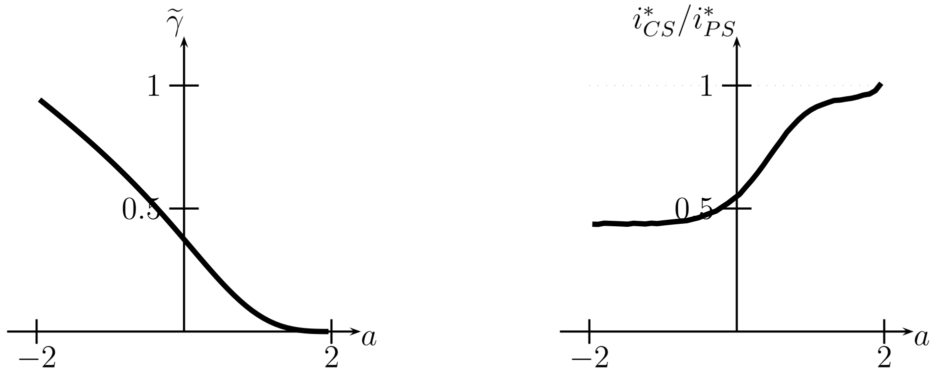
(i) For $\gamma \to \infty$ we have $i^*_{PS} \to 0$ and $i^*_{CS} \to 0$.

(ii) If $\mathbf{E}[w(\theta_E)] < w(v^{-1}(\mathbf{E}[v(\theta_E)]))$, the PS is better at inducing investment when γ is sufficiently high.[18] If the inequality is reversed, the CS is better.

Intermediate Expensive Investment

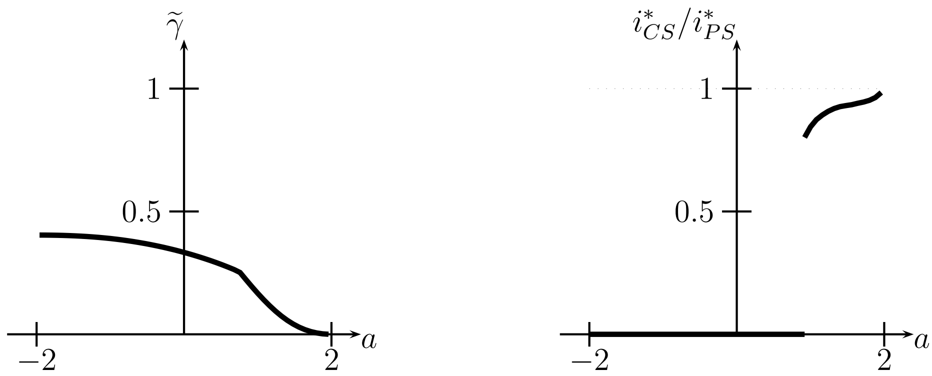
In both examples there is an interval of intermediate cost parameters for which investment is higher in the PS. However, the structure of optimal investment may differ more drastically for intermediate cost parameters. Particularly important is the relative curvature of costs and revenue, i.e. the relative slope of marginal revenue and marginal

[18] We need to define what we mean with the statement "system 1 is better at inducing investment than system 2 when γ is sufficiently high": If $C'(0) = 0$, then there exists a γ' such that investment in system 1 is higher for all $\gamma > \gamma'$. If $C'(0) > 0$, then there exist values $\gamma' < \gamma''$ such that there is no investment in both system for $\gamma > \gamma''$ and there is investment in system 1 but not in system 2 for $\gamma \in (\gamma', \gamma'')$.



(a) highest γ for which $i^*_{PS} = \overline{i}_{PS}$ (b) relative investment for $\gamma = \widetilde{\gamma}$

Figure 4.8: Property 4.1 (ϕ linear with support $[0, 1]$ and slope a, $n = 2$, $C_\gamma(i) = \gamma i^2/2$)



(a) highest γ for which $i^*_{PS} = \overline{i}_{PS}$ (b) relative investment for $\gamma = \widetilde{\gamma}$

Figure 4.9: Property 4.1 (ϕ linear with support $[0, 1]$ and slope a, $n = 2$, $C_\gamma(i) = \gamma i$)

costs. Figure 4.6(a) shows the case of quadratic costs in which the cost function is for $i < \overline{i}_{PS}$ more convex than the revenue functions. This induces interior optima and optimal investment varies continuously in the cost parameter. In contrast, for linear costs revenue is everywhere in the PS and partly in the CS more convex than costs, as depicted in figure 4.7(a). This causes corner solutions and jumps in the optimal investment level.

Crucial for the comparison of the two procurement systems in section 4.5 will be that marginal revenue being higher in the PS for intermediate investment levels (see Lemma 4.3) carries over to investment being higher in the PS for intermediate cost parameters. More specifically, we will need that for decreasing cost parameters investment level $\overline{i}_{PS}$ is reached in the PS before it is reached in the CS:

Property 4.1 *There exists a $\tilde{\gamma}$ for which $i^*_{PS} = \bar{i}_{PS} \geq i^*_{CS}$.*

Figures 4.8 and 4.9 show that Property 4.1 does indeed hold for any distribution with linear density when costs are either quadratic or linear and there are two entrants.[19] In figure 4.8(a) we plot the highest γ, say $\tilde{\gamma}$, for which it is still optimal for the incumbent to choose investment level $\bar{i}_{PS}$ in the PS when costs are quadratic. In figure 4.9(a) we do the same for linear costs. Figures 4.8(b) and 4.9(b) display the optimal investment level in the CS relative to the optimal investment level in the PS for cost parameter $\tilde{\gamma}$. It can be seen that the optimal investment level is in the PS always higher than in the CS, i.e. $i^*_{CS}/i^*_{PS} \leq 1$. Hence, Property 4.1 holds in these cases.

Although Property 4.1 seems to hold quite generally, the non–concavity of revenue makes it difficult to find general conditions on the primitives of the model which ensure that it holds. When we compare the two procurement systems in section 4.5, we will impose Property 4.1 directly.

4.4.3 Efficiency of Investment

For the comparison of the procurement systems in section 4.5 it will be helpful to know whether there is over– or under–investment in the two systems.

We can determine the jointly efficient investment level and contract allocation rule in a two–step procedure: First, we have to determine the efficient contract allocation for a given investment level (as we have already done in subsection 4.3.6), then we have to determine the efficient investment choice given the efficient contract allocation rule. For the CS we obtain the following under–investment result:[20]

Proposition 4.8 (Efficiency of investment)
If it is efficient to have a positive level of investment, then there is under–investment in the CS.

The result is driven by two effects: First, to obtain the benefits from investment, the relationship with the incumbent has to be continued. By the analysis in subsection 4.3.6,

[19] The plots are derived numerically using Maple.

[20] Proposition 4.8 is consistent with findings in Tirole (1986) and Dasgupta (1990) who analyze settings in which a procurer is also not able to commit himself to a mechanism prior to an investment decision. In these papers opportunistic behavior by the procurer leads to under–investment.

this happens too rarely in the CS. Second, it follows from the analysis in subsection 4.3.5 that even if the relationship with the incumbent is continued, the incumbent is (marginally) under–compensated for investment. Both effects together imply that the incumbent's incentive to invest is too low such that the efficient investment level is not achieved in the CS.

An analogous reasoning as for the CS does not go through for the PS since the incumbent does not necessarily obtain the contract too rarely in this system. For instance, if assumption 4.1 is violated, he may obtain the contract for sure although this is not efficient.

4.5 Comparison of the Procurement Systems

After having discussed the incumbent's optimal investment decision, we can now turn to discussing the procurer's preferences over the two procurement systems. We do this for the cases in which repetition is not and very important.

4.5.1 Repetition Unimportant

When the discount factor goes to zero, only the payoffs in the current period are relevant. We obtain the procurer's expected revenue from expression (4.3) by setting the continuation values to zero:

$$\mathbf{E}[t_I(\cdot) + \sum_j t_{Ej}(\cdot)] = \mathbf{E}[y_I(\cdot) \cdot (v(\theta_I) + i) + (1 - y_I(\cdot)) \cdot v(\theta_E)]. \tag{4.5}$$

$y_I(\cdot)$ is as described in Proposition 4.2 (i) and (ii) in the CS and the PS, respectively. The determinants of the procurer's revenue are the level of investment and the degree of extraction of the information rents in the current period.

Due to his commitment problem, the procurer takes investment as given und chooses a procurement mechanism in order to extract as much rents from the suppliers as possible. I.e., he chooses $y_I(\cdot)$ to maximize just expression (4.5). Since the procurer is in the CS, in contrast to the PS, not restricted in his mechanism choice, rent extraction is clearly better in this system. Thus, the PS could only be preferable for the procurer if investment was higher there. However, for small values of cost parameter γ investment is higher in the CS (Proposition 4.6) and for large values it is negligible in both systems

determinant		γ low	γ intermediate	γ high
level of investment	(+)	CS better	ambiguous	negligible
rent extraction	(+)	CS better		
optimal system		CS	ambiguous	CS

Table 4.1: Determinants of the procurer's preferences over the systems when repetition is unimportant

(Proposition 4.7). Hence, only for intermediate values the PS might perform better. This result is stated in the following Proposition, the relevant effects are summarized in table 4.1.

Proposition 4.9 (Optimal system, repetition unimportant)
If repetition is unimportant, then the procurer either never prefers the PS over the CS, or he does so only for intermediate cost parameters γ. If he prefers the PS, this is due to a higher investment there.

Since for intermediate cost parameters investment is likely to be higher in the PS, it is not unlikely that an intermediate region of cost parameters exists for which the procurer prefers the PS.

4.5.2 Repetition Very Important

When the discount factor goes to one, the procurer's ability to extract future rents matters. He is able to extract in each period a supplier's advantage of being the incumbent instead of one of the entrants in the next period, $V_I - V_E$. From expressions (4.1) and (4.2) we obtain the incumbency advantage

$$V_I - V_E = -C(i) + \mathbf{E}[y_I(\cdot)(\theta_I - v(\theta_I)) - \frac{1}{n}(1 - y_I(\cdot))(\theta_E - v(\theta_E))]. \tag{4.6}$$

In the CS the incumbency advantage is always positive by a revealed preferences argument: If the incumbent does not invest, he is treated like an entrant. If he nevertheless decides to invest, he must be better off. An analogous reasoning does not work in the PS, because the incumbent is in this system not treated like an entrant if he refrains from investing. It turns out that in the PS the incumbent can indeed be worse off than an entrant such that the *incumbency advantage* may actually be a *disadvantage*.[21]

Inserting the incumbency advantage (4.6) in (4.3), we obtain the procurer's expected per period revenue:

$$\begin{aligned}\mathbf{E}[t_I(\cdot) + \sum_j t_{Ej}(\cdot)] &= -C(i) + \mathbf{E}[y_I(\cdot) \cdot (\theta_I + i) + (1 - y_I(\cdot)) \cdot \theta_E] \\ &\quad -(n+1)\mathbf{E}[\frac{1}{n}(1 - y_I(\cdot)) \cdot (\theta_E - v(\theta_E))]. \qquad (4.7)\end{aligned}$$

$y_I(\cdot)$ is as described in Proposition 4.2 (i) and (ii) in the CS and the PS, respectively. Except for the expected rent of an entrant, i.e. the expectation term in the second line of (4.7), which the procurer has to leave to each of the $n+1$ suppliers because of their outside options, he can extract the entire value that is created in each period. Therefore he is interested in investment *and* contract allocation being efficient, and in the rent of an entrant being as small as possible.

We now compare efficiency of investment, efficiency of contract allocation and the rent of an entrant in the two systems for small, high and intermediate investment cost parameters.

If γ is sufficiently small, investment is higher in the CS (Proposition 4.6). Since, there is never over–investment in the CS (Proposition 4.8), investment is also more efficient in the CS. Since investment is such that the contract is (almost) with certainty continued in both systems, the contract allocation is (almost) efficient in both systems (Proposition 4.5 (iv)) and entrants obtain (almost) no rent in both systems. Combining these effects, we obtain that the higher investment in the CS makes this system preferable when the cost parameter is low.

If γ is sufficiently high, there is (almost) no investment in both systems (Proposition 4.7), but the contract allocation induced by the CS is (almost) efficient, whereas that induced by the PS is not (Proposition 4.5 (i)). Thus, it would only be possible that the procurer preferred the PS over the CS if the entrants obtained a smaller rent in the PS. However, heuristically, the stronger competition in the CS drives the rent an entrant obtains in the CS below what he obtains in the PS. We can prove this formally for the

[21] For the entire class of distributions with linear densities the incumbency advantage is negative if there is no investment. Heuristically, the procurer is in the PS better than in the CS at screening the incumbent. This lowers the incumbent's information rent in the PS relative to the CS. Although the incumbent has an advantage regarding the probability of winning if the slope parameter is positive, the negative information rent effect is strong enough to make him worse off than an entrant.

determinant	γ low	γ intermediate	γ high
efficiency of (+) investment	CS better	equal or PS better	negligible in both
efficiency of (+) allocation	both efficient	PS efficient	CS efficient
an entrant's (−) rent	none in both	none in the PS, positive in the CS	CS better
optimal system	CS	PS	CS

Table 4.2: Determinants of the procurer's preferences over the systems when repetition is important

case in which the entrants have in the CS a disadvantage regarding the probability of obtaining the contract relative to the PS,[22] but it is also likely to be true otherwise. For instance, in our example with linear densities it is true in the entire class. Hence, when the cost parameter is high, the procurer prefers the CS because it induces fiercer competition.

If Property 4.1 holds, there exists an intermediate γ for which investment level $\bar{i}_{PS}$ is reached in the PS, an investment level (weakly) below this is obtained in the CS, and there is no over–investment in the PS. The reasoning why there is no over–investment for the γ in question is involved and can be found in the proof of Proposition 4.10. If Assumption 4.1 is satisfied, the incumbent obtains the contract in the PS for sure but not in the CS such that the contract allocation is more efficient in the PS (Proposition 4.5 (iii)). Moreover, the entrants obtain positive rents in the CS but not in the PS. Hence, if Property 4.1 and Assumption 4.1 hold, the procurer clearly prefers the PS for a region of intermediate cost parameters.

These results are stated in the following Proposition, the relevant effects are summarized in table 4.2.

Proposition 4.10 (Optimal system, repetition important)

(i) If γ is sufficiently small, the procurer prefers the CS.

(ii) If Assumption 4.1 is satisfied and Property 4.1 holds, there exists an interval of intermediate cost parameters γ for which the procurer prefers the PS.

[22] This is, for instance, the case for $a > 0$ in our example with linear densities.

(iii) *If γ is sufficiently large and Prob(I wins in the PS$|i = 0) \leq 1/(n+1)$, the procurer prefers the CS.*

We can conclude that none of the two procurement systems in question is preferable for the procurer for all cost parameters. Instead the procurer will prefer the CS, if investment is cheap and thus very effective / important compared to the procurement value or very expensive and thus very ineffective / unimportant compared to the procurement value. In–between there will be a range of intermediate costs, for which the PS proves to be preferable for the procurer. Thus, there seems to be an edge for both systems depending on the investment properties of the part procured and a procurer who has to purchase a great variety of different parts should engage in both systems at a time.

4.6 Empirical Evidence

This paper was developed in the context of broader research on supply networks in the automotive industry. One centerpiece is a series of deep case interviews with suppliers and car manufacturers on their strategic supply and purchasing behavior.[23] The authors interviewed 15 suppliers and three car manufacturers on issues concerning the parts supplied, the organization of purchasing, the supply strategies, the information about other players in the market, the contractual arrangements, and the competitive situation. We draw on the empirical evidence of this case study to comment on the assumptions and results of the theoretical model.

In the automotive industry, the structure of one incumbent and multiple entrant suppliers is very common. The incumbency advantage due to idiosyncratic process knowledge is sizeable. A switch from one supplier to another is legally possible without complications, as the manufacturer mostly owns the tools to produce the parts. However, in practice switching is very expensive, as the tools come together with very specific process knowledge, that cannot easily be replicated by a new supplier. Thus, the bigger part of the incumbency advantage is constituted by a special type of switching cost. The level of switching cost is endogenous, and can be affected by the incumbent. Furthermore the case study suggests, that modeling infinite repetition of a stage game is adequate. The manufacturers as well as the suppliers are long–lived and the possible interactions

[23] See Mueller, Stahl, and Wachtler (2006).

should not induce end game behavior. In the same sense also the commitment to a certain purchasing strategy (i.e. to a certain procurement system) is long-term. The purchasing strategy only becomes effective if the manufacturer can build up reputation for a certain strategy. Thus, we model long–term strategies concerning the choice of a purchasing system. By contrast, the case study shows evidence for a widely opportunistic behavior of all players in the short–run. Consequently we set up the model without ex–ante commitment on decisions within the stages. Usually, supply contracts become binding only when the first part has been delivered, long after sizable investments in development, capacity, and idiosyncratic tools have been made.

4.7 Conclusion

The presented model tries to make a contribution to the literature by modeling a procurement process while adopting features from observed procurement processes in reality. It features infinite repetition, relationship–specific investment with the associated hold–up problem, asymmetric information, and short–term opportunistic behavior of the procurer.

We analyze the main trade–off for the procurer between exploiting competitive pressure and creating investment incentives for the incumbent. The two proposed procurement systems, resembling the Western and the Asian procurement procedures, differ in how the procurer uses competition to exercise direct market pressure on his incumbent supplier. By choosing the PS, the procurer weakens up his ability to use direct market pressure, but he maintains the ability to exercise indirect market pressure by threatening to replace the incumbent for an entrant.

We show that the optimal procurement depends on characteristics of the part procured: The procurer will prefer the CS, if investment is cheap thus very effective / important compared to the procurement value or very expensive thus very ineffective / unimportant compared to the procurement value and in–between there will be a range of intermediate investment cost, for which the PS proves to be preferable for the procurer. Concerning direct versus indirect market pressure we can see, that depending on the investment characteristics of the part, indirect pressure may serve the procurer better. Furthermore, we find that the frequency of observed market interactions, like public tender auctions, might be a bad indicator for the total pressure a supplier experiences.

Appendix A

Appendix to Chapter 2

A.1 Proofs of Section 2.2

Proof of Lemma 2.1

(i) In footnote 6 we show that decision function $d_1^*(\theta_1, \ldots, \theta_n) = \frac{1}{n} \sum_i \theta_i$ maximizes the sum of utilities. Since welfare is a positive linear transformation thereof (see footnote 7), it possesses the same maximizers.

(ii) The optimal decision function that uses only the median information maximizes $-n\mathbf{E}[(d - d_1^*(\cdot))^2 | \theta_{m:n}]$. The maximizer is directly obtained from FOC $-2n\mathbf{E}[d - d_1^*(\cdot)|\theta_{m:n}] = -2n(d - \frac{1}{n}\theta_{m:n} - \frac{1}{n}\frac{n-1}{2}\mathbf{E}[\theta_i|\theta_i < \theta_{m:n}] - \frac{1}{n}\frac{n-1}{2}\mathbf{E}[\theta_i|\theta_i > \theta_{m:n}]) \stackrel{!}{=} 0$.

(iii) The optimal decision function using no information maximizes $-n\mathbf{E}[(d - d_1^*(\cdot))^2]$. The maximizer is directly obtained from FOC $-2n\mathbf{E}[d - d_1^*(\cdot)] = -2n(d - 0) \stackrel{!}{=} 0$.

Proof of Lemma 2.2

(i) The definition of welfare, (2.1), and Lemma 2.1 (i) imply

$$\mathcal{W}(d_1^*) = -n\mathbf{E}[(d_1^*(\cdot) - d_1^*(\cdot))^2] = 0.$$

(ii) The definition of welfare, (2.1), and Lemma 2.1 (ii) imply

$$\begin{aligned} \mathcal{W}(d_2^*) &= -n\mathbf{E}[(d_2^*(\theta_1, \ldots, \theta_n) - d_1^*(\theta_1, \ldots, \theta_n))^2] \\ &= -n\mathbf{E}[(\delta_2^*(\theta_{m:n}) - d_1^*(\theta_1, \ldots, \theta_n))^2] \end{aligned}$$

$$= -n\mathbf{E}[\delta_2^*(\theta_{m:n})^2] + 2n\mathbf{E}[\delta_2^*(\theta_{m:n})d_1^*(\theta_1,\ldots,\theta_n)] \\ -n\mathbf{E}[d_1^*(\theta_1,\ldots,\theta_n)^2].$$

Independence of signals implies $n\mathbf{E}[d_1^*(\theta_1,\ldots,\theta_n)^2] = n\mathbf{E}[(\frac{1}{n}\sum_i \theta_i)^2] = \mathbf{E}[\theta_i^2] = \sigma^2$. Hence, to prove the result it remains only to show that $\mathbf{E}[\delta_2^*(\theta_{m:n})d_1^*(\theta_1,\ldots,\theta_n)] = \mathbf{E}[\delta_2^*(\theta_{m:n})^2]$.

$$\begin{aligned}
&\mathbf{E}[\delta_2^*(\theta_{m:n})d_1^*(\theta_1,\ldots,\theta_n)] \\
= &\int_\Theta \frac{n!}{(m-1)!^2}\delta_2^*(\theta_m) \\
&\quad \int_{\substack{\theta_1,\ldots,\theta_{m-1}\le\theta_m \\ \theta_m\le\theta_{m+1},\ldots,\theta_n}} \left(\frac{1}{n}\sum_{i=1}^n \theta_i\right)\prod_{j\neq m}\mathrm{d}\Phi(\theta_j) \\
&\mathrm{d}\Phi(\theta_m) \\
= &\int_\Theta \frac{n!}{(m-1)!^2}\delta_2^*(\theta_m)\Phi(\theta_m)^{m-2}(1-\Phi(\theta_m))^{m-2} \\
&\quad \int_{\substack{\theta_1\le\theta_m \\ \theta_m\le\theta_n}} \left(\frac{1}{n}\theta_m + \frac{1}{n}(m-1)(\theta_1+\theta_n)\right)\prod_{j=1,n}\mathrm{d}\Phi(\theta_j) \\
&\mathrm{d}\Phi(\theta_m) \\
= &\int_\Theta \frac{n!}{(m-1)!^2}\delta_2^*(\theta_m)\Phi(\theta_m)^{m-1}(1-\Phi(\theta_m))^{m-1} \\
&\quad \left(\frac{1}{n}\theta_m + \frac{1}{n}(m-1)(\mathbf{E}[\theta_1|\theta_1<\theta_m] + \mathbf{E}[\theta_n|\theta_n>\theta_m])\right) \\
&\mathrm{d}\Phi(\theta_m) \\
= &\int_\Theta \delta_2^*(\theta_m)^2\frac{n!}{(m-1)!^2}\Phi(\theta_m)^{m-1}(1-\Phi(\theta_m))^{m-1}\mathrm{d}\Phi(\theta_m) \\
= &\mathbf{E}[\delta_2^*(\theta_{m:n})^2]
\end{aligned}$$

(iii) The definition of welfare, (2.1), and Lemma 2.1 (iii) imply

$$\mathcal{W}(d_3^*) = -n\mathbf{E}[(0-d_1^*(\cdot))^2] = -\mathbf{E}[\theta_i^2] = -\sigma^2.$$

A.2 Proofs of Section 2.3

Proof of Proposition 2.1

(i) The proof proceeds in two steps. First, we derive all equilibria of mean mechanism $\Gamma = ([\underline{v},\overline{v}], x_1)$, then we show that there is a unique symmetric equilibrium.

Notation

In this proof we denote agent i's ad interim expected utility if he has signal θ, votes for $v \in [\underline{v}, \overline{v}]$, and all other agents vote according to voting rule $v_j(\theta_j)$ by

$$\widetilde{U}(\theta, v) := \mathbf{E}\Big[-\Big(\frac{1}{n}v + \frac{1}{n}\sum_{j\neq i} v_j(\theta_j) - \theta_i^*(\theta_i, \theta_{-i})\Big)^2 \Big| \theta_i = \theta\Big].$$

Step 1: Derivation of all equilibria of $\Gamma = ([\underline{v}, \overline{v}], x_1)$

Since $\widetilde{U}(\theta, v)$ is strictly concave in v, local incentive compatibility is necessary and sufficient for global incentive compatibility.

If $v < \overline{v}$, the upward incentive compatibility constraint

$$\widetilde{U}(\theta, v) \geq \widetilde{U}(\theta, v') \quad \Leftrightarrow \quad 0 \geq \frac{\widetilde{U}(\theta, v') - \widetilde{U}(\theta, v)}{v' - v}$$

must hold for any $v' \in (v, \overline{v}]$. Taking limit $v' \to v$, the constraint becomes

$$\begin{aligned} 0 &\geq -2\frac{1}{n}\Big(\frac{1}{n}v + \frac{1}{n}\sum_{j\neq i}\mathbf{E}[v_j(\theta_j)] - (1-\alpha)\theta\Big) \\ \Leftrightarrow \quad v &\geq n(1-\alpha)\theta - \sum_{j\neq i}\mathbf{E}[v_j(\theta_j)]. \end{aligned}$$

Analogously, if $\underline{v} < v$, the downward incentive compatibility constraint

$$v \leq n(1-\alpha)\theta - \sum_{j\neq i}\mathbf{E}[v_j(\theta_j)]$$

must hold. Let $c_i := \sum_{j\neq i}\mathbf{E}[v_j(\theta_j)]$. Both incentive compatibility constraints are satisfied at the same time if and only if

$$v_i(\theta) = \begin{cases} \overline{v} & \text{if } \overline{v} < n(1-\alpha)\theta - c_i \\ n(1-\alpha)\theta - c_i & \text{if } \underline{v} \leq n(1-\alpha)\theta - c_i \leq \overline{v} \\ \underline{v} & \text{if } n(1-\alpha)\theta - c_i < \underline{v} \end{cases}.$$

Step 2: The unique symmetric equilibrium of $\Gamma = ([\underline{v}, \overline{v}], x_1)$

Symmetry allows us to omit indices. Parameter c is in the symmetric case implicitly specified by $c = (n-1)\mathbf{E}[v(\theta_i)]$. Using this we obtain

$$v(\theta) := \begin{cases} \delta_1(\overline{\tau}) - c & \text{if } \overline{\tau} < \theta \\ \delta_1(\theta) - c & \text{if } \underline{\tau} \leq \theta \leq \overline{\tau} \\ \delta_1(\underline{\tau}) - c & \text{if } \theta < \underline{\tau} \end{cases}$$

with $\underline{\tau} := \delta_1^{-1}(\underline{v} + c)$ and $\overline{\tau} := \delta_1^{-1}(\overline{v} + c)$. Since $\mathbf{E}[v(\theta_i)]$ is weakly decreasing in c, the solution to $c = (n-1)\mathbf{E}[v(\theta_i)]$ is unique.

(ii) Now we derive the welfare level attained in the unique symmetric equilibrium of $\Gamma = ([\underline{v}, \overline{v}], x_1)$:

$$\mathcal{W}(d_1) = -n\mathbf{E}[(d_1(\cdot) - d_1^*(\cdot))^2] = -n\mathbf{E}[(\frac{1}{n}\sum_i (v(\theta_i) - \theta_i))^2]$$

Using independence of signals this becomes

$$\ldots = -\mathbf{E}[(v(\theta_i) - \theta_i)^2].$$

Proof of Proposition 2.2

The interval $[\delta_1(\underline{\theta}), \delta_1(\overline{\theta})]$ is the smallest set of admissible votes for which each agent is able to pick his preferred vote such that voting is effectively not restricted. In the following we show that it is beneficial to choose at least a slightly smaller interval.

Consider $V = [\delta_1(\underline{\theta} + \underline{\epsilon}), \delta_1(\overline{\theta} - \overline{\epsilon})]$ with $\underline{\epsilon}, \overline{\epsilon} \geq 0$ but small. For $\underline{\epsilon} = \overline{\epsilon} = 0$ we have $c = 0$, but for most other combinations of $\overline{\epsilon}$ and $\underline{\epsilon}$ this is not true. Suppose, however, that it was true for the time being. Then from Proposition 2.1 (ii) we obtain

$$\begin{aligned} \mathcal{W}(d_1) &= -\int_{\underline{\theta}}^{\underline{\theta}+\underline{\epsilon}} (\delta_1(\underline{\theta} + \underline{\epsilon}) - \theta)^2 \, \mathrm{d}\Phi(\theta) - \int_{\underline{\theta}+\underline{\epsilon}}^{\overline{\theta}-\overline{\epsilon}} (\delta_1(\theta) - \theta)^2 \, \mathrm{d}\Phi(\theta) \\ &\quad - \int_{\overline{\theta}-\overline{\epsilon}}^{\overline{\theta}} \Big(\delta_1(\overline{\theta} - \overline{\epsilon}) - \theta\Big)^2 \, \mathrm{d}\Phi(\theta) \end{aligned}$$

and applying Leibnitz's rule we get

$$\frac{\mathrm{d}\mathcal{W}}{\mathrm{d}\underline{\epsilon}}(d_1) = \int_{\underline{\theta}}^{\underline{\theta}+\underline{\epsilon}} 2n(1-\alpha)\,(\theta - \delta_1(\underline{\theta} + \underline{\epsilon}))\,\mathrm{d}\Phi(\theta) \tag{A.1}$$

and

$$\frac{\mathrm{d}\mathcal{W}}{\mathrm{d}\overline{\epsilon}}(d_1) = \int_{\overline{\theta}-\overline{\epsilon}}^{\overline{\theta}} 2n(1-\alpha)\Big(\delta_1(\overline{\theta} - \overline{\epsilon}) - \theta\Big)\,\mathrm{d}\Phi(\theta). \tag{A.2}$$

Since we have $\delta_1(\underline{\theta}) < \underline{\theta}$ and $\delta_1(\overline{\theta}) > \overline{\theta}$ if preferences are not common, both derivatives are strictly positive when $\underline{\epsilon}$ and $\overline{\epsilon}$ are sufficiently small (but strictly positive). Thus, forbidding extreme votes enhances welfare.

Recall that the preceding argument is only valid when $\underline{\epsilon}$ and $\overline{\epsilon}$ are such that $c = 0$. Although this is not generally true, we can always construct a particular restriction of the set of admissible votes for which it is. Consider the restriction specified by $\overline{\epsilon}$ and $\underline{\epsilon} = \gamma\overline{\epsilon}$. Since c is strictly increasing and continuous in $\underline{\epsilon}$, and strictly decreasing and continuous in $\overline{\epsilon}$, it is always possible to find a positive γ such that the marginal effects of $\overline{\epsilon}$ and of $\underline{\epsilon}$ on c just cancel out. In this case c stays marginally constant when $\overline{\epsilon}$ changes rendering derivatives (A.1) and (A.2) valid. This proves that there exists always a *particular* restriction of the set of admissible votes which is beneficial.

Proof of Proposition 2.3

We already know that it is optimal to have no (binding) restriction of V when preferences are common. For this case the condition in the Proposition becomes $\mathbf{E}[\theta_i|\theta_i > \overline{v}] = \overline{v}$ and possesses the solution $\overline{v} = \overline{\theta}$. Since $\overline{v} = \overline{\theta}$ imposes no binding restriction on the agents' voting behavior, the condition in the Proposition is consistent with the optimal restriction of votes. It remains to show that the condition is also valid for the case when preferences are not common. The argument proceeds in four steps.

Step 1: There is at least one interior local extremum
Symmetry of distribution and voting behavior imply $c = 0$ and

$$\mathcal{W}([-\overline{v},\overline{v}],x_1) = -2\left(\int_0^{\tau}(\delta_1(\theta_i)-\theta_i)^2\mathrm{d}\Phi(\theta_i) + \int_{\tau}^{\overline{\theta}}(\delta_1(\tau)-\theta_i)^2\mathrm{d}\Phi(\theta_i)\right)$$

with $\tau \in [0,\overline{\theta}]$ and $\overline{v} = \delta_1(\tau)$. Applying Leibnitz's rule we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mathcal{W}([-\overline{v},\overline{v}],x_1) = 2n(1-\alpha)\underbrace{\int_{\tau}^{\overline{\theta}}(\theta-\delta_1(\tau))\,\mathrm{d}\Phi(\theta)}_{:=\xi(\tau)}. \tag{A.3}$$

Since $\xi(0) > 0$, $\xi(\tau) < 0$ for τ close to $\overline{\theta}$, and since ξ is continuous, the first–order condition possesses at least one interior solution by the Intermediate Value Theorem.

Step 2: There is at most one interior local extremum
Sufficient for the existence of at most one interior solution to the first–order condition is that $\xi(\tau)$ changes sign only once on $(0,\overline{\theta})$. We prove that $\xi(\tau)$ has at most one zero by showing that $\xi(\tau) < 0$ for τ close to $\overline{\theta}$, and that ξ is first decreasing and then increasing. We already know from step 1 that the first part of this statement is true, now we show that this is also the case for the second part.

Applying Leibnitz's rule we obtain

$$\begin{aligned}\xi'(\tau) &= -(1-n(1-\alpha))\tau\phi(\tau) - n(1-\alpha)(1-\Phi(\tau)) \\ &= (1-\Phi(\tau))\left[(n(1-\alpha)-1)\tau\frac{\phi(\tau)}{1-\Phi(\tau)} - n(1-\alpha)\right].\end{aligned}$$

The term in front of the bracketed expression is always positive for $\tau \in (0, \overline{\theta})$, the term inside the bracketed expression is strictly increasing in τ by the increasing hazard rate property. Furthermore, $\xi'(0) < 0$ and $\xi'(\overline{\theta}) > 0$. This implies that $\xi(\tau)$ is first decreasing and then increasing.

Step 3: The unique interior local extremum is a global maximum
We already know that the first–order condition changes sign only once on $(0, \overline{\theta})$. If this change is from positive to negative, the unique interior local extremum is a global maximum. Since $\xi(0) > 0$ and $\xi(\tau) < 0$ for τ close to $\overline{\theta}$, this is the case.

Step 4: The condition characterizing the optimum
From (A.3) and the reasoning in steps 1, 2 and 3 we obtain that the global maximum is characterized by $\xi(\tau) = 0$. Or, equivalently, by $\mathbf{E}[\theta|\theta > \tau] = \delta_1(\tau) \Leftrightarrow \mathbf{E}[\theta|\theta > \tau] = \overline{v}$.

Proof of Proposition 2.4

We derive the optimal welfare level attainable by mechanisms belonging to a subclass of mean mechanisms. This welfare level serves as a lower bound on what the generally optimal mean mechanism achieves. The proof proceeds in three steps. In the first two steps we derive the unique symmetric equilibrium of mechanism $\Gamma = (\{-\overline{v}, \overline{v}\}, x_1)$ and the welfare level attained by this equilibrium. In the third step we determine the optimal $\overline{v}$ and compute the respective welfare level.

Notation
For agent i's ad interim expected utility if he has signal θ, votes for $v \in [\underline{v}, \overline{v}]$, and all other agents vote according to voting rule $v_j(\theta_j)$ we use the same notation as in the proof of Proposition 2.1:

$$\widetilde{U}(\theta, v) := \mathbf{E}\Big[-\Big(\frac{1}{n}v + \frac{1}{n}\sum_{j\neq i} v_j(\theta_j) - \theta_i^*(\theta_i, \theta_{-i})\Big)^2 \Big| \theta_i = \theta\Big].$$

Step 1: The symmetric equilibrium of mechanism $\Gamma = (\{-\overline{v}, \overline{v}\}, x_1)$
Incentive compatibility requires that an agent votes for $\overline{v}$ if $\widetilde{U}(\theta, \overline{v}) > \widetilde{U}(\theta, -\overline{v})$. I.e., if

$$\begin{aligned}
& \mathbf{E}_{-i}\Big[-\Big(\Big(\frac{1}{n}\overline{v} - (1-\alpha)\theta\Big) + \sum_{j\neq i}\Big(\frac{1}{n}v_j(\theta_j) - \frac{\alpha}{n-1}\theta_j\Big)\Big)^2\Big] \\
& > \mathbf{E}_{-i}\Big[-\Big(\Big(-\frac{1}{n}\overline{v} - (1-\alpha)\theta\Big) + \sum_{j\neq i}\Big(\frac{1}{n}v_j(\theta_j) - \frac{\alpha}{n-1}\theta_j\Big)\Big)^2\Big] \\
\Leftrightarrow\ & -\Big(\frac{1}{n}\overline{v} - (1-\alpha)\theta\Big)^2 - 2\frac{1}{n}\overline{v}\sum_{j\neq i}\frac{1}{n}\mathbf{E}[v_j(\theta_j)] \\
& > -\Big(-\frac{1}{n}\overline{v} - (1-\alpha)\theta\Big)^2 + 2\frac{1}{n}\overline{v}\sum_{j\neq i}\frac{1}{n}\mathbf{E}[v_j(\theta_j)] \\
\Leftrightarrow\ & \theta > \frac{1}{n(1-\alpha)}\sum_{j\neq i}\mathbf{E}[v_j(\theta_j)].
\end{aligned}$$

In the symmetric equilibrium this condition becomes

$$\theta > \frac{(n-1)}{n(1-\alpha)}\mathbf{E}[v(\theta_j)].$$

Suppose now $\mathbf{E}[v(\theta_j)]$ is positive. Then agent i votes only for $\overline{v}$ if his signal is strictly positive, i.e. with a probability smaller than 1/2. As consequence, his expected vote is negative. This, however, contradicts symmetry. Since an analogous statement goes through if $\mathbf{E}[v(\theta_j)]$ is negative, $\mathbf{E}[v(\theta_j)] = 0$ is true in any symmetric equilibrium.

To sum up, except for the behavior of an agent with type zero, equilibrium is unique. Each agent votes for $\overline{v}$ if his signal is positive and for $\underline{v}$ if it is negative.

Step 2: Welfare attained in the symmetric equilibrium
Using independence of signals again, we obtain

$$\begin{aligned}
\mathcal{W}(\{-\overline{v}, \overline{v}\}, x_1) &= -\mathbf{E}[(v(\theta_i) - \theta_i)^2] \\
&= -\Big(\int_{\underline{\theta}}^{0}(-\overline{v} - \theta_i)^2 \mathrm{d}\Phi(\theta_i) + \int_0^{\overline{\theta}}(\overline{v} - \theta_i)^2 \mathrm{d}\Phi(\theta_i)\Big).
\end{aligned}$$

Using symmetry, this becomes

$$\ldots = -\Big(2\int_0^{\overline{\theta}}(\overline{v} - \theta_i)^2 \mathrm{d}\Phi(\theta_i)\Big) = -\mathbf{E}[(\overline{v} - \theta_i)^2 | \theta_i > 0].$$

Step 3: The optimal $\overline{v}$ and the optimal welfare level
Since

$$\frac{\mathrm{d}^2}{\mathrm{d}\overline{v}^2}\mathcal{W}(\{-\overline{v}, \overline{v}\}, x_1) = -2 < 0,$$

utility is globally concave in $\overline{v}$ such that there is a unique global maximum. The first–order condition

$$\frac{\mathrm{d}}{\mathrm{d}\overline{v}}\mathcal{W}(\{-\overline{v},\overline{v}\},x_1) = -\Big(4\int_0^{\overline{\theta}}(\overline{v}-\theta_i)\mathrm{d}\Phi(\theta_i)\Big) \stackrel{!}{=} 0$$

implies $\overline{v}^* = \mathbf{E}[\theta_i|\theta_i > 0]$. Thus,

$$\mathcal{W}(\{-\overline{v}^*,\overline{v}^*\},x_1) = -\left(\mathbf{E}[\theta_i^2|\theta_i > 0] - \mathbf{E}[\theta_i|\theta_i > 0]^2\right).$$

Since we have $\mathbf{E}[\theta_i^2|\theta_i > 0] = \mathbf{E}[\theta_i^2] = \sigma^2$ by symmetry of the distribution, welfare can be written in the following way:

$$\mathcal{W}(\{-\overline{v}^*,\overline{v}^*\},x_1) = -\sigma^2 + \mathbf{E}[\theta_i|\theta_i > 0]^2.$$

Proof of Proposition 2.5

(i) **Step 1: The stated behavior specifies an equilibrium**

Since, by construction, each agent is required to vote for his preferred decision from set V conditional on what he can infer from the other agents' equilibrium behavior, no agent has an incentive to deviate.

Step 2: This is the only equilibrium in which the agents' voting behavior is strictly monotonic on $[\delta_2^{-1}(\underline{v}),\delta_2^{-1}(\overline{v})]$

For agents with signals in $[\delta_2^{-1}(\underline{v}),\delta_2^{-1}(\overline{v})]$, the same reasoning as in Lemma 2 and Lemma 3 in Grüner and Kiel (2004) applies. If such an agent, say agent i, votes for some decision in $[\underline{v},\overline{v}]$ *and* it happens that he is pivotal, symmetry of equilibrium and the monotonicity property imply that he can infer that half of the agents have smaller signals and half of the agents have higher signals. Note that this is also true for votes which are chosen with strictly positive probability in equilibrium. Such votes do occur because of the pooling at the upper and the lower end of the distribution. By the reasoning in Grüner and Kiel (2004) the stated monotonicity property can only be satisfied (locally) if agent i votes for $\delta_2(\theta_i)$. Hence, there is no other possibility of getting the stated monotonicity property as by the stated voting behavior.

(ii) **Welfare attained in this equilibrium**

We use (2.1) to compute the welfare level attained by the equilibrium in (i):

$$\begin{aligned}\mathcal{W}([\underline{v},\overline{v}],x_2) &= -n\mathbf{E}[(d_2(\theta_1,\ldots,\theta_n)-d_1^*(\theta_1,\ldots,\theta_n))^2] \\ &= -n\mathbf{E}[(v(\theta_{m:n})-d_1^*(\theta_1,\ldots,\theta_n))^2] \\ &= -n\mathbf{E}[v(\theta_{m:n})^2]+2n\mathbf{E}[v(\theta_{m:n})d_1^*(\theta_1,\ldots,\theta_n)] \\ &\quad -n\mathbf{E}[d_1^*(\theta_1,\ldots,\theta_n)^2].\end{aligned}$$

Since we have $\mathbf{E}[d_1^*(\theta_1,\ldots,\theta_n)^2]=\mathbf{E}[(\frac{1}{n}\sum_i\theta_i)^2]=\frac{1}{n}\mathbf{E}[\theta_i^2]$ we obtain

$$\begin{aligned}\mathcal{W}([\underline{v},\overline{v}],x_2) &= \mathcal{B}_2-n\mathbf{E}[\delta_2^*(\theta_{m:n})^2]-n\mathbf{E}[v(\theta_{m:n})^2] \\ &\quad +2n\mathbf{E}[v(\theta_{m:n})d_1^*(\theta_1,\ldots,\theta_n)].\end{aligned}$$

It remains only to show that $\mathbf{E}[v(\theta_{m:n})d_1^*(\theta_1,\ldots,\theta_n)]=\mathbf{E}[v(\theta_{m:n})\delta_2^*(\theta_{m:n})]$.

$$\begin{aligned}&\mathbf{E}[v(\theta_{m:n})d_1^*(\theta_1,\ldots,\theta_n)] \\ =\;& \int_\Theta\frac{n!}{(m-1)!^2}v(\theta_m)\int_{\substack{\theta_1,\ldots,\theta_{m-1}\le\theta_m\\ \theta_m\le\theta_{m+1},\ldots,\theta_n}}\left(\frac{1}{n}\sum_{i=1}^n\theta_i\right)\prod_{j\neq m}\mathrm{d}\Phi(\theta_j)\mathrm{d}\Phi(\theta_m) \\ =\;& \int_\Theta\frac{n!}{(m-1)!^2}v(\theta_m)\Phi(\theta_m)^{m-2}(1-\Phi(\theta_m))^{m-2} \\ &\qquad\int_{\substack{\theta_1\le\theta_m\\ \theta_m\le\theta_n}}\left(\frac{1}{n}\theta_m+\frac{1}{n}(m-1)(\theta_1+\theta_n)\right)\prod_{j=1,n}\mathrm{d}\Phi(\theta_j) \\ &\mathrm{d}\Phi(\theta_m) \\ =\;& \int_\Theta\frac{n!}{(m-1)!^2}v(\theta_m)\Phi(\theta_m)^{m-1}(1-\Phi(\theta_m))^{m-1} \\ &\qquad\left(\frac{1}{n}\theta_m+\frac{1}{n}(m-1)(\mathbf{E}[\theta_1|\theta_1<\theta_m]+\mathbf{E}[\theta_n|\theta_n>\theta_m])\right) \\ &\mathrm{d}\Phi(\theta_m) \\ =\;& \int_\Theta v(\theta_m)\delta_2^*(\theta_m)\frac{n!}{(m-1)!^2}\Phi(\theta_m)^{m-1}(1-\Phi(\theta_m))^{m-1}\mathrm{d}\Phi(\theta_m) \\ =\;& \mathbf{E}[v(\theta_{m:n})\delta_2^*(\theta_{m:n})]\end{aligned}$$

Proof of Proposition 2.6

We show in this proof that the seller can improve welfare by forbidding large votes. From Proposition 2.5 we know

$$\mathcal{W}([\underline{v},\overline{v}],x_2)=\mathcal{B}_2-n\mathbf{E}[(v(\theta_{m:n})-\delta_2^*(\theta_{m:n}))^2].$$

Define $\underline{v} = \delta_2(\underline{\theta})$ and $\overline{v} = \delta_2(\tau)$. Then we get

$$\ldots = \mathcal{B}_2 - n\Big[\int_{\underline{\theta}}^{\tau} (\delta_2(\theta) - \delta_2^*(\theta))^2 \mathrm{d}\Phi_{m:n}(\theta) + \int_{\tau}^{\overline{\theta}} (\delta_2(\tau) - \delta_2^*(\theta))^2 \mathrm{d}\Phi_{m:n}(\theta)\Big].$$

Applying Leibnitz's rule we get

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mathcal{W}([\underline{v},\overline{v}], x_2) = -2n\int_{\tau}^{\overline{\theta}} \delta_2'(\tau)(\delta_2(\tau) - \delta_2^*(\theta))\mathrm{d}\Phi_{m:n}(\theta).$$

Since δ_2 and δ_2^* are continuous and since $\delta_2' > 0$, it suffices to show that $\delta_2(\overline{\theta}) > \delta_2^*(\overline{\theta})$ in order to prove that it is strictly beneficial to decrease τ below $\overline{\theta}$:

$$\begin{aligned}\delta_2(\overline{\theta}) - \delta_2^*(\overline{\theta}) \;&=\; (1-\alpha)\overline{\theta} + \frac{\alpha}{2}(\mathbf{E}[\theta_i|\theta_i \le \overline{\theta}] + \mathbf{E}[\theta_i|\theta_i \ge \overline{\theta}]) \\ &\quad -\frac{1}{n}\overline{\theta} - \frac{1}{2}\frac{n-1}{n}(\mathbf{E}[\theta_i|\theta_i \le \overline{\theta}] + \mathbf{E}[\theta_i|\theta_i \ge \overline{\theta}]).\end{aligned}$$

Using $\mathbf{E}[\theta_i|\theta_i \le \overline{\theta}] = \mathbf{E}[\theta_i] = 0$ and $\mathbf{E}[\theta_i|\theta_i \ge \overline{\theta}] = \overline{\theta}$, this becomes

$$\ldots \;=\; \frac{1}{2}\left(\frac{n-1}{n} - \alpha\right)\overline{\theta}.$$

This expression is strictly positive if preferences are not common, i.e. if $\alpha < \overline{\alpha}_n = \frac{n-1}{n}$. This proves that for all but common preferences it is optimal to restrict voting.

A.3 Proofs of Section 2.4

Proof of Proposition 2.7

The proof proceeds in four steps. First, we derive the welfare level attained by the unrestricted mean mechanism. Then, we show that the welfare level attained by the unrestricted median mechanism is increasing in the degree of interdependence, which allows us to derive a tractable lower bound on this level in the third step. Finally, we show that for any degree of interdependence except for common preferences the lower bound on median welfare is higher than mean welfare when the number of agents is sufficiently large.

Step 1: Welfare attained by the unrestricted mean mechanism

From Proposition 2.1 we obtain

$$\mathcal{W}(\mathbf{R}, x_1) = -\mathbf{E}[(n(1-\alpha)\theta_i - \theta_i)^2] = -(n(1-\alpha)-1)^2\sigma^2.$$

Note that if $(n(1-\alpha)-1)^2 > 1$, i.e. if $n > 2/(1-\alpha)$, the unrestricted mean mechanism performs worse than the best uninformed mechanism.

Step 2: Welfare attained by the unrestricted median mechanism is (weakly) increasing in α

By Proposition 2.5 we have

$$\mathcal{W}(\mathbf{R}, x_2) = \mathcal{B}_2 - n\mathbf{E}[(\delta_2(\theta_{m:n}) - \delta_2^*(\theta_{m:n}))^2].$$

Taking the derivative with respect to α we get

$$\begin{aligned} & \frac{\mathrm{d}}{\mathrm{d}\alpha}\mathcal{W}(\mathbf{R}, x_2) \\ = \quad & -2n\mathbf{E}\left[\frac{\mathrm{d}\delta_2(\theta_{m:n})}{\mathrm{d}\alpha}(\delta_2(\theta_{m:n}) - \delta_2^*(\theta_{m:n}))\right] \\ = \quad & 2n\left(\frac{n-1}{n} - \alpha\right)\mathbf{E}[(\theta_{m:n} - \frac{1}{2}(\mathbf{E}[\theta_i|\theta_i \leq \theta_{m:n}] + \mathbf{E}[\theta_i|\theta_i \geq \theta_{m:n}]))^2]. \end{aligned}$$

This expression is strictly positive if preferences are not common and zero if they are.

Step 3: Lower bound on welfare attained by the unrestricted median mechanism

Using the definition of welfare, (2.1), directly, we obtain a second way of describing welfare attained by the unrestricted median mechanism:

$$\mathcal{W}(\mathbf{R}, x_2) = -n\mathbf{E}[(\delta_2(\theta_{m:n}) - d_1^*(\cdot))^2].$$

We already know from step 2 that welfare attained by the unrestricted median mechanism is (weakly) increasing in α. Hence, welfare is for any degree of interdependence at least as high as welfare for $\alpha = 0$, i.e. as welfare in the case in which $\delta_2(\theta) = \theta$. Thus,

$$\mathcal{W}(\mathbf{R}, x_2) \geq -n\mathbf{E}[(\theta_{m:n} - \frac{1}{n}\sum_i \theta_i)^2] = -\mathbf{E}[\theta_i^2] + 2\mathbf{E}[\theta_{m:n}\sum_i \theta_i] - n\mathbf{E}[\theta_{m:n}^2].$$

Since the median type and the mean type are positively correlated, we have $\mathbf{E}[\theta_{m:n}\sum_i \theta_i] > 0$ such that we can write

$$\mathcal{W}(\mathbf{R}, x_2) > -\mathbf{E}[\theta_i^2] - n\mathbf{E}[\theta_{m:n}^2].$$

By a standard result in order statistics (see, e.g., David and Nagaraja (2003) page 69) we have $\mathbf{E}[\theta_{m:n}^2] \leq \sigma^2$. Hence,

$$\mathcal{W}(\mathbf{R}, x_2) > -(n+1)\sigma^2.$$

Step 4: Comparison of the lower bound on welfare attained by the unrestricted median mechanism with welfare attained by the unrestricted mean mechanism

Sufficient for the unrestricted median mechanism to be preferable over the unrestricted mean mechanism is

$$\begin{aligned} -(n+1)\sigma^2 \geq -(n(1-\alpha)-1)^2\sigma^2 \quad &\Leftrightarrow \quad (n+1) \leq (n(1-\alpha)-1)^2 \\ &\Leftrightarrow \quad n \geq \frac{1+2(1-\alpha)}{(1-\alpha)^2}. \end{aligned}$$

Since the right hand side is finite for any $\alpha < 1$, we obtain that the unrestricted median mechanism is preferable over the unrestricted mean mechanism when n is sufficiently large.

Proof of Proposition 2.8

The proof proceeds in four steps. First, we compute an analytical expression of the upper bound on welfare attainable by a median mechanism. Then, for private preferences, we derive a lower bound on welfare obtained by the optimally restricted mean mechanism and we show in the third step that this lower bound is even higher when preferences are not private. Finally, we show that the lower bound on mean welfare lies above the upper bound on median welfare for any degree of interdependence.

Without loss of generality we can consider a uniform distribution on a normalized support of length 2, i.e. $\Theta = [-1, 1]$. In this case we have $\phi(\theta) = 1/2$, $\Phi(\theta) = (1+\theta)/2$ and

$$\begin{aligned} \phi_{m:n}(\theta) &= \frac{n!}{(m-1)!^2}\frac{1}{2^n}(1+\theta)^{m-1}(2-(1+\theta))^{m-1} \\ &= \frac{n!}{(m-1)!^2}\frac{1}{2^n}(1-\theta^2)^{m-1}. \end{aligned} \tag{A.4}$$

Step 1: Computation of an analytical expression of the upper bound on welfare attainable by a median mechanism

For uniformly distributed signals we have enough structure to obtain a tractable expression of bound $\mathcal{B}_2$ (for a definition of this bound see Lemma 2.2).

The optimal decision conditional on using only the median information is

$$\delta_2^*(\theta) = \frac{1}{n}\theta + \frac{1}{2}\frac{n-1}{n}\left(\frac{\underline{\theta}+\theta}{2} + \frac{\theta+\overline{\theta}}{2}\right) = \frac{1}{2}\frac{n+1}{n}\theta.$$

Using this we get

$$\mathcal{B}_2 = -\sigma^2 + n\mathbf{E}[\delta_2^*(\theta_{m:n})^2] = -\sigma^2 + \frac{1}{4}\frac{(n+1)^2}{n}\int_{-1}^{1}\theta^2\phi_{m:n}(\theta)\mathrm{d}\theta.$$

By using (A.4) and rearranging the constants this becomes

$$\ldots \quad = \quad -\sigma^2 + \frac{1}{4}\frac{n+1}{n+2}\frac{n+1}{n}\frac{(n+2)!}{m!^2 2^{n+2}}\int_{-1}^{1}(-\theta)(-2m\theta)(1-\theta^2)^{m-1}\mathrm{d}\theta,$$

and after applying partial integration we obtain

$$\ldots \quad = \quad -\sigma^2 + \frac{1}{4}\frac{n+1}{n+2}\frac{n+1}{n}\frac{(n+2)!}{m!^2 2^{n+2}}\int_{-1}^{1}(1-\theta^2)^{m}\mathrm{d}\theta.$$

Using a formula for the integral from a formulary and using properties of the Gamma–function $\widehat{\Gamma}$,[1] we can make the transformation

$$\int_{-1}^{1}(1-\theta^2)^m\mathrm{d}\theta = \frac{\widehat{\Gamma}(m+1)\sqrt{\pi}}{\widehat{\Gamma}(m+1+\frac{1}{2})} = \frac{m!\sqrt{\pi}}{\frac{(2(m+1))!\sqrt{\pi}}{(m+1)!2^{2(m+1)}}} = \frac{m!^2 2^{n+2}}{(n+2)!}$$

such that we finally obtain

$$\mathcal{B}_2 = -\sigma^2 + \frac{1}{4}\frac{n+1}{n+2}\frac{n+1}{n}.$$

Step 2: Welfare attained by mechanism $\Gamma = ([-1/2, 1/2], x_1)$ for private preferences

Although we can explicitly compute the optimal restriction of votes when signals are uniformly distributed, we derive welfare for the non–optimal set of admissible votes $V = [-1/2, 1/2]$. This welfare level serves as a lower bound on optimal welfare. For the considered set of admissible votes all agents with signals $\theta \in [-1/(2n), 1/(2n)]$ are not restricted in voting. From Proposition 2.1 (ii) we obtain

$$\begin{aligned}
\mathcal{W}([-1/2,1/2],x_1) &= -\mathbf{E}[(v(\theta_i)-\theta_i)^2] \\
&= -\sigma^2 + 2\int_0^1 v(\theta)(2\theta - v(\theta))\frac{1}{2}\mathrm{d}\theta \\
&= -\sigma^2 + \int_0^{\frac{1}{2n}} n\theta(2\theta - n\theta)\mathrm{d}\theta + \int_{\frac{1}{2n}}^1 \frac{1}{2}(2\theta - \frac{1}{2})\mathrm{d}\theta \\
&= -\sigma^2 + \frac{1}{4}\left(\frac{6n^2+2n-1}{6n^2}\right).
\end{aligned}$$

[1] In particular, $\widehat{\Gamma}(m+1) = n!$ and $\widehat{\Gamma}(m+1+\frac{1}{2}) = \frac{(2(m+1))!\sqrt{\pi}}{((m+1)/2)!2^{2(m+1)}}$. See (Bronstein, Semendjajew, Musiol, and Mühlig, 2001, p. 478).

Step 3: Welfare attained by a mechanism $\Gamma = ([-\underline{v}, \overline{v}], x_1)$ is increasing in the degree of interdependence

We have

$$\begin{aligned} & \mathcal{W}([-\overline{v}, \overline{v}], x_1) \\ = & -2\left(\int_0^{\delta_1^{-1}(\overline{v})} (\delta_1(\theta_i) - \theta_i)^2 \mathrm{d}\Phi(\theta_i) + \int_{\delta_1^{-1}(\overline{v})}^{\overline{\theta}} (\overline{v} - \theta_i)^2 \mathrm{d}\Phi(\theta_i) \right). \end{aligned}$$

Taking the derivative with respect to α and using that $\mathrm{d}\delta_1(\theta_i)/\mathrm{d}\alpha = -n\theta_i$, we obtain by applying Leibnitz's rule

$$\frac{\mathrm{d}\mathcal{W}}{\mathrm{d}\alpha}([-\overline{v}, \overline{v}], x_1) = -2 \int_0^{\delta_1^{-1}(\overline{v})} 2(-n\theta_i)(\delta_1(\theta_i) - \theta_i) \mathrm{d}\Phi(\theta_i) > 0.$$

Step 4: Comparison of the upper bound on median welfare with the lower bound on mean welfare

Note that the upper bound on median welfare computed in step 1 is independent of the degree of interdependence and that the lower bound on mean welfare is increasing by step 3. Hence, if the lower bound is higher than the upper bound for private preferences, then it is also higher for any other degree of interdependence. We conclude this proof by showing that even for private preferences the lower bound is higher:

$$\begin{aligned} \mathcal{W}([-1/2, 1/2], x_1) \overset{?}{>} \mathcal{B}_2 \;\; \Leftrightarrow & \;\; -\sigma^2 + \frac{1}{4}\left(\frac{6n^2 + 2n - 1}{6n^2}\right) \overset{?}{>} -\sigma^2 + \frac{1}{4}\frac{n+1}{n+2}\frac{n+1}{n} \\ \Leftrightarrow & \;\; (2n+1)(n-2) \overset{?}{>} 0 \end{aligned}$$

Since the last inequality is true for $n \geq 3$, we are done.

Proof of Proposition 2.9

If preferences are private, we have $\delta_2(\theta_{m:n}) = \theta_{m:n}$ and the level of welfare attained by the unrestricted median mechanism is

$$\begin{aligned} \mathcal{W}(\mathbf{R}, x_2) &= -n\mathbf{E}[(\theta_{m:n} - d_1^*(\cdot))^2] \\ &= -\mathbf{E}[(\sqrt{n}\theta_{m:n})^2] + 2\mathbf{E}[(\sqrt{n}\theta_{m:n})(\sqrt{n}\frac{1}{n}\sum_i \theta_i)] - \mathbf{E}[(\sqrt{n}\frac{1}{n}\sum_i \theta_i)^2]. \end{aligned}$$

Using a result stated, e.g., in Ferguson (1999), the asymptotic distribution of the median signal and of the average signal is jointly normal. Under our assumptions we obtain

$$\sqrt{n}\begin{pmatrix}\frac{1}{n}\sum_i \theta_i \\ \theta_{m:n}\end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\mathbf{E}[\theta_i^2] & \frac{\mathbf{E}[\theta_i|\theta_i>0]}{2\phi(0)} \\ \frac{\mathbf{E}[\theta_i|\theta_i>0]}{2\phi(0)} & \frac{1}{4\phi(0)^2}\end{pmatrix}).$$

Hence,

$$\begin{aligned}\lim_{n\to\infty}\mathcal{W}(\mathbf{R},x_2) &= -\frac{1}{4\phi(0)^2}+2\frac{\mathbf{E}[\theta_i|\theta_i>0]}{2\phi(0)}-\mathbf{E}[\theta_i^2] \\ &= -\sigma^2-\left(\frac{1}{2\phi(0)}\right)^2+2\left(\frac{1}{2\phi(0)}\right)\mathbf{E}[\theta_i|\theta_i>0]. \qquad (A.5)\end{aligned}$$

By Proposition 2.4 we have for any number of agents

$$\max_V \mathcal{W}(V,x_1) \geq \max_{\overline{v}} \mathcal{W}(\{-\overline{v},\overline{v}\},x_1) = -\sigma^2+\mathbf{E}[\theta_i|\theta_i>0]^2. \qquad (A.6)$$

Comparing (A.5) with (A.6) we obtain

$$\begin{aligned} & \max_{\overline{v}}\mathcal{W}(\{-\overline{v},\overline{v}\},x_1) \overset{?}{\geq} \lim_{n\to\infty}\mathcal{W}(\mathbf{R},x_2) \\ \Leftrightarrow\quad & -\sigma^2+\mathbf{E}[\theta_i|\theta_i>0]^2 \overset{?}{\geq} -\sigma^2-\left(\frac{1}{2\phi(0)}\right)^2+2\left(\frac{1}{2\phi(0)}\right)\mathbf{E}[\theta_i|\theta_i>0] \\ \Leftrightarrow\quad & \left(\mathbf{E}[\theta_i|\theta_i>0]-\frac{1}{2\phi(0)}\right)^2 \overset{?}{\geq} 0.\end{aligned}$$

Since this inequality is true, we obtain that the optimally restricted mean mechanism is asymptotically preferable.

Proof of Proposition 2.10

Step 1: $\delta_2(\theta_{\mathbf{Med}}) \neq 0$

Since $\mathbf{E}[\theta] = \frac{1}{2}\mathbf{E}[\theta_i|\theta_i < \theta_{\text{Med}}] + \frac{1}{2}\mathbf{E}[\theta_i|\theta_i > \theta_{\text{Med}}] \overset{!}{=} 0$, we obtain $\delta_2(\theta_{\text{Med}}) = (1-\alpha)\theta_{\text{Med}}$. This expression differs from zero for $\alpha < 1$.

Step 2: Welfare attained by the optimally restricted mean and the unrestricted median mechanism

The welfare level attained by the unrestricted median mechanism is

$$\mathcal{W}(\mathbf{R},x_2) = -n\mathbf{E}[(\delta_2(\theta_{m:n})-d_1^*(\cdot))^2].$$

Asymptotically, $\delta_1^*(\cdot)$ differs only on a set of measure zero (significantly) from zero and $\delta_2(\theta_{m:n})$ differs only on a set of measure zero (significantly) from $(1-\alpha)\theta_{\text{Med}}$. Hence, we get

$$\lim_{n\to\infty} \mathcal{W}(\mathbf{R}, x_2) = \lim_{n\to\infty} -n(1-\alpha)^2\theta_{m:n}^2 = -\infty.$$

Since the level of welfare attained by the optimally restricted mean mechanism is necessarily higher than the level of welfare attained by the best uninformed mechanism, i.e. higher than $-\sigma^2$, we obtain the result stated in the Proposition.

A.4 Implementable Decision Functions

Proposition A.1 (Implementable decision functions)
Let $U_i(\theta) := \mathbf{E}[-(d(\cdot)-\theta_i^*(\cdot))^2|\theta_i = \theta]$. *A decision function* $d : \Theta^n \to \mathbf{R}$ *is implementable if and only if*

(I1) $\mathbf{E}[d(\theta_1, \ldots, \theta_n)|\theta_i = \theta]$ *is weakly increasing in* θ,

(I2) $U_i'(\theta) = 2(1-\alpha)(\mathbf{E}[d(\theta_1, \ldots, \theta_n)|\theta_i = \theta] - (1-\alpha)\theta)$ *for all* θ *at which* $\mathbf{E}[d(\theta_1, \ldots, \theta_n)|\theta_i = \theta]$ *is continuous in* θ, *and*

(I3) $U_i(\theta)$ *is continuous.*

Proof. By a Revelation Principle we can restrict attention to *direct* mechanisms $\Gamma = (\Theta, d)$ and the case in which each agent reveals his type truthfully in equilibrium. Note that for direct mechanisms decision rule and decision function coincide.

Notation

To simplify notation within this proof we introduce a special notation for an agent's expected utility from a unilateral deviation. We denote by

$$U_i(\theta, \widehat{\theta}) := \mathbf{E}[-(d(\widehat{\theta}, \theta_{-i}) - \theta_i^*(\theta_i, \theta_{-i}))^2|\theta_i = \theta]$$

agent i's ad interim expected utility if he has signal θ, reveals signal $\widehat{\theta}$ and all other agents reveal their signals truthfully. Furthermore, we denote the partial derivatives of $U_i(\theta, \widehat{\theta})$ with respect to its first and to its second argument by $U_i^{[1]}(\theta, \widehat{\theta})$ and $U_i^{[2]}(\theta, \widehat{\theta})$, respectively, and we denote by $U_i(\theta) := U_i(\theta, \theta)$ agent i's ad interim expected utility from truth–telling.

Necessity: IC implies (I1), (I2) and (I3)

Step 1: (I1). $\mathbf{E}[d(\theta_1, \ldots, \theta_n)|\theta_i = \theta]$ **is weakly increasing in** θ**.**
IC implies $U_i(\theta, \theta) \geq U_i(\theta, \theta')$ and $U_i(\theta', \theta') \geq U_i(\theta', \theta)$. From this it follows

$$\begin{aligned}
& U_i(\theta, \theta) + U_i(\theta', \theta') \geq U_i(\theta, \theta') + U_i(\theta', \theta) \\
\Rightarrow\ & \mathbf{E}_{-i}[2d(\theta, \theta_{-i})\theta_i^*(\theta, \theta_{-i})] + \mathbf{E}_{-i}[2d(\theta', \theta_{-i})\theta_i^*(\theta', \theta_{-i})] \\
& \geq \mathbf{E}_{-i}[2d(\theta', \theta_{-i})\theta_i^*(\theta, \theta_{-i})] + \mathbf{E}_{-i}[2d(\theta, \theta_{-i})\theta_i^*(\theta', \theta_{-i})] \\
\Rightarrow\ & \mathbf{E}_{-i}[d(\theta, \theta_{-i})](\theta - \theta') \geq \mathbf{E}_{-i}[d(\theta', \theta_{-i})](\theta - \theta') \\
\Rightarrow\ & \mathbf{E}[d(\theta_i, \theta_{-i})|\theta_i = \theta](\theta - \theta') \geq \mathbf{E}[d(\theta_i, \theta_{-i})|\theta_i = \theta'](\theta - \theta').
\end{aligned}$$

This condition can only be satisfied if (I1) holds.

Step 2: (I2). $U_i'(\theta) = U_i^{[1]}(\theta, \theta)$ **for all points** θ **at which** $\mathbf{E}[d(\theta_1, \ldots, \theta_n)|\theta_i = \theta]$ **is continuous in** θ**.**
IC implies

$$\begin{aligned}
U_i(\theta) &\geq U_i(\theta, \theta') \\
&= \mathbf{E}_{-i}[-(d(\theta', \theta_{-i}) - \theta_i^*(\theta', \theta_{-i}) + (1-\alpha)(\theta' - \theta))^2] \\
&= U_i(\theta') - 2(1-\alpha)(\theta' - \theta)(\mathbf{E}[d(\theta_i, \theta_{-i})|\theta_i = \theta'] - (1-\alpha)\theta') \\
&\quad -(1-\alpha)^2(\theta' - \theta)^2. \qquad \text{(A.7)}
\end{aligned}$$

If $\theta' < \theta$, this becomes

$$\frac{U_i(\theta) - U_i(\theta')}{\theta - \theta'} \geq 2(1-\alpha)(\mathbf{E}[d(\theta_i, \theta_{-i})|\theta_i = \theta'] - (1-\alpha)\theta') - (1-\alpha)^2(\theta - \theta').$$

For $\theta' > \theta$ the inequality is reversed. Thus, by taking the limit $\theta' \to \theta$, we obtain for any θ at which $\mathbf{E}[d(\theta_i, \theta_{-i})|\theta_i = \theta]$ is continuous in θ

$$\begin{aligned}
& \lim_{\theta' \to \theta} \frac{U_i(\theta) - U_i(\theta')}{\theta - \theta'} = 2(1-\alpha)(\mathbf{E}[d(\theta_i, \theta_{-i})|\theta_i = \theta] - (1-\alpha)\theta) \\
\Rightarrow\ & U_i'(\theta) = U_i^{[1]}(\theta, \theta).
\end{aligned}$$

This is (I2).

Step 3: (I3). $U_i(\theta)$ **is continuous.**
By taking the limit $\theta \to \theta'$ in (A.7), we obtain $\lim_{\theta \to \theta'} U_i(\theta) \geq U_i(\theta')$. By changing the roles of θ and θ' and taking the limit again, we obtain $U_i(\theta') \geq \lim_{\theta \to \theta'} U_i(\theta)$. This establishes $\lim_{\theta \to \theta'} U_i(\theta) = U_i(\theta')$, i.e. continuity.

Sufficiency: (I1), (I2) and (I3) imply IC

Assume to the contrary that (I1), (I2) and (I3) hold, but that IC is violated. Then there exist θ and $\widehat{\theta}$ such that some agent i can profit from a unilateral deviation, i.e.

$$\begin{aligned} & U_i(\theta,\widehat{\theta}) - U_i(\theta,\theta) > 0 \\ \Rightarrow \quad & [U_i(\theta,\widehat{\theta}) - U_i(\widehat{\theta},\widehat{\theta})] + [U_i(\widehat{\theta},\widehat{\theta}) - U_i(\theta,\theta)] > 0. \end{aligned}$$

Since $U_i(\theta,\widehat{\theta})$ is partially differentiable in its first argument we can rewrite the first bracketed expression as $\int_{\widehat{\theta}}^{\theta} U_i^{[1]}(t,\widehat{\theta})\mathrm{d}t$. By (I1) and (I2) $U_i(\theta)$ is differentiable almost everywhere and by (I3) it exhibits no jumps such that we can rewrite the second bracketed expression as $\int_{\theta}^{\widehat{\theta}} U_i'(t)\mathrm{d}t$. Hence, we have

$$\int_{\theta}^{\widehat{\theta}} -U_i^{[1]}(t,\widehat{\theta})\mathrm{d}t + \int_{\theta}^{\widehat{\theta}} U_i'(t)\mathrm{d}t > 0.$$

Using the definition of $U_i(\theta,\widehat{\theta})$ and (I2), we obtain

$$\begin{aligned} & \int_{\theta}^{\widehat{\theta}} -2(1-\alpha)(\mathbf{E}[d(\theta_1,\ldots,\theta_n)|\theta_i=\widehat{\theta}] - (1-\alpha)t)\mathrm{d}t \\ & + \int_{\theta}^{\widehat{\theta}} 2(1-\alpha)(\mathbf{E}[d(\theta_1,\ldots,\theta_n)|\theta_i=t] - (1-\alpha)t)\mathrm{d}t > 0 \\ \Rightarrow \quad & \int_{\theta}^{\widehat{\theta}} (\mathbf{E}[d(\theta_1,\ldots,\theta_n)|\theta_i=t] - \mathbf{E}[d(\theta_1,\ldots,\theta_n)|\theta_i=\widehat{\theta}])\mathrm{d}t > 0. \end{aligned}$$

This contradicts (I1). q.e.d.

A.5 Remark on Stochastic Mechanisms

In an earlier version of this paper we considered also stochastic mechanisms. If a stochastic mechanism is used, there are two sources of uncertainty. First, the decision (conditional on the agents' signals), and second, the signals itself. Since utility is quadratic, the first source of uncertainty matters only through its variance, $\mathbf{Var}[d|\theta_1,\ldots,\theta_n]$, and its expectation, $\mathbf{E}[d|\theta_1,\ldots,\theta_n]$. In principle, any expected decision function $\mathbf{E}[d|\theta_1,\ldots,\theta_n]$ can be implemented through a stochastic mechanism since incentive compatibility can be restored by adding an ad interim expected variance $\mathbf{Var}[d|\theta_i]$ having the right marginal behavior. However, while it is easy to construct some function $\mathbf{Var}[d|\theta_1,\ldots,\theta_n]$ which supports any marginal behavior of $\mathbf{Var}[d|\theta_i]$, it is involved to find the one inducing the lowest expected variance.

Appendix B

Appendix to Chapter 3

B.1 Proofs of Section 3.3

Proof of Proposition 3.1

This result can be obtained as special case from Proposition 3.5 by assuming that the seller has value $\theta_0 = 0$ with probability one, i.e. by assuming $p = 1$, or from any textbook on auction theory.

Proof of Proposition 3.2

This is Example 23.F.2 in Mas-Colell, Whinston, and Green (1995).

Proof of Proposition 3.3

This follows directly from payoffs and Assumption 3.1.

Proof of Proposition 3.4

Notation

In this proof we use notation

$$U(\theta, \widehat{\theta}) := H(b^*(\widehat{\theta}))(\theta - b^*(\widehat{\theta})) \tag{B.1}$$

for the expected payoff of a buyer who unilaterally deviates from equilibrium by behaving as if he had type $\widehat{\theta} \in P$ although he has type θ. Furthermore, we denote the expected equilibrium payoff of a buyer with type θ by $U(\theta) := U(\theta, \theta)$.

Necessity, i.e. IC+IR for some B imply (3.3) and (3.4)

Step 1: Bids are weakly increasing
Let $\widehat{\theta} < \theta$. IC implies

$$U(\theta) \geq U(\theta, \widehat{\theta}) = U(\widehat{\theta}) + H(b^*(\widehat{\theta}))(\theta - \widehat{\theta})$$

and

$$U(\widehat{\theta}) \geq U(\widehat{\theta}, \theta) = U(\theta) + H(b^*(\theta))(\widehat{\theta} - \theta).$$

Thus,

$$H(b^*(\theta)) \geq \frac{U(\theta) - U(\widehat{\theta})}{\theta - \widehat{\theta}} \geq H(b^*(\widehat{\theta})). \tag{B.2}$$

This implies that $H \circ b^*$ is weakly increasing. If b^* was somewhere strictly decreasing, then there would exist $\theta > \widehat{\theta}$ with $H(b^*(\theta)) \geq H(b^*(\widehat{\theta}))$ but $b^*(\theta) < b^*(\widehat{\theta})$. Since this contradicts IC, we obtain (3.3).

Step 2: $b^*(r) = r$ and $U(r) = 0$
Consider first $r = 0$. If $b^*(0) > 0$, step 1 implies that a buyer with value $b^*(0)/2$ bids at least $b^*(0)$ and that his bid is not lower than the bids of all lower buyer types. Thus, he obtains the object with a probability of at least $G(b^*(0))[0 + \frac{1}{2}F(b^*(0)/2)]$ which is strictly positive since we assumed $G(\epsilon) > 0$ for any positive ϵ. Since he obtains the object with positive probability and since his payoff conditional on obtaining it is negative, he is better off not participating. Thus, we obtain a contradiction to $r = 0$ as long as $b^*(0) > 0$. Since we assumed bids to be positive, we must have $b^*(0) = 0$. This implies $U(0) = 0$.

Consider now $r \in (0, 1]$. If $b^*(r) = 0$, type $\theta = r/2$ would have a strict incentive to participate contradicting that the lowest participating type is r. Thus, $b^*(r) > 0$. Because we assumed $G(\epsilon) > 0$ for $\epsilon > 0$, the lowest participating type obtains the object with positive probability, i.e. $H(b^*(r)) > 0$. Therefore type r only participates if his payoff conditional on obtaining the object is non–negative, i.e. if $b^*(r) \leq r$. If $b^*(r) < r$, also some types in $\theta \in (b^*(r), r)$ would have a strict incentive to participate. Hence, $b^*(r) = r$. This implies $U(r) = 0$.

Step 3: $U(\theta)$ is continuous
This follows from multiplying (B.2) by $(\theta - \overline{\theta})$ and letting $\widehat{\theta} \to \theta$.

Step 4: The condition characterizing the equilibrium bid function
Since $H \circ b^*$ is non-decreasing by (B.2), $H \circ b^*$ is differentiable almost everywhere and thus also continuous almost everywhere. Letting $\widehat{\theta} \to \theta$ we obtain $U'(\theta) = H(b^*(\theta))$ at all points at which $H \circ b^*$ is continuous. Since $U(\theta)$ is continuous by step 3, we can write

$$U(\theta) = U(r) + \int_r^\theta H(b^*(s))\mathrm{d}s \overset{\text{Step 2}}{=} 0 + \int_r^\theta H(b^*(s))\mathrm{d}s. \tag{B.3}$$

Setting (B.1) equal to (B.3) and dividing by $H(b^*(\theta))$ (which is possible since $H(b^*(\theta))$ is strictly positive for $\theta > r$), we obtain (3.4).

Sufficiency, i.e. (3.3) and (3.4) imply IC+IR for some B

Any bid function $b^* : P \to \mathbf{R}_+$ that is implementable by some set of admissible bids is in particular implementable by the set $B = b^*(P)$. Adding further elements to B adds only additional incentive compatibility constraints and makes incentive compatibility harder to be satisfied. Thus, it suffices to consider the smallest set of admissible bids that is consistent with b^*, i.e. $B = b^*([r, 1])$. Since only bids which are actually chosen by some buyer type belong to B, it suffices to show that no type has an incentive to imitate any other type in order to prove IC.

$$\begin{aligned}
& U(\theta,\theta) - U(\theta,\widehat{\theta}) \\
\overset{(B.1)}{=} \quad & H(b^*(\theta))(\theta - b^*(\theta)) - H(b^*(\widehat{\theta}))(\theta - b^*(\widehat{\theta})) \\
= \quad & (H(b^*(\theta)) - H(b^*(\widehat{\theta})))\theta + H(b^*(\widehat{\theta}))b^*(\widehat{\theta}) - H(b^*(\theta))b^*(\theta) \\
\overset{(3.4)}{=} \quad & \int_{\widehat{\theta}}^{\theta} (H(b^*(s)) - H(b^*(\widehat{\theta})))\mathrm{d}s \overset{(3.3)}{\geq} 0
\end{aligned}$$

This proves that no type has a strict incentive to imitate the bidding behavior of any other type.

Since $b^*(r) = r$ by (3.4), the buyer with value r is just indifferent between participating and not doing so. Because type r obtains the object with positive probability and because payoffs are strictly increasing in value, all types below r have a strict incentive to not participate, all higher types have a strict incentive to participate. This is IR.

B.2 Proofs of Section 3.4

Proof of Proposition 3.5

Necessity, i.e. IC+IR for $B = [r, \infty)$ imply $P = [r, 1]$ and (3.6) with $\widehat{\theta} = \sigma_r^{-1}(x)$

In contrast to Proposition 3.4 we now use the specific structure imposed by Assumption 3.2 and are only interested in bid functions implementable by a connected set $B = [r, \infty)$. If we do not impose any restrictions on B, we can choose for any function b^* the set of admissible bids $B = b^*(P)$ which contains only bids chosen by some buyer type. Thus, IC is satisfied for b^* if no type has an incentive to imitate any other type. If B has to be connected, it may contain bids not chosen by any type. This adds additional incentive compatibility constraints. Thus, the necessary conditions from Proposition 3.4 are still necessary, but the set of implementable bid functions becomes smaller. For instance, no bid functions which exhibit pooling are implementable.

Lemma B.1 *If b^* is implementable by $B = [r, \infty)$, then b^* is strictly increasing.*

Proof. If b^* was not strictly increasing, some of the pooling types would have an incentive to increase their bids marginally. This would increase their probabilities of obtaining the object discretely, while it would increase their payments conditional on obtaining the object only marginally. q.e.d.

We now use the specific structure imposed by Assumption 3.2:

Lemma B.2 *If b^* is implementable and strictly increasing, then there exists a $\theta' \geq r$ such that*

$$b^*(\theta) = \begin{cases} \beta_r(\theta) & \text{if } \theta < \theta' \\ \beta_r(\theta) + \frac{F(\theta')}{F(\theta)}(\sigma_r(\theta') - \beta_r(\theta')) & \text{if } \theta \geq \theta' \end{cases} \tag{B.4}$$

with $b^*(\theta) < x$ for $\theta \in [r, \theta')$ and $b^*(\theta) \geq x$ for $\theta \in [\theta', 1]$.

Proof. Using the specific structure imposed by Assumption 3.2 and using that b^* is strictly increasing the necessary condition (3.4) becomes

$$b^*(\theta) = \theta - \int_r^\theta \frac{G(b^*(s))}{G(b^*(\theta))} \frac{F(s)}{F(\theta)} \mathrm{d}s$$

$$= \begin{cases} \theta - \int_r^\theta \frac{p}{p}\frac{F(s)}{F(\theta)}\mathrm{d}s & \text{if } b^*(\theta) < x \\ \theta - \int_{\{s\in[r,\theta]\|b^*(s)<x\}} \frac{p}{1}\frac{F(s)}{F(\theta)}\mathrm{d}s - \int_{\{s\in[r,\theta]\|b^*(s)\geq x\}} \frac{1}{1}\frac{F(s)}{F(\theta)}\mathrm{d}s & \text{if } b^*(\theta) \geq x \end{cases}.$$

This characterization is still only implicit, but we can obtain explicit conditions by distinguishing three cases:

Case 1: $b^*(r) = r < x$ and $b^*(1) \geq x$.
Since b^* is strictly increasing, there exists a unique $\theta' \in (r, 1]$ such that $b^*(\theta) < x$ for $\theta < \theta'$ and $b^*(\theta) > x$ for $\theta \geq \theta'$. Thus,

$$\begin{aligned} b^*(\theta) &= \begin{cases} \beta_r(\theta) & \text{if } \theta < \theta' \\ \theta - \int_r^{\theta'} p\frac{F(s)}{F(\theta)}\mathrm{d}s - \int_{\theta'}^\theta \frac{F(s)}{F(\theta)}\mathrm{d}s & \text{if } \theta \geq \theta' \end{cases} \\ &= \begin{cases} \beta_r(\theta) & \text{if } \theta < \theta' \\ \beta_r(\theta) - \frac{F(\theta')}{F(\theta)}(\sigma_r(\theta') - \beta_r(\theta')) & \text{if } \theta \geq \theta' \end{cases}. \end{aligned}$$

Case 2: $b^*(\theta) \geq x$ for all $\theta \in [r, 1]$.
In this case we have $b^*(\theta) = \beta_r(\theta)$. Using $\beta_r(r) = \sigma_r(r) = r$, we can write this in a more complicated way:

$$b^*(\theta) = \begin{cases} \beta_r(\theta) & \text{if } \theta < \theta' \\ \beta_r(\theta) + \frac{F(\theta')}{F(\theta)}(\sigma_r(\theta') - \beta_r(\theta')) & \text{if } \theta \geq \theta' \end{cases}$$

with $\theta' = r$.

Case 3: $b^*(\theta) < x$ for all $\theta \in [r, 1]$.
In this case we have $b^*(\theta) = \beta_r(\theta)$, or more complicated:

$$b^*(\theta) = \begin{cases} \beta_r(\theta) & \text{if } \theta < \theta' \\ \beta_r(\theta) + \frac{F(\theta')}{F(\theta)}(\sigma_r(\theta') - \beta_r(\theta')) & \text{if } \theta \geq \theta' \end{cases}$$

with $\theta' > 1$.

Thus, there is always a θ' such that b^* can be written as in (B.4). q.e.d.

In the subsequent two Lemmas we show how θ' must look like:

Lemma B.3 *If b^* is implementable and strictly increasing, (B.4) holds with $\theta' \geq \sigma_r^{-1}(x)$.*

Proof. Assume to the contrary that $\theta' < \sigma_r^{-1}(x)$ is true. By Lemma B.2 we obtain $b^*(\theta') = \beta_r(\theta') + \frac{F(\theta')}{F(\theta')}(\sigma_r(\theta') - \beta_r(\theta')) = \sigma_r(\theta') < x$. However, by Lemma B.2 type θ' has also to bid above x. Contradiction. q.e.d.

Lemma B.4 *If b^* is implementable by $B = [r, \infty)$, (B.4) holds with $\theta' = \sigma_r^{-1}(x)$.*

Proof. Assume to the contrary that $\theta' > \sigma_r^{-1}(x)$, i.e. $\sigma_r(\theta') > x$. Since $b^*(\theta') = \sigma_r(\theta')$ by Lemma B.2, type θ' could decrease his bid from $b^*(\theta')$ to x without changing the probability of obtaining the object. This contradicts IC. q.e.d.

The four Lemmas imply $P = [r, 1]$ and (B.4) with $\theta' = \sigma_r^{-1}(x)$.

Sufficiency, i.e. $P = [r, 1]$ and (3.6) with $\widehat{\theta} = \sigma_r^{-1}(x)$ imply IC+IR for $B = [r, \infty)$

Condition (3.6) with $\widehat{\theta} = \sigma_r^{-1}(x)$ is a special case of condition (3.4) in Proposition 3.4. Since we have already proven in this Proposition that no type has an incentive to imitate any other type and that IR is satisfied, this must also be true here too. It remains to show that no type has a strict incentive to choose a bid not chosen by any other type, i.e. a bid from $B \backslash b^*([r, 1])$. There are three cases to consider:

Case 1: $\theta' = \sigma_r^{-1}(x) > 1$.
In this case all types bid below x in equilibrium. Admissible bids not chosen in equilibrium are those in $(b^*(1), \infty)$. However, all bids in $(b^*(1), x)$ are dominated by bid $b^*(1)$ because those bids only decrease payoffs without increasing the probability of obtaining the object. Analogously, all bids in (x, ∞) are dominated by bid x. It remains only to show that no type has a strict incentive to choose bid x. If we had $x = \sigma_r(1)$, type $\theta_i = 1$ was indifferent between bidding $b^*(1)$ and bidding x. But since we assume $\sigma_r^{-1}(x) > 1 \Leftrightarrow x > \sigma_r(1)$, type $\theta_i = 1$ has a strict incentive not to choose x. By a monotonicity argument, we obtain that all lower types have also no strict incentive to choose x.

Case 2: $\sigma_r^{-1}(x) = r$.
In this case all types bid above x. Bids not chosen are those in $(b^*(1), \infty)$. However, all these bids are dominated by bid $b^*(1)$.

Case 3: $\theta' = \sigma_r^{-1}(x) \in (r, 1]$.

In this case some types bid below and others bid above x. Admissible bids not chosen are those in $(b^*(1), \infty)$ and in $[\beta_r(\sigma_r^{-1}(x)), x)$. However, any bid in $(b^*(1), \infty)$ is dominated by bid $b^*(1)$ and any bid in $(\beta_r(\sigma_r^{-1}(x)), x)$ is dominated by bid $\beta_r(\sigma_r^{-1}(x))$. Thus, it remains only to show that no type has a strict incentive to choose bid $\beta_r(\sigma_r^{-1}(x))$. Since the lowest type bidding above x is just indifferent between bidding $\beta_r(\sigma_r^{-1}(x))$ and bidding x, we obtain by a monotonicity argument that no higher type can have a strict incentive to choose bid $\beta_r(\sigma_r^{-1}(x))$. Finally, types below $\theta' = \sigma_r^{-1}(x)$ have no strict incentive to choose bid $\beta_r(\sigma_r^{-1}(x))$ for the same reason they have no such incentive in the standard case.

Proof of Proposition 3.6

Necessity, i.e. IC+IR for some B and b^* strictly increasing imply that (3.6) holds with $\widehat{\theta}' \in [\sigma_r^{-1}(x), \beta_r^{-1}(x)]$

Lemma B.2 and Lemma B.3 in the proof of Proposition 3.5 rely only on IC and b^* being strictly increasing such that they remain valid under the assumptions made here. From this we know that any strictly increasing implementable bid function must satisfy condition (3.6) with $\widehat{\theta}' \geq \sigma_r^{-1}(x)$. However, in contrast to the case in which B is connected, $\widehat{\theta}' = \sigma_r^{-1}(x)$ needs not to be true here, but we can derive an upper bound for $\widehat{\theta}'$:

Lemma B.5 *If b^* is implementable and strictly increasing, (3.6) holds with $\widehat{\theta}' \leq \beta_r^{-1}(x)$.*

Proof. Assume to the contrary that $\widehat{\theta}' > \beta_r^{-1}(x)$. Then, condition (3.6) and strict monotonicity of bids imply that any type $\theta \in (\beta_r^{-1}(x), \widehat{\theta}')$ submits a bid $\beta_r(\theta) > \beta_r(\beta_r^{-1}(x)) = x$. This contradicts that types on the left of $\widehat{\theta}'$ bid below x (see Lemma B.2 in the proof of Proposition 3.5). q.e.d.

Sufficiency, i.e. (3.6) with $\widehat{\theta}' \in [\sigma_r^{-1}(x), \beta_r^{-1}(x)]$ implies IC+IR for some B and b^* strictly increasing

Condition (3.6) with $\widehat{\theta}' \in [\sigma_r^{-1}(x), \beta_r^{-1}(x)]$ is a special case of condition (3.4) in Proposition 3.4. From this it follows that no type has an incentive to imitate any other type and that IR is satisfied. Since we only have to find some set of admissible bids for which IC and IR is satisfied, we can simply choose $B = b^*([r, 1])$ and are done.

B **must not contain bids from** $[x, \sigma_r(\widehat{\theta}'))$

Lemma B.6 *If b^* is implementable by set B and b^* induces separating type $\widehat{\theta}' > \sigma^{-1}(x)$, then B is non–connected and does not contain bids from $[x, \sigma_r(\widehat{\theta}'))$.*

Proof. Assume to the contrary that B contained bids from $[x, \sigma_r(\widehat{\theta}'))$. Then type $\widehat{\theta}'$ would have a strict incentive to choose such a bid. This would increase his payoff conditional on obtaining the object without decreasing his probability of obtaining it. Contradiction. q.e.d.

Proof of Proposition 3.7

We denote the distribution function of the highest of two independently F–distributed random variables by $\Phi := F^2$ and the respective density function by $\phi := \Phi' = 2fF$. Using this notation we obtain the following expression for the seller's expected utility as function of the separating type:

$$\begin{aligned}
\mathbf{E}[u_0] &= p\Big[0 + \int_r^{\widehat{\theta}} \beta_r(\theta)\mathrm{d}\Phi(\theta) + \int_{\widehat{\theta}}^1 (\beta_r(\theta) + \frac{F(\widehat{\theta})}{F(\theta)}(\sigma_r(\widehat{\theta}) - \beta_r(\widehat{\theta})))\mathrm{d}\Phi(\theta)\Big] \\
&\quad +(1-p)\Big[\int_0^{\widehat{\theta}} x\mathrm{d}\Phi(\theta) + \int_{\widehat{\theta}}^1 (\beta_r(\theta) + \frac{F(\widehat{\theta})}{F(\theta)}(\sigma_r(\widehat{\theta}) - \beta_r(\widehat{\theta})))\mathrm{d}\Phi(\theta)\Big] \\
&= \int_{\widehat{\theta}}^1 \frac{F(\widehat{\theta})}{F(\theta)}(\sigma_r(\widehat{\theta}) - \beta_r(\widehat{\theta}))\mathrm{d}\Phi(\theta) \\
&\quad +p\Big[\int_r^1 \beta_r(\theta)\mathrm{d}\Phi(\theta)\Big] + (1-p)\Big[\int_0^{\widehat{\theta}} x\mathrm{d}\Phi(\theta) + \int_{\widehat{\theta}}^1 \beta_r(\theta)\mathrm{d}\Phi(\theta)\Big] \\
&= 2F(\widehat{\theta})(\sigma_r(\widehat{\theta}) - \beta_r(\widehat{\theta}))(1 - F(\widehat{\theta})) \\
&\quad +p\Big[\int_r^1 \beta_r(\theta)\mathrm{d}\Phi(\theta)\Big] + (1-p)\Big[\int_0^{\widehat{\theta}} x\mathrm{d}\Phi(\theta) + \int_{\widehat{\theta}}^1 \beta_r(\theta)\mathrm{d}\Phi(\theta)\Big]
\end{aligned}$$

Note that $F(\widehat{\theta})(\sigma_r(\widehat{\theta}) - \beta_r(\widehat{\theta})) = (1-p)\int_r^{\widehat{\theta}} F(s)\mathrm{d}s$ such that $\frac{\mathrm{d}}{\mathrm{d}\widehat{\theta}}\Big(F(\widehat{\theta})(\sigma_r(\widehat{\theta}) - \beta_r(\widehat{\theta}))\Big) = (1-p)F(\widehat{\theta})$. Using this we obtain

$$\begin{aligned}
\frac{\mathrm{d}\mathbf{E}[u_0]}{\mathrm{d}\widehat{\theta}} &= \Big[(1-p)\frac{1 - F(\widehat{\theta})}{f(\widehat{\theta})} - (\sigma_r(\widehat{\theta}) - \beta_r(\widehat{\theta}))\Big]2f(\widehat{\theta})F(\widehat{\theta}) \\
&\quad +\Big[(1-p)(x - \beta_r(\widehat{\theta}))\Big]2f(\widehat{\theta})F(\widehat{\theta})
\end{aligned}$$

and using $\sigma_r(\widehat{\theta}) - \beta_r(\widehat{\theta}) = (1-p)\int_r^{\widehat{\theta}} \frac{F(s)}{F(\widehat{\theta})}\mathrm{d}s$ as well as $(1-p)\beta_r(\widehat{\theta}) = (1-p)[\widehat{\theta} - \int_r^{\widehat{\theta}} \frac{F(s)}{F(\widehat{\theta})}\mathrm{d}s]$ we get

$$\ldots = (1-p)\Big[x - \widehat{\theta} + \frac{1-F(\widehat{\theta})}{f(\widehat{\theta})}\Big]2f(\widehat{\theta})F(\widehat{\theta}) = (1-p)\Big[x - v(\widehat{\theta})\Big]2f(\widehat{\theta})F(\widehat{\theta}).$$

Since v is strictly increasing the sign of $\mathrm{d}\mathbf{E}[u_0]/\mathrm{d}\widehat{\theta}$ can only change from positive to negative such that the solution to $x - v(\widehat{\theta}) = 0$ specifies a global maximum. However, we have to respect that $\widehat{\theta}$ can only be chosen from $[\sigma_r^{-1}(\widehat{\theta}), \beta_r^{-1}(\widehat{\theta})]$. Therefore we obtain

$$\widehat{\theta}^* = \begin{cases} \sigma_r^{-1}(x) & \text{if } \rho(x) < \sigma_r^{-1}(x) \\ \rho(x) & \text{if } \sigma_r^{-1}(x) \leq \rho(x) \leq \beta_r^{-1}(x) \\ \min\{\beta_r^{-1}(x), 1\} & \text{if } \rho(x) > \beta_r^{-1}(x) \end{cases} .$$

Since $\rho(x) \leq 1$, it cannot happen that $\rho(x) > \beta_r^{-1}(x) > 1$. This leads to condition (3.7).

That the optimum can only be implemented by a non–connected set of admissible bids if $\widehat{\theta}^* > \sigma_r^{-1}(x)$ follows from Lemma B.6 in the proof of Proposition 3.6.

Proof of Proposition 3.8

(i) We have to show that there exists a set of x–values such that

$$\rho(x) \leq \beta_{\rho(0)}^{-1}(x) \Leftrightarrow x \geq \beta_{\rho(0)}(\rho(x)) = \rho(x) - \int_{\rho(0)}^{\rho(x)} \frac{F(\rho(s))}{F(\rho(x))}\mathrm{d}s$$

and

$$\rho(x) > \sigma_{\rho(0)}^{-1}(x) \Leftrightarrow x < \sigma_{\rho(0)}(\rho(x)) = \rho(x) - p\int_{\rho(0)}^{\rho(x)} \frac{F(\rho(s))}{F(\rho(x))}\mathrm{d}s.$$

Since $p < 1$, we have $\beta_{\rho(0)}(\rho(x)) < \sigma_{\rho(0)}(\rho(x))$. Hence, if the first condition holds with equality, the second one holds strict. We now show that there are always x–values for which the first condition holds with equality, i.e.

$$\rho(x) - \int_{\rho(0)}^{\rho(x)} \frac{F(\rho(s))}{F(\rho(x))}\mathrm{d}s - x = 0. \tag{B.5}$$

Because the left–hand side of (B.5) is continuous in x, we can apply an Intermediate Value Theorem: Since the left-hand side is positive for $x = 0$ (it is $\rho(0) > 0$) and negative for $x = 1$ (it is $\rho(1) - \int_{\rho(0)}^{\rho(1)} \frac{F(\rho(s))}{F(\rho(1))} - 1 < 0$), there exists a $x' \in (0,1)$ such that (B.5) holds and such that the left–hand side of (B.5) is negative for slightly higher values of x. This and the second condition being also continuous in x and holding strictly imply that both inequalities must hold for an entire interval of x–values.

(ii) Suppose the seller knew his value upfront and he could commit himself to any mechanism. Following standard reasoning, the optimal general mechanism is such that the seller keeps the object if the virtual valuation of the highest buyer type is smaller than his reservation value and he sells it to the highest buyer otherwise. One way of implementing this mechanism is by holding a first-price auction with reserve price $\rho(0)$ if the seller's value is zero and a first–price auction with reserve price $\rho(x)$ if the seller's value is x.

Hence, if the seller cannot commit to selling and he learns his value only after designing the auction, he can still implement the generally optimal mechanism if $\rho(x) \in (\sigma^{-1}_{\rho(0)}(x), \beta^{-1}_{\rho(0)}(x)]$ (Proposition 3.6).

B.3 Proofs of Section 3.5

Proof of Proposition 3.9

If the seller postpones the design of the auction until he is informed, bid function

$$\beta_{r''}(\theta) = \theta - \int_{r''}^{\theta} \frac{F(s)}{F(\theta)} \mathrm{d}s \tag{B.6}$$

is implemented if the seller's value turns out to be x and bid function

$$\beta_{r'}(\theta) = \theta - \int_{r'}^{\theta} \frac{F(s)}{F(\theta)} \mathrm{d}s \tag{B.7}$$

is implemented if it turns out to be zero.

In contrast, if he does not wait, bid function

$$b^*(\theta) = \begin{cases} \theta - \int_{r'}^{\theta} \frac{F(s)}{F(\theta)} \mathrm{d}s & \text{if } \theta \in [r', r'') \\ \theta - p\int_{r'}^{r''} \frac{F(s)}{F(\theta)} \mathrm{d}s - \int_{r''}^{\theta} \frac{F(s)}{F(\theta)} \mathrm{d}s & \text{if } \theta \in [r'', 1] \end{cases} \tag{B.8}$$

is implemented.

We now compare obtaining bid function (B.6) if the seller's value is $\theta_0 = x$ and obtaining bid function (B.7) if it is $\theta_0 = 0$ with always obtaining bid function (B.8): We have $b^*(\theta) = p\beta_{r'}(\theta) + (1-p)\beta_{r''}(\theta)$ for any $\theta \in [\rho(0), 1]$. Strict concavity of ν implies

$\nu(b^*(\theta)) > p\nu(\beta_{r'}(\theta)) + (1-p)\nu(\beta_{r''}(\theta))$ for any $\theta \in [\rho(0), 1]$. Thus, the seller strictly prefers staying uninformed over waiting until he gets informed.

Appendix C

Appendix to Chapter 4

C.1 Proofs of Section 4.3

Proof of Proposition 4.1

This is Proposition 23.D.2 in Mas-Colell, Whinston, and Green (1995) with an individual rationality constrained added and a specific structure of payoffs.

We first consider incentive compatibility for a player who obtains utility

$$U(\theta, \widehat{\theta}) := \overline{y}_k(\widehat{\theta})(\theta + c_k^1) + c_k^2 - \overline{t}_k(\widehat{\theta})$$

if he has value θ but announces having value $\widehat{\theta}$. Let $U(\theta) := U(\theta, \theta)$ be his utility from truth–telling. We show that incentive compatibility is satisfied if and only if (i) $\overline{y}_k(\theta)$ is weakly increasing and (ii) $U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} \overline{y}_k(s)\mathrm{d}s$.

Remark: We obtain the utility of the incumbent ($k = I$) by choosing $c_I^1 = i + V_I - V_E$ and $c_I^2 = V_E$, and the utility of an entrant ($k = Ej$) by choosing $c_{Ej}^1 = V_I - V_E$ and $c_{Ej}^2 = V_E$.

IC implies (i) and (ii).
Let $\widehat{\theta} < \theta$. Then IC implies

$$U(\theta) \geq U(\theta, \widehat{\theta}) = U(\widehat{\theta}) + \overline{y}_k(\widehat{\theta})(\theta - \widehat{\theta}) \tag{C.1}$$

and

$$U(\widehat{\theta}) \geq U(\widehat{\theta}, \theta) = U(\theta) + \overline{y}_k(\theta)(\widehat{\theta} - \theta). \tag{C.2}$$

Thus,

$$\overline{y}_k(\theta) \geq \frac{U(\theta) - U(\widehat{\theta})}{\theta - \widehat{\theta}} \geq \overline{y}_k(\widehat{\theta}). \tag{C.3}$$

This is (i).

Suppose $U(\theta)$ is not continuous. Then there exists a sequence $\{\theta_l\}_{l=1}^{\infty} \to \theta$ with $\lim_{l\to\infty} \theta_l = \theta$ but $U(\theta) \neq \lim_{l\to\infty} U(\theta_l)$. However, taking limits in (C.1) and (C.2) we obtain $U(\theta) \geq \lim_{l\to\infty} U(\theta_l)$ as well as $U(\theta) \leq \lim_{l\to\infty} U(\theta_l)$. Contradiction.

Letting $\widehat{\theta} \to \theta$ in (C.3) we obtain for any θ at which $\overline{y}_k$ is continuous $U'(\theta) = \overline{y}_k(\theta)$. Since it follows from (i) that $\overline{y}_k$ is continuous almost everywhere and since $U(\theta)$ is continuous, we can write (ii).

(i) and (ii) imply IC.

If $\theta > \widehat{\theta}$, then

$$U(\theta) - U(\widehat{\theta}) \overset{(ii)}{=} \int_{\widehat{\theta}}^{\theta} \overline{y}_k(s)\mathrm{d}s \overset{(i)}{\geq} \int_{\widehat{\theta}}^{\theta} \overline{y}_k(\widehat{\theta})\mathrm{d}s = (\theta - \widehat{\theta})\overline{y}_k(\widehat{\theta}).$$

From this it follows $U(\theta) \geq U(\widehat{\theta}) + (\theta - \widehat{\theta})\overline{y}_k(\widehat{\theta}) = U(\theta, \widehat{\theta})$. If $\theta > \widehat{\theta}$, then

$$U(\theta) - U(\widehat{\theta}) \overset{(ii)}{=} \int_{\widehat{\theta}}^{\theta} \overline{y}_k(s)\mathrm{d}s \overset{(i)}{\leq} \int_{\widehat{\theta}}^{\theta} \overline{y}_k(\theta)\mathrm{d}s = (\theta - \widehat{\theta})\overline{y}_k(\theta).$$

From this it follows $U(\widehat{\theta}) \geq U(\theta) + (\widehat{\theta} - \theta)\overline{y}_k(\theta) = U(\widehat{\theta}, \theta)$. Both conditions together are IC.

Individual rationality.

Since any supplier can ensure himself a zero–probability of obtaining the object and a zero–payment, individual rationality is satisfied for player k if $U(\underline{\theta}) \geq c_k^2$. If we choose $\overline{t}_k(\theta)$ such that it is just satisfied for player k's worst type, condition (ii) becomes

$$\overline{t}_k(\theta) = \overline{y}_k(\theta)(\theta + c_k^1) - \int_{\underline{\theta}}^{\theta} \overline{y}_k(s)\mathrm{d}s.$$

Proof of Lemma 4.1

This is basically Proposition 23.D.3 in Mas-Colell, Whinston, and Green (1995).

Using notation $c_I^1 = i + V_I - V_E$ and $c_{Ej}^1 = V_I - V_E$ again, we obtain from Proposition 4.1

$$\mathbf{E}[t_k(\theta_I, \theta_{E1}, \ldots, \theta_{En})] = \mathbf{E}[\overline{t}_k(\theta_k)] = \int_{\underline{\theta}}^{\overline{\theta}} \left[\overline{y}_k(\theta_k)(\theta_k + c_k^1) - \int_{\underline{\theta}}^{\theta_k} \overline{y}_k(s)\mathrm{d}s\right] \mathrm{d}\Phi(\theta_k).$$

Applying integration by parts we get

$$\begin{aligned}\int_{\underline{\theta}}^{\overline{\theta}}\int_{\underline{\theta}}^{\theta_k}\overline{y}_k(s)\mathrm{d}s\phi(\theta_k)\mathrm{d}\theta_k &= \left[\int_{\underline{\theta}}^{\theta_k}\overline{y}_k(s)\mathrm{d}s\Phi(\theta_k)\right]_{\underline{\theta}}^{\overline{\theta}} - \int_{\underline{\theta}}^{\overline{\theta}}\overline{y}_k(\theta_k)\Phi(\theta_k)\mathrm{d}\theta_k \\ &= \int_{\underline{\theta}}^{\overline{\theta}}\overline{y}_k(\theta_k)\frac{1-\Phi(\theta_k)}{\phi(\theta_k)}\mathrm{d}\Phi(\theta_k).\end{aligned}$$

Hence,

$$\mathbf{E}[t_k(\theta)] = \int_{\underline{\theta}}^{\overline{\theta}}\overline{y}_k(\theta_k)\left[\theta_k + c_k^1 - \frac{1-\Phi(\theta_k)}{\phi(\theta_k)}\right]\mathrm{d}\Phi(\theta_k) = \int_{\underline{\theta}}^{\overline{\theta}}\overline{y}_k(\theta_k)(v(\theta_k)+c_k^1)\mathrm{d}\Phi(\theta_k).$$

Proof of Proposition 4.2

(i) This follows directly from (4.3).

(ii) Using $y_I(\theta_I, \theta_{E1}, \ldots, \theta_{En}) = y_I(\theta_I)$ in condition (4.3), we see that the procurer's expected revenue is maximized if for any θ_I expression

$$y_I(\theta_I)(v(\theta_I)+i) + \mathbf{E}[\textstyle\sum_j y_{Ej}(\theta_I, \theta_{E1}, \ldots, \theta_{En})v(\theta_{Ej})|\theta_I]$$

is maximized. Since, for a given value of $y_I(\theta_I)$, it is clearly optimal to award the contract with probability $1 - y_I(\theta_I)$ to the entrant with the highest virtual valuation, the optimal $y(\theta_I)$ maximizes

$$y_I(\theta_I)(v(\theta_I)+i) + (1-y_I(\theta_I))\mathbf{E}[\max_j v(\theta_{Ej})].$$

Thus, it is optimal to award the contract to the incumbent if $v(\theta_I) + i > \mathbf{E}[v(\theta_E)]$ and to the entrant with the highest value otherwise.

Proof of Proposition 4.3

The incumbent obtains in the CS the contract for sure if we have $v(\theta_I)+i \geq v(\theta_E)$ for all θ_I and all θ_E. This is satisfied if $v(\underline{\theta})+i \geq v(\overline{\theta})$. Analogously, he obtains it in the PS for sure if we have $v(\theta_I)+i \geq \mathbf{E}[v(\theta_E)]$ for all θ_I. This is equivalent to $v(\underline{\theta})+i \geq \mathbf{E}[v(\theta_E)]$. Finally, we obtain $\overline{i}_{PS} < \overline{i}_{CS}$ from $\mathbf{E}[v(\theta_E)] < v(\overline{\theta})$.

Proof of Proposition 4.4

If the incumbent does not invest, he obtains the contract in the CS if $v(\theta_I) > v(\theta_E)$, i.e. if $\theta_I > \theta_E$. Hence, he obtains it if his value is higher than the value of each of the n entrants. Since the values of all suppliers are iid, this happens with probability $1/(n+1)$. To compare this probability with that in the PS, it is helpful to write it down in a more complicated way:

$$\begin{aligned}\text{Prob(I wins in the CS)} &= \text{Prob}(\theta_I > v^{-1}(v(\theta_E))) \\ &= \int_{\underline{\theta}}^{\overline{\theta}} \int_{v^{-1}(v(\theta_E))}^{\overline{\theta}} 1 \mathrm{d}\Phi(\theta_I) \mathrm{d}\Psi(\theta_E) = \mathbf{E}[1 - \Phi(v^{-1}(v(\theta_E)))].\end{aligned}$$

The incumbent obtains the contract in the PS if $v(\theta_I) > \mathbf{E}[v(\theta_E)]$. Thus,

$$\begin{aligned}\text{Prob(I wins in the PS)} &= \text{Prob}(\theta_I > v^{-1}(\mathbf{E}[v(\theta_E)])) \\ &= \int_{v^{-1}(\mathbf{E}[v(\theta_E)])}^{\overline{\theta}} 1 \mathrm{d}\Phi(\theta_I) = 1 - \Phi(v^{-1}(\mathbf{E}[v(\theta_E)])).\end{aligned}$$

Hence,

$$\begin{aligned}&\text{Prob(I wins in the PS)} > \text{Prob(I wins in the CS)} \\ \Leftrightarrow\ &\Phi \circ v^{-1}(\mathbf{E}[v(\theta_E)]) < \mathbf{E}[\Phi \circ v^{-1}(v(\theta_E))].\end{aligned}$$

A sufficient condition for this inequality to be true is by Jensen's Inequality that $\Phi \circ v^{-1}$ is strictly convex. Analogously, we obtain that the probability is equal (smaller), if $\Phi \circ v^{-1}$ is linear (concave).

Proof of Proposition 4.5

(i) For $i = 0$ the incumbent obtains the contract in the CS if $v(\theta_I) \geq v(\theta_E)$, i.e. if $\theta_I \geq \theta_E$. This is just the condition for efficiency. In the PS, the contract allocation differs from that in the CS and thus cannot be efficient.

(ii) *Step 1: For $i > 0$ the incumbent obtains the contract too seldom in the CS, i.e. $\theta_I \geq v^{-1}(v(\theta_E) - i)$ implies $\theta_I > \theta_E - i$.* This is equivalent to $v^{-1}(v(\theta_E) - i) > \theta_E - i$ and to $v(\theta_E) > v(\theta_E - i) + i$. Since both sides of the inequality are equal for $i = 0$, it suffices to show that the right hand side is strictly decreasing in i. The derivative of the right hand side with respect to i is $-v'(\theta_E - i) + 1$. It is negative since the derivative of the virtual valuation function is larger than one by the increasing hazard rate property.

Step 2: If $i < \bar{i}_{PS}$, the functions describing the contract allocations in the CS and in the PS intersect in the interior of the joint support. The functions describing the contract allocations are $\widetilde{\theta}_I(\theta_E) = v^{-1}(v(\theta_E) - i)$ in the CS and $\widetilde{\theta}_I(\theta_E) = v^{-1}(\mathbf{E}[v(\theta_E)] - i)$ in the PS. Sufficient for an intersection is that there exists a θ'_E such that $v(\theta'_E) = \mathbf{E}[v(\theta_E)]$. Since we have $v(\underline{\theta}) < \mathbf{E}[v(\theta_E)]$, $v(\overline{\theta}) > \mathbf{E}[v(\theta_E)]$ and since v is continuous there exists by the Intermediate Value Theorem a $\theta'_E \in (\underline{\theta}, \overline{\theta})$ for which $v(\theta'_E) = \mathbf{E}[v(\theta_E)]$. Since $v^{-1}(\mathbf{E}[v(\theta_E)] - i) \in (\underline{\theta}, \overline{\theta})$ for the relevant investment levels, the intersection is in the interior of the joint support.

Step 3. From step 2 it follows that there are regions close to the point of intersection in which the incumbent obtains the contract in the CS but not in the PS and vice versa. From step 1 we know that the incumbent should always obtain the contract close to the point of intersection. Hence, there exist regions in which he should obtain the contract and he obtains it in the CS but not in the PS and vice versa.

(iii) For an investment of $\bar{i}_{PS} = \mathbf{E}[v(\theta_E)] - v(\underline{\theta})$, it is efficient that the incumbent obtains the contract for sure if $\underline{\theta} + \bar{i}_{PS} \geq \overline{\theta}$, i.e. if $\mathbf{E}[v(\theta_E)] - v(\underline{\theta}) \geq \overline{\theta} - \underline{\theta}$. This is just Assumption 4.1. Thus, for $i \in [\bar{i}_{PS}, \bar{i}_{CS})$ it is efficient that the incumbent obtains the contract for sure and this happens in the PS but not in the CS.

(iv) For $i \in [\bar{i}_{CS}, \infty)$ the incumbent obtains the contract for sure in both systems. It remains to show that this is efficient. I.e., we have to show that $\underline{\theta} + \bar{i}_{CS} \geq \overline{\theta}$ is true. By rewriting this inequality as $\underline{\theta} - v(\underline{\theta}) \geq \overline{\theta} - v(\overline{\theta}) \Leftrightarrow \frac{1}{\phi(\underline{\theta})} \geq 0$ we see that this is the case.

C.2 Proofs of Section 4.4

Proof of Lemma 4.2

(i) Using the specific structure of the lowest winning type in the PS (see subsection 4.3.5) in expression (4.4) we obtain

$$R'_{PS}(i) = -\frac{\mathrm{d}\widetilde{\theta}_I}{\mathrm{d}i}(1 - \Phi(\widetilde{\theta}_I)) = \begin{cases} \frac{1-\Phi(\widetilde{\theta}_I)}{v'(\widetilde{\theta}_I)} = w(\widetilde{\theta}_I) & \text{if } i < \bar{i}_{PS} \\ 0 & \text{if } i > \bar{i}_{PS} \end{cases}$$

with $\widetilde{\theta}_I = v^{-1}(\mathbf{E}[v(\theta_E)] - i)$.

Note that $w(\widetilde{\theta}_I)$ is strictly positive for $\widetilde{\theta}_I < \overline{\theta}$. Since we have $v^{-1}(\mathbf{E}[v(\theta_E)]-i) << \overline{\theta}$ for all i, we obtain for $i < \overline{i}_{PS}$ that $R'_{PS}(i)$ is strictly positive and bounded away from zero. Furthermore, R'_{PS} is for $i < \overline{i}_{PS}$ continuous as composition of continuous functions.

(ii) Using the specific structure of the lowest winning type in the CS (see subsection 4.3.5) in expression (4.4) we obtain

$$\begin{aligned} R'_{CS}(i) &= -\int_{\underline{\theta}}^{\overline{\theta}} \frac{\mathrm{d}\widetilde{\theta}_I(\theta_E)}{\mathrm{d}i}(1-\Phi(\widetilde{\theta}_I(\theta_E)))\mathrm{d}\Psi(\theta_E) \\ &= \begin{cases} \int_{\widetilde{\theta}_I^{-1}(\underline{\theta})}^{\overline{\theta}} \frac{1-\Phi(\widetilde{\theta}_I(\theta_E))}{v'(\widetilde{\theta}_I(\theta_E))}\mathrm{d}\Psi(\theta_E) & \text{if } i < \overline{i}_{CS} \\ 0 & \text{if } i > \overline{i}_{CS} \end{cases} \\ &= \begin{cases} \int_{\widetilde{\theta}_I^{-1}(\underline{\theta})}^{\overline{\theta}} w(\widetilde{\theta}_I(\theta_E))\mathrm{d}\Psi(\theta_E) & \text{if } i < \overline{i}_{CS} \\ 0 & \text{if } i > \overline{i}_{CS} \end{cases} \end{aligned}$$

with $\widetilde{\theta}_I(\theta_E) = v^{-1}(v(\theta_E) - i)$.

The integrand is for the same reason as in part (i) strictly positive and bounded away from zero. However, for $i \to \overline{i}_{CS}$, the set over which integration happens converges to a set of measure zero. This has two consequences: First, for any interval $(0, i')$ with $i' < \overline{i}_{CS}$ marginal revenue is positive and bounded away from zero, for instance by bound $(1 - \Psi(\widetilde{\theta}_I^{-1}(\underline{\theta}))) \inf_{i<i',\theta_E\in\Theta} w(\widetilde{\theta}_I(\theta_E))$. However, for $i \to \overline{i}_{CS}$ the probability term converges to zero. Since $w(\cdot)$ is bounded from above, also marginal revenue converges to zero. This leads to the second consequence, marginal revenue is everywhere continuous, even at $\overline{i}_{CS}$.

Proof of Lemma 4.3

(i) Using Lemma 4.2 we obtain

$$\begin{aligned} R'_{CS}(\overline{i}_{PS}) &= \int_{v^{-1}(v(\underline{\theta})+\overline{i}_{PS})}^{\overline{\theta}} w(v^{-1}(v(\theta_E) - \overline{i}_{PS}))\mathrm{d}\Psi(\theta_E) \\ &\overset{w \text{ decr.}}{\leq} (1 - \Psi(v^{-1}(v(\underline{\theta}) + \overline{i}_{PS})))w(\underline{\theta}). \end{aligned}$$

Since for investment $i = \overline{i}_{PS}$ the incumbent does not obtain the contract for sure in the CS, we have $(1 - \Psi(v^{-1}(v(\underline{\theta}) + \overline{i}_{PS}))) < 1$. Thus,

$$R'_{CS}(\overline{i}_{PS}) < w(\underline{\theta}) = \lim_{i\uparrow\overline{i}_{PS}} R'_{PS}(i).$$

Since the inequality is strict and marginal revenue is in both systems continuous, there is an open interval $(\underline{i}, \bar{i}_{PS})$ such that $R'_{PS}(i) > R'_{CS}(i)$ for any $i \in (\underline{i}, \bar{i}_{PS})$.

(ii) We want to show that $R'_{CS}(i) < R'_{PS}(i)$ for any $i \in (0, \bar{i}_{PS})$.

From Lemma 4.2 we obtain

$$R'_{CS}(i) = \mathbf{E}[w(v^{-1}(v(\theta_E) - i))] - \Psi(v^{-1}(v(\underline{\theta}) + i))w(\underline{\theta}).$$

If $w \circ v^{-1}$ is weakly concave on the support of $v(\theta_E) - i$, i.e. on $[v(\underline{\theta}) - \bar{i}_{PS}, v(\bar{\theta})]$, we can apply Jensen's Inequality to obtain

$$\begin{aligned} R'_{CS}(i) &\leq w(v^{-1}(\mathbf{E}[v(\theta_E)] - i)) - \Psi(v^{-1}(v(\underline{\theta}) + i))w(\underline{\theta}) \\ &< w(v^{-1}(\mathbf{E}[v(\theta_E)] - i)) = R'_{PS}(i). \end{aligned}$$

Proof of Proposition 4.6

Since $R'_{PS}(i)$ is on $(0, \bar{i}_{PS})$ bounded from below by a strictly positive number (Lemma 4.2 (i)), there exists a strictly positive γ' such that for all $\gamma < \gamma'$ and for all $i < \bar{i}_{PS}$ marginal costs are smaller than marginal revenue. Hence, we obtain the corner solution $i^*_{PS} = \bar{i}_{PS}$ already for strictly positive values of γ.

$R'_{CS}(i)$ is bounded from below by a strictly positive number on any set $(0, i')$ with $i' < \bar{i}_{CS}$ (Lemma 4.2 (ii)). However, marginal revenue at i' converges to zero as $i' \to \bar{i}_{CS}$. Thus, for any $i' < \bar{i}_{CS}$ we can find a γ' such that marginal costs lie below marginal revenue on $(0, i')$, but the marginal costs function never lies completely below the marginal revenue function. Hence, i^*_{CS} converges to $\bar{i}_{CS}$, but it does not reach it for positive values of γ.

Proof of Proposition 4.7

(i) This follows from marginal revenue being finite for investments close to $i = 0$, but marginal costs becoming arbitrarily large as $\gamma \to \infty$.

(ii) Note that we have $R'_{CS}(0) = \mathbf{E}[w(\theta_E)]$ and $R'_{PS}(0) = w(v^{-1}(\mathbf{E}[v(\theta_E)]))$.

Case 1: $C'(0) = 0$. Since marginal revenue is continuous in investment close to $i = 0$, $R'_{CS}(0) = \mathbf{E}[w(\theta_E)] < (>) R'_{PS}(0) = w(v^{-1}(\mathbf{E}[v(\theta_E)])$ implies that optimal investment is strictly larger (smaller) in the PS for $\gamma \to \infty$.

Case 2: $C'(0) > 0$. If $R'_{CS}(0) < (>) R'_{PS}(0)$, then there exists a γ' such that there is no investment in either of the systems for $\gamma > \gamma'$, but investment is larger (smaller) in the PS for values of γ just below γ'.

Proof of Proposition 4.8

Marginal social revenue

When the efficient contract allocation rule is used, the joint revenue from investment is $R_e(i) = \mathbf{E}[\mathbf{1}_{\theta_I+i\geq\theta_E}(\theta_I + i) + (1 - \mathbf{1}_{\theta_I+i\geq\theta_E})\theta_E]$. For $i > \overline{\theta} - \underline{\theta}$ we have $R'_e(i) = 1$. For $i < \overline{\theta} - \underline{\theta}$ we have

$$\begin{aligned} R_e(i) &= \int_{\underline{\theta}}^{\underline{\theta}+i} \int_{\underline{\theta}}^{\overline{\theta}} (\theta_I + i) \mathrm{d}\Phi(\theta_I) \mathrm{d}\Psi(\theta_E) \\ &+ \int_{\underline{\theta}+i}^{\overline{\theta}} \Big[\int_{\underline{\theta}}^{\theta_E - i} \theta_E \mathrm{d}\Psi(\theta_I) + \int_{\theta_E - i}^{\overline{\theta}} (\theta_I + i) \mathrm{d}\Phi(\theta_I) \Big] \mathrm{d}\Psi(\theta_E) \end{aligned}$$

and by applying Leibnitz's rule we get

$$R'_e(i) = \int_{\underline{\theta}}^{\underline{\theta}+i} 1 \mathrm{d}\Psi(\theta_E) + \int_{\underline{\theta}+i}^{\overline{\theta}} (1 - \Phi(\theta_E - i)) \mathrm{d}\Psi(\theta_E).$$

Note that when the efficient contract allocation rule is used, marginal joint revenue is for any i just the probability with which the incumbent obtains the contract.

Marginal revenue in the CS

From Lemma 4.2 we obtain for $i \in (0, \bar{i}_{CS})$

$$\begin{aligned} R'_{CS}(i) &= \int_{v^{-1}(v(\underline{\theta})+i)}^{\overline{\theta}} \frac{1 - \Phi(v^{-1}(v(\theta_E) - i))}{v'(v^{-1}(v(\theta_E) - i))} \mathrm{d}\Psi(\theta_E) \\ &\overset{v'>1}{<} \int_{v^{-1}(v(\underline{\theta})+i)}^{\overline{\theta}} (1 - \Phi(v^{-1}(v(\theta_E) - i))) \mathrm{d}\Psi(\theta_E) \\ &< \int_{\underline{\theta}}^{v^{-1}(v(\underline{\theta})+i)} 1 \mathrm{d}\Psi(\theta_E) \\ &\quad + \int_{v^{-1}(v(\underline{\theta})+i)}^{\overline{\theta}} (1 - \Phi(v^{-1}(v(\theta_E) - i))) \mathrm{d}\Psi(\theta_E). \end{aligned}$$

The latter expression is just the probability with which the incumbent actually obtains the contract in the CS.

Comparison

For $i > \bar{i}_{CS}$ we have $R'_{CS}(i) = 0 < 1 = R'_e(i)$.

For $i \in (0, \bar{i}_{CS})$ we obtain from Proposition 4.5 that the set of type combinations for which the incumbent obtains the contract in the CS is a strict subset of the set of type combinations for which it is efficient that he obtains the contract. Thus, $R'_{CS}(i) < R'_e(i)$.

Hence, when the efficient contract allocation rule is used, marginal joint revenue is for $i > 0$ strictly larger than marginal revenue in the CS. This implies that if it is efficient to have a positive level of investment, investment is smaller than efficient in the CS.

C.3 Proofs of Section 4.5

Proof of Proposition 4.9

We have to compare (4.5) for both systems. By construction of the systems, the extracted information rent is for a given investment strictly higher in the CS (Proposition 4.2). Thus, it could only be that the procurer preferred the PS if it induced a strictly higher investment. However, if γ goes to zero, investment is higher in the CS (Proposition 4.6 and Proposition 4.3). Furthermore, if γ tends to infinity, investment converges towards zero in both systems (Proposition 4.7). Hence, the PS is at best preferable for intermediate cost parameters.

Proof of Proposition 4.10

We have to compare (4.7) for both systems. This expression consists of two parts, the expected value that is generated, and the rent that has to be left to the entrants. To prove the results we will make use of the following two properties regarding the expected value that is generated:

> [A] If the contract allocation is efficient in system 2 and if $i_1^* < i_2^* \leq$ efficient level of investment, then the expected value that is generated is larger in system 2.
>
> [B] If investment is equal in system 1 and system 2 and the contract allocation in system 2 is efficient but that in system 1 is not, then the expected value that is generated is larger in system 2.

(i) Consider $\gamma \to 0$.

(a) Expected value that is generated. By Proposition 4.6 we have $i^*_{PS} = \bar{i}_{PS}$ and $i^*_{CS} \to \bar{i}_{CS}$. By Proposition 4.5 (iv) the contract allocation is efficient in the PS and converges towards the efficient allocation in the CS. Since Proposition 4.3 implies that investment is strictly higher in the CS and since by Proposition 4.8 there is never over–investment in the CS, the expected value that is generated is larger in the CS by property [A].

(b) An entrant's rent. By Proposition 4.3 the incumbent obtains the contract with probability one in the PS and a probability converging towards one in the CS. Thus, the entrants obtain no rent in the PS and a rent converging towards zero in the CS.

Hence, in the limit, the procurer prefers the CS because it induces a strictly higher investment.

(ii) Note that we have $i^*_{PS} \geq i^*_{CS}$ for $\gamma = \tilde{\gamma}$ (by Property 4.1) and $i^*_{PS} < i^*_{CS}$ for γ close to zero (by Proposition 4.6). If optimal investment was continuous in the cost parameter, we could apply an Intermediate Value Theorem to obtain that there must also be a cost parameter for which investment is equal in both systems. Since we don't have this continuity property, we obtain only that there exists either a

$$\tilde{\gamma}' \in (0, \tilde{\gamma}] \text{ such that } i^*_{PS} = \bar{i}_{PS} = i^*_{CS}, \tag{C.4}$$

or a

$$\tilde{\gamma}'' \in (0, \tilde{\gamma}] \text{ such that } \lim_{\gamma \uparrow \tilde{\gamma}''} i^*_{CS} > i^*_{PS} = \bar{i}_{PS} > \lim_{\gamma \downarrow \tilde{\gamma}''} i^*_{CS}. \tag{C.5}$$

Consider $\gamma = \tilde{\gamma}'$ first.

(a) Expected value that is generated. Investment is equal in both systems by (C.4). From Assumption 4.1 and Proposition 4.5 (iii) we obtain that the contract allocation is efficient in the PS but not in the CS. Thus, the expected value that is generated is larger in the PS by property [B].

(b) An entrant's rent. By Proposition 4.3 the incumbent obtains the contract with certainty in the PS but not in the CS. Thus, entrants obtain a positive rent in the CS but a rent of zero in the PS.

Hence, both (a) and (b) are better in the PS such that the PS is clearly superior.

Consider now $\gamma = \tilde{\gamma}'' + \epsilon$.

(a) Expected value that is generated. Since γ scales marginal costs monotonically down, investment is weakly decreasing in γ. Hence, (C.5) implies that for cost parameter $\tilde{\gamma}'' + \epsilon$ we have $i^*_{CS} < \bar{i}_{PS}$. Since investment is in the PS flat for $\gamma \leq \tilde{\gamma}$, we still have $i^*_{PS} = \bar{i}_{PS}$. Thus, investment is higher in the PS.

For cost parameter $\tilde{\gamma}''$ the incumbent is in the CS indifferent between an investment level strictly above $\bar{i}_{PS}$ and one strictly below of $\bar{i}_{PS}$. Since there is never over–investment in the CS (Proposition 4.8), there is even no over–investment if the incumbent chooses the investment strictly above $\bar{i}_{PS}$. Because the efficient investment level varies continuously in the cost parameter γ, there can also be no over–investment in the PS when the incumbent invests $\bar{i}_{PS}$ for cost–parameter $\tilde{\gamma}'' + \epsilon$ as long as ϵ is sufficiently small. Thus, for cost parameter $\gamma'' + \epsilon$ investment is clearly more efficient in the PS.

Furthermore, Assumption 4.1 and Proposition 4.5 (iii) imply that the contract allocation is efficient in the PS. Thus, the expected value that is generated is larger in the PS by property [A].

(b) An entrant's rent. By Proposition 4.3 the incumbent obtains the contract with certainty in the PS but not in the CS. Thus, entrants obtain a positive rent in the CS but a rent of zero in the PS.

Hence, both (a) and (b) are better in the PS such that the procurer clearly prefers this system.

(iii) Consider $\gamma \to \infty$.

(a) Expected value that is generated. By Proposition 4.7 investment is negligible in both systems. However, the contract allocation converges towards the efficient allocation in the CS, but not in the PS (Proposition 4.5). Thus, by property [B], the expected value that is generated is larger in the CS.

(b) An entrant's rent. The highest entrant obtains the contract in the CS but not in the PS if $\mathbf{E}[v(\theta_E)] \leq v(\theta_I) < v(\theta_E)$ and he obtains the contract in the PS but not in the CS if $v(\theta_E) < v(\theta_I) \leq \mathbf{E}[v(\theta_E)]$. Thus, relative to the PS, the

highest entrant obtains the contract in the CS in additional cases when his value is above $\mathbf{E}[v(\theta_E)]$, but he does not obtain the contract in some cases when it is below $\mathbf{E}[v(\theta_E)]$. From (4.2) and the increasing hazard rate assumption we obtain that it is more profitable for the entrant to obtain the contract in cases in which his value is low. Hence, the cases in which the entrant obtains the contract in the CS but not in the PS are less profitable for him than those in which the reverse is the case. If, in addition, the probability with which an entrant obtains the contract is lower in the CS, an entrant is clearly worse off there. By Proposition 4.4 this is just the case if Prob(I wins in PS$|i = 0) \leq 1/(n+1)$.

Hence, (a) is always better in the CS and (b) is at least better in the CS if Prob(I wins in PS$|i = 0) \leq 1/(n+1)$.

Bibliography

ALONSO, R., AND N. MATOUSCHEK (2007): "Optimal Delegation," *The Review of Economic Studies*, forthcoming.

AROZAMENA, L., AND E. CANTILLON (2004): "Investment Incentives in Procurement Auctions," *The Review of Economic Studies*, 71, 1–18.

ASHENFELTER, O., AND K. GRADDY (2003): "Auctions and the Price of Art," *Journal of Economic Literature*, 41(3), 763–787.

BAG, P. K. (1997): "Optimal auction design and R&D," *European Economic Review*, 41, 1655–1674.

BERGEMANN, D., AND S. MORRIS (2006): "Ex Post Implementation," *Games and Economic Behavior*, forthcoming.

BRONSTEIN, I., K. SEMENDJAJEW, G. MUSIOL, AND H. MÜHLIG (2001): *Taschenbuch der Mathematik*. Verlag Harri Deutsch, 5. edn.

BULOW, J., AND J. ROBERTS (1989): "The Simple Economics of Optimal Auctions," *The Journal of Political Economy*, 97(5), 1060–1090.

CASSADY, R. J. (1967): *Auctions and Auctioneering*. University of California Press, Berkeley and Los Angeles.

DASGUPTA, S. (1990): "Competition for Procurement Contracts and Underinvestment," *International Economic Review*, 31(4), 841–865.

DAVID, H. A., AND H. N. NAGARAJA (2003): *Order Statistics*. Wiley-Interscience, third edn.

Dyer, J. H. (1996): "Does Governance Matter? Keiretsu Alliances and Asset Specificity As Sources of Japanese Competitive Advantage," *Organization Science*, 7(6), 649–666.

Elyakime, B., J. J. Laffont, P. Loisel, and Q. Vuoung (1994): "First-Price Sealed-Bid Auctions With Secret Reservation Prices," *Annales d'Économie et de Statistique*, 34, 115–142.

Ferguson, T. S. (1999): "Asymptotic Joint Distribution of Sample Mean and a Sample Quantile," Unpublished manuscript.

Fudenberg, D., and J. Tirole (1991): *Game Theory*. The MIT Press.

Greenstein, S. M. (1993): "Did Installed Base Give an Incumbent any (Measuraeable) Advantages in Federal Computer Procurement," *The RAND Journal of Economics*, 24(1), 19–39.

——— (1995): "Sole-sourcing versus Competitive Bidding: US Government Agencies' Procedural Choices for Mainframe Computer Procurement," *The Journal of Industrial Economics*, 43(2), 125–140.

Grüner, H. P., and A. Kiel (2004): "Collective decisions with interdependent valuations," *European Economic Review*, 48, 1147 – 1168.

Hahn, C. K., K. H. Kim, and J. S. Kim (1986): "Costs of Competition: Implications for Purchasing Strategy," *Journal of Purchasing and Materials Management*, 22(3), 2–7.

Hendricks, K., R. H. Porter, and G. Tan (1993): "Optimal Selling Strategies for Oil and Gas Leases with an Informed Buyer," *The American Economic Review*, 83(2), 234–239.

Krishna, V. (2002): *Auction Theory*. Academic Press, San Diego.

Laffont, J.-J., and J. Tirole (1988): "Repeated Auctions of Incentive Contracts, Investment, and Bidding Parity with an Application to Takeovers," *The RAND Journal of Economics*, 19(4), 516–537.

Lewis, T. R., and H. Yildirim (2005): "Managing Switching Costs in Multiperiod Procurements with Strategic Buyers," *International Economic Review*, 46(4), 1233 – 1269.

Liker, J. K., and T. Y. Choi (2004): "Building Deep Supplier Relationships," *Harvard Business Review*, 82(12), 104–113.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995): *Microeconomic Theory*. Oxford Univ. Press, New York.

Maskin, E., and J. Tirole (2001): "Markov Perfect Equilibrium," *Journal of Economic Theory*, 100, 191 – 219.

McAfee, R. P., and D. Vincent (1997): "Sequentially Optimal Auctions," *Games and Economic Behavior*, 18, 512–518.

McMillan, J. (1990): "Managing Suppliers: Incentive Systems in Japanese and U.S. Industry," *California Management Review*, 32(4), 38–55.

Mueller, F., K. O. Stahl, and F. Wachtler (2006): "Automotive Upstream – a Case Study," Working Paper.

Myerson, R. B. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6(1), 58–73.

Riley, J. G., and W. F. Samuelson (1981): "Optimal Auctions," *The American Economic Review*, 7(3), 381–392.

Skreta, V. (2004): "Optimal Auction Design under Non-Commitment," Working Paper.

Taylor, C. R., and S. N. Wiggins (1997): "Competition and Compensation: Supplier Incentives Under the American and Japanese Subcontracting Systems," *The American Economic Review*, 87(4), 598–618.

Tirole, J. (1986): "Procurement and Renegotiation," *The Journal of Political Economy*, 94(2), 235–259.

Vartiainen, H. (2007): "Auction Design without Commitment," Working Paper.

Waehrer, K., R. M. Harstad, and M. H. Rothkopf (1998): "Auction Form Preferences of Risk-Averse Bid Takers," *The RAND Journal of Economics*, 29(1), 179–192.

Erklärung

Hiermit erkläre ich, dass ich die Dissertation selbständig angefertigt und mich anderer als der in ihr angegebenen Hilfsmittel nicht bedient habe, insbesondere, dass aus anderen Schriften Entlehnungen, soweit sie in der Dissertation nicht ausdrücklich als solche gekennzeichnet und mit Quellenangaben versehen sind, nicht stattgefunden haben.

Mannheim, den 15.10.2007 Frank Rosar

Curriculum Vitae

1979	Born in Wadern, Germany
1998	Graduation from Secondary School, Hochwald-Gymnasium, Wadern, Germany
1998-2001	Undergraduate Studies in Business Administration, Saarland University, Saarbrücken, Germany
2002-2007	Graduate Studies in Economics at the Center for Doctroal Studies in Economics and Management (CDSEM), University of Mannheim, Germany
2003-2007	Research and Teaching Assistant at the Chair for Applied Microeconomics, University of Mannheim, Germany
2006-2007	Graduate Studies in Economics at the MPSE (visiting), University of Toulouse I, France