# Four Essays in Microeconomic Theory

Inaugural-Dissertation

zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften

durch die

Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität
Bonn

vorgelegt von
Anne-Katrin Roesler
aus Mainz

Bonn 2015

Dekan: Prof. Dr. Rainer Hüttemann
Erstreferent: Prof. Dr. Benny Moldovanu
Zweitreferent: Prof. Dirk Bergemann, Ph.D.

Tag der mündlichen Prüfung: 07.08.2015
Ersatzprüfer: Prof. Dr. Daniel Krähmer

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn http://hss.ulb.uni-bonn.de/diss/_online elektronisch publiziert.

# Acknowledgements

I would like to thank everyone who has supported and encouraged me during the course of preparing my dissertation.

I'm greatly indebted to my supervisor Benny Moldovanu for his guidance and support throughout the last years. I have learned a lot from our discussions, and his critique and comments on my work. I'm particularly grateful that he offered me countless opportunities to learn more about Economics, get in touch with different views and perspectives, and to develop my own research agenda.

I am grateful to Dirk Bergemann for his advice and steady encouragement throughout my year at Yale and beyond. I cannot appreciate enough our insightful conversations, and his encouragement to believe in my work and to not give up on ideas.

It was a pleasure to have Benny and Dirk as my advisors.

I greatly benefited from the steady support and many conversations with Daniel Krähmer - thank you. Special thanks to Johannes Hörner for many insightful, critical, and inspiring discussions.

I would also like to thank my fellow students, colleagues and everyone in the community who made the time of my doctoral studies so fun and enjoyable (most of the time). Special thanks to my office partner and friend Max Conze for the good times and also for calming me down whenever I struggled with work.

I cannot thank enough my friends and my family for all their support. I especially thank my parents for providing me a world full of opportunities, my sister for always being there for me, and Clara and Jacob - who always manage to put a smile on my face.

# Contents

# Introduction

This thesis contributes to the area of Microeconomic Theory, by addressing questions on informational effects and information design in game-theoretic settings. In the first chapter, I analyze the implications of rational inattention for monopoly pricing. Chapters 2 and 3 cover topics on information acquisition and disclosure in auctions, contests, and matching markets. The focus of Chapter 4 is on understanding how the composition of committees affects collective decisions.

Even though, the applications and settings of the models that I study in this thesis may seem diverse at first, they all address different aspects of the following question: How does the information of agents influence the outcome in strategic situations, and what are the resulting implications for information design?

Chapter 1 contributes to the relatively young research area of information design. The recent interest in this topic can (to some extent) be attributed to the rapid advance of information technologies, which has changed our access to information, and thus the nature of decision making. One could argue that it is not the main challenge anymore to gather enough information to make an informed decision. The bottleneck is rather the limited capacity of a decision maker to pay attention to all of the available information. The literature on *rational inattention*, initiated by Sims (1998, 2003, 2006), captures this idea. It assumes that information is fully and freely available, but agents encounter costs or face a constraint to process this information. This raises the question how a decision maker should optimally divide his limited attention, in order to acquire information when facing a specific decision problem. Here, the decision maker designs his own information environment. The complementary question, how to optimally design the information environment of another agent in order to influence his actions in a way that is profitable for the information designer (or sender), is studied in the *Bayesian persuasion* literature (see Kamenica and Gentzkow, 2011, and subsequent papers). Typically, the literature studies settings in which agents face decision problems, i.e., non-strategic environments.

In Chapter 1, I study the implications of a rationally inattentive consumer on optimal pricing. The analysis considers a strategic environment, in which the consumer chooses an information structure and makes a purchasing decision, the seller chooses an optimal price. The consumer may face a capacity constraint, which limits the amount of information that he can process.

In this setting, the consumer's information choice not only determines his value estimate or willingness to pay for the product. It also affects the effective, interim demand function that the seller faces, and hence may influence the seller's (optimal) pricing behavior. In the essay, I show that by being inattentive (i.e. ignoring some information), the consumer can induce or persuade the seller to offer better terms of trade, that is, to charge a lower price.

I provide a completely analytical solution to this problem.[1] The first main result is, that the equilibrium is determined by the consumer's information structure, and has the following three characterizing properties: (1) The consumer-optimal information structure is *monotone partitional*. This means that the space of consumer's valuations can be partitioned into subintervals, and the consumer only learns in which range his true valuation falls. (2) In equilibrium, all gains from trade are realized. (3) Every equilibrium satisfies a property that I refer to as *seller-indifference*. This means, that the induced demand function that the seller faces is an *equal revenue curve*. The seller is indifferent between charging prices equal to any of the value estimates induced by the equilibrium information structure. By leveling the seller to a revenue level, the consumer leaves her with just enough surplus in order to guarantee that she does not want to deviate to a higher price level.

The second main result shows that even in the absence of information constraints and costs, the consumer does not want to become perfectly informed. This means that even if the consumer can learn his true valuation at no cost, he prefers to commit to ignoring information about low values, whereas the information about high values is chosen to be more precise and may be perfectly informative. By not paying attention to information that would separate low valuations, the consumer commits to buy at an intermediate price, that is, at a price that will sometimes be higher than his true valuation. Thus, the consumer offers the seller a higher probability of trade. Given that the consumer can design his information environment, he has a lot of power – he effectively designs the demand curve that the seller faces. By dividing his attention optimally he can thus induce the seller to offer a lower price. In return for the increased probability of trade, the seller offers better terms of trade, that is, a lower price. Hence, the positive effect of additional gains from trade reverberates back to the consumer. The consumer obtains a higher expected surplus.

Chapters 2 and 3 of this thesis contribute to the literature studying informational effects in markets, and endogenous information in Mechanism Design.

Mechanism or Market Design considers settings in which a number of players who hold private information interact. The task of the mechanism or market designer is to create a game that provides the players with the right incentives to reveal their private information, such that the objective – for example the efficient allocation of resources or revenue maximization – is best met. Typical examples are auctions, contests, and matching markets. In auctions, bidders have private information about their valuation for the object for which they bid. In contests, contestants are privately informed about their skills and exert effort to compete for prizes. In

[1] By contrast, the results in the rational inattention literature often focus on specific information environments (e.g. normal experiments), or are based on numerical solutions.

matching markets, agents signal their characteristics to compete for match partners.

In the traditional mechanism design setting, the private information of agents is exogenously given. Recently, there are various papers that relax this assumption. This research analyzes how the precision of private information held by players influences the equilibrium outcome in a market, and identifies the optimal level of information for players. The next step is then to include the information acquisition decision of players into the mechanism design problem, and to design a mechanism that creates both, the right incentives to acquire information, and to then reveal this private information truthfully.

Much of this research focuses on the auction setting and studies the informational effects on the allocation and on the seller's revenue.[2] Extending this line of research, in Chapter 2, I analyze how the level of information of participants in a two-sided matching market affects the match outcome and welfare. The analysis is based on the matching model in Hoppe et al. (2009), in which agents have to invest in non-productive signaling to compete for match partners. The model incorporates as special cases auctions and contests, which can be considered as one-sided matching markets. Auctions match objects to buyers, and bids serve as signals, whereas contests match prizes to contestants, and efforts serve as signals.

The analysis shows that providing more precise information to market participants has two main effects: First, increased information of market participants allows for a better allocation and thus for a higher expected match output. Second, increased information of market participants may yield higher investments in (wasteful) signaling due to amplified competition among agents. The effects depend on whether the additional or more precise information is provided to the short or to the long side of the market. The second effect on signaling investments may dominate. For matching markets, where signaling investments are assumed to be wasteful, it may thus happen that increased precision of information reduces welfare. Specifically, the increased competition within groups may eat up all of the additional match surplus made possible by the higher information level of market participants.

An application of the results are promotion contests and feedback systems in companies. The results of Chapter 2 suggest, that the effect of providing more precise information to contestants on aggregate effort depends on the ratio of contestants and prizes in the promotion contest. Hence, different feedback systems are optimal, depending on the organizational structure of a company. In organizations with steeper hierarchies or an up-or-out system, strong feedback systems should optimally be implemented, for example by a high frequency of periodical performance reports. By contrast, for organizations with flat hierarchies or promotion by seniority practices, the results suggest less sophisticated feedback structures. These predictions of optimal feedback policies seem to be in line with common practices. For example, large consulting firms with an up-or-out policy are known to have a very rigorous feedback structure.

The methodological contribution of Chapter 2 lies in suggesting *single-crossing precision* as a new informativeness criterion. This concept establishes a link between the information orders in

[2]See e.g. Persico (2000), Bergemann and Pesendorfer (2007), Ganuza and Penalva (2010), Shi (2012), and the summary of Bergemann and Välimäki (2007).

Ganuza and Penalva (2010) and the heterogeneity order in Hoppe et al. (2009). The connection allows to use the powerful tools from majorization theory that Hoppe et al. (2009) adopt in their analysis, to study informational effects in matching markets, contests and auctions.

The results and applications of Chapter 2 suggest information management as an addition to the toolbox of a (mechanism) designer. It may sometimes be infeasible or complicated for the designer to change certain structures or rules of a game. For example, in a promotion contest the number and value of prizes (jobs on the next higher level) may be fixed. In these situations, it can be easier to influence or design the information environment of agents, for example by introducing more or less precise feedback systems in a company.

The analysis of Chapter 3 is a technical contribution to the literature on mechanism design with endogenous information.

In traditional mechanism design settings with exogenously given private information, regularity conditions (such as *increasing virtual valuations* or a *monotone hazard rate*) are often imposed in order to simplify the analysis, and to avoid technicalities, specifically ironing-out procedures. Moreover, they guarantee that (optimal) mechanisms have a specific, regular form.

In a mechanism design setting with endogenous information, the distribution of posterior types of agents emerges from the information acquisition or disclosure choices of the agents. In such settings, conditions that guarantee regularity of the distribution of posterior types are essential for tractability. Without these conditions a circular effect could arise: small changes in the information level of agents could result in significant changes of the structure of the optimal mechanism, which would change the incentives to acquire or disclose information. This effect would render the model fragile, complicate the analysis tremendously, and make the model untractable.

In Chapter 3, I identify sufficient conditions on the primitives of an information structure that guarantee that certain regularity properties of the prior distribution – an increasing hazard rate, increasing virtual valuations or costs – translate to the distribution of posterior estimates. These characterization results make it possible to study mechanism design problems with endogenous information, without imposing regularity conditions on the interim stage or restricting attention to specific information structures. Applications to information acquisition and disclosure in optimal auctions, and to allocation problems without money are discussed.

Chapter 4 considers a committee voting situation. The novel aspect of this model is that, in an interdependent values environment, it introduces a second dimension of private information, about the *preference type* of agents. The goal of the analysis is to understand how private preference types, preference uncertainty, and the composition of the committee affect collective decisions.

In the essay, I study a situation in which a committee faces a binary decision: whether to accept a proposal or to stay with the status quo. Proposals that are put to vote are complex and committee members can only assess the quality of a particular dimension of it. Agents care about every aspect of the proposal, but are biased towards the factor that they can evaluate best. One could think for example about a board committee, which has to decide on whether to adopt a proposal for a product update or not. Committee members can best assess the quality of the proposal in their dimension of expertise – for example technological or design aspects. They will typically be biased towards this dimension but also be aware of the importance of the other dimensions and hence care about all of them.

The extent of such partisanship is typically intrinsic in nature and can be regarded as part of the personality of an individual. In the model, committee members hence hold two-dimensional private information: about a quality criterion of the proposal, and about their individual preference type.

In the essay, I first establish the existence of a Nash equilibrium in undominated strategies. In equilibrium, agents adopt cutoff-strategies: An agent accepts an alternative whenever his private quality signal is above a certain threshold. Moreover, it is shown that an agent's private preference type is reflected in the cutoff that he adopts. Strongly partisan agents base their votes mostly on their own observed signal. More socially-oriented types adjust their acceptance thresholds based on the information that they can derive from the event of being pivotal. Hence, for less partisan preference types, acceptance standards move away from the sincere voting threshold. For example, under unanimity voting, partisan agents adopt higher acceptance standards than their more socially-oriented colleagues.

Based on the equilibrium characterization, I then address the questions how private preference types and the composition of the committee affect collective decisions. It is shown that as the partisanship level of the population of committee members increases, agents adjust their acceptance standards more. Equilibrium cutoffs move away from the sincere voting threshold. By contrast, agents who believe to find themselves in a committee with a more heterogeneous distribution of preference types are more uncertain about the preference types of the other committee members. Hence, they will base their vote more on their own private signal. Equilibrium cutoffs move towards the sincere voting threshold.

Applications of this work include decisions of corporate boards on how to invest, whom to hire, and whether or not to adopt a new technology. Further examples are the allocation of research grants or the approval of new drugs by the FDA. The results allow to make predictions, regarding the acceptance standards and acceptance sets of such committees. For example, for decisions that require unanimity, one should expect that committees whose members display a high level of heterogeneity of preference types only accept alternatives that are of sufficiently high quality in every dimension. Moreover, if committee members over-estimate the level of partisanship of fellow committee members they will adopt too low acceptance standards.

Chapter 1

# Is Ignorance Bliss? Rational Inattention and Optimal Pricing

A rationally inattentive consumer processes information about his valuation prior to making his purchasing decision. In a monopoly pricing problem, I study the case in which information processing constraints restrict the consumer to finite information structures. The limiting, unconstrained case is analyzed as well.

Any finite equilibrium information structure satisfies three properties: It is *partitional*, guarantees *seller-indifference*, and induces *efficient trade*. The consumer strictly benefits from having access to information structures with more signal realizations. Every equilibrium information structure yields only a coarse perception about low values, whereas the information about high values is more precise and may be perfectly informative. In equilibrium, trade is efficient and the consumer is strictly better off than under monopoly pricing for a fully attentive consumer. Surprisingly, even in the absence of information processing constraints and costs, the consumer does not want to become perfectly informed.

## 1 Introduction

The rapid advance of information technologies and access to large data has changed the nature of decision making. As the gathering of information has become easier, the processing of information has become increasingly challenging. A rational consumer, who may only have limited cognitive capacity, can now choose which information to process and what to learn about his valuation for a good. In this situation, how should a consumer optimally process information? And what are the implications for optimal pricing?

In this essay, I study a rationally inattentive consumer who faces this information processing problem in a monopoly pricing model. The consumer has to decide how much to learn about his valuation for the good, prior to observing the price and making the purchasing decision. The consumer faces the following trade-off: His information choice determines the estimate of his valuation for the product, and the interim demand function that the seller faces. The seller might be induced to charge a higher price when facing a more informed consumer. Hence, the consumer may be better off by knowing less.

The aim of this essay is to identify the economic effects that arise as a result of this new feature of the model. The analysis provides answers to questions such as: What are the optimal information processing and price setting choices for the consumer and the seller? What are the implications for the market allocation, the price, and the consumer's and seller's expected surplus? Can the seller exploit the consumer's limited capacity to process information, or are there benefits for the consumer from being selectively, but not perfectly informed? Is ignorance bliss?

A broadly observed phenomenon is that, when faced with a complex product, consumers use heuristics or rules of thumb to reach a decision (Gabaix and Laibson, 2003; Gabaix et al., 2006; Shah and Oppenheimer, 2008). Such behavior is observed in a multitude of markets, including the market for electronics and the used car market (see Yee et al., 2007; DellaVigna, 2009, and references therein). This may affect prices. Lacetera et al. (2012), Busse et al. (2013), and Englmaier et al. (2013) all provide empirical evidence that links price discontinuities to consumers being inattentive to features that influence the value of a used car. For instance, consumers display a left-digit bias. This means that they only focus on the left-most digits, when they evaluate the mileage or registration year of a car. Should such behavior be interpreted as a mistake or limitation in the consumer's information processing, or could it be rational for the consumer to be partially inattentive? The analysis in this essay will provide insights into how prices are influenced by the information processing structure of the consumer. As will be shown, it can be rational for the consumer to not become perfectly informed and only have a coarse perception of the world. Partial ignorance can be bliss.

The essay analyzes a monopoly pricing model with one seller (she) who wants to sell a good to a consumer (he). The consumer chooses how to process information about his valuation prior to observing the price charged by the seller and making his purchasing decision. The information processing decision of the consumer corresponds to a selection of an information structure that provides him with an (imperfect) signal about his valuation. Capacity constraints to process information impose a restriction on the set of *accessible information structures* from which the consumer can choose. Before setting a price, the seller observes the information structure but not the private signal realization of the consumer.

I analyze two cases. In the first case, the consumer only has access to information structures with a finite number of signal realizations. That is, he can only form a limited number of categories of valuations, on which he can condition his purchasing decision. Hence, the consumer has limited capacity to process information, which may be due to limited cognitive abilities. In the second case, the consumer has no information processing constraints, and there are no restrictions beyond standard feasibility and consistency requirements on the consumer's choice set of information structures.

Information structures within the accessible set are assumed to be free, while all other information structures are infinitely costly. By working with this simple cost structure, it is possible to identify which information is the most valuable for the consumer. These insights can be used to make predictions for the case with cost differentiation among accessible information structures.

In the monopoly pricing model that I study, both agents have strategic influence. This property has two important implications. (1) When obtaining more information, the consumer faces a trade-off between being able to make a more informed decision and securing information rents. The consumer's choice of an information structure determines his interim valuations, and thus the interim demand curve faced by the seller. Hence, information acquisition by the consumer can have adverse effects on the informational rents that he can secure, because the seller might be induced to set a higher price when facing a more informed consumer. (2) Even if the consumer can process enough information to make an optimal purchasing decision for a given price, his information processing constraint can be "*strategically binding*." This means that, the capacity of the consumer may not suffice to process enough information to take an optimal purchasing decision for any price in the seller's action set.

An essential contribution of this essay is to provide a completely analytical solution to this problem. The equilibrium is determined by the consumer's information structure, and has the following three characterizing properties:

The equilibrium information structure is *monotone partitional*. This means that the space of possible valuations is split into sub-intervals and the consumer learns which of these intervals his true valuation falls into.

The equilibrium information processing structure induces *efficient trade*, that is, all possible gains from trade are realized. Among these information structures, the consumer adopts the one that grants him the largest possible share of the realized surplus.

Every equilibrium satisfies the *seller-indifference* property: The interim demand function that is induced by the equilibrium information structure is an equal revenue curve. This means that the seller is indifferent between charging a price equal to any of the value estimates that are induced by the equilibrium information structure. In equilibrium, the seller charges the lowest of these prices. By adopting an information structure with these properties, the consumer induces the seller to charge a price that yields efficient trade while – at that price – leaving the seller with just enough revenue in order to guarantee that she does not want to deviate and charge a higher price.

The main features of the equilibrium information structure also persist in the unconstrained case. In this case, there exists an equilibrium, in which the equilibrium information structure is monotone partitional, the induced interim demand function is an equal revenue curve, and all gains from trade are realized. There typically exist multiple equilibria in the unconstrained case, but they are all outcome equivalent. Remarkably, even without information processing constraints or costs, the consumer does not choose to become perfectly informed. Instead, it is optimal for the consumer to ignore information that would separate low values, whereas his perception of higher values is finer and may be perfectly informative.

Finally, I discuss the implications of optimal information processing on the consumer's and the seller's expected surplus. In the present model, the expected surplus of a rationally inattentive consumer is always higher than in the case in which the consumer knows his true valuation. Moreover, the consumer's expected surplus strictly increases if he has access to information structures with more signal realizations. If the consumer has no information processing constraints, the seller's expected revenue is bounded above by the monopoly revenue. I provide examples for which the seller's expected revenue is strictly lower than the monopoly revenue.

In the absence of information constraints, the present problem has similarities to the literature on Bayesian persuasion. In Kamenica and Gentzkow (2011), the sender designs the information environment of the receiver in order to persuade the receiver to take the sender's preferred action. By contrast, in this essay, the consumer designs his own information environment in order to induce the seller to charge his preferred price. Just as in the literature on Bayesian persuasion, I make the assumption that the sender, here the consumer, can commit to an information structure. A discussion of the specific modeling choices and the robustness of the results is provided in Section 7.

A significant difference in the analyses is the following. In Kamenica and Gentzkow (2011) it is possible to identify each posterior belief with a value for the sender. By contrast, in the monopoly model, the consumer's value of a posterior belief depends on the price charged by the seller, and hence the full information structure. Consequently, the concavification approach from Aumann and Maschler (1995) that Kamenica and Gentzkow (2011) use in order to obtain their results is not applicable in the strategic environment that I study. The methods that are used to establish the results in this essay are mostly constructive.

**Related Literature** This essay contributes to the recent research on information design. This topic is addressed by various strands of literature, such as the literature on rational inattention, bounded rationality, and Bayesian persuasion.

A closely related paper is Gul et al. (2015), who study a model of an exchange economy. Consumers have limited cognitive abilities and can only choose coarse consumption plans. The authors introduce the concept of a *coarse competitive equilibrium* and find that the limited cognitive abilities of consumers lead to more price variation than in the standard competitive equilibrium. This property is a result of the new function of the market mechanism, which now also serves to allocate the agents' scarce attention. The way in which the behavioral limitations are modeled in this essay resembles the approach in Gul et al. (2015). In the competitive market analyzed in Gul et al. (2015), none of the agents has strategic influence. This is precisely the opposite of what is assumed in this essay; I consider a setting in which both agents have strategic influence. Hence, the role of rational inattention and prices is reversed to the one identified in Gul et al. (2015). In their paper, market prices serve to allocate attention, whereas in the present model the allocation of attention determines the induced price.

This essay contributes to the literature on information acquisition. Previous literature has mostly focused on how much information agents should acquire (Kessler, 1998; Shi, 2012; Bergemann and Välimäki, 2002), whereas the model studied here can be considered as one of flexible information acquisition:[1] I not only discuss how much information a consumer should acquire, but also identify which pieces of information are the most valuable to him. This interpretation links the analysis to Bergemann and Pesendorfer (2007). In an auction setting, they identify the seller-optimal information structure and selling mechanism. The seller has full flexibility in his choice of information structures and information is costless. The results in this essay identify the consumer-optimal information structure, if the seller best-responds with a revenue-maximizing mechanism.

This essay analyzes how the information environment of the consumer affects prices and the market outcome. A related question is studied in Bergemann et al. (2015). In their model, the consumers' valuation is private information. They describe the set of market outcomes that are achievable for different informational environments of the seller. Bergemann et al. (2015) identify an *outcome triangle* and show that any pair of consumer and producer surplus within this triangle is achievable. They find that the seller's expected surplus is bounded below by the monopoly profit. By contrast, I show that if the consumer can choose his information structure, then the seller's surplus may fall below the monopoly level.

This essay is also connected to the rational inattention literature. Starting with the seminal papers by Sims (1998, 2003, 2006), this literature studies the question of how an agent should optimally divide his attention if information is fully and freely available, but information processing is costly. Several papers analyze pricing models with rationally inattentive consumers. The most significant paper in this context is Matejka (2015). He studies a dynamic model with a consumer who is rationally inattentive to prices. He finds that rational inattention leads to rigid pricing, since such a pricing structure yields more prior knowledge and is easier to assess for the consumer.

The cost structures that are used in the rational inattention literature and in this paper differ strongly. Much of the rational inattention literature models information costs as a function of entropy reduction,[2] whereas I limit the number of categories that agents can distinguish and assume that all of these information structures have zero costs. Similar approaches to model cognitive limitations are taken by Wilson (2014) and de Clippel et al. (2014).

**Outline** The rest of the essay is organized as follows. The model is introduced in Section 2. In Section 3, an illustrative example is discussed. The main results are presented in Section 4 and Section 5. Section 4 covers the case of a consumer with information processing constraints. The unconstrained case is discussed in Section 5. The implications of optimal information processing and capacity constraints for the consumer's and the seller's profits are addressed in

[1] Some recent papers, such as Yang (2015b,a), also study flexible information acquisition in other contexts, for instance asset pricing and coordination problems.

[2] As discussed in Gentzkow and Kamenica (2014), it is more generally possible to define an information cost function based on a given *measure of uncertainty* (Ely et al., 2014). Woodford (2012) suggests an alternative cost function based on a different entropy-based measure.

Section 6. Section 7 provides a discussion of the specific modeling choices of the timing, and of the observability of the information processing structure. Section 8 and concludes. Unless stated otherwise, all proofs are in the appendix.

# 2 The Model

## 2.1 Payoffs and Information

A seller (she) wants to sell one object to a consumer (he). Both players are risk-neutral. The consumer's valuation for the object, $v$, is drawn from a distribution $F$ with support on the unit interval, $[0,1]$. The distribution $F$ is twice continuously differentiable, atomless, $F(0)=0$, with full support, $f>0$ on $(0,1)$, and mean $\mu_0$. The seller's marginal costs are zero. The consumer's true valuation is ex-ante unknown to both agents. The distribution $F$ and the seller's valuation are common knowledge.

If a consumer with valuation $v$ and the seller trade the object at price $p$, then the seller's payoff (*revenue*) is $r=p$, and the consumer's net payoff (*surplus*) is $u=v-p$.

**Information processing.** Information processing of the consumer corresponds to him choosing an information structure that determines how and what the consumer learns about his valuation for the object. An information structure,

$$\pi = \left(S,\, \{G(\cdot|v)\}_{v\in[0,\,1]}\right),$$

is given by a set of signal realizations $S \subseteq \mathbb{R}$ and a family of conditional distributions $\{G(\cdot|v)\}_{v\in[0,\,1]}$, where $G(s|v)$ is the probability that the consumer observes a signal realization less or equal to $s$ if his true valuation is $v$. The corresponding density or mass functions are denoted by $g\left(\cdot|v\right)$.

The consumer updates his beliefs according to Bayes' Rule. For a given information structure, each signal realization $s$ induces a posterior belief $F(\cdot|s) \in \Delta([0,\,1])$ of the consumer, given by

$$F(v|s) = \frac{\int_0^v g(s|v)f(v)\,\mathrm{d}v}{\int_0^1 g(s|v)f(v)\,\mathrm{d}v},$$

as well as a value estimate

$$V_s := \mathbb{E}\left[v|s\right] = \int_0^1 v\,\mathrm{d}F(v|s).$$

Moreover, an information structure $\pi$, induces a distribution $F_\pi \in \Delta\left([0,\,1]\right)$ over value estimates of the consumer.

For an information structure, *feasibility* requires that for every $v \in [0,\,1]$, $G(\cdot|v)$ is well-defined as a distribution function.

Bayesian updating implies that every information structure is *Bayes consistent*. That is, the induced posterior beliefs are consistent with the prior:

$$\mathbb{E}_S\left[F\left(v|s\right)\right] = F\left(v\right) \quad \forall\, v \in [0,\, 1]\,. \tag{1.1}$$

A consumer, who has no capacity constraints to process information can choose every information structure that satisfies feasibility.

**Capacity constraints.** In this essay, the consumer's information processing constraint is modeled as an upper bound $n \in \mathbb{N}$ on the number of signals or "categories" that he can distinguish. A capacity constrained consumer has only access to information structures with at most $n$ signal realizations. This set of information structures is called the *accessible set*. Information structures within this set are not differentiated by information costs. For all information structures in the accessible set, the information processing costs are zero, the costs for all other information structures can be considered to be infinite.

This approach to model cognitive limitations is similar to those in Gul et al. (2015) and Wilson (2014). Of course there are alternatives to model capacity constraints of agents, for example, by introducing information processing cost proportional to the entropy reduction, or some other measure of uncertainty.[3]

## 2.2 Strategies and Timing

**Action sets.** The consumer's action sets are the set of information structures $\mathcal{S}$, with typical elements $\pi$, that he can choose from, and the decision set $A = \{0,\, 1\}$, where $a = 1$ represents the case in which the consumer buys the object, and $a = 0$ the case in which the consumer makes no purchase. The action set of the seller is the set of prices $\mathbb{R}_0^+$.[4]

**Timing.** The consumer moves first. He chooses an information structure $\pi$, subject to his capacity constraint, and privately observes a signal realization $s \in S$. The seller observes the information structure of the consumer, but not the private signal realization. She then sets a price $p$, and the consumer decides whether to purchase the object at the price $p$ or not. The timing of the game is illustrated in Figure 1.1. The timing of the private signal and the price setting decision can be interchanged, or be simultaneous.

**Strategies and Solution Concept.** Every information structure $\pi$ induces a distribution over value estimates of the consumer.

For the consumer, a strategy is a tuple, $(\pi, \phi(\cdot,\cdot))$ of an information structure $\pi$ and a

[3]For examples of alternative modeling choices see Gentzkow and Kamenica (2014), Ely et al. (2014), Sims (1998, 2003) and Woodford (2012).

[4]One could more generally let the seller choose a selling mechanism. In the setting studied here, the result of Riley and Zeckhauser (1983) applies and the posted price mechanism is an optimal selling mechanism. Hence, for the sake of brevity of the exposition, I directly reduce the action set of the seller to prices. This is without loss of generality.

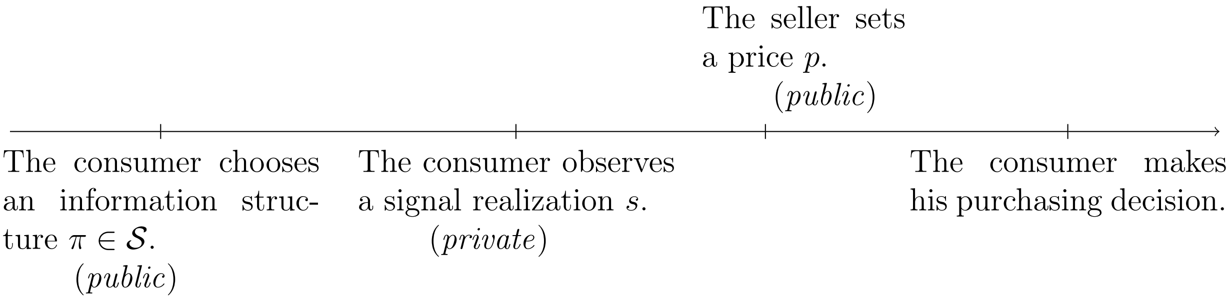



Figure 1.1: Timing in the monopoly model with a rationally inattentive consumer.

mapping from value estimates and prices to a purchasing decision,

$$\phi: [0,\, 1] \times \mathbb{R}_0^+ \to [0,\, 1]\,.$$

That is, $\phi(V,\, p)$ is the probability that the consumer will buy the object if his value estimate is $V$ and the price is $p$.[5]

A strategy for the seller is a mapping from the set of information structures, $\mathcal{S}$, that the consumer may choose from to the set of price distributions,

$$\sigma_S : \mathcal{S} \to \Delta([0,\, 1]).$$

Under strategy $\sigma_S$, if the seller observes that the consumer chooses information structure $\pi$, she chooses the price distribution $\sigma_S(\pi) \in \Delta([0,\, 1])$.

The solution concept is perfect Bayesian equilibrium.

The information structure $\pi$ is said to *induce* the price $p$, if $p$ is a best-response for the seller to the information structure $\pi$. Say that the information structure $\pi$ *induces* the expected surplus from trade $T(\pi)$, the seller's expected revenue $R(\pi)$, and the consumer's expected surplus $U(\pi)$, if these are the resulting values, if the seller plays a best-response to the information structure $\pi$, and the consumer best-responds to this.

# 3 Illustrative Example: The Uniform Prior Case

In order to fix ideas and to illustrate the fundamental effects in the monopoly pricing model with a rationally inattentive consumer, I start with an example. Throughout this section, the consumer's valuations are assumed to be uniformly distributed on the unit interval, $v \sim U\,[0,\, 1]$.

**Benchmarks: Uninformed and fully informed consumer.** The two relevant benchmarks are the case in which the consumer has no information about his valuation, and the case in

[5]In the linear setting with risk-neutral agents considered in this essay, the distribution over value estimates captures all information about $\pi$ that is relevant for the consumer's purchasing decision and the seller's pricing decisions. This observation is used in order to reduce the problem, and to simplify the strategy sets that have to be considered. Similar reductions are used, for example, by Kamenica and Gentzkow (2011) and Caplin et al. (2014). They reduce the problem to posterior beliefs. The model in this essay considers risk-neutral agents and linear utilities. Hence, a reduction of the problem to value estimates is possible.

which he privately knows his true valuation.

In the first case, both the consumer and the seller are uninformed. They only know the distribution of the consumer's valuation. In equilibrium the seller charges a price equal to the expected value of the prior distribution $p = \mu_0 = \frac{1}{2}$, and the consumer always buys the good. Trade is efficient, that is, the potential gains from trade are fully realized. The seller extracts all surplus from trade. Her expected revenue is $R^{(0)} = \frac{1}{2}$, the consumer obtains zero surplus.

The latter case, in which the consumer is fully informed and privately knows his true valuation, is the standard monopoly pricing problem. In equilibrium, the seller will charge the monopoly price $p^M = \frac{1}{2}$, and the consumer only buys the good if his true valuation is greater (or equal) to the price.[6] A consumer with a lower valuation is excluded from trade, and hence trade is not efficient. The resulting expected surplus from trade is $T^M = \frac{3}{8}$, the seller's expected revenue is $R^M = \frac{1}{4}$, and the consumer's expected surplus is $U^M = \frac{1}{8}$.

The benchmark cases with an uninformed and a fully informed consumer are illustrated in Figure 1.2.

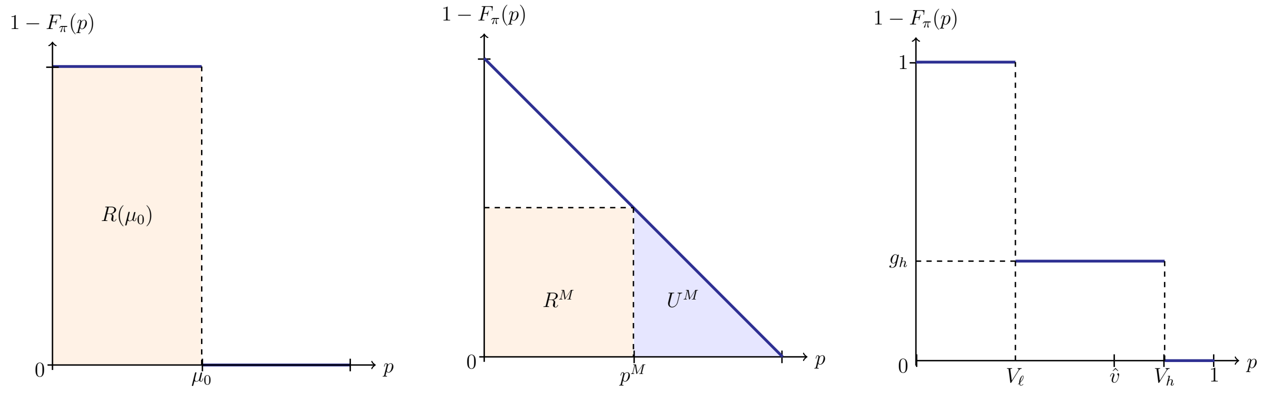
**Monotone partitional two-signal information structure.** Suppose that information is fully and freely available, and that the consumer has to decide how to process this information. Consider the case in which the consumer can only distinguish two categories, one of which he interprets as "good" and the other one as "bad". If the consumer has full flexibility in designing these two categories, then how should he define them?

If the consumer forms two categories, he can condition his purchasing decision only on these categories and the realized price. Each category induces a willingness to pay of the consumer, that is, a region of prices for which he would buy the good. The information processing choice of the consumer can be modeled as the consumer observing a signal realization that informs him in which of the two categories his true valuation falls. The high signal realization $s_h$ indicates that the consumer's true valuation is in the good category, and hence increases the consumer's willingness to pay. The induced value estimate $V_h$ is larger than the prior mean. By contrast, a realization of the low signal $s_\ell$ decreases the consumer's willingness to pay, $V_\ell \leq \mu_0$.

The resulting interim demand function that the seller faces is a step function. For prices smaller or equal to the willingness to pay of a consumer who observes a low signal, the probability of trade is one. Upon passing this value, the probability of trade drops to $g_h$, which is the probability that the high signal realizes. The probability of trade is zero for prices above the value estimate of a consumer who observes a high signal. This is illustrated in Figure 1.2(c).

The seller's objective is to maximize her expected revenue. It is straightforward, that the seller never charges a price on the flat, inelastic region of the interim demand curve. Hence, the seller's problem is to decide whether to charge the *inclusive price* $V_\ell$ and to sell with probability one, or to charge the *exclusive price* $V_h$ and to only sell to a consumer who receives a high signal, that is, with probability $g_h$. An information structure, respectively choice of categories, thus

[6] It is irrelevant whether the consumer buys the good or not if he is indifferent, since the event that the consumer's valuation is equal to the price is a zero-probability event.




(a) Uninformed consumer.

(b) Fully informed consumer.

(c) Two-signal information structure.

Figure 1.2: Demand function, expected seller's revenue and consumer's surplus for (a) an uninformed consumer, and (b) a fully informed consumer. Demand function induced by a two-signal information structure (c).

determines both, the possible price realizations $V_\ell$ and $V_h$, as well as the corresponding demand or probability of trade for the exclusive price, $g_h$.

The consumer only obtains a positive surplus if the seller charges an inclusive price.[7] Hence, the only way in which the consumer can secure information rents is to choose an information structure that induces the seller to charge an inclusive price. For a given information structure $\pi$, the seller charges the inclusive price, $V_\ell(\pi)$, if it yields a weakly higher revenue than the exclusive price, $V_\ell(\pi) \geq g_h(\pi)V_h(\pi)$.

The consumer's problem reduces to:

$$\max_{\pi \in \mathcal{S}^{(2)}} \{(V_h(\pi) - p^*(\pi)) \cdot g_h(\pi) \ s.t. \ p^*(\pi) = V_\ell(\pi)\}, \tag{1.2}$$

where $\mathcal{S}^{(2)}$ is the set of all feasible, two-signal information structures, and $p^*(\pi)$ is the revenue-maximizing price for the seller given information structure $\pi$.

For simplicity, assume that the consumer chooses a monotone partitional information structure. That is, he splits the interval of true valuations into two subintervals, such that the high signal realizes whenever the true valuation is in the upper subinterval, otherwise the low signal realizes. Such an information structure $\pi_{\hat{v}}$ is determined by the threshold $\hat{v}$, which is the boundary of the two subintervals.

For example, a used car could fall in the category "good" if its mileage is below 25000, or if it is less than three years old. Apartments could be categorized based on the number of rooms, or certain aspects of the location. The criteria do not need to be one dimensional as long as they reduce to two categories.[8] The induced willingnesses to pay of the consumer for a

[7] Given an exclusive price, the consumer only buys the object if the high value estimate realizes. In this case, the consumer obtains the object, but at a price that is equal to his expected valuation; his expected surplus is zero.

[8] For example, the category "good" could consist of apartments in a preferred location with at least three

high or low signal realization depend on the threshold that determines the categorization. For instance, if at least five rooms are required to sort an apartment in the "good" category, then the willingness to pay of a consumer is higher for both of the signal realizations, compared to the case in which all apartments with at least three rooms are classified as "good". Formally, the high and the low value estimate, are both increasing in the threshold.

It is easy to see that the expected revenue of the seller from charging the inclusive price is increasing in the threshold $\hat{v}$. The inclusive price is increasing in the threshold and the probability of trade is constant and equal to one. The effect of a threshold shift on the expected revenue from the exclusive price is less obvious. The exclusive price is increasing in the threshold, whereas the probability that the high signal realizes and hence the probability of trade is decreasing in the threshold.

If the seller charges an exclusive price, then the consumer's expected surplus is zero, and the seller can extract all gains from trade.[9] Hence, the seller's expected revenue from charging the exclusive price is equal to the gains from trade in the region where the high signal realizes. It follows that the expected revenue from charging the exclusive price is decreasing in the threshold $\hat{v}$.

It follows that there exists a critical threshold at which the seller is indifferent between charging the inclusive or the exclusive price, $V_\ell(\pi_{\hat{v}}) = g_h(\pi_{\hat{v}})V_h(\pi_{\hat{v}})$. For all thresholds above the critical one, the seller charges the inclusive price, else he charges the exclusive price. For a uniform prior, the critical threshold is $\frac{1}{2}\left(\sqrt{5}-1\right)$.

The consumer's expected surplus, if the seller charges an inclusive price, is

$$U\left(\pi_{\hat{v}}\right) = \mu_0 - p^*\left(\pi_{\hat{v}}\right), \tag{1.3}$$

which is decreasing in the price. It follows that the equilibrium information structure is the one that induces the lowest price among all information structures that induce the seller to charge an inclusive price (cf. 1.2). The equilibrium threshold is equal to the critical threshold $v^* = \frac{1}{2}\left(\sqrt{5}-1\right)$. In equilibrium, the seller charges the inclusive price, but is indifferent between charging the inclusive or the exclusive price. The equilibrium two-signal information structure is illustrated in Figure 1.3.

The identification of the equilibrium information structure in the example was based on the assumption that the set of accessible information structures consists of all monotone partitional two-signal information structures. Certainly, the question arises whether this information structure is optimal for the consumer among larger classes of accessible information structures. Is a partitional information structure optimal, or can the consumer benefit from noise in the

rooms, and all apartments with at least five rooms.

[9]Indeed,

$$R(V_h) = (1 - F(\hat{v})) \cdot \mathbb{E}\left[v|v \geq \hat{v}\right] = \int_{\hat{v}}^{1} v \, \mathrm{d}F(v).$$

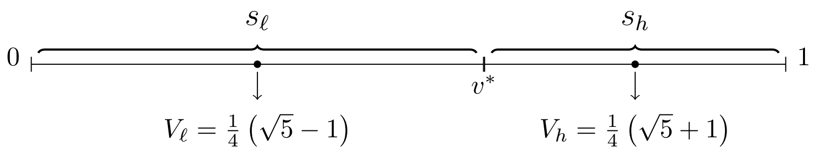



Figure 1.3: Equilibrium two-signal information structure, $\pi_2^*$, with threshold $v^* = \frac{1}{2}\left(\sqrt{5}-1\right)$. The high signal, $s_h$, realizes if the consumer's valuation is in the good category, otherwise the low signal, $s_\ell$, realizes.

information? And are two signal realizations enough, or can the consumer strictly benefit from having access to information structures with more signal realizations? These and further question will be answered in the course of the general analysis in the following sections.

# 4 Finite Information Structures

After the illustrative example, this section returns to the analysis of the general model. Suppose that there is a limit on the maximal number, $n \in \mathbb{N}$, of signal realizations or categories that the consumer can distinguish. This constraint may for example represent the limited cognitive ability of the consumer to process the available information. In this section, the equilibrium information structure is identified, which determines how a capacity constrained consumer should optimally process or categorize information. First, some preliminary observations are discussed, then properties are derived that any equilibrium information structure must satisfy, if it exists. These properties are then used in order to reduce the problem of finding an equilibrium information structure, and to prove existence.

For any information structure $\pi$, with $n$ signal realization $s_1, \dots, s_n$, these signals are indexed such that the induced value estimates, $V_i := \mathbb{E}\left[V|s_i\right]$, are arranged in an ascending order, $V_1 \leq \dots \leq V_n$.

## 4.1 Preliminary Observations

To begin the analysis, some observations about optimal purchasing, price-setting and information processing decisions are established.

**Consumer's purchasing decision.** It is immediate that for any given information structure $\pi$, signal realization $s_i$, and price $p$, a best-response for the consumer is to buy the object if and only if his value estimate $V_i$ is greater or equal to the price,

$$\phi^*(V_i, p) = 1 \;\Leftrightarrow\; V_i \geq p, \quad \forall i \in \{1, \dots, n\}. \tag{1.4}$$

**Price-setting by the seller.** For a given information structure $\pi = \left(S,\, \{G(\cdot|v)\}_{v\in[0,1]}\right)$, the situation for the seller is as if he faces a consumer whose valuation is drawn from distribution $F_\pi$ with support

$$\operatorname{supp}\left(F_\pi\right) = \{V_i = \mathbb{E}\left[v|s_i\right] :\, s_i \in S\}.$$

The seller charges the price $p$ that maximize her expected revenue, $R(p) = p \cdot (1 - F_\pi(p))$. As discussed in the example in Section 3, the induced interim demand function is a step function, and the seller charges a price equal to one of the value estimates of the consumer.

**Lemma 1.1** (Selling Mechanism).
*For an information structure $\pi$ that induces the distribution of value estimates $F_\pi$, a best-response of the seller is to sell the object by a posted-price mechanism with price*

$$p^*(\pi) = \arg\max_{p \in [0,\,1]} \{p \cdot (1 - F_\pi(p))\}.$$

*This price is equal to one of the induced value estimates of the consumer, $p^*(\pi) \in supp\ (F_\pi)$.*

The expected revenue that the seller can extract by charging the price $V_i$ is denoted by $R(V_i)$. The seller sets a price equal to the value estimate that maximizes her expected revenue.

It is useful to classify prices and distinguish between *exclusive* and *inclusive* prices. For a given information structure $\pi$, say that price $p$ is an *exclusive* price, if the consumer buys the object only if the highest value estimate realizes. Price $p$ is *(partially) inclusive*, if there are at least two signal realizations that induce distinct value estimates, and for which the consumer will buy the object at the price $p$. A price is called *fully inclusive*, if, under the given information structure, the consumer always buys the good at that price, irrespective of the signal realization. If the seller charges an exclusive price, only a consumer with the highest value estimate buys the good, and demand is low. By contrast, for a fully inclusive price the probability of trade is one. Given that marginal costs are zero, this implies that any *fully inclusive* price yields efficient trade, that is, potential gains from trade are fully realized.

**Information processing by the consumer.** The consumer faces the following trade-off in his choice of an information structure: His information processing decision does not only determine the distribution over his value estimates, but also the interim demand function that the seller faces, and hence the price.

For the problem of identifying the equilibrium information structure, the following complication arises. Any change in the information structure may induce the seller to switch from charging an inclusive price to charging an exclusive price. Hence, a seemingly small effect on the distribution and realizations of the value estimates, may have a large effect on the consumer's expected surplus. The strategic interaction among the consumer and the seller results in discontinuities in the consumer's problem.

As discussed in the example of Section 3, the consumer only obtains a positive surplus if the seller charges an inclusive price. It follows that the consumer will always choose an information structure $\pi$ that induces the seller to charge an inclusive price.

**Lemma 1.2** (Inclusive Prices).
*The consumer chooses an information structure that induces the seller to charge a (partially or fully) inclusive price.*

## 4.2 Seller-Indifference

A central result in the characterization of the equilibrium information structure is a property that I refer to as *seller-indifference*. This property requires that, if the equilibrium information structure induces the seller to charge a price equal to the $i^{th}$-lowest of the induced value estimates, $V_i$, then the seller is indifferent between charging this price and charging a price equal to any of the higher value estimates.[10]

**Proposition 1.1** (Seller-Indifference).
*Suppose that $\pi$ is an equilibrium information structure that induces the seller to charge the price $p = V_i$. Then, the seller's revenue $R(V_i)$, is equal to the revenue that he could extract by charging a price equal to any of the higher value estimates in the support of $F_\pi$:*

$$R(V_i) = R(V_j) \quad \forall j \geq i. \tag{1.5}$$

*The seller is indifferent between charging any of these prices.*

For any finite information structure $\pi$, the optimal price for the seller is equal to one of the value estimates induced by $\pi$ (Lemma 1.1). Roughly put, the seller has to choose a price-level for his product, where the set of possibly optimal price-levels is given by the support of $F_\pi$. Proposition 1.1 establishes that the distribution over value estimates induced by the equilibrium information structure is an equal revenue curve above the equilibrium price charged by the seller. This means that charging a price equal to any of the value estimates above the equilibrium price yields the same expected revenue for the seller. By leveling the seller to a revenue level, the consumer leaves her with just enough surplus to guarantee that she does not want to deviate to a higher price-level.

Let me now sketch the idea of the proof. The result is proven by an indirect argument. For any information structure $\pi$ that does not satisfy the seller-indifference property (1.5), a new information structure, $\widetilde{\pi}$, is constructed that makes the consumer better off.

Suppose that $\pi$ is an information structure that does not satisfy (1.5). Let $k > i$ be an index such that the seller strictly prefers charging the price $V_i$ over charging the price $V_k$, that is, $R(V_i) > R(V_k)$. The idea is to construct a new information structure as follows: Mass is taken from the upper part of the support of $F(\cdot|s_i)$. This will reduce the value of $V_i$, and hence the seller's expected revenue if he charges the price induced by signal $s_i$, $p = \mathbb{E}[v|s_i]$. Similarly, mass is taken from the supports of posterior distributions induced by signals $s_j$, for which the seller is initially indifferent between charging the prices $V_j$ and $V_i$. The distribution of mass among signals $s_j$ with $j \geq i$ is re-adjusted such that the seller is still induced to charge a price equal to the $i^{th}$-lowest value estimate $\mathbb{E}[v|s_i]$. In particular, mass will be added to signal

[10] This result is reminiscent of the indifference result of Proposition 5 in Kamenica and Gentzkow (2011). In the interpretation of their leading example, this result states that under the sender-optimal signal the judge is certain of the innocence of the defendant if he chooses the action "*acquit*", and indifferent if he chooses "*convict*".

$s_k$, which increases the seller's expected revenue from charging a price equal to $\mathbb{E}[v|s_k]$, the posterior estimate induced by signal $s_k$.

Such a construction has the following properties: (1) The probability of trade and the expected surplus from trade remain the same, since mass is only re-distributed among types that participate in trade. (2) The price charged by the seller decreases. The consumer's expected surplus is given by the difference of expected total surplus and expected revenue.

$$U = \mathbb{P}(trade) \cdot \mathbb{E}[v|trade] - \mathbb{P}(trade) \cdot p.$$

Hence, under the new information structure, the consumer's expected surplus will be higher than before. The formal details about the existence and the construction of such an information structure are relegated to the appendix.

## 4.3 Optimal Information Processing Induces Efficient Trade

A central property of the equilibrium information structure is that it induces efficient trade.[11]

**Proposition 1.2** (Efficient Trade)**.**
*Any finite equilibrium information structure induces efficient trade. All gains from trade are realized, and the exclusion region is empty*

The intuition for this result is as follows. If the potential gain from trade are not fully realized, then there are consumer types that are excluded from trade, but for whom it would be profitable to trade. In the present model, the consumer has a lot of power. He can design his information environment and thus the demand curve that the seller faces. Hence, he can influence the seller's pricing behavior. In particular, the consumer can switch to an information structure that yields efficient trade and makes him better off. Under the equilibrium information structure, the consumer does not pay attention to information that would separate low valuations. This means that he effectively commits to buy at an intermediate price that may be higher than his true valuation. He thus offers the seller a higher probability of trade at this intermediate price. By dividing his attention optimally, the consumer can induce the seller to offer better terms of trade in return for the increased probability of trade. The seller charges a lower price and the consumer obtains a higher expected surplus. The positive effect of the additionally realized gains from trade reverberates back to the consumer.

Let me sketch the proof of Proposition 1.2; the details are in the appendix. The proof consists of two steps. First, the problem is reduced by showing that any equilibrium information structure is outcome-equivalent to an information structure for which at most one value estimate lies in the exclusion region. Here, two information structures $\pi$ and $\widetilde{\pi}$ are said to be *outcome equivalent*, if the realized price, the expected surplus from trade, the seller's expected revenue, and the consumer's expected surplus induced by $\pi$ and $\widetilde{\pi}$ coincide.

[11] Recall that, by assumption, the consumer's valuation is always greater or equal to the seller's cost. In this case, efficient trade means that trade must occur with probability one.

Suppose that there is more than one value estimate in the exclusion region, which means that there are various signal realizations that will result in the consumer not buying the good. For a given price, it would not make a difference if the consumer had a coarser perception of these values, and would obtain only one signal that informs him that he should not buy the good. From the seller's perspective, such an adjustment of the information structure reduces the dispersion in the exclusion region part of the demand curve. It creates a single mass point on the expected value of the types in the exclusion region. This change in the demand curve either has no effect on the seller's pricing decision, or incentivizes the seller to switch to charging a price equal to the expected value of the types in the former exclusion region. In the latter case, trade is efficient and the consumer is better off.

The second step is to prove that for any information structure that induces a partially inclusive price, there exists an information structure that induces a fully inclusive price, and yields a weakly higher expected surplus for the consumer.

## 4.4 Monotone Partitional Information Structures

The previous findings have established properties (*seller-indifference* and *efficient trade*) of the outcome induced by the equilibrium information structure. But how should the consumer optimally process information? As the next result shows, any finite equilibrium information structure is monotone partitional. The consumer only learns in which range his true valuation falls.[12] He cannot profit from adding noise to the signal.

**Proposition 1.3** (Equilibrium Structures are Monotone Partitional).
*For every $n \in \mathbb{N}$, any equilibrium information structure, $\pi_n^*$ is monotone partitional.*

## 4.5 Equilibrium Existence

Building on the necessary properties for an equilibrium information structure that were identified in the previous sections, equilibrium existence is established. First, the problem is reduced by using that any equilibrium information structure must be monotone partitional, induce seller-indifference, and efficient trade. Then, using this simplification of the problem, equilibrium existence is proven.

**Theorem 1.1** (Equilibrium Existence).
*For every $n \in \mathbb{N}$ there exists an essentially unique equilibrium. The equilibrium is determined by the equilibrium information structure $\pi_n^*$, which is characterized by the following properties:*

*(i)* $\pi_n^*$ *is* monotone partitional*,*

*(ii) induces* efficient trade*, and*

*(iii) guarantees* seller-indifference*.*

[12] This result is reminiscent of the optimality of coarse information structures in the literature on communication or strategic information transmission based on the seminal paper by Crawford and Sobel (1982). As pointed out in Sobel (2012) complexity in communication is an alternative explanation for limited communication.

## 4.6 More Signals are Better

It is obvious that having access to more information structures increases the choice set of the consumer, and hence he will be weakly better off. But will he strictly benefit from having access to information structures with more signal realizations? Or is there a maximal number of signal realizations such that the consumer cannot profit from additional signal realizations?

For a given price, the consumer faces a binary decision – whether to buy the good or not.[13] For this decision, it suffices for the consumer to know whether his valuation is above or below the price. He can obtain this information with a binary information structure.

The situation is different, if the consumer cannot condition his information processing choice on an observed price, and has to take into account that this choice will influence the price charged by the seller. In this case, having access to information structures with more signal realizations allows the consumer to better react to different prices that the seller may charge. Hence, the consumer can better influence the price setting strategy of the seller.

The following example illustrates how the consumer can use additional signal realizations to secure a strictly higher profit.

**Example 1.1** (Two signal realizations are not enough)**:**
Reconsider the uniform prior example discussed in Section 3. The equilibrium monotone partitional two-signal information structure $\pi_2^*$ was identified (illustrated in Figure 1.3). By Theorem 1.1, this is indeed the equilibrium two-signal information structure.

Suppose now that the consumer has access to one more signal realization and can distinguish three categories, say "good", "intermediate" and "bad". Can the consumer strictly benefit from the enlarged set of feasible information structures and improve upon the case with only two-signal realizations?

The consumer can use the additional signal to identify an interval of intermediate values, such that this interval covers true valuations that have previously resulted in a good signal, as well as some that have resulted in a bad signal. Figure 1.4 illustrates such an information structure, $\widetilde{\pi}_3$.

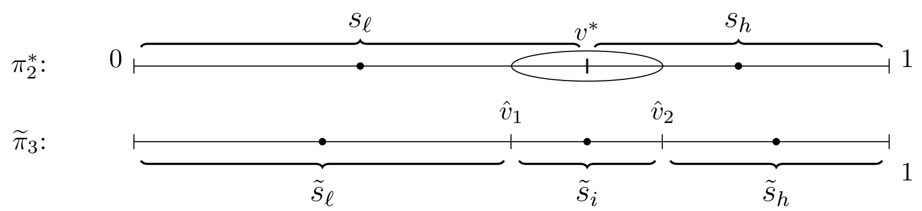



Figure 1.4: Illustration of the use of an additional signal.

[13]In Kamenica and Gentzkow (2011) it is shown that one can restrict attention to signals with at most as many signal realizations as available actions. This feature is an implication of the revelation principle as discussed in Myerson (1997). In the present model, the consumer's action set is binary. This would suggest that two signal realizations are enough. However, the seller can set any real-valued price, which implies that the action set that the consumer has to consider when choosing the information processing structure is large.

Under this new information structure, good and bad signals both provide stronger evidence that the true valuation is high, respectively low. Hence, the value estimate induced by the good (bad) signal increases (decreases). For the price induced by the bad signal, the demand remains the same whereas the price that the seller can charge decreases. Hence, the seller's expected revenue decreases. For the good signal, the price increases and the probability of trade decreases. The expected revenue that the seller can extract by charging a price equal to the value estimate induced by the good signal is equal to the realized surplus from trade for values in the good category. Hence, if this category gets smaller, the corresponding expected revenue of the seller decreases.

What can be said about the revenue that the seller can extract by charging a price equal to the value estimate induced by the new, intermediate signal? Compare the seller's revenue if she adopts a price equal to the intermediate value estimate to the seller's revenue under the optimal two-signal information structure. Given the seller-indifference property for the latter, one can consider the seller's revenue if she charges the exclusive price, induced by the original good signal. There is a *demand effect* and a *price effect*. From the perspective of the seller, the demand effect is positive: the probability to sell if she charges a price equal to the intermediate value estimate is higher than for the exclusive price, induced by the original good signal. By contrast, the price effect is negative: the price is lower than the original exclusive price.

The demand effect only depends on the probability that the true valuation falls in the intermediate or good category, but not on the realization of the good category. The price effect, by contrast, depends on how the region that yields an intermediate or a good signal is split between the intermediate and the good categories. All else equal, the larger the fraction of the good category, the stronger is the price effect. If the good category is sufficiently large, then the seller's expected revenue from charging a price equal to the intermediate value estimate is below the revenue level of the two-signal case. Hence, for appropriately chosen categories, under the new information structure, trade is still efficient but the seller obtains a smaller share of the total expected surplus. The consumer is strictly better off.

The feature that the consumer strictly benefits from having access to more signal realizations is a general property. This result is formally established in Proposition 1.4. For $n \in \mathbb{N}$, let $p_n^*$ be the minimal price that can be induced as a fully inclusive price by an information structure with at most $n$ signal realizations. Any finite equilibrium information structure induces efficient trade, and hence the equilibrium $n$-signal information structure induces the minimal price $p_n^*$. The question whether the consumer strictly benefits from having access to information structures with more signal realizations is equivalent to the question whether the sequence $p_n^*$ is strictly decreasing in $n$.

**Proposition 1.4** (More Signals are Better).
*The consumer strictly profits from having access to information structures with more signal realizations. Equilibrium prices $p_n$ are strictly decreasing in $n$, and the consumer's expected surplus $U_n^*$ is strictly increasing in $n$.*

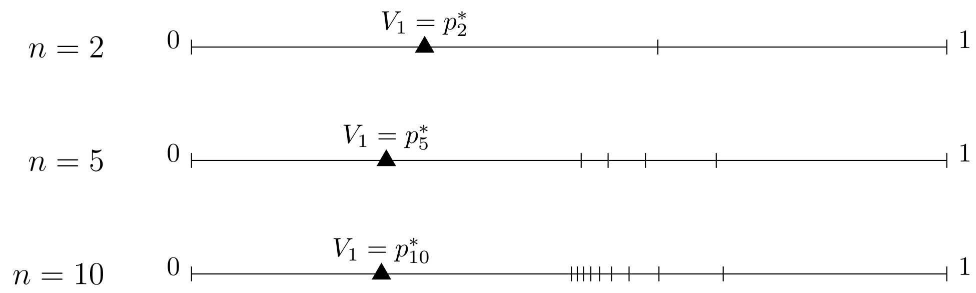



Figure 1.5: Thresholds and induced prices for the equilibrium $n$-signal information structures.

**Evolution of optimal information structures and thresholds with $n$**

How should a capacity constrained consumer optimally allocate his attention? As established in Theorem 1.1, the consumer chooses a monotone partitional information structure. Still, the question remains, which pieces of information are the most valuable for the consumer. Should the consumer pay more attention to information about low or about high valuations?

For the uniform prior example, the evolution of optimal information structures and prices is illustrated in Figure 1.5.[14] As can be seen, if more signal realizations are available, the information structure gets finer around valuations close to the lowest threshold. This is the threshold that determines the value estimate that corresponds to the induced price. The consumer pays more attention to values closer to the threshold that is relevant for the price.

# 5 The Unconstrained Case

In the absence of any information constraints or costs, the consumer can choose freely how to process the available information. Which pieces of information should the consumer acquire in this case? Should he learn his valuation perfectly, or are there benefits from remaining partially uninformed?

The main result of this section shows that it is in the consumer's best interest to remain partially uninformed. Every equilibrium information structure pools all values below a certain threshold. The consumer benefits from committing to ignore information that would separate low valuations.

**Theorem 1.2** (Unconstrained Equilibrium).
*Without information constraints, there exists a threshold* $0 < \underline{v} < 1$ *such that every equilibrium information structure is outcome-equivalent to the following information structure* $\pi^*$*:*

*The information structure* $\pi^*$ *pools all values below the threshold* $\underline{v}$*, and the induced distribution of value estimates,* $F_{\pi^*}$*, is an equal revenue distribution.*

Why is it optimal for the consumer not to get perfectly informed but to pool an interval of low values into one signal? By obtaining only a coarse perception about low values, the consumer can induce the seller to charge a lower price, which increases the consumer's expected surplus.

[14] A more detailed and formal discussion is provided in Appendix B.

To obtain intuition for this result, consider the benchmark, in which the consumer knows his true valuation. In this case, the seller charges the monopoly price. Hence, the seller excludes low values from trade in order to maximize her expected revenue. This means that potential gains from trade are not fully realized. Moreover, the consumer cannot obtain information rents for valuations that fall in the non-trade region. An information structure that pools low valuations creates a mass point on an intermediate value estimate. For a sufficiently large pooling region, the seller will be induced to charge a price equal to the value estimate of this pooling region. This information structure induces efficient trade, and the induced price is below the monopoly price in the case with a fully informed consumer. The consumer can secure more information rents, and hence has a higher expected surplus.

In the unconstrained case, multiplicity of equilibrium information structures, respectively equilibria, arises. The optimal information structure $\pi^*$ identified in Theorem 1.2, is the limit of the finite equilibrium information structures. In the absence of information processing constraints, the number of signal realizations is unconstrained. Consequently, the number of signal realizations that are used to obtain information about some interval of the valuation space does not affect the number of the signals that are available to acquire information about other valuations. Hence, instead of pooling types within intervals in which the prior distribution has positive virtual valuations, the consumer can also learn his true valuation perfectly. Such an adjustment of the information structure, which is only possible if the consumer has no information constraints, does not affect the equilibrium outcome.

In order to illustrate the properties of the equilibrium information structures in more detail, I discuss three examples. Each of these examples illustrates a typical equilibrium information structure. The discussion shall provide some intuition how specific features of optimal information structures depend on properties of the prior distribution.

**Example 1.2:** Consider the uniform prior case, $v \sim U[0, 1]$. In the benchmark with a fully informed consumer, the monopoly price is $p^M = \frac{1}{2}$. As just discussed, in the absence of information processing constraints or costs, there are multiple equilibria. Two examples of equilibrium information structures are the following:

a) The information structure $\pi^*_{max}$ pools all types in the interval $\left[0, \frac{1}{2}\right)$, and is perfectly informative on the interval $\left[\frac{1}{2}, 1\right]$. This information structure is illustrated in Figure 1.6.

b) The information structure $\pi^*_{\infty}$ pools all types in the interval $\left[0, \frac{1}{2}\right)$. On $\left[\frac{1}{2}, 1\right]$, and the induced distribution of value estimates is and equal revenue distribution. This information structure is the limit of the finite equilibrium information structures,

$$\pi^*_{\infty} = \lim_{n \to \infty} \pi^*_n.$$

The information structure $\pi^*_{max}$ is the finest equilibrium information structure whereas $\pi^*_{\infty}$ is the coarsest equilibrium information structure. The distributions of value estimates induced by these two information structures are illustrated in Figure 1.7.

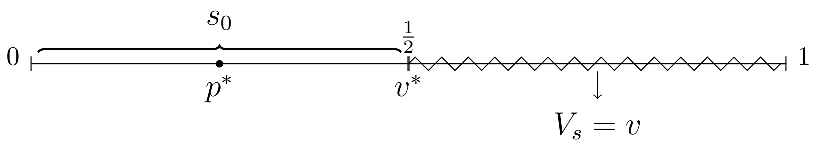



Figure 1.6: The equilibrium information structure $\pi^*_{max}$ for the uniform prior case, $v \sim U\,[0,\,1]$.

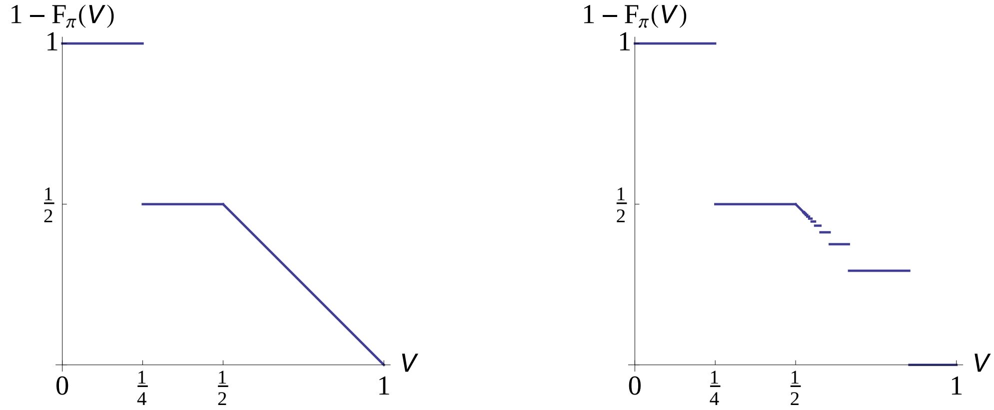



Figure 1.7: Distributions over value estimates $F_{\pi^*}$ that are induced by the equilibrium information structures $\pi^*_{max}$ and $\pi^*_\infty$, for the uniform prior case, $v \sim U\,[0,\,1]$.

In the uniform prior case, $\mathbb{E}\left[v|v \leq p^M\right] = \frac{1}{4} = p^M\left(1 - F(p^M)\right)$. Hence, pooling of the types within the region $\left[0,\,\frac{1}{2}\right)$, makes the seller indifferent between charging the fully inclusive price $p^* = \frac{1}{4}$ and the monopoly price, which is the price that maximizes the seller's revenue among all prices in $\text{supp}\,(F_{\pi^*})\setminus\{p^*\}$. In equilibrium, the seller will charge the fully inclusive price.

**Example 1.3:** Suppose now, that the valuation of the consumer is distributed on $[0,\,1]$ with the linearly decreasing density $f(v) = 1-2v$. For this distribution, pooling all valuations smaller than the monopoly price, induces the seller to charge the fully inclusive price. Moreover, she strictly prefers this price over the monopoly price, $\mathbb{E}\left[v|v \leq p^M\right] > R^M$. It thus suffices to pool a smaller interval $[0,\,\hat{v}]$ of low valuations to make the seller indifferent between charging the fully inclusive price and the monopoly price. Notice, that the seller prefers to charge these two prices over any price in the region $\left(\hat{v},\,p^M\right)$. Hence, there is still some slack in this information structure.

For the given prior distribution, any equilibrium information structure $\pi^*$ pools all types within an interval $[0,\,\underline{v})$, with $\underline{v} \leq \hat{v} < p^M$, and has a "*sweeping up*" region $[\underline{v},\,\overline{v}]$. In this region the information structure induces a distribution over value estimates with constant virtual valuation equal to zero. The *sweeping up function* $H$ on $[\underline{v},\,\overline{v}]$ is given by:

$$H(V) = 1 - \frac{\underline{v}}{V}\left(1 - F(\underline{v})\right).$$

Above $\overline{v}$, the information structure can be perfectly informative. This equilibrium information structure $\pi^*_{max}$ is illustrated in Figure 1.8. It induces efficient trade, $T^* = \frac{1}{3}$, and the induced expected revenue of the seller is below the monopoly revenue $R^* < R^M$.

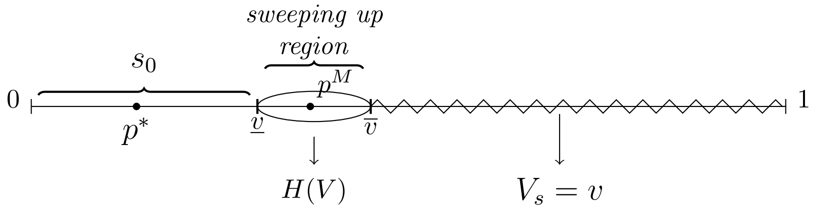



Figure 1.8: Equilibrium information structures $\pi^*_{max}$ for a decreasing prior density.

**Example 1.4:**
Suppose that the consumer's valuations are distributed according to the beta-distribution $v \sim \beta\left(1, \frac{1}{2}\right)$. For this distribution, the monopoly price is $p^M = \frac{2}{3}$. Moreover, it holds that $\mathbb{E}\left[v|v \leq p^M\right] < R^M$. Under the information structure that pools the values in the interval $\left[0, p^M\right)$ and is perfectly informative otherwise, the seller still charges the monopoly price. By moving the threshold of the pooling region up, the value estimate of the values within this region increases. This increases the revenue that the seller can extract by charging the fully inclusive price. Moreover, some types for which the marginal revenue is positive are now included in the pooling region. Hence, the revenue that the seller can extract by charging the optimal price in the separating region decreases. The critical threshold $v^*$ at which the seller is indifferent between charging the partially inclusive price and charging the fully inclusive price determines the equilibrium information structure. The seller's expected revenue induced by this information structure is smaller than the monopoly revenue. This equilibrium information structure is illustrated in Figure 1.9.

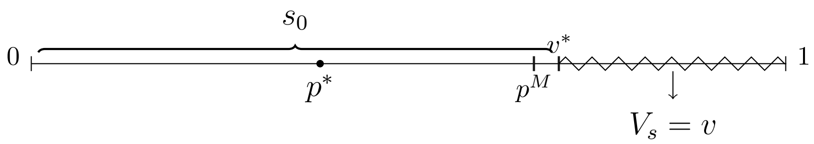



Figure 1.9: The equilibrium information structure $\pi^*_{max}$ for, $v \sim \beta\left(\frac{1}{2}, 1\right)$.

# 6 Seller's Revenue and Consumer's Surplus

This section discusses the implications of a rationally inattentive consumer in a monopoly pricing problem for welfare, the seller's expected revenue and the consumer's expected surplus. In order to analyze these effects, the equilibrium outcomes that are induced by the equilibrium information structures identified in Section 4 and Section 5 are compared to the outcome in the monopoly pricing problem, in which the consumer is privately informed about his true valuation. The latter situation corresponds to the benchmark with a fully informed consumer (cf. Section 3).

Given the discussion in the previous sections, the effect on aggregate welfare is straightforward. Every equilibrium information structure induces efficient trade. By contrast, under monopoly pricing for a fully informed consumer, some types are excluded from trade, and hence not all possible gains from trade are realized. Implementing the equilibrium information structure for a rationally inattentive consumer thus improves welfare.

The implications of monopoly pricing for a rationally inattentive consumer for the seller's expected revenue and the consumer's expected surplus are less obvious. Denote the expected surplus of the consumer and expected revenue of the seller induced by the equilibrium information structure by $U_n^*$ and $R_n^*$ for the finite case, and by $U_\infty^*$ and $R_\infty^*$ in the unconstrained case. In Proposition 1.4 it was shown that the consumer strictly profits from having access to more signal realizations. The following corollary generalizes this result.

**Corollary 1.1.** *The expected surplus of the consumer $U_n^*$ is increasing in $n$, whereas the expected revenue of the seller $R_n^*$ is decreasing in $n$. Both sequences approach the respective values of the unconstrained case in the limit.*

$$\lim_{n\to\infty} U_n^* = U_\infty^*, \quad \textit{and} \quad \lim_{n\to\infty} R_n^* = R_\infty^*.$$

In the unconstrained case, a rationally inattentive consumer is always strictly better off than a consumer who knows his true valuation a priori. The seller's expected revenue is bounded above by the monopoly revenue in the benchmark with a fully informed consumer. As illustrated in Example 1.3 and Example 1.4 it may be strictly smaller.

In each of the above examples, it can be shown that the outcome induced by the equilibrium two-signal information structure is a Pareto improvement compared to the benchmark case with a fully informed consumer. Numerical values are provided in Table 1.1.

| | $p^M$ | $p_2^* = R_2^*$ | $p_\infty^* = R_\infty^*$ | $R^M$ | $U^M$ | $U_2^*$ | $U_\infty^*$ |
|---|---|---|---|---|---|---|---|
| $v \sim U[0, 1]$ | $\frac{1}{2}$ | $\frac{1}{4}\left(\sqrt{5}-1\right)$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{4}\left(3-\sqrt{5}\right)$ | $\frac{1}{4}$ |
| $f(v) = 1 - 2v$ | $\frac{1}{3}$ | $\approx 0.19808$ | $\approx 0.14782$ | $\frac{4}{27}$ | $\frac{8}{81}$ | $\approx 0.13526$ | $\approx 0.18552$ |
| $v \sim \beta\left(1, \frac{1}{2}\right)$ | $\frac{2}{3}$ | $\approx 0.43593$ | $\approx 0.38349$ | $\frac{2}{3}\frac{1+\sqrt{3}}{3+\sqrt{3}}$ | $\frac{2}{9}\frac{1+\sqrt{3}}{3+\sqrt{3}}$ | $\approx 0.28318$ | $\approx 0.23074$ |

Table 1.1: Equilibrium prices, consumer's expected surplus and seller's revenue for the benchmark with a fully informed consumer, the optimal two-signal information structure, and the unconstrained optimal information structure.

It is instructive to illustrate the results on the consumer's expected surplus and the seller's expected revenue in a surplus triangle. By Proposition 1.2, all equilibrium outcomes lie on the efficient frontier. The locations of the equilibrium outcomes in the surplus triangle are illustrated in Figure 1.10.

The case, in which the consumer cannot process any information and the seller extracts all surplus from trade, corresponds to the upper extreme point of the efficient frontier. As the

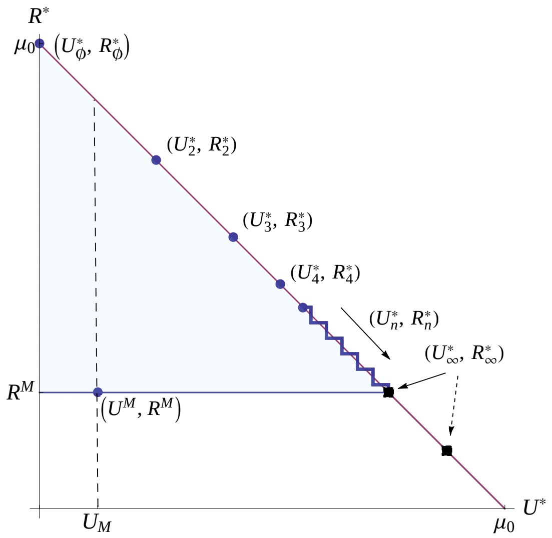



Figure 1.10: Surplus Triangle

number of available signal realizations $n$ increases, the points that mark the induced outcomes move down on the efficient frontier. The limiting case either corresponds to the intersection of the efficient frontier with the monopoly revenue level, or may lie on the efficient frontier below this point. In the first case, trade is efficient, the seller obtains the monopoly revenue and the consumer obtains the remaining surplus from trade. The latter case corresponds to the situation illustrated in Example 1.3 and Example 1.4, in which the seller's expected revenue is strictly below the monopoly revenue.

# 7 Discussion: Timing and Observability

In strategic situations, the timing influences whether the information processing constraints of the consumer are *strategically binding* or not. In the model studied in this essay, the timing is such that the consumer commits to an information structure prior to observing the price. The seller observes the information structure, but not the private signal realizations, before he chooses which price to charge. Hence, the choice of an informational environment of the consumer is a non-contingent choice. The analysis in this essay highlights the role of rational inattentiveness of the consumer in such a setting. By committing to not pay attention (ignore) certain information, the consumer can induce the seller to charge a lower price. This increases the realized gains from trade, and benefits the consumer. It often even leads to a Pareto improvement compared to the case in which the consumer knows his true valuation.

There are alternative modeling choices of timing and observability, and depending on the application that one has in mind any of these modeling choices may be the most natural.

First, one could consider the case in which the consumer chooses his information structure contingent on the price that he observes. Without information costs, this case is trivial, since even with two signal realizations, the consumer can choose an information structure that lets him take an optimal purchasing decision. By learning whether his valuation is above or below the price, the consumer obtains all information that is relevant for his purchasing decision. In equilibrium, the seller charges the standard monopoly price. Hence, the result is outcome equivalent to the benchmark with a fully informed consumer from Section 3.

The second alternative is to assume that the seller and the consumer choose the price and the information structure simultaneously. The consumer then observes the private signal realization and the price, and makes his purchasing decision. This case is equivalent to the model in which the consumer first chooses an information structure, but this is not observed by the seller.

The idea that the consumer faces limitation to process information[15] is central to the analysis in this essay. The equilibrium characterizations for the case in which the seller cannot observe the consumer's information structure differs strongly from the one in the present model. In the constrained case, if the seller cannot observe the information structure of the consumer, there exists no equilibrium in pure strategies. However, the main effects of a rationally inattentive consumer on monopoly pricing and the realized outcome are robust, and do not depend on the assumption that the seller can observe the information structure of the consumer. The effects established in this essay for the observable case also exists in the unobservable case, even though they are weaker.

To briefly illustrate the features of a mixed-strategy equilibrium in the unobservable case, reconsider the uniform prior example of Section 3. The consumer is restricted to monotone partitional two-signal information structures. In this setting, the consumer will mix over two information structures with thresholds $v_A^* = \frac{1}{3}$ and $v_B^* = \frac{2}{3}$, respectively. These are also the two prices over which the seller mixes. Both players mix with probability $\frac{1}{2}$. Notice that each information structure of the buyer is a best-response to one of the seller's prices and vice versa. For instance, for the price $p_A = \frac{1}{3}$, it is optimal for the consumer to learn whether his valuation is above or below this threshold. If the consumer adopts the information structure with threshold $v_A^* = \frac{1}{3}$, then the best response of the seller would be to charge a price equal to $\frac{2}{3}$. In this equilibrium, not all gains from trade are realized. However, an important insight gained from the analysis in this essay, remains. There are again positive welfare effects of a rationally inattentive consumer, even though they are weaker than in the observable case. In the equilibrium with a two-signal information structure, the expected realized gains from trade, as well as the expected surplus of the consumer and the expected revenue of the seller are higher than in the benchmark with a fully informed consumer.

The equilibrium values of total surplus, consumer's expected surplus and seller's revenue for the three modeling choices, the benchmark with a fully informed consumer, the observable and the unobservable case, are summarized in Table 1.2. It can be seen that both, the consumer and the seller are best off in the case in which the consumer chooses the information structure first, and this is observed by the seller.

The relation between the results for the three modeling choices is different in the absence of information constraints. In this case, if the seller cannot observe the information structure of the

[15] Those may be self-imposed, due to cognitive limitations or external regulation.

|  | price(s) | $U_2^*$ | $R_2^*$ | $T^*$ |
|---|---|---|---|---|
| Full-information benchmark | $\frac{1}{2}$ | $\frac{1}{12}$ | $\frac{1}{4}$ | $\frac{1}{3}$ |
| Observable case | $\frac{1}{4}(\sqrt{5}-1)$ | $\frac{1}{4}(3-\sqrt{5})$ | $\frac{1}{4}(\sqrt{5}-1)$ | $\frac{1}{2}$ |
| Unobservable case | $(\frac{1}{2} \circ \frac{1}{3}, \frac{1}{2} \circ \frac{2}{3})$ | $\frac{1}{9}$ | $\frac{4}{15}$ | $\frac{17}{45}$ |

Table 1.2: Equilibrium prices, and consumer's, seller's, and total expected surplus, for each of the three timings.

consumer, then in equilibrium the consumer learns his true valuation and the seller charges the monopoly price. Hence, the outcome in the unobservable case is the same as in the benchmark with a fully informed consumer.

This relation shows that in the absence of information constraints, the effect that rational inattention can reduce prices and yield welfare improvements strongly depends on the assumption that the seller can make his pricing decision contingent on the information choice of the consumer. This observation suggests that it is desirable to either make the consumer's choice of an information structure observable for the seller or to have an intermediary or regulator control the information structure of consumers and recommend pricing strategies for sellers.

# 8 Concluding Remarks

The objective of this essay was to understand the implications of a rationally inattentive consumer in a monopoly pricing model. The consumer can freely decide which pieces of information about his valuation to process, but may face information processing constraints. These constraints can be due to limited cognitive abilities or other sources that restrict the flexibility of information acquisition of the consumer.

A main contribution of this essay is to identify a "*persuasion through rational inattentiveness*" effect. By choosing his information structure, the consumer designs the information environment of the monopoly pricing problem. Any given information structure together with a signal realization induces a belief of the consumer, which determines his reaction to realized prices. Moreover, this information structure also affects the demand curve that the seller faces, and hence her pricing strategy.

The analysis identified the effects of persuasion through rational inattention on the outcome in a monopoly pricing model. It was shown that, by committing to ignore some information that would separate low valuations, the consumer can induce the seller to charge a lower price. The consumer unambiguously benefits from the lower realized price, whereas the seller's expected revenue may decrease. For every equilibrium information structure all possible gains from trade are realized. Moreover, it is often the case that pricing for a rationally inattentive consumer yields a Pareto improvement, compared to the case in which the consumer is privately informed about his true valuation. A Pareto improvement is more likely, if the consumer faces information processing constraints.

Based on the insights gained from the analysis in this essay, there are various interesting directions for future research. Examples include the analysis of optimal design and pricing of product lines, and the effects of competition among sellers (inter- or cross-market competition) or buyers (auction setting). Moreover, it would also be interesting to explore persuasion through rational inattentiveness in other strategic environments such as contract theory or collective decision making.

# Appendix

## A Proofs

*Proof of Lemma 1.1.* Suppose that the consumer chooses the $n$-signal information structure $\pi$, that induces the value estimates $V_1 < V_2 < \cdots < V_n$, and the corresponding probability masses $g_1, \ldots, g_n$. For the seller, the situation is as if he faces a consumer whose valuation is drawn from distribution $F_\pi$. If the seller charges a price $p \in (V_k,\, V_{k+1}]$ then by (1.4) the consumer will buy the object if and only if his value estimate is at least $p$. Hence, the probability of trade is $\sum_{i=k}^{n} g_i$, and the expected revenue for the seller is $R(p) = p \cdot \sum_{i=k}^{n} g_i$. Among all prices $p \in (V_k,\, V_{k+1}]$, the revenue maximizing price for the seller is $p = V_k$, and it follows that the seller's best-response function is given by

$$p^*(\pi) := \arg \max_{p \in Supp\,(F_\pi)} \{p \cdot \sum_{i=1}^{n} g_i \cdot \mathbb{1}_{[p,1]}(V_i)\}.$$

□

*Proof of Lemma 1.2.* If the seller charges the exclusive price, $p = V_n$, the consumer's expected surplus is 0. By choosing an informaiton structure that induces the seller to charge an inclusive price $p < p_n$, the consumer can guarantee himself a positive expected surplus $U \geq g_n\,(V_n - p)$.

□

**Notation:** The sets of information structures with exactly $n$, respectively at most $n$, signal realizations are denoted by

$$\mathcal{S}^{(n)} = \{\pi \in \mathcal{S} : \, |S| = n\} \subset \mathcal{S}, \text{ and } \mathcal{S}^{[n]} \cup_{k=1}^{n} \mathcal{S}^{(k)}.$$

For a given information structure that induces the distribution $F_\pi$ of value estimates, let $V_i \sim_S V_j$ denote the case in which the seller is indifferent between charging any of the prices $V_i,\, V_j \in \text{supp}\,(F_\pi)$. If the seller strictly prefers to charge price $V_i$ over $V_j$, write $V_i \succ V_j$.

Let $\mathcal{S}_i^{(n)} \subset \mathcal{S}^{(n)}$ be the subset of information structures with $n$ signal realizations that induce the seller to charge a price equal to the $i^{th}$-lowest value estimate, $V_i = \mathbb{E}\,[v|s_i]$.

*Proof of Proposition 1.1 (Seller-Indifference).* Suppose that the equilibrium information structure $\pi$ induces the seller to charge a price equal to the $i^{th}$-lowest value estimate, that is, $p = V_i$ and $\pi \in \mathcal{S}_i^{(n)}$. Suppose moreover that there exists some $k > i$ such that $R(V_i) > R(V_k)$. That is, the seller strictly prefers to charge the price $p = V_i$ over charging a price $\hat{p} = V_k$. It will be shown that in this case it is possible to construct an information structure that makes the consumer better off.

*Some technical preliminaries:* For every $j \in \mathcal{I}$, let $\mathcal{V}_j := \text{supp}\,(F(\cdot|s_j)) \subseteq \mathcal{V}$ be the support of the posterior distribution induced by the signal realization $s_j$. Since $\mathcal{V}_j \subseteq [0,\, 1]$, it is bounded above and below, and $\overline{v}_j := \sup \mathcal{V}_j$ exists. Moreover, for every $v \in [0,\, 1]$, $\{v\}$ is a zero-probability event.[16] Hence, w.l.o.g. we can assume that $\overline{v}_j \in \mathcal{V}_j$, that is, $\overline{v}_j = \sup \mathcal{V}_j = \max \mathcal{V}_j$. It follows that for every $j \in \mathcal{I}$, there exists some $\delta_j > 0$ such that $[\overline{v}_j - \delta_j,\, \overline{v}_j] \subseteq \mathcal{V}_j$, that is, $f(v|s_j) > 0$ for all $v \in [\overline{v}_j - \delta_j,\, \overline{v}_j]$. Finally, given that $f(v)$, $g(s_j) > 0$ it holds that $f(v|s_j) = 0$ if and only if $g(s_j|v) = 0$, which implies that $g(s_j|v) > 0$ if and only if $v \in \mathcal{V}_j$. Using analogous arguments establishes that every $j \in \mathcal{I}$, $\underline{v}_j := \min \mathcal{V}_j$ exists, as well as some $\delta_j > 0$ such that $\big[\underline{v}_j,\, \underline{v}_j + \delta_j\big) \subseteq \mathcal{V}_j$.

*Formal construction of an information structure that yields a higher expected surplus for the consumer than* $\pi$*:*

**Construction 1.**

Let $k \in \mathcal{I}$ be the largest index such that $R(V_k) < R(V_i)$. For every

$$\begin{aligned}
&\delta = (\delta_i, \ldots, \delta_n) \;\text{ with } \delta_j > 0 \;\forall\, j = 1, \ldots, n, \;\text{ such that} \\
&\big[\underline{v}_j,\, \underline{v}_j + \delta_j\big) \subseteq \big(\mathcal{V}_j \cup \big[\underline{v}_{j+1},\, \underline{v}_{j+1} + \delta_{j+1}\big]\big) \;\forall\, j > k, \text{ and} \\
&(\overline{v}_j - \delta_j,\, \overline{v}_j] \subseteq (\mathcal{V}_j \cup (\overline{v}_{j-1} - \delta_{j-1},\, \overline{v}_{j-1}]) \;\forall\, i \le j < k,
\end{aligned} \tag{1.6}$$

it is possible to define a new information structure $\widetilde{\pi}$ as follows. Set $g(s_{n+1}|v) := 0$ and define the family of conditional distributions that characterize $\widetilde{\pi}$ by:

$$\begin{aligned}
&\text{For every } j > k: \\
&\widetilde{g}(s_j|v) = \begin{cases} g(s_j|v) + g(s_{j+1}|v) & \text{for } v \in \big[\underline{v}_{j+1},\, \underline{v}_{j+1} + \delta_{j+1}\big) \\ 0 & \text{for } v \in \big[\underline{v}_j, \underline{v}_j + \delta_j\big) \\ g(s_j|v) & \text{otherwise,} \end{cases} \\
&\text{For every } j, \text{ s.t. } i \le j < k: \\
&\widetilde{g}(s_j|v) = \begin{cases} 0 & \text{for } v \in (\overline{v}_j - \delta_j,\, \overline{v}_j] \\ g(s_j|v) + g(s_{j-1}|v) & \text{for } v \in (\overline{v}_{j-1} - \delta_{j-1},\, \overline{v}_{j-1}] \\ g(s_j|v) & \text{otherwise,} \end{cases} \\
&\widetilde{g}(s_k|v) = \begin{cases} g(s_k|v) + g(s_{k+1}|v) & \text{for } v \in \big[\underline{v}_{k+1},\, \underline{v}_{k+1} + \delta_{k+1}\big] \\ g(s_k|v) + g(s_{k-1}|v) & \text{for } v \in (\overline{v}_{k-1} - \delta_{k-1},\, \overline{v}_{k-1}] \;\text{ and} \\ g(s_k|v) & \text{otherwise,} \end{cases} \\
&\widetilde{g}(s_j|v) = g(s_j|v) \quad \forall\, j < i.
\end{aligned} \tag{1.7}$$

For every $v \in [0,\, 1]$, this construction satisfies $\widetilde{g}(s_j|v) \ge 0$ for all $s_j \in S$, and $\sum_{j=1}^{n} \widetilde{g}(s_j|v) = 1$. Hence, $\widetilde{\pi}$ is well-defined as an information structure. This construction is illustrated in Figure 1.11.

Construction 1 takes some mass off the support $\mathcal{V}_n$ of $F(\cdot|s_n)$. This reduces the revenue that

[16]By assumption $F$ is continuous and has no atoms.

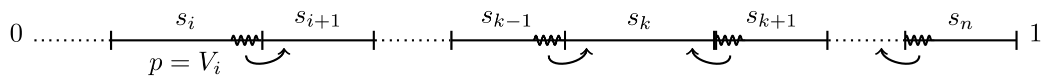



Figure 1.11: Illustration of Construction 1.

the seller could extract by charging a price equal to $\mathbb{E}[v|s_n]$. Indeed,

$$\begin{aligned}
R(\widetilde{V}_n) =& \widetilde{V}_n \cdot \widetilde{g}(s_n) \\
=& \left( \int_0^1 v f(v|s_n) \, \mathrm{d}v \right) \cdot \widetilde{g}(s_n) \\
=& \int_0^1 v \widetilde{g}(s_n|v) f(v) \, \mathrm{d}v \\
=& R(V_n) - \int_{\underline{v}_n}^{\underline{v}_n + \delta_n} v g(s_n|v) f(v) \, \mathrm{d}v < R(V_n).
\end{aligned}$$

The construction then adds the mass taken from the support of of $F(\cdot|s_n)$ to the support of $F(\cdot|s_{n-1})$, which increases the revenue that the seller could extract by charging the price $\mathbb{E}[v|s_{n-1}]$. The probability to sell remains the same whereas the seller can charge a higher price. However, Construction 1 also takes mass off the lower part of the support of $F(\cdot|s_{n-1})$, which reduces the revenue that the seller can extract by charging price $\mathbb{E}[v|s_{n-1}]$.

For signal $s_{k-1}$, mass is taken off some high values $v \in (\overline{v}_{k-1} - \delta_{k-1}, \overline{v}_{k-1}]$ in the support $\mathcal{V}_{k-1}$. This reduces the revenue that the seller can extract by charging the price $\mathbb{E}[v|s_{k-1}]$, since this price reduces whereas the probability to sell remains the same. Construction 1 then adds mass to the support of $F(\cdot|s_{k-1})$, which increases the revenue that the seller can extract by charging the price $\mathbb{E}[v|s_{k-1}]$.

Moveover, for any $\delta$ that satisfies (1.10), the revenue that the seller can extract by charging the price $\mathbb{E}[v|s_k]$ increases.

For a continuous prior distribution, there exists a $\delta$ that satisfies (1.10), such that Construction 1 yields $R(\widetilde{V}_j) < R(V_j)$ for all $j \geq i$, $j \neq k$, and $R(\widetilde{V}_k) = R(\widetilde{V}_j)$ for all $j \geq i$. For the resulting information structure $\widetilde{\pi}$ there are two cases to be considered:

*Case (1):* Given information structure $\pi$, the seller still charges a price equal to the $i^{th}$-lowest value estimate $\mathbb{E}[v|s_i]$. In this case, for every $v \in [0, 1]$ the probability of trade is not affected by Construction 1. Hence, the probability of trade remains the same, $\sum_{j=i}^{n} \widetilde{g}(s_j) = \sum_{j=1}^{n} g(s_j)$. The same is true for the expected value of types that participate in trade, and hence for the expected total surplus from trade, $T(\pi) = T(\widetilde{\pi})$. Moreover, by Construction 1 $\widetilde{V}_i < V_i$, and hence $R(\widetilde{V}_i) < R(V_i)$. It follows that the expected surplus of the consumer is higher than before, $U(\widetilde{\pi}) = T(\widetilde{\pi}) - R(\widetilde{\pi}) < U(\pi)$.

*Case (2):* Under the information structure $\widetilde{\pi}$, the seller wants to charge a price other than $\mathbb{E}[v|s_i]$. If this is the case, the seller must have an incentive to switch to a lower price $V_j < \widetilde{V}_i$. Again, the consumer is better off than under information structure $\pi$.

This contradicts the assumption that there exists a finite equilibrium information structure $\pi$ that does not satisfy the seller-indifference property. □

**Proof of Proposition 1.2:** First, two lemmas are established, which are then combined in order to prove Proposition 1.2.

**Lemma 1.3** (Small Exclusion Region).
*Any equilibrium information structure $\pi^*$ is outcome-equivalent to an information structure with at most one value estimate in the exclusion region. That is, there exists at most one value estimate $V \in supp\,(F_{\pi^*})$ such that $V < p$.*

*Proof of Lemma 1.3.* Suppose that $\pi^*$ is an optimal information structure. Let $\mathcal{V} := \text{supp}\,(F_{\pi^*})$ be the support of value estimates that are induced by this information structure. By Lemma 1.1, the seller will charge a price $p^* \in \mathcal{V}$, say $p^* = V_p$. Let $\mathcal{V}_E \subseteq \mathcal{V}$ be the subset of value estimates that will be excluded from trade in equilibrium. That is, $V \in \mathcal{V}_E$ if and only if $V < V_p = p^*$.

*Claim* 1.1. The set $\mathcal{V}_E$ is either empty, or is a singleton $\mathcal{V}_E = \{V_1\}$.

Suppose not, that is, suppose that $|\mathcal{V}_E| > 1$, say $V,\, V' \in \mathcal{V}_E$. That is, $V,\, V' < V_p$, and w.l.o.g. assume that $V < V'$. The seller charges price $p^* = V_p$, and hence it must hold that $R(V_p) \geq V,\, V'$. Let $s,\, s'$ be the signal realizations that yield $V$, respectively $V'$. The following construction defines a new information structure $\widetilde{\pi}$ that is obtained from $\pi$ by merging the signal realizations $s$ and $s'$ into one signal realization $\overline{s}$.

**Construction 2.** Let $\widetilde{\pi} := \left(\widetilde{S},\, \{\widetilde{g}(\cdot|v)\}_{v\in[0,\,1]}\right)$ with

$$\begin{aligned}\widetilde{S} &:= (S \cup \{\overline{s}\}) \setminus \{s,\, s'\},\\ \widetilde{g}(\overline{s}|v) &:= g(s|v) + g(s'|v) \quad \forall\, v \in [0,\,1]\,, \text{ and}\\ \widetilde{g}(s_j|v) &:= g(s_j|v) \quad \forall\, v \in [0,\,1]\,,\; s_j \neq s, s'.\end{aligned}$$

Construction 2 implies

$$\overline{V} := \mathbb{E}_{\widetilde{\pi}}\,[v|\overline{s}] = \frac{g(s)V + g(s')V'}{g(s) + g(s')} \in (V,\, V')\,.$$

There are two cases to be considered. (1) Under $\widetilde{\pi}$, the seller still wants to set a price equal to $V_p$, and (2) $\widetilde{\pi}$ induces the seller to charge a price equal to $p = \overline{V}$.

*Case (1):* Switching from $\pi$ to $\widetilde{\pi}$ has no effect on the outcome, $\pi$ and $\widetilde{\pi}$ are *outcome-equivalent.*

*Case (2):* Under the new information structure $\widetilde{\pi}$, the seller is induced to charge a price $\widetilde{p} = \overline{V} < V' < V_p$. Notice that Construction 2 does not affect the value estimates that are larger than $V_p$ – neither their value $V$, nor the probability that they arise, $g_V$. However, by Construction 2 the price charged by the seller reduces. Hence, it is easy to see that the consumer's surplus increases. Indeed,

$$U(\pi) = \sum_{\{V \in \mathcal{V}:\, V \geq V_p\}} (V - p^*) \cdot g_V < \sum_{\{V \in \mathcal{V}:\, V \geq V_p\}} (V - \widetilde{p}) \cdot g_V = U(\widetilde{\pi}).$$

The consumer is strictly better off under information structure $\widetilde{\pi}$ than under $\pi$ – a contradiction to the assumption that $\pi$ is a equilibrium information structure. □

**Lemma 1.4** (Implementability by Fully Inclusive Prices).
*Suppose that the price $p_0$ is inducible by the n-signal information structure $\pi$, and $U_\pi$ is the resulting expected surplus of the consumer. Then the price $p := \mu_0 - U_\pi$ is inducible as a fully inclusive price by an n-signal information structure, which yields the expected surplus $U_\pi$ for the consumer.*

*Proof of Lemma 1.4.* If the exclusion region is empty, $\mathcal{V}_E = \emptyset$, then the result follows trivially. In this case, $p_0 = V_1$ and trade is efficient. The realized total expected surplus from trade is $T = \mu_0$, the seller's expected revenue is $R = p_0$, and the consumer's expected surplus is $U_\pi = T - R = \mu_0 - p_0$. In this case $p = p_0$, which is implemented as a fully inclusive price by $\pi$.

Suppose now that the exclusion region is non-empty $\mathcal{V}_E \neq \emptyset$. By Lemma 1.3, one can assume w.l.o.g. that there is only one value estimate in the exclusion region, $\mathcal{V}_E = \{V_1\}$. It follows that $p_0 = V_2 > V_1$. The mass of the exclusion region is $g_1$, and the mass on the price charged by the seller, $p_0 = V_2$, is $g_2$. In order for $p_0$ to be a best-response of the seller, it must hold that

$$R(V_2) \geq R(V_1), \quad \text{and } R(V_2) \geq R(V_j) \quad \forall j > 2.$$

*Case (1):* Suppose that the first relation holds with equality, $R(V_2) = R(V_1)$, then by merging signal realizations $s_1$ and $s_2$ into one joint signal, say $\overline{s}$, one obtains a merged value estimate $\overline{V} \in (V_1, V_2)$. For the resulting information structure, charging the price $\overline{V}$ yields an expected revenue of $R(\overline{V}) = \overline{V} > R(V_1)$, and hence the seller would strictly prefer to charge a price equal to $\overline{V}$ over any other price. This contradicts the assumption that $\pi$ is an equilibrium information structure.
*Case (2):* Suppose that $R(V_2) > R(V_1)$. For $p_0 = V_2$, expected total surplus from trade is[17]

$$T(p_0) = \mu_0 - g_1 V_1.$$

The expected revenue of the seller is:

$$R(p_0) = (1 - g_1) \cdot p_0,$$

[17]Here the following property is used: For every non-zero probability event $s$, it holds that

$$F(s) \cdot \mathbb{E}[v|s] + (1 - F(s)) \cdot \mathbb{E}[v|\neg s] = \mu_0,$$

where $\neg s$ is the complementary event to $s$.

and the expected surplus of the consumer is:

$$U_\pi(p_0) = (1 - g_1) \cdot \left(\mathbb{E}\left[v|\neg s_1\right] - p_0\right).$$

It follows that

$$\begin{aligned} p =& \mu_0 - \left[(1-g_1)\mathbb{E}\left[v|s\neg s_1\right] - (1-g_1)p_0\right] \\ =& g_1 V_1 + (1-g_1)p_0. \end{aligned} \tag{1.8}$$

*Claim* 1.2. The price $p$, given by (1.8), can be implemented by an $n$-signal information structure as a fully inclusive price.

Construct a new information structure $\widetilde{\pi}$ from $\pi$ as follows: Add all of $g_2$ and a fraction $\alpha = \frac{g_2}{1-g_1}$ to the signal realization $\widetilde{s}_1$, the resulting value estimate is $\widetilde{V}_1 = p$. This is well-defined, given that Bayesian consistency requires that $1 - g_1 > g_2$. Hence, unless $g_s = 1 - g_1$, there will be some mass left on type $V_E$. Add all of the remaining mass on $V_E$ to the signal realization $\widetilde{s}$, and add mass from higher value estimates, $V_j$, $j > 2$ until $\widetilde{V}_1 = p$. Taking mass from higher value estimates will only reduce the expected revenue that the seller can extract from setting a price equal to any of these types, that is, reduce the seller's incentives to do so. It follows that

$$R(\widetilde{V}_1) = p = g_1 V_1 + (1-g_1)p_0 > R(V_2 = p_0) \geq R(\widetilde{V}_j)\ \forall j > 2,$$

which shows that the price $p$ is implementable as a fully inclusive price by an $n$-signal information structure.[18]

Under the new information structure the expected surplus of the consumer is the same as before. However, under the new information structure there is at least one $j > 1$ such that the seller strictly prefers to charge price $p = \widetilde{V}_1$ over charging a price equal $\widetilde{V}_j$. Hence, by Proposition 1.1, there exists another information structure that makes the consumer better off. □

*Proof of Proposition 1.2 (Efficient Trade).*
Suppose not, that is, suppose that for some $n \in \mathbb{N}$, there exists an equilibrium information structure $\pi \in \mathcal{S}^{(n)}$ such that $\mathcal{V}_E \neq \emptyset$. W.l.o.g. assume that $\mathcal{V}_E$ is a singleton (Lemma 1.3), that is $\mathcal{V}_E = \{V_1\}$. Consequently, it must hold that the seller charges the price $p = V_2 = \min \mathcal{V}\backslash\{V_1\}$. By Lemma 1.4 there exists an information structure $\widetilde{\pi}$ that induces the seller to charge a fully inclusive price, induces efficient trade, and yields the same surplus for the consumer as the initial information structure $\pi$. Recall that the construction in Lemma 1.4 of the information structure $\widetilde{\pi}$, leads to $R(\widetilde{V}_1) > R(\widetilde{V}_j)$ for all $j \geq 2$. That is, for the newly constructed information structure $\widetilde{\pi}$, the seller has a strict preference to charge the fully inclusive price, $\widetilde{V}_1$. But in this case, Proposition 1.1 implies that there exist an $n$-signal information structure that yields a

[18]To be precise, it is only possible to reach prices $p \leq \mu_0$ with this construction. However, this is without loss of generality, since it is never an optimal action for the consumer to induce a price greater than $\mu_0$.

strictly higher surplus for the consumer than the surplus induced by $\widetilde{\pi}$, which is also higher than the surplus induced by $\pi$. A contradiction. □

*Proof of Proposition 1.3 (Equilibrium Structures are Monotone Partitional).*
The result is established by an indirect proof. It proceeds as follows: Suppose that there exists some $n \in \mathbb{N}$ and an equilibrium information structure $\pi_n^*$ that is not partitional. Provide a construction of an $n$-signal information structure that induces a strictly higher surplus for the consumer. This contradicts the assumption that the initial non-partitional information structure is part of an equilibrium.

Suppose that $\pi$ is a non-partitional equilibrium information structure, and let $s_k$ be the highest signal realization for which the partitional property fails. That is, suppose that there exists an index $k \in \{1, \ldots, n\}$ such that $g(s_k|v) \in (0,\, 1)$ for a set $\widehat{\mathcal{V}}_k \subseteq [0,\, 1]$, with non-empty interior $int(\widehat{\mathcal{V}}_k) \neq \emptyset$. As before, let $\mathcal{V}_k$ denote the support of $F(\cdot|s_k)$.

**Construction 3.**
*Step 1:* (Adding the "missing mass" to the values in $\widehat{\mathcal{V}}_k$.)
Let $s_j < s_k$ be a signal realization such that $int\,(\text{supp}F(\cdot|s_j) \cap \text{supp}F(\cdot|s_k)) \neq \emptyset$. Define a new information structure $\widehat{\pi}$ by

$$\hat{g}(s_k|v) = \begin{cases} g(s_k|v) + g(s_j|v) & \text{for } v \in \mathcal{V}_k \\ 0 & \text{otherwise,} \end{cases} \tag{1.9}$$

$$\hat{g}(s_j|v) = \begin{cases} 0 & \text{for } v \in \mathcal{V}_k \\ g(s_j|v) & \text{otherwise, and} \end{cases}$$

$$\hat{g}(s_\ell|v) = g(s_\ell|v) \quad \forall\, \ell \neq j, k.$$

This construction adds "missing mass" to the values $v \in \widehat{\mathcal{V}}_k$. [19] The information structure $\widehat{\pi}$ is obtained from $\pi$ by shifting masses across signal realizations. Hence, by construction, $\widehat{\pi}$ is well-defined as an information structure.

For the construction in (1.9), $\hat{g}(s_k|v) \geq g(s_k|v)$ for all $v \in [0,\, 1]$. Moreover, the probability that signal $s_k$ realizes increases, $\hat{g}(s_k) = \int_{[0,\,1]} \hat{g}(s_k|v) f(v)\,\mathrm{d}v > g(s_k)$, and the revenue that the seller can extract by charging a price equal to the value estimate induced by signal $s_k$ increases, $R(\widehat{V}_k) > R(V_k)$.

*Step 2:* (Re-leveling the seller's expected revenue.)
The second step of the construction induces a "re-leveling" of the seller's revenue. That is, mass is taken off the lower part of the support $\mathcal{V}_k$ of $\widehat{F}(\cdot|s_k)$. For every $i \in \{1, \ldots, n\}$, let $\underline{v}_i := \min \mathcal{V}_i$, the minimum of the support of $F(\cdot|s_i)$. Again one can assume that $\underline{v}_i$ exists, given that $\mathcal{V}_i \subset [0,\, 1]$ is bounded below and, moreover, $\{v\}$ is a zero-probability event for all $v \in [0,\, 1]$.

[19] Notice that it may happen that the construction yields $\hat{g}(s_k|v) = 1$ for all $v \in [0,\, 1]$, which would imply that signal realization $s_k$ occurs with probability 1.

The probability mass that was added to signal realization $s_k$ in *Step 1* was taken from signal realization $s_j$. Now, for every

$$\begin{aligned} &\delta = (\delta_{j+1}, \ldots, \delta_k, 0) \ \text{ with } \delta_\ell > 0 \ \text{ such that} \\ &[\underline{v}_\ell, \, \underline{v}_\ell + \delta_\ell] \subseteq \big(\mathcal{V}_\ell \cup [\underline{v}_{\ell+1}, \, \underline{v}_{\ell+1} + \delta_{\ell+1}]\big) \ \forall \, \ell \in \{j+1, \ldots, k\}, \end{aligned} \tag{1.10}$$

it is possible to define a new information structure $\widetilde{\pi}$ by:

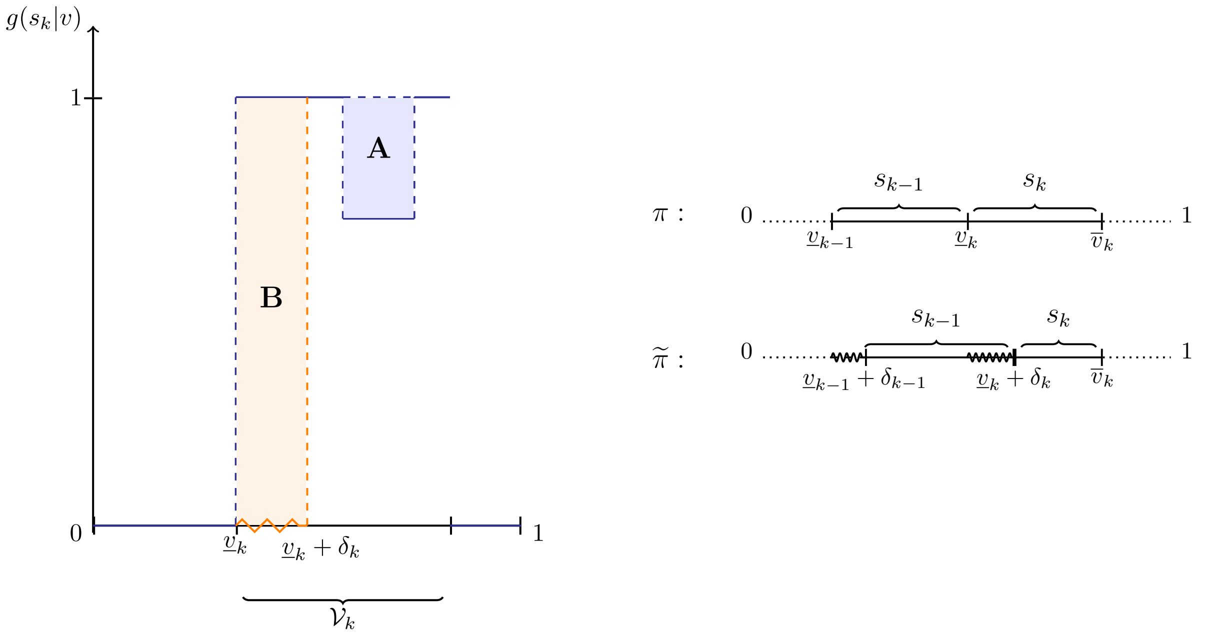
$$\widetilde{g}(s_k|v) = \begin{cases} 0 & \text{for } v \in [\underline{v}_k, \, \underline{v}_k + \delta_k) \\ \hat{g}(s_k|v) & \text{otherwise,} \end{cases} \tag{1.11}$$

$$\widetilde{g}(s_\ell|v) = \begin{cases} \hat{g}(s_\ell|v) + \hat{g}(s_{\ell+1}|v) & \text{for } v \in \big[\underline{v}_{\ell+1}, \, \underline{v}_{\ell+1} + \delta_{\ell+1}\big) \\ 0 & \text{for } v \in [\underline{v}_\ell, \, \underline{v}_\ell + \delta_\ell) \\ \hat{g}(s_\ell|v) & \text{otherwise,} \end{cases} \quad \forall \, \ell \in \{j+1, \ldots, k-1\}, \ \text{ and}$$

$$\widetilde{g}(s_j|v) = \begin{cases} \hat{g}(s_j|v) + \hat{g}(s_{j+1}|v) & \text{for } v \in \big[\underline{v}_{j+1} + \delta_{j+1}, \, \underline{v}_{j+1}\big) \\ \hat{g}(s_j|v) & \text{otherwise,} \end{cases}$$

$$\widetilde{g}(s_i|v) = \hat{g}(s_i|v) \quad \forall i \notin \{j, \ldots, k\}.$$

This construction in illustrated in Figure 1.12.




(a) Conditional probability of signal realization $s_k$ (b) Supports of signal realizations under $\pi$ and $\widetilde{\pi}$.

Figure 1.12: Illustration of Construction 3. In (a), area **A** is the mass that is added to signal realization $s_k$ in *Step 1*, area **B** is the probability mass that is taken off the lower part of the support of $F(\cdot|s_k)$.

For every $s_\ell$ with $\ell \in \{s_{j+1}, \ldots, s_k\}$, Construction 3 increases the probability that signal

$s_\ell$ realizes for some high values, which increases the revenue that the seller could extract by charging the price $\mathbb{E}[v|s_\ell]$. Construction 3 then takes probability mass off values in the lower part of the support $\mathcal{V}_\ell$, which again reduces the revenue that the seller could extract by charging the price $\mathbb{E}[v|s_\ell]$.

For a continuous distribution function, it is possible to find a $\delta$ that satisfies (1.10) such that:

$$R(\widetilde{V}_\ell) = R(V_\ell) \quad \forall \ell \in \{j+1, \ldots, k\}. \tag{1.12}$$

*Claim* 1.3. For any $\ell \in \{j+1, \ldots, k\}$ such that $\widetilde{g}_\ell \le g_\ell$, the distribution of values in the support of $\widetilde{F}(\cdot|s_\ell)$ first-order stochastically dominates those in $F(\cdot|s_\ell)$.

Indeed,

$$\widetilde{f}(v|s_\ell) = \begin{cases} f(v|s_\ell) = 0 & \forall\, v \in [0,\, \underline{v}_\ell) \\ 0 < f(v|s_\ell) & \forall\, v \in [\underline{v}_\ell,\, \underline{v}_\ell + \delta_\ell) \\ \frac{\widetilde{g}(s_\ell|v)f(v)}{\widetilde{g}(s_\ell)} \ge \frac{g(s_\ell|v)f(v)}{g(s_\ell)} = f(v|s_\ell) & \forall\, v \in [\underline{v}_\ell + \delta_\ell, 1]\,. \end{cases}$$

Since $\widetilde{F}(\cdot|s_\ell)$ and $F(\cdot|s_\ell)$ are both distribution functions, thus monotone increasing, and satisfy $\widetilde{F}(1|s_\ell) = F(1|s_\ell) = 1$, it must be that $\widetilde{F}(v|s_\ell) \le F(v|s_\ell)$ for all $v \in [0,\, 1]$. This is exactly the defining property for first-order stochastic dominance and it follows that $\widetilde{F}(\cdot|s_\ell) \ge_{FOSD} F(\cdot|s_\ell)$.

*Claim* 1.4. Construction 3 with $\delta$ such that (1.12) is satisfied yields:
(i) $\sum_{m=\ell}^{n} \widetilde{g}_m < \sum_{m=\ell}^{n} g_m$,
(ii) $\widetilde{V}_\ell > V_\ell$, and
(iii) $\sum_{m=\ell}^{n} \widetilde{g}_m \widetilde{V}_m \ge \sum_{m=\ell}^{n} g_m V_m$,
for all $\ell \in \{j+1, \ldots, k\}$.

For $\widetilde{g}_k$ this is easy to see. Suppose that $\widetilde{g}_k = g_k$. Then, by Claim 1.3, $\widetilde{F}(\cdot|s_k) \ge_{FOSD} F(\cdot|s_k)$, which would imply:

$$\widetilde{V}_k = \mathbb{E}_{\widetilde{\pi}}[v|s_k] = \int_0^1 v \cdot \mathrm{d}\widetilde{F}(v|s_k) > \int_0^1 v \cdot \mathrm{d}F(v|s_k) = V_k,$$

and

$$R(\widetilde{V}_k) = \widetilde{V}_k \left( \sum_{m=k}^{n} \widetilde{g}_m \right) > V_k \left( \sum_{m=k}^{n} g_k \right) = R(V_k).$$

$R(\widetilde{V}_k)$ is decreasing in the lower threshold, $\underline{v}_k + \delta_k$. That is, in order to obtain $R(\widetilde{V}_k) = R(V_k)$, one has to increase $\delta_k$. It follows that $\widetilde{g}_k < g_k$ and (1.12) implies $\widetilde{V}_k > V_k$.

The statements of (i) and (ii) of Claim 1.4 for $\ell \in \{j+1, \ldots, k-1\}$ follow by induction. They are verified by applying the same arguments as for $\widetilde{g}_k$ and $\widetilde{V}_k$ to $\sum_{m=\ell}^{n} \widetilde{g}_m$ and $\widetilde{V}_\ell$ using that by (1.12)

$$\left( \sum_{m=\ell}^{n} \widetilde{g}_m \right) \widetilde{V}_\ell = \left( \sum_{m=\ell}^{n} g_m \right) V_\ell.$$

Part (iii) of Claim 1.4 is still left to show. Notice that the value of $\sum_{m=\ell}^{n} g_m V_m$ is just the expected value of the values in the supports of $F(\cdot|s_m)$ for signals $s_\ell, \dots, s_n$ (cf. Figure 1.12). Hence, the result of (iii) follows directly from Construction 3.

The last step of the proof is to analyze the effect of Construction 3 for the signal realization $s_j$. The probability that signal $s_j$ realizes increases to one, for all values $v \in \left[\underline{v}_{j+1}, \underline{v}_{j+1} + \delta_{j+1}\right)$, and decreases to zero, for all values $v \in \widehat{\mathcal{V}}_k$. Whether this re-allocation of probability mass results in an increase or decrease of the expected revenue that the seller can extract by charging a price equal to the value estimate induced by the signal realization $s_j$ is in general not obvious. The next claim establishes that it is decreasing for the construction that satisfies (1.12).

*Claim* 1.5. Construction 3 with $\delta$ such that (1.12) is satisfied implies $R(\widetilde{V}_j) < R(V_j)$.

The construction only re-distributes mass among signals $s_\ell$ with $\ell \geq j$. This implies that

$$\sum_{m=j}^{n} \widetilde{g}_m = \sum_{m=j}^{n} g_m.$$

That is, the probability of trade if the seller charges the price equal to the value estimate induced by $s_j$ remains the same under the construction. Combined with the result of Claim 1.4, it follows that $\widetilde{g}_j > g_j$.

The information structure $\widetilde{\pi}$ satisfies Bayes consistency by construction. Construction 3 only re-allocates mass among signals $s_\ell$ with $\ell \geq j$. Hence, it must hold that:

$$\begin{aligned}
\widetilde{g}_j \widetilde{V}_j + \left( \sum_{k=j+1}^{n} \widetilde{g}_k \widetilde{V}_k \right) &= g_j V_j + \left( \sum_{k=j+1}^{n} g_k V_k \right) \\
\Rightarrow \quad \widetilde{V}_j = \underbrace{\frac{g_j}{\widetilde{g}_j}}_{<1} V_j &+ \frac{1}{\widetilde{g}_j} \underbrace{\left( \left( \sum_{k=j+1}^{n} g_k V_k \right) - \left( \sum_{k=j+1}^{n} \widetilde{g}_k \widetilde{V}_k \right) \right)}_{<0} < V_j.
\end{aligned}$$

It follows that

$$R(\widetilde{V}_j) = \widetilde{V}_j \left( \sum_{\ell=j}^{n} \widetilde{g}_\ell \right) < V_j \left( \sum_{\ell=j}^{n} g_\ell \right) = R(V_j),$$

which verifies Claim 1.5.

*The case $j > 1$:*
Under information structure $\widetilde{\pi}$, $R(\widetilde{V}_j) < R(V_j)$ and $R(\widetilde{V}_\ell) = R(V_\ell)$ for all $\ell \neq j$. The initial information structure was assumed to be an equilibrium structures. Hence, by Proposition 1.2 must induce the price $p = V_1$. It follows directly that under information structure $\widetilde{\pi}$, the seller charges price $\widetilde{V}_1 = V_1$. However, under information structure $\widetilde{\pi}$, the seller strictly prefers to charge price $\widetilde{V}_1$ over charging the price $\widetilde{V}_j$. It follows that, $\widetilde{\pi}$ does not satisfy seller-indifference and hence, by Proposition 1.1, there exists another information structure that yields a strictly higher expected surplus for the consumer. A contradiction to the assumption that $\pi$ is a non-

partitional equilibrium information structure.

*The case $j = 1$:*
Suppose that $j = 1$. Then, under the constructed information structure $\widetilde{\pi}$, the seller does not charge the fully inclusive price $\widetilde{V}_1$ anymore but prefers to charge a higher price, which yields a higher expected revenue for him $R(\widetilde{V}_\ell) > R(\widetilde{V}_1)$ for all $\ell > 1$.

In this case, one can simply add mass from signal realization $s_2$ to $s_1$. This re-distribution of mass increases the revenue that the seller can extract by charging a price equal to $\mathbb{E}[v|s_1]$ and decreases the revenue from charging the price $\mathbb{E}[v|s_2]$. Adding mass to $s_1$ until reaching the revenue level $R(\check{V}_1) = R(\widetilde{V}_\ell)$, $\ell > 2$, results in $R(\check{V}_2) < R(\check{V}_1)$. Hence, the seller strictly prefers to charge the price $\check{V}_1$ over price $\check{V}_2$. By Proposition 1.1 there exists an information structure that yields a strictly higher expected surplus for the consumer, which contradicts the assumption that $\pi$ is consumer optimal. □

*Proof of Theorem 1.1 (Equilibrium Existence).*
Consider any $n \in \mathbb{N}$. From Proposition 1.2 it follows that if an equilibrium information structure exists, then it must induce efficient trade, that is, must be an element of $S_1^{(n)}$. For any $\pi \in S_1^{(n)}$, the consumer's expected surplus is:

$$U(\pi) = \mu_0 - p^*(\pi).$$

Consequently, within the set $S_1^{(n)}$, the information structure that induces the minimal price is optimal. Define $P_n^* := \{p^*(\pi) : \pi \in S_1^{(n)}\}$, the set of prices that are inducible as fully inclusive prices by an $n$-signal information structure. Given that $P_n^* \subseteq [0, 1]$, it is bounded below, and hence $\inf P_n^*$ exists.

Existence of $\min P_n^*$ follows from Proposition 1.1 and Proposition 1.3. The equilibrium price is $p^* = \min P_n^*$. It is now possible to construct the (equilibrium) information structure that induces this price:

Consider any $p \in [0, 1]$. If $p$ is inducible as a fully inclusive price, it must hold that $V_1 = p$ and $R(V_1) = p$. Every equilibrium information structure must be partitional (Proposition 1.3). Hence, the lowest threshold $\hat{v}_1$ is determined by:

$$\mathbb{E}[v|v \leq \hat{v}_1] = p,$$

which implies $g_1 := F(\hat{v}_1)$.

By Proposition 1.1, it must hold that $R(V_2) = (1 - g_1)V_2 = R(V_1) = p$, which determines the value of $V_2$. Using Proposition 1.3, the threshold $\hat{v}_2$ is determined by

$$\mathbb{E}[v|\hat{v}_1 \leq v \leq \hat{v}_2] = V_2,$$

and the realization probability by $g_2 = F(\hat{v}_2)$. This procedure can be used to iteratively determine the thresholds $\hat{v}_1, \ldots, \hat{v}_{n-1}$. In order to establish Bayes consistence, set $V_n := \mathbb{E}[v|v \geq \hat{v}_{n-1}]$, and $g_n := 1 - F(\hat{v}_{n-1})$.

The above construction is well-defined for every $p \in [0, 1]$, as long as one stops the construction if at some point $\hat{v}_k > 1$. In this case, the price is certainly not the minimal price that is implementable as a fully inclusive price. Notice that, since $F$ is twice continuously differentiable, for every $i = 1, \ldots, n$, $V_i(p)$ is continuous, and moreover strictly increasing in $p$. Now for every $p \in [0, 1]$ such that $V_n(p) \leq 1$ there are three possibilities:

a) $R(V_n(p)) > R(V_1(p))$. In this case, $p$ is not inducible as a fully inclusive price.

b) $R(V_n(p)) < R(V_1(p))$. In this case, $p$ is not the minimal price that is inducible as a fully inclusive price.

c) $R(V_n(p)) = R(V_1(p))$. This is the minimal price that is inducible as a fully inclusive price, $p^* = \min P_n^*$. The partitional information structure that is obtained from the above construction is the optimal $n$-signal information structure $\pi_n^*$.

□

*Proof of Proposition 1.4 (More Signals are Better).*
The proof is by induction. For $n \in \mathbb{N}$, let $\pi_n^*$ be the equilibrium information structure if the consumer is restricted to information structures with at most $n$ signals.

**Base case:** It has already been established that for $n = 2$ the equilibrium information structure makes use of both signal realizations.

**Induction hypotheses:** Suppose that the result was already proven for all $m \leq n$. That is, suppose that for $m \leq n$, the equilibrium information structure uses all signal realizations, the consumer strictly profits from having access to information structures with more signal realizations, $U\left(\pi_{m-1}^*\right) < U\left(\pi_m^*\right)$, and the equilibrium price is strictly decreasing in the number of signal realizations, $p_{m-1}^* > p_m^*$.

**Induction step:** $(n \to n+1)$
Consider the equilibrium $n$-signal information structure $\pi_n^*$, and let $V_1^{(n)}, \ldots V_n^{(n)}$ be the value estimates induced by this structure. By the induction hypothesis, the probability mass of each of these value estimates is strictly positive.
*To show:* If the consumer has access to information structures with at most $n+1$ signal realizations, then $\pi_{n+1}^* \in \mathcal{S}^{(n+1)}$, and $U\left(\pi_{n+1}^*\right) > U\left(\pi_n^*\right)$.

Let $\widehat{\pi}$ be an information structure with $n+1$ signal realizations $\{s_1, \ldots, s_{n-1}, \hat{s}_n, \hat{s}_{n+1}\}$ that is constructed from $\pi_n^*$ by splitting the mass of the value estimate $V_n$, respectively of the signal realization $s_n$, between two signal realizations $\hat{s}_n$ and $\hat{s}_{n+1}$. That is, $g(\hat{s}_n|v) + g(\hat{s}_{n+1}|v) = g(s_n|v)$, for all $v \in [0, 1]$. The construction can be chosen, such that[20]

$$V_{n-1} < \widehat{V}_n < V_n < \widehat{V}_{n+1}, \tag{1.13}$$

[20]Notice that $\pi_n^*$ is an equilibrium information structure and thus partitional. Hence, the desired construction can be obtained by choosing $\hat{v}_{n+1}$ slightly above the highest threshold $v_n$ and let the additional signal realize whenever $v \in [v_n, \hat{v}_{n+1}]$. This construction implies $R(V_n) > R(\widehat{V}_{n+1})$, since the revenue that the seller can extract by charging a price equal to the hightest value estimate is decreasing in the highest threshold.

and $R(V_n) > R(\widehat{V}_{n+1})$. If mass from $V_n$ is split such that (1.13) is satisfied, it is always the case that $R(V_n) > R(\widehat{V}_n)$, since the probability to sell $g(s_n) = g(\hat{s}_n) + g(\hat{s}_{n+1})$ is the same and $\widehat{V}_n < V_n$. Under this construction, the revenue that the seller can extract by charging a price equal to $V_j$ remains the same for all $j < n$. That is, $R(V_j) = R(\widehat{V}_j)$, $\forall\, j < n$.

It follows that the information structure $\widehat{\pi}$ does not satisfy the seller-indifference property. Hence, by Proposition 1.1 there exists an information structure with at most $(n+1)$-signals that yields a strictly higher surplus for the consumer. Consequently, the equilibrium information structure $\pi^*_{n+1}$ yields a strictly higher surplus than the expected surplus that the consumer would obtain under $\pi^*_n$. Hence, the surplus $U(\pi^*_{n+1})$ is not achievable with an $n$-signal information structure, which implies that $\pi^*_{n+1}$ has to make use of all $n+1$ available signal realizations.

The result on equilibrium prices follows directly from the observation that the expected surplus of the consumer is strictly greater under $\pi^*_{n+1}$ than under $\pi^*_n$, and the result that any equilibrium information structure induces efficient trade (Proposition 1.2): $U^*_n = \mu_0 - p^*_n$, and it follows that $p^*_{n+1} < p^*_n$. □

The following concept will be used in the proof of the next theorem.

**Definition 1.1:** Say that the sequence of finite information structures $\{\pi_n\}_{n\in\mathbb{N}}$ converges to the information structure $\widetilde{\pi}$, if the sequence of distributions over value estimates $\{F_n\}_{n\in\mathbb{N}}$ induced by $\pi_n$ weakly converges to the distribution $\widetilde{F}$ induced by $\widetilde{\pi}$. Weak convergence of distribution functions, respectively convergence of information structures is denoted by:

$$F_n \Rightarrow \widetilde{F}, \quad \text{and} \quad \pi_n \Rightarrow \widetilde{\pi}.$$

*Proof of Theorem 1.2 (Unconstrained Equilibrium).*

*Claim* 1.6. The sequence of finite equilibrium information structures $\pi^*_{(n)}$ converges.

The support of true valuations is $I = [0,\, 1] \subseteq \mathbb{R}$, which is a compact metric space. Let $\mathcal{B}(I)$ be the Borel algebra on $I$. Then, the space $\mathcal{P}(I)$ of probability measures on $(I,\, \mathcal{B}(I))$ is metrizable by the Levy-Prokhorov metric. That is, since $I$ is separable, weak convergence of measures is equivalent to convergence of measures in the Levy-Prokhorov metric. $\mathcal{P}(I)$ is compact in the weak topology, and hence sequentially compact. It follows that for the sequence $\{F^*_n\}_{n\in\mathbb{N}}$ induced by $\{\pi^*_n\}_{n\in\mathbb{N}}$, there exists a convergent subsequence $F^*_{n_k} \Rightarrow F_\infty$.

Every finite equilibrium information structure $\pi^*_n$ induces the seller to charge a fully inclusive price. The sequence $\{\pi^*_n\}_{n\in\mathbb{N}}$ induces a sequence of prices $\{p^*_n\}_{n\in\mathbb{N}}$, which is strictly decreasing in $n$ (cf. Proposition 1.4). Moreover, every finite equilibrium information structure is partitional and satisfies the seller-indifference property (Proposition 1.3 and Proposition 1.1). The thresholds of the interval partition that characterizes a finite equilibrium information structure are already determined by the lowest threshold through the seller-indifference condition. It follows that the weak convergence of distributions of values estimates is not only satisfied for a subsequence but that the sequence $\{F^*_n\}_{n\in\mathbb{N}}$ converges, $F^*_n \Rightarrow F_\infty$. This establishes convergence of the sequence of information structures, $\{\pi^*_n\}_{n\in\mathbb{N}}$. Define $\pi^*_\infty := \lim\limits_{n\to\infty} \pi^*_n$.

Convergence in the Levy-Prokhorov metric implies convergence of prices, that is, $p_n^* \overset{n\to\infty}{\longrightarrow} p_\infty^*$. Every finite equilibrium information structures satisfies the seller-indifference condition (Proposition 1.1), which implies

$$p_n^* = \mathbb{E}\left[v | v \leq \hat{v}_{1,n}\right] = \mathbb{E}\left[v | \hat{v}_{1,n} \leq v \leq \hat{v}_{2,n}\right] (1 - F(\hat{v}_{1,n})), \tag{1.14}$$

where $\hat{v}_{1,n}$, and $\hat{v}_{2,n}$ are the lowest and next to lowest threshold of the interval-partition that characterizes the information structure $\pi_n^*$. The left-hand side of (1.14) is increasing in $\hat{v}_{1,n}$. Moreover, for a given $\hat{v}_{1,n}$, the right-hand side of (1.14) is increasing in $\hat{v}_{2,n}$, and bounded below by $\hat{v}_{1,n} \cdot (1 - F(\hat{v}_{1,n}))$. It follows that, for the information structure $\pi_\infty^*$, the lowest threshold $\underline{v}$ is given by

$$\underline{v} := \min\{\hat{v} \in (0, 1] : \, (1 - F(\hat{v}))\, \hat{v} = \mathbb{E}\left[v | v \leq \hat{v}\right]\}. \tag{1.15}$$

The property that $F_{\pi_\infty^*}$ is an equal revenue curve, directly follows from the feature that all finite equilibrium information structures satisfy seller-indifference (Proposition 1.1), and that this property is preserved in the limit.

It is still left to verify that $\pi_\infty^*$ is indeed an equilibrium information structure. Suppose not, that is, suppose that there exists an information structure $\widetilde{\pi}^*$ that makes the consumer strictly better off $U(\widetilde{\pi}^*) > U(\pi_\infty^*)$. Let $\{\widetilde{\pi}_n\}_{n\in\mathbb{N}}$ be a sequence of finite information structures that converges to $\widetilde{\pi}^*$. This also implies that $U(\widetilde{\pi}_n) \overset{n\to\infty}{\longrightarrow} U(\widetilde{\pi}^*)$. It follows that there exists some $\epsilon > 0$ such that for every $n_\epsilon \in \mathbb{N}$, there exists some $n > n_\epsilon$ such that $U(\widetilde{\pi}_n) - U(\pi_n^*) > \epsilon$. A contradiction to the assumption that $\pi_n^*$ is an equilibrium information structure. □

*Proof of Corollary 1.1.* As shown in Proposition 1.4, the consumer's expected equilibrium surplus $U_n^*$ is strictly increasing in the number of signal realizations $n$. Given that every equilibrium information structure induces efficient trade (Proposition 1.2), the induced expected total surplus from trade $T_n^*$ is constant in $n$. The total surplus is split between the consumer and seller. It follows that the seller's expected revenue $R_n^* = T_n^* - U_n^*$ is strictly decreasing in $n$. In Theorem 1.2, weak convergence of the sequence of finite equilibrium information structures has been established, which implies

$$\lim_{n\to\infty} U(\pi_n^*) = U(\pi_\infty^*) \quad \text{and} \quad \lim_{n\to\infty} R(\pi_n^*) = R(\pi_\infty^*).$$

□

## B Capacity, Equilibrium Information Structures and Thresholds

Reconsider the uniform prior example. The aim of this section is to show how to determine the thresholds of the finite equilibrium information structures, and to illustrate how they evolve with the number of signals.

Any finite equilibrium information structure is partitional (Proposition 1.3). A partitional information structure with $n$ signal realizations is determined by a vector of thresholds

$\mathbf{a} = (a_0, a_1, \dots, a_n) \in \mathbb{R}^{n+1}$, with $0 = a_0 < a_1 < \cdots < a_n = 1$. In the uniform prior case, the value estimates and realization probabilities are given by:

$$V_i = \frac{a_i + a_{i-1}}{2}, \quad \text{and} \quad g_i := a_i - a_{i-1}, \ \forall\, i = 1, \dots, n.$$

Bayes consistency and feasibility are satisfied by construction.

The equilibrium information structure satisfies two more properties. It induces *efficient trade* and yields *seller-indifference*. These two properties are satisfied, if and only if

$$V_1 = \left( \sum_{i=k}^{n} g_i \right) V_k \quad \forall\, k = 2, \dots, n.$$

It follows that the threshold vector $\mathbf{a} \in \mathbb{R}^{n+1}$ defines the equilibrium $n$-signal information structure $\pi_n^*$, if the thresholds satisfy the following system of equations:

$$\begin{aligned} a_k =& \frac{a_1}{(1 - a_{k-1})} - a_{k-1}, \text{ for } k = 1, \dots, n \\ a_0 =& 0 \text{ and } a_n = 1. \end{aligned} \tag{1.16}$$

The thresholds of the equilibrium information structure for $n \in \{2, 5, 10\}$ are illustrated in Figure 1.5.

Observe that highest threshold $a_n$ increases in $n$. This threshold is determined by

$$R_n^* = \frac{1 - a_n^2}{2}.$$

In the limit, the lowest threshold approaches $\frac{1}{2}$, yielding a value estimate and induced price of $\frac{1}{4}$. The seller's revenue also approaches this value from above, $\lim_{n \to \infty} R_n^* = \frac{1}{4}$. For the highest threshold $a_n$, it follows that it is bounded above by $\frac{1}{2}\sqrt{2}$ and approaches this value from below, $\lim_{n \to \infty} a_n = \frac{1}{2}\sqrt{2}$.

# Chapter 2

# Information Disclosure in Markets: Auctions, Contests, and Matching Markets

We study the impact of information disclosure on equilibrium properties in a model of a two-sided matching market that incorporates a large class of market design environments. In this model, each agent first privately observes an informative, but potentially noisy, signal about his private type. The agents then enter a matching stage in which they choose signaling investments to compete for match partners. In order to study the impact of information disclosure, we introduce a novel criterion that orders signals in terms of their informativeness. We show that information disclosure increases the expected total match output, but may also increase wasteful signaling investments due to amplified competition within groups. The second effect may dominate, leading to a decrease in expected welfare. Disclosure effects on equilibrium properties depend on whether information is disclosed to agents on the short or on the long side of the market. Applications to auctions, contests, and matching markets are discussed.

## 1 Introduction

In most market environments the information available to market participants significantly influences agents' behavior and the market outcome. Examples include auctions, contests, and various matching markets, among them school choice, college admission and labor markets. In auctions, bidding behavior as well as the revenue of an auctioneer depend on the information available to bidders about the object being sold. Feedback provided by a company to their workers affects workers' effort in promotion tournaments, and the precision of grading systems in high-schools influences the outcome of college admission.

In the recent auction literature, the effects of the precision level of information available to bidders on bidding behavior, efficiency, and expected revenue in an auction have been studied extensively. The resulting implications for auction design have been discussed.[1] It is therefore

[1] Examples include Persico (2000) and Shi (2012) who study information acquisition, whereas the focus in Bergemann and Pesendorfer (2007), Esö and Szentes (2007), Ganuza and Penalva (2010) and Ganuza and

surprising that this important topic has been little studied and is still not well understood in other market design environments.

In this essay we address this problem and analyze the effects of information disclosure on agents' behavior, the resulting assignment, and welfare implications in a two-sided matching market model. We explain how our results map to various of the aforementioned market design environments, and discuss the implications for auctions, contests and matching markets.

We consider a model of a two-sided, one-to-one matching market with a finite number of agents on each side of the market. We refer to the two groups of agents as *workers* and *firms*. These terms are only used to distinguish the two groups of agents. They can represent for example workers and firms, students and schools, competitors and prizes, or bidders and objects in auctions.

Our model is a modification of the marriage market formulated in Becker (1973), with two-sided incomplete information. Firms have private information about their types, whereas workers are a priori uncertain about their own types. The model has two stages: an *information stage* followed by a *matching stage*. In the information stage, workers obtain an informative, but typically noisy, private signal about their individual type and update their beliefs accordingly.

The signal realizations in the information stage determine the private information of workers. Information disclosure means that workers obtain more informative signals, which results in a higher information level of workers. In order to study the effects of information disclosure on equilibrium properties in the second-stage matching game, we introduce a criterion, which we call *single-crossing precision*, that orders signals in terms of their informativeness. This precision criterion is similar to those introduced in Ganuza and Penalva (2010). It also uses the insight that a more informative signal yields a more dispersed distribution of posterior estimates.

In the matching stage, agents take part in a matching game, in which they choose investments to compete for match partners. Investments are non-productive and serve as observable signals about the private types of agents. This matching game was introduced and studied by Hoppe et al. (2009) who refer to it as a *matching tournament*.[2] They prove the existence of a separating equilibrium in which agents are matched positively assortatively according to their investments. In our analysis, we focus on this equilibrium.

We find that increasing the level of information available to workers increases the expected total match output as well as the expected investments of firms, whereas the expected investments of workers may decrease. Workers always profit from information disclosed to them, whereas this disclosure may negatively affect the welfare of firms. We show that the second effect may be so strong that expected aggregate welfare is decreased, but also identify conditions that guarantee that the expected aggregate welfare is increased by a higher information

Penalva (2014) is on information disclosure.

[2]Such a tournament is a generalization of a contest in which prizes are replaced by matching opportunities. A similar model with a continuum of agents is discussed in Hopkins (2012).

level available to workers. For the case in which disclosing more precise information is costly, we characterize the worker-optimal and the socially optimal levels of information. The socially optimal level of information maximizes expected aggregate welfare, whereas the worker-optimal level only takes into account the utilities of workers and maximizes the welfare of this group.

The results are driven by two, possibly opposing, effects of information. On the one hand, a higher information level of market participants allows for a better assignment in the matching game, which increases expected total match output. On the other hand, disclosing information to agents also amplifies competition within groups, which may result in increased (wasteful) investments in the matching game. Our results indicate that the second effect may dominate, resulting in decreased welfare.

These two effects are based on the following feature of two-sided markets in which agents have private information: Agents impose externalities not only on agents within their group, but also on agents on the other side of the market. In particular, a worker imposes a positive externality on firms by providing a match opportunity. However, he also imposes a negative externality on them, since more or better match opportunities lead to increased competition among firms, which results in higher expected investments.

We apply the results that we establish in our general framework, to discuss the implications of information disclosure in various market design settings, focusing on auctions, contests, and matching markets. Depending on the application, agents' investments are interpreted as wasteful signaling costs or as (monetary or non-monetary) transfers to a third party. In each of the applications we highlight certain features of our results.

The most straightforward application of our model are two-sided matching markets, and the implications of our results yield new insights for these settings. Even though there seems to be a broad agreement that the level of information of agents in two-sided matching markets is an important factor which influences agents' behavior and the market outcome, these informational effects are poorly understood theoretically.[3] Interpreting our results as they apply to two-sided matching markets yields one of our main contributions: To our knowledge, we provide the first study of the impact of information disclosure, and the level of information available to market participants, on the outcome in a two-sided matching market. The main observations are that disclosing information may decrease expected aggregate welfare, and that the effects depend on whether information is disclosed to the agents on the short or on the long side of the market.

Our results can also be applied to illustrate the impact of information disclosure in contests or rank-order tournaments, for example through feedback systems in organizations. The flexibility of our framework provides two distinct ways to project our general model to contests,

[3] A reason for this may be that most of the theoretical analysis of matching markets studies complete information models, in which agents know their preferences over potential match alternatives. The two aspects, that agents have private information about their characteristics and may moreover be uncertain about their own characteristics, are hardly captured by the theoretical models in the literature. Our model incorporates both of these aspects and therefore takes a first step towards a theoretical analysis of the effects of private information and the information level of market participants on the equilibrium outcome in matching markets.

each yielding different insights and predictions. This feature allows us to obtain some of the existing results in the contest literature as special cases of our results. More importantly, our analysis provides new insights for the role of feedback systems in contest. We show that the effects of information disclosure in contests depend on the ratio of competitors to prizes, and on the prize-structure. For example, in promotion tournaments with a large pool of workers and only a few available positions, providing information to workers increases overall effort, whereas this is not necessarily the case if the ratio of workers and promotions is relatively balanced. Our results moreover suggest that information management through feedback systems may serve as an powerful element of contest design. We derive predictions about which feedback systems are to be expected in different organizational structures.

Our methodological contribution is highlighted in the application of our results to auctions. We discuss how the statistical methods that we use in this essay provide an alternative way to prove the results of Ganuza and Penalva (2010) on information disclosure in auctions. To further illustrate the potential of these statistical methods, we show how they can be used to strengthen and generalize the results of Ganuza and Penalva (2010).

**Outline** The rest of the essay is organized as follows. In Section 2 we present the model. In Section 3 we characterize equilibrium properties for an exogenously given level of information. In Section 4, we introduce single-crossing precision, a novel criterion to measure the informational content of signals. The effects of information disclosure and the resulting higher information level of market participants are discussed in Section 5. Section 6 contains our results on the worker-optimal and socially optimal levels of information. The implications of our results for auctions, contests, and matching markets are presented in Section 7. Related literature is discussed in Section 8, and Section 9 concludes. All proofs are relegated to the appendix. In Appendix C, we also briefly discuss different precision criteria.

# 2 The Model

Consider a two-sided matching market with a finite number of agents. We refer to the two groups of agents, constituting the two sides of the market, as *workers*, $\mathcal{I} = \{1, \ldots, n\}$, and *firms*, $\mathcal{J} = \{1, \ldots, k\}$. These terms are only used to distinguish the two groups. Depending on the application they may represent for example, workers and firms, men and women, students and colleges, or competitors and prizes.

The types of workers, $x_i$, and firms, $y_i$, are determined by iid draws from the interval $[0, \overline{x}]$, respectively $[0, \overline{y}]$.[4] Agents' types are independently distributed with prior distribution $F_X$ for workers and $F_Y$ for firms. We assume throughout the essay that $F_X(0) = F_Y(0) = 0$ and $F_X$ and $F_Y$ are continuously differentiable with positive densities, $f_X > 0$ and $f_Y > 0$, on the support.

There is incomplete information on both sides of the market: Firms' types $y_j$, are private information to the firms, whereas workers' do not know their types ex-ante. The distributions

[4]If $\overline{x}$ or $\overline{y}$ equal infinity, types are drawn from $[0, \infty)$.

$F_X$ and $F_Y$ are common knowledge.[5]

If worker $i$ is matched with firm $j$, each agent obtains match payoff $x_i y_j$, unmatched agents produce zero output. In other words, the match value function is $v(x, y) = 2xy$ and match output is split equally among match partners.

In this model, under complete information, all workers agree on the ranking of firms and vice versa. Our match value function is supermodular, which implies that *positive assortative matching* is the allocation that maximizes expected match output and, moreover, is the only stable matching.

We consider a two-period model which consists of an *information stage* followed by a *matching stage*. In the first period, workers obtain an informative, private signal about their individual types and update their beliefs accordingly. Agents then enter the matching stage, in which they compete for match partners.

## 2.1 Information Stage

In the first period, the *information stage*, every worker observes a private signal realization from an information technology $S$.

An information technology is a signal $S$, with typical realizations $s \in [0, \bar{s}]$, which is characterized by a family of conditional distributions $\{G(\cdot|x)\}_{x \in X}$ of signal realization.

$$G(s|x) := Pr(S \leq s|X = x)$$

is the probability that a worker with type $x$ receives a signal realization $s' \leq s$. We assume that for every $x \in X$, $G(\cdot|x)$ is absolutely continuous, that is, admits a density function $g(\cdot|x)$ almost everywhere. Together with the prior distribution $F_X$, an information technology induces a joint distribution on $(X, S)$, a so-called *information structure*. Agents update their beliefs according to Bayes' rule. With a slight abuse of notation, the posterior distribution of $X$ conditional on $S = s$ is $G(\cdot|s)$, and the resulting conditional expectation is

$$\widehat{X}(s) = E[X|s] = \int_{\mathcal{X}} x \, \mathrm{d}G(x|s).$$

We denote the marginal distribution of $S$ by $G$.

We assume that high signals are *more favorable* than low signals in the sense of Milgrom (1981). This condition implies that workers with high types are more likely to observe a high signal realization than workers with low types. A high signal thus indicates a higher underlying type of the agent than a low signal.

[5] We consider this model in order to simplify notation. It is straightforward to extend the analysis to the case in which agents on both sides of the market are uncertain about their types.

**Assumption 2.1** (Monotone Signals)**:** For all signal realizations $s, s' \in S$ with $s' > s$, signal realization $s'$ is *more favorable* than $s$. That is, for every non-degenerate prior distribution $F$ on $X$, if $s' > s$, then the posterior distribution $G(\cdot|s')$ dominates $G(\cdot|s)$ in terms of first-order stochastic dominance, $G(\cdot|s') \geq_{FOSD} G(\cdot|s)$.

This assumption implies that posterior estimates are strictly increasing in signal realizations:

$$E[X|s'] > E[X|s], \text{ for every } s' > s.$$

If signal $S$ is characterized by conditional densities $\{g(\cdot|x)\}_{x \in X}$, then Assumption 2.1 is equivalent to the strict monotone likelihood ratio property.[6]

For a given prior distribution $F$, every information technology $S$ results in a distribution of posterior estimates, represented by a random variable $\widehat{X} := E[X|S]$. Given Assumption 2.1, the function $\widehat{X} : S \to \mathbb{R}^+$, with $\widehat{X}(s) = E[X|s]$ is strictly increasing in $s$, which implies that there exists an inverse function, $\widehat{X}^{-1}$. The distribution function of the posterior estimates is

$$H(\hat{x}) = G\left(\widehat{X}^{-1}(\hat{x})\right) = \int_{\mathcal{X}} G\left(\widehat{X}^{-1}(\hat{x})|x\right) \mathrm{d}F(x),$$

with quantile function $H^{-1}(u) = \inf\{\hat{x}|H(\hat{x}) \geq u\}$ for $u \in [0, 1]$.

We provide two examples of information technologies that are commonly used in the literature.

**Example 2.1** (Truth-or-noise technology)**:**
Suppose that $X$ is the set of states with prior distribution $F_X$. A truth-or-noise technology provides with some probability $\alpha \in [0, 1]$ a perfectly informative signal $s = x$ and with probability $(1 - \alpha)$ pure noise, independently drawn from prior distribution $F_X$. The receiver cannot distinguish which kind of signal he observes. For signal realization $s$, the conditional expected value is $E[X|s] = \alpha s + (1 - \alpha)E[X]$. △

**Example 2.2** (Normal Experiments)**:**
Suppose that workers' types are normally distributed, $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$, and that signal $S$ is given by $S = X + \epsilon$, with a normally distributed noise term, $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. It follows that signals are normally distributed, $S \sim \mathcal{N}(\mu_X, \sigma_X^2 + \sigma_\epsilon^2)$. The posterior estimates are given by

$$\widehat{X}(s) = \frac{\sigma_\epsilon^2}{\sigma_X^2 + \sigma_\epsilon^2}\mu + \frac{\sigma_X^2}{\sigma_X^2 + \sigma_\epsilon^2}s,$$

thus linear in $S$, and again normally distributed. △

In our setting, workers do not know their types a priori. Thus, in the matching stage, workers can only condition their decisions on the information obtained in the information stage. For a given prior distribution, $F_X$, the information technology $S$ determines the distribution of con-

[6]The collection $\{g(\cdot|x)\}_{x \in X}$ has the strict *monotone likelihood ratio property* (MLRP) if for every $x > x'$, $\frac{g(s|x)}{g(s|x')}$ is strictly increasing in $s$.

ditional expected types of workers in the matching stage, $H^S(\hat{x})$, which is common knowledge. The individual expected types conditional on the private signal realizations are $\hat{x}_i = E[X|s_i]$. With a slight abuse of terminology, we refer to them as workers' *posterior types*. They are private information of the workers. Similarly, for firms, type $y_j$ is private information to firm $j$. Given the linearity of our model, if a worker with posterior type $\hat{x}_i$ is matched to a firm with type $y_j$, the expected match output for each of the match partners is $\hat{x}_i y_j$.

## 2.2 Matching Stage

In the second period, the *matching stage*, all agents simultaneously choose an individual investment that serves as a costly signal of their type. Agents on each side of the market are ranked based on their investments to then be matched positively assortatively. In case of equal investments, we assume random tie-breaking. Under this assignment rule the worker with the highest investment will be matched to the firm with the highest investment, the agents with the second highest investments in each of the groups will be matched, and so on.[7] If worker $i$ invests $b$ and is matched with firm $j$, his payoff is

$$u_i\left((x_i, b), y_j\right) := x_i y_j - b.$$

A (pure) strategy for a firm is a measurable function from the set of types $Y$ to non-negative investments $\mathbb{R}_0^+$. For workers it is a mapping from signal realizations $S$ to investments. The solution concept is Bayesian Nash equilibrium.

**Remark.** The discussion at the end of Section 2.1 illustrates that after the information stage the situation is as if agents on both sides of the market have private information about their types. We explicitly model the information stage because we are interested in the comparative statics effects that correspond to changes in the informativeness of the information technology of workers. One contribution of the essay is to identify the effect of a more precise information technology on the distribution of posterior types of workers (Section 4), to then study the comparative static effects that result from these changes (Section 5), and discuss the implications for various applications (Section 7).

## 2.3 Market Design Settings Captured by the Model

We now briefly discuss how the model captures various important market design settings. Table 2.1 provides a summary.

**Matching Markets** The model represents a two-sided, one-to-one matching market, in which agents on each side have homogeneous preferences about match partners. There is two-sided incomplete information and agents invest in non-productive signaling à la Spence (1973) to

[7] A matching mechanism which would yield this outcome is, for example, the worker-, or firm-proposing deferred acceptance algorithm, assuming that agents rank their potential match-partners according to the observed investments. The assignment rule is also the natural extension to auctions, in which the highest bidder obtains the object.

compete for match partners.[8] Investments are wasteful. They solely serve as an observable signal of the agent's unobservable type but do not have an effect on the match output. Agents of each group are ranked according to their observable signaling investments to then be matched positively assortatively.

**Auctions** The standard private values auction setting can be mapped into a special case of our model. There is only one firm, with type $Y \equiv 1$, interpreted as the auction platform which sells an object of commonly known quality 1. The workers represent the bidders in the auction, with their types corresponding to their valuations of the object, and investments corresponding to their bids. The seller is represented by a third party, which collects the bids. This interpretation yields an all-pay auction. By the revenue equivalence theorem, the results which we present in Section 5 and Section 6 apply to all standard auctions that implement the efficient allocation.

**Contests** The model also captures a rank-order tournament or contest setting in the following sense: Consider the firms as passive agents, who are the prizes in a contest, with commonly known values $\eta_{1:k} \geq \cdots \geq \eta_{k:k}$. Workers represent the competitors who participate in the contest, workers' types correspond to their abilities, and investments capture the exerted effort.

| | workers | firms | types | investments |
|---|---|---|---|---|
| auctions | bidders | *–passive–*[9] | valuations | bids |
| contests/ tournaments | workers, competitors | promotions, prizes | productivity, abilities | effort |
| matching markets | students, workers | schools, jobs | characteristics | signaling investments |

Table 2.1: Examples of environments captured by the model.

# 3 Equilibrium Characterization

In this section we characterize the equilibrium in the second-period matching game on which we will focus in our analysis. It is easy to see that there exist multiple equilibria, among them a pooling equilibrium in which all agents choose zero investments and the assignment is random. We say that an equilibrium is *symmetric* if all workers adopt the same strategy and so do all firms. Strategies are *monotone* if they are given by continuously differentiable, strictly increasing functions. Under monotone strategies, firms' with higher types choose higher investments and workers' investments are increasing in signal realizations. The existence of a

[8]In contrast to Spence (1973), here agents do not have different cost-types but high-type agents receive a higher payoff from a particular match than low-type agents.

[9]The passive firm could for example represent an auction platform. The seller is a third party who collects the investments.

symmetric separating equilibrium in monotone strategies follows directly by adapting the results of Hoppe et al. (2009) to our setting. Given Assumption 2.1 there exists a unique equilibrium of this type.

**Theorem 2.1** (Hoppe et al. 2009)**.** *Given the assumptions in Section 2, in the second-period matching game, there exists a unique symmetric separating equilibrium in monotone strategies.*

In our analysis, we will focus on this separating equilibrium. In this equilibrium, the positive assortative matching with respect to the (posterior) types of agents is implemented, all agents of a group adopt the same strategies, and high-type agents choose higher investments than low-type agents.

**Remark.** There are various reasons why it is natural to focus on the equilibrium of Theorem 2.1. This separating equilibrium is the unique equilibrium that is monotone in signal realizations. Moreover, it implements the unique stable (and core) matching, given the information available in the market after the information stage. It is also the natural extension to the efficient allocation in auctions and contests.

Before we can provide some intuition for the equilibrium and the formulas for expected total output, investments and welfare in equilibrium, we need to introduce some more notation.

For a sample $X_1, \ldots, X_n$ let

$$X_{1:n} \geq_{FOSD} \cdots \geq_{FOSD} X_{n:n}$$

be the corresponding order statistics, where $\geq_{FOSD}$ indicates first-order stochastic dominance. The random variable $X_{i:n}$ represents the distribution of the $i^{th}$ highest among $n$ iid draws. In particular, $X_{1:n} = \max\{X_1, \ldots, X_n\}$.[10]

For given market sizes $n$, $k$, priors $F_X$, $F_Y$ and information technology $S$ set

$$\mu^S_{i:n} := E\left[\widehat{X}_{i:n}\right] \quad \text{and} \quad \eta_{i:k} := E\left[Y_{i:k}\right].$$

That is, $\mu^S_{i:n}$ denotes the expected value of the $\mathrm{i}^{\mathrm{th}}$ order statistics of the posterior types, given information technology $S$.

To obtain some intuition for this equilibrium it is instructive to analyze the reduced game faced by agents on either side of the market separately, and relate equilibrium investments to Vickrey-payments.[11] This interpretation also more precisely demonstrates how the model is a natural extension of a standard auction setting.

Suppose firms adopt separating strategies, and consider the situation for workers after the information stage. In this case, the problem faced by workers in the second-period matching

[10] Hereby, we adopt the notation which is used in most of the economics literature. It should be noted that, by contrast, the standard convention in statistics is to denote the highest order statistic by $X_{n:n}$.

[11] This was also pointed out by Hoppe et al. (2009). Adapting their results to our model yields the equilibrium properties summarized in Table 2.2.

game is as if they are in a contest competing for $k$ heterogeneous prizes, where the values of the prizes are determined by an agent's posterior type and the expected values of the highest, second-highest, third-highest,... types of firms. To be precise, for a worker with posterior type $\hat{x}$, the values of the prizes are $\hat{x} \cdot \eta_{1:k}, \ldots, \hat{x} \cdot \eta_{k:k}$. The assignment in the matching stage is positive assortative with respect to investments. In the corresponding contest faced by workers, this allocation rule thus prescribes that prizes be allocated to the agents in order of their investments, where the agent with the highest investment receives the highest price. Our model is linear, and it is well-known that in such an environment expected payoffs of agents are fully specified by the allocation rule, and the expected payoff of the lowest type.[12] By the revenue equivalence theorem, it follows that expected investments must be the same as in a VCG-mechanism.[13] In the VCG-mechanism, each worker must pay the amount equal to the negative externality he imposes on the other workers. For a profile of signal realizations $s_1, \ldots, s_n$, after appropriate relabeling, let the corresponding posterior types be $\hat{x}_1 \geq \cdots \geq \hat{x}_n$. We refer to the worker receiving the $i^{th}$-highest signal, as *(posterior) type i*. The presence of type $i$ does not affect workers who receive a higher signal than himself, but he imposes a negative externality on all workers receiving a lower signal realization. Each of those workers would be assigned to a higher match-partner if type $i$ were not present. It follows that the expected investment, $t_i$, of the $i^{th}$ type is:

$$t_i = \sum_{j=i}^{\min\{k,n\}} \mu_{j+1:n} \cdot (\eta_{j:k} - \eta_{j+1:k}). \tag{2.1}$$

Summing up over all $i$ we obtain the formula for expected total investments of workers,

$$T_w = \sum_{i=1}^{\min\{k,n\}} i\big(\eta_{i:k} - \eta_{i+1:k}\big)\mu^S_{i+1:n}.$$

In the separating equilibrium of Theorem 2.1 the assignment in the matching stage is positive assortative with respect to agents (posterior) types. It is easy to see that the resulting expected total match output is $O = 2 \sum_{i=1}^{\min\{k,n\}} \mu^S_{i:n} \cdot \eta_{i:k}$. Expected total welfare of workers is $W_w = \frac{1}{2}O - T_w$ and the formulas for equilibrium expected total investments and welfare of firms are derived in a similar fashion. Table 2.2 provides a summary of the formulas.

# 4 Precision of Information Technologies

We now introduce a novel criterion, which we call *single-crossing precision*, to compare signals in terms of their informational content. In Section 5, we apply this concept to discuss the effects of a higher level of information of workers on total match output, total investments, and welfare.

Given our assumption that signals are monotone, the natural informativeness criterion to

[12] A worker who receives signal realization 0 does not invest and is matched to the lowest firm with certainty. His expected payoff is $E[X|0] \cdot \eta_{n:k}$ (which is 0 if $n > k$).

[13] The well-known Vickrey-Clarke-Groves (VCG) mechanisms due to Vickrey (1961), Clarke (1971) and Groves (1973)

| | workers | firms |
|---|---|---|
| expected total output | $O = 2\sum_{i=1}^{\min} \mu^S_{i:n} \cdot \eta_{i:k}$ | |
| expected total investments | $T_w = \sum_{i=1}^{\min} i\big(\eta_{i:k} - \eta_{i+1:k}\big)\mu^S_{i+1:n}$ | $T_f = \sum_{i=1}^{\min} i\big(\mu^S_{i:n} - \mu^S_{i+1:n}\big)\eta_{i+1:k}$ |
| expected welfare of workers/firms | $W_w = \sum_{i=1}^{\min} i \cdot \big(\mu^S_{i:n} - \mu^S_{i+1:n}\big)\eta_{i:k}$ | $W_f = \sum_{i=1}^{\min} i \cdot \big(\eta_{i:k} - \eta_{i+1:k}\big)\mu^S_{i:n}$ |
| expected aggregate welfare | $W = \sum_{i=1}^{\min} i \cdot \big(2\mu^S_{i:n}\eta_{i:k} - \mu^S_{i+1:n}\eta_{i:k} - \mu^S_{i:n}\eta_{i+1:k}\big)$ | |

Table 2.2: Formulas for expected total output, expected total investments and welfare for workers and firms, respectively, and expected aggregate welfare, for given prior distributions, information technology $S$, and market sizes, $n$, $k$. Here, $\min := \min\{k, n\}$.

use is the concept of *effectiveness* introduced by Lehmann (1988) – Persico (2000) calls this concept *accuracy*. The basic idea behind this concept is that, for a more accurate signal the conditional distributions that characterize the signal are more dependent on the state than for a less accurate signal.[14] Mizuno (2006) shows that for a more accurate signal the resulting distribution of posterior estimates is more dispersed.

As Ganuza and Penalva (2010), we use this observation to define a novel precision criterion in terms of properties of the resulting distribution of posterior estimates.[15] It is based on the following insight: for a completely random signal nothing can be inferred from the signal realization and the resulting posterior estimate is always the ex-ante mean. For a more informative signal, the resulting distribution of posterior estimates will be more responsive to the signal realization and thus result in a more variable distribution of posterior estimates.

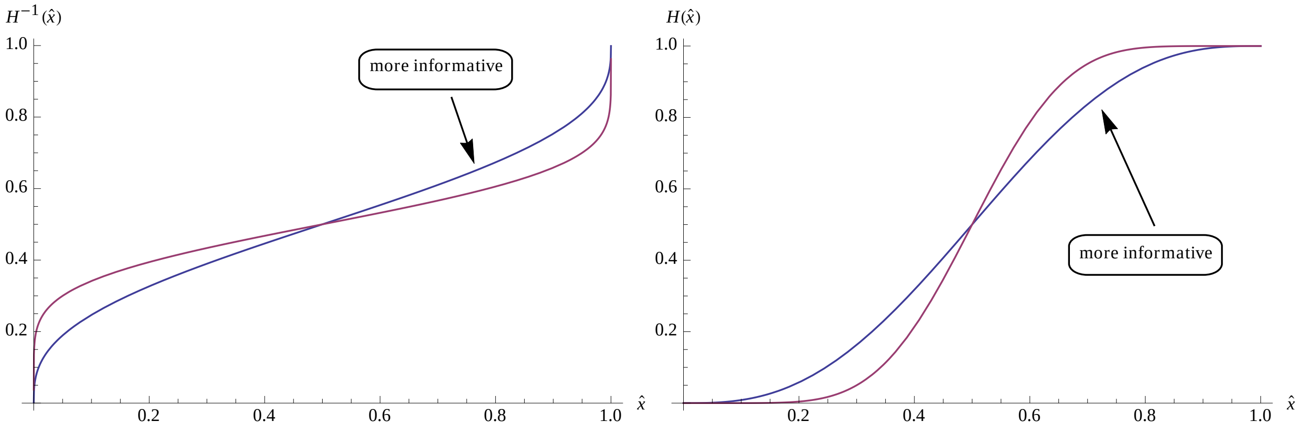
The information criterion that we use requires that a more precise signal leads to a more dispersed distribution of posterior estimates in terms of a mean-preserving spread. Moreover, for signals ordered in terms of *single-crossing precision*, we require single-crossing of the quantile functions.

**Definition 2.1:** For a given prior $F_X$ and signals $S_1$, $S_2$, let $H_1^{-1}$ and $H_2^{-1}$ be the quantile functions of $E[X|S_1]$ and $E[X|S_2]$. Say that signal $S_2$ is more *single-crossing precise* than $S_1$, denoted $S_2 \succsim_* S_1$, if

$$\frac{H_2^{-1}(u)}{H_1^{-1}(u)} \text{ is increasing in } u \in (0,1)\,.$$

[14]Signal $S_1$ is more effective than signal $S_2$ if $G_{S_1}\left(G_{S_2}^{-1}(s|x)|x\right)$ is increasing in $x$. Effectiveness applies to monotone decision problems and requires less restrictive conditions than *sufficiency* (Blackwell, 1951) to compare signals in terms of their informativeness.

[15]Our concept is slightly stronger than the concept of *integral precision* in Ganuza and Penalva (2010). Their concept of *supermodular precision* and our concept of *single-crossing precision* are not nested – a formal discussion is provided in Appendix C.




(a) *Quantile functions.* (b) *Distribution functions.*

Figure 2.1: Relation of the quantile and distribution functions of posterior estimates for signals ordered in terms of single-crossing precision.

We say that agents have a *higher information level* if the private signal realizations that agents receive in the information stage originate from a more (single-crossing) precise signal. Single-crossing precision implies that the distribution of posterior estimates resulting from the more precise signal crosses the one resulting from the less precise signal only once and from above. Our criterion is therefore slightly more restrictive than the ordering induced by accuracy.[16] However, many commonly used information structures are ordered in terms of *single-crossing precision*, among them those in Example 2.1 and Example 2.2. For truth-or-noise technologies, signal $S_\alpha$ is more precise than $S_\beta$ if and only if $\alpha \geq \beta$. For normal experiments a signal with less noise is more precise.

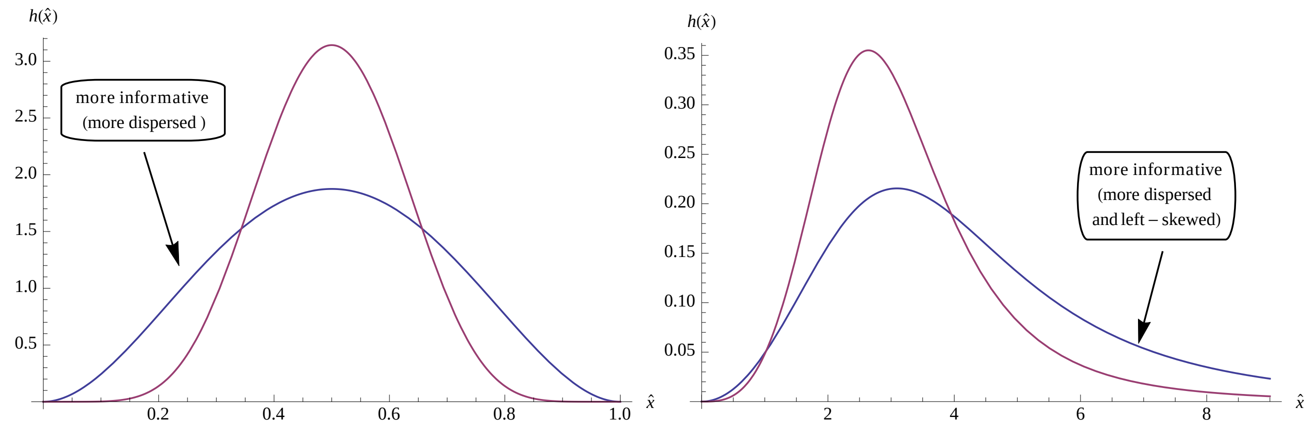
Increasing the precision of a signal in terms of single-crossing precision has two main implications. First, it results in a more dispersed distribution of posterior estimates in terms of a mean-preserving spread. Second, if the distributions of posterior estimates exhibit different levels of skewness, then the more precise signal results in a more left-skewed distribution of posterior estimates.[17] Figure 2.1 and Figure 2.2 illustrate properties of the distribution, density and quantile functions of the posterior estimates from signals that are ordered in terms of single-crossing precision.

In order to establish our results, we use the *quantile function representation* for order statistics, which establishes a close link between the properties of the quantile functions and the vector of expected order-statistics.[18] For signals ordered in terms of single-crossing precision the vec-

[16]A more accurate signal results in a mean-preserving spread of the distribution of posterior estimates, a property that does not exclude multiple crossings. Signals which are characterized by more or less fine partitions of the state space are typically not ordered in terms of single-crossing precision, since the distributions of posterior estimates may cross multiple times.

[17]This feature of our information order is in line with the well-documented observation in the empirical finance literature that many asset return distributions exhibit negative skewness (e.g. Beedles, 1979; Alles and Kling, 1994, and subsequent papers). This property is often attributed to standard practices adopted to release information. Companies tend to release good news immediately (more frequently), whereas bad news are released in clumps. This was first pointed out in Damodaran (1985) and recently discussed in Acharya et al. (2011).

[18]See Arnold et al. (1992).




(a) *Densities – the symmetric case.*

(b) *Densities – skewed distributions.*

Figure 2.2: Density functions of posterior estimates for signals ordered in terms of single-crossing precision. (a) illustrates the case for symmetric functions, (b) the case with different skewness.

tors of the expected order statistics of the posterior types satisfy a "single-crossing condition". Switching to a more precise information technology results in an increase of the expected value of the highest order statistics of posterior types, whereas the expected values of lower order statistics will decrease. The following lemma formally states this property.

**Lemma 2.1.** *If information technology $S_2$ is more single-crossing precise than $S_1$, then for all $n \in \mathbb{N}$*

$$\frac{\mu_{i:n}^{S_1}}{\mu_{i:n}^{S_2}} \text{ is increasing in } i.$$

# 5 The Comparative Statics Effects of Higher Precision

How does a change in the information level of workers affect equilibrium behavior of agents and the resulting match outcome? In this section, we address this question and characterize the effects on expected total output, investments and welfare.

In our model workers are a priori uncertain about their own types, and the outcome depends on the private signals that they receive in the information stage. Thus, from the ex-ante perspective, the matching mechanism yields a lottery over all possible matchings of workers to firms. In the two extreme cases, workers either receive no information about their own types, which results in random matching, or they observe a perfectly informative signal and are matched positive assortatively in equilibrium. Since signals are monotone (Assumption 2.1), workers with high types are more likely to receive a high signal in the information stage than low-type workers. Consequently, if the information level of workers increases, then the probability that workers with high types are matched with firms of a similar ranking increases. This results in a higher expected match output for pairs of high-ranking agents, but may result in a decrease of expected match output for lower-ranked pairs.[19]

[19]This result holds if $E[X|S_2]$ is a mean-preserving spread of $E[X|S_1]$. That is, is suffices to require that signals are ordered in terms of *integral precision* (Ganuza and Penalva, 2010).

**Proposition 2.1.** *Let signal $S_2$ be more precise than signal $S_1$. Then, expected total match output is increasing in signal precision.*

The result establishes that, at the aggregate level, a higher information level of workers always results in an increase in total match output. This is intuitive. A higher information level of workers allows on average for a better matching.

The effect on expected investments and welfare of workers is harder to characterize. If workers obtain more precise information, the resulting distribution of posterior types is more dispersed. The expected value of the low posterior-type workers decreases. Consequently, for these workers the (marginal) benefit in match output from being matched with a firm of higher ranking reduces and thus also the expected externalities imposed on them by other workers.[20] For high-type workers the effect is reversed and externalities imposed on them are increasing.

At the aggregate level, it is not clear a priori, which of these effects dominates, and some of the following results will depend on the sizes of the two sides of the market, or distributional properties of firms' types.[21]

**Theorem 2.2** (Workers' expected total investments and welfare)**.** *Let signals be ordered in terms of single-crossing precision. Then:*

*(i) Expected total investments:*

- *a) There exists some $\hat{n} \geq k$ such that for all $n \geq \hat{n}$, expected total investments of workers are increasing in information precision.*
- *b) If $n \leq k$, and the distribution of firms' types, $F_Y$, has an increasing hazard rate, then expected total investments of workers are decreasing in information precision.*

*(ii) Workers' expected welfare is increasing in information precision.*

The result shows that, if workers are on the long side of the market and the number of workers is sufficiently large, then a more single-crossing precise signal always results in an increase of workers' expected total investments. In this case, only the high-ranked workers are matched in equilibrium. The expected types of these workers are increasing in the level of information and so are the externalities imposed on them. If the ratio of workers to firms is large enough only the effect on high-type workers matter. Consequently, the expected total investments, which capture the externalities workers impose on each other, will increase. Competition among workers may even be so strong that the expected investment of each individual worker is nondecreasing in information.[22]

[20]To be precise, the expected externality imposed on the low posterior-type workers is non-increasing, since these workers may not be matched in equilibrium. Workers with zero investments do not adjust their investments.

[21]The results of Theorem 2.2 incorporate as special cases both, the comparative statics results on heterogeneity of Hoppe et al. (2009) and the results of Ganuza and Penalva (2010) on the expected valuation and the informational rent of the winning bidder, and the seller's expected revenue. More details on this are provided in Section 7.

[22]This result is easily established by combining Lemma 2.4 with (2.1).

By contrast, if there are more firms than workers, all workers are matched in equilibrium. We know that if the information level of workers increases, then the externalities imposed on the low-type workers decrease. If this effect drives the effect on expected total investments of workers, then they decrease as the information level of workers increases. This is the case if the distribution of firms' types has an *increasing hazard rate*, which implies that the externalities imposed on low-type workers have a higher impact on aggregate expected investments than those imposed on high-type workers (cf. Lemma 2.2 and Table 2.2).[23]

If workers hold private information, workers' expected welfare $W_w$ captures the informational rents of workers. In the model that we consider, the matching mechanism is fixed. Hence, as expected, for a higher information level of workers, more information rent is left to the workers and the expected welfare of workers is increasing.

The next result establishes the effects of a higher information level of workers for the other side of the market, that is, on firms' expected total investments and welfare.

**Theorem 2.3** (Firms' expected total investments and welfare)**.** *Let signal $S_2$ be more single-crossing precise than signal $S_1$.*

*(i) Expected total investments of firms are increasing in information precision.*

*(ii) Firms' expected welfare:*

- *a) There exists some $\hat{n} > k$ such that firms' expected welfare is increasing in information precision for all $n \geq \hat{n}$.*
- *b) If $n \leq k$ and $F_Y$ has an increasing hazard rate, then firms' expected welfare is decreasing in information precision.*
- *c) If $F_Y$ has a decreasing hazard rate, firms' expected welfare is always increasing in information precision.*

If workers have a higher information level, firms face a sample of potential match partners with a more heterogeneous distribution of posterior types. Consequently, the expected difference of the match outputs from being paired with one of two workers whose ranking differs only by one increases. Competition among firms increases, which results in firms increasing their expected investments. This is also true at the individual level – every firm will increase its expected investment.

Among the firms that are matched in equilibrium, the match output of high ranked firms is increasing in the information level of workers. For lower ranked firms it will typically be decreasing, unless there is a much larger number of workers than firms. Thus, lower ranked firms will be worse off if workers' hold more precise information whereas high ranked firms may profit. The effect on expected welfare of firms depends on which of these effects is dominant, and hinges on the distribution of firms' types and the sizes of the two sides of the market. If firms constitute the long side of the market and their distribution of types has an increasing

[23] $F_Y$ has an *increasing hazard rate* if $\frac{f_Y(y)}{1-F_Y(y)}$ is increasing in $y$. This property is a common assumption in mechanism design and satisfied by a large class of distribution functions, including the uniform, normal, and exponential distribution. For a detailed discussion see Bagnoli and Bergstrom (2005) and Ewerhart (2013).

failure rate, firms' expected welfare decreases the information level of workers increases. In this case, the increased competition among firms will eat up all of the additional match surplus made possible by the higher information level of workers.

Is it always welfare improving to provide workers with additional and hence more precise information? The answer is not obvious. For a higher information level of workers, there is a trade-off between a higher expected total match output and a possibly increase in (wasteful) investments. From the previous analysis we know that providing more information to workers always increases workers' expected welfare whereas firms' expected welfare may be decreasing. A higher information level of workers has a more direct effect on expected welfare of workers than on that of firms. This property may suggest that the first effect is stronger, which would imply that expected aggregate welfare is always increasing in the information level of workers.

However, this intuition is not correct. The following example shows that increasing the information level of workers may result in a decrease of aggregate welfare.

**Example 2.3** (Expected aggregate welfare)**:**
Consider a matching market with three workers and three firms, $n = k = 3$. Workers' types are standard uniformly distributed, $X_i \overset{iid}{\sim} U[0,1]$, and the information technology is a truth-or-noise technology $S_\alpha$ with precision level $\alpha$. In this setting, the posterior types of workers are uniformly distributed on $\left[\frac{1}{2}(1-\alpha), \frac{1}{2}(1+\alpha)\right]$, and the corresponding vector of posterior mean-order statistics is $(\mu^\alpha_{1:3}, \mu^\alpha_{2:3}, \mu^\alpha_{3:3}) = \left(\frac{1}{2} + \frac{1}{4}\alpha, \frac{1}{2}, \frac{1}{2} - \frac{1}{4}\alpha\right)$. From an ex-ante perspective, the expected posterior type of the highest worker is $\frac{1}{2} + \frac{1}{4}\alpha$. Suppose firms' types are represented by the vector $(\eta_{1:3}, \eta_{2:3}, \eta_{3:3}) = \left(\frac{2}{3}, \frac{1}{2}, \frac{1}{3}\right)$. Table 2.3 illustrates expected output, expected total investments and welfare of workers and firms for the given specifications.

<table>
<tr><td colspan="2">$O = \frac{3}{2} + \frac{1}{12}\alpha$</td></tr>
<tr><td>$T_w = \frac{1}{4} - \frac{1}{12}\alpha$</td><td>$T_f = \frac{7}{24}\alpha$</td></tr>
<tr><td>$W_w = \frac{1}{2} + \frac{1}{6}\alpha$</td><td>$W_f = \frac{3}{4} - \frac{5}{24}\alpha$</td></tr>
<tr><td colspan="2">$W = \frac{5}{4} - \frac{1}{24}\alpha$</td></tr>
</table>

Table 2.3: Expected total output, investments and welfare for workers and firms, for $n = k = 3$, $X_i \overset{iid}{\sim} U[0,1]$, firms' types $(\eta_{1:3}, \eta_{2:3}, \eta_{3:3}) = \left(\frac{2}{3}, \frac{1}{2}, \frac{1}{3}\right)$, and a truth-or-noise technology of precision level $\alpha$.

In this setting, as the information level $\alpha$ increases, the externalities imposed on the lowest-ranked worker, $2 \cdot \left(\frac{1}{2} - \frac{1}{4}\alpha\right) \cdot \frac{1}{6}$, decrease whereas those imposed on the middle-ranked worker are constant. Consequently, workers' expected total investments, $T_w = \frac{1}{4} - \frac{1}{12}\alpha$, is decreasing in the information level $\alpha$. Moreover, as the information level of workers increases, welfare of firms, $W_f = \frac{3}{4} - \frac{5}{24}\alpha$, decreases. In total, we obtain that the negative effect on firms is stronger than the positive effect on workers:

$$\begin{aligned}\frac{\partial W}{\partial \alpha} &= \frac{\partial o}{\partial \alpha} + \frac{\partial W_f}{\partial \alpha} - \frac{\partial T_w}{\partial \alpha} \\ &= \frac{1}{12} - \frac{5}{24} + \frac{1}{12} = -\frac{1}{24} < 0.\end{aligned}$$

A notable feature of this example is, that even though workers' expected total investments are decreasing in the level of information, the negative effect on firms' welfare is so strong that expected aggregate welfare is decreasing. △

**Remark** This observation that increasing the information level of agents may not be welfare enhancing complements and strengthens existing results, which show that random matching may be welfare superior to assortative matching because it allows to avoid wasteful signaling or screening costs.[24] Random matching requires that neither side adopts separating strategies. In many settings this is unlikely to be true, be it because there is some information about a ranking of agents available in the market, or because it is simply infeasible.[25] In this case, our example shows that for a higher information level of workers, the increased competition among firms may be so strong that the increased investments of firms may eat up all gains from increased match output and decreased wasteful investments of workers. As a result, overall expected welfare may be decreasing in the information level of workers.

To better understand the informational effects on aggregate welfare, it helps to decompose aggregate welfare as $W = o + (W_f - T_w)$; the sum of total match output of workers, $o = \frac{1}{2}O$, and aggregate externalities imposed by workers on other agents, $W_f - T_w$. Here, $T_w$ captures the aggregate externalities workers impose on each other, whereas $W_f$ captures the aggregate externalities imposed on firms, i.e., agents on the other side of the market. Thus, the effect of a higher level of information of workers on total welfare consists of the effect on workers' match output and the change in the aggregate externalities workers impose on all agents. By Proposition 2.1 we know that total match output is increasing in workers' information level whereas $W_f - T_w$ may be decreasing. In this case the effect on aggregate welfare depends on which of the two effects is dominant.

We conclude this section by identifying conditions, which each individually guarantee that aggregate welfare is increasing in workers' information level.

**Theorem 2.4.** *Aggregate welfare is increasing in information precision if one of the following conditions is satisfied:*

*(i) $F_Y$ has a decreasing hazard rate, or*

*(ii) $n < k$ and $f_Y$ is monotone decreasing.*[26]

[24] See for example Hoppe et al. (2009), Condorelli (2012), and Chakravarty and Kaplan (2013).

[25] Of course, a way to implement the random matching is to ignore any investments of agents. However, if agents have some private information this is not a stable matching.

[26] Every absolute continuous random variable with a decreasing hazard rate has a decreasing density function.

# 6 Optimal Level of Precision

We now take the analysis one step further and assume that, in the information stage, the precision of the information technologies is not exogenously given, but can be chosen before the information stage. We call this game the *precision and matching game.*

In a precision and matching game, first, workers' information technology $S_\alpha$ is chosen from a set $\mathcal{S}$ of feasible information technologies, either collectively by one group of agents, or by a social planner. The information technology $S_\alpha$ is then implemented and the rest of the game proceeds as described in Section 2. In the information stage, every worker obtains a private signal from $S_\alpha$, agents then update their beliefs according to Bayes' rule before they enter the matching stage. The timing in the precision and matching game is illustrated in Figure 2.3.

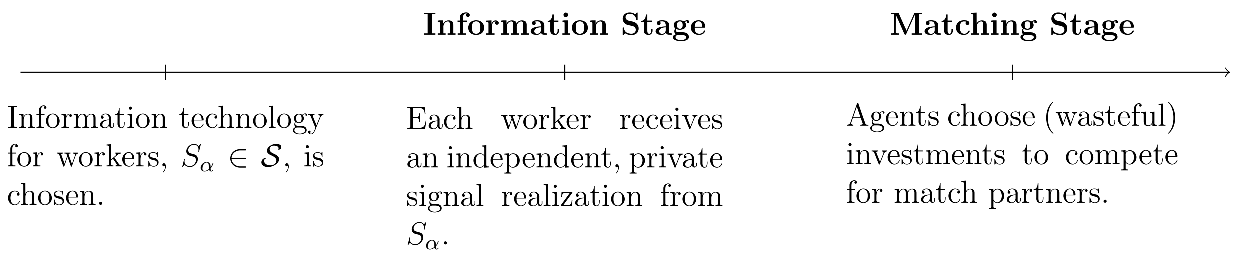



Figure 2.3: Timing in the *precision and matching game.*

We characterize the socially optimal information level, $\alpha^{so}$, which maximizes aggregate welfare of all agents in the market, and compare it to the *worker-optimal* information level, $\alpha^{ao}$, which maximizes workers' welfare. This is the optimal information level in a one-sided market in which only workers are active agents. It is also the information level that a designer who only cares about the well-being of workers would want to implement in a two-sided market.[27] Focusing on these two information levels allows to isolate the effects which originate from the two-sidedness of the matching market and are not prevalent in one-sided markets.

If information is costless it is easy to see from our previous discussion that the worker-optimal information level is to be perfectly informed (cf. Theorem 2.2). However, this is not necessarily the socially optimal information level since aggregate welfare may be decreasing in information precision (cf. Example 2.3). The same is true for the *firm-optimal* information level, i.e., the precision of workers' information which maximizes firms' welfare. In any case, if information is costless, the firm-optimal and the socially optimal level of information will always be extreme, that is, either full information or no information.

But there exist also distributions with an increasing hazard rate and monotone decreasing density functions (cf. Bagnoli and Bergstrom, 2005).

[27]This applies to setting in which there is a lobby group representing agents on one side of the market. Examples include the parent empowerment movement or labor unions. It should be noted that workers would also choose $\alpha^{ao}$ if they could coordinate on a common information level. e.g. by collectively choosing an information technology.

## Costly Precision

We now consider the case when information is costly. To formally analyze this case, let $\mathcal{S}$ be a set of feasible information technologies which is totally ordered in terms of strict precision. That is, there exists some $\mathcal{A} \subseteq [0, \infty)$ such that $\mathcal{S} = \{S_\alpha\}_{\alpha \in \mathcal{A}}$ and $S_\alpha$ is more precise than $S_{\alpha'}$ if and only if $\alpha > \alpha'$. Information technology $S_\alpha$ is characterized by $\{G^\alpha(\cdot|x)\}_{x \in \mathcal{X}}$. For ease of presentation, we restrict attention to linear information models, with $E[X|S] = \alpha S + (1-\alpha)E[X]$, $\alpha \in [0,1]$. The natural indexation in this case is to denote by $S_\alpha$ the information technology that results in $E[X|S] = \alpha S + (1-\alpha)E[X]$.

The following condition on the distribution of signals guarantees that for all precision levels $\alpha \in (0,1)$, the distribution and density functions of the posterior estimates, $H^\alpha$ and $h^\alpha$, are continuously differentiable in the precision level $\alpha$:

**Assumption 2.2** (Differentiable Signals)**:**
The marginal distribution of signal realizations $G$ is twice continuously differentiable in $s$.

We assume that information costs have a 'pay per signal' structure. For information technology $S_\alpha \in \mathcal{S}$ of precision $\alpha \in [0,1]$ every worker who receives a signal from $S_\alpha$ has to pay $c(\alpha) \in \mathbb{R}^+$. Precision costs are increasing in $\alpha$ and capture for example investments in time or resources to generate or collect information. The precision-cost-function $c : [0,1] \to [0,\infty)$ is increasing and continuously differentiable with $c(0) = 0$ and $c'(0) = 0$. Since every worker obtains exactly one signal, total costs for information technology $S_\alpha$ are $C(\alpha) := nc(\alpha)$.

Given Assumption 2.2 expected welfare of workers and expected aggregate welfare are continuously differentiable in the level of information of workers (cf. Lemma 2.6). In order to identify and compare the worker-optimal and socially optimal information levels, we impose the following single-crossing conditions.

**Assumption 2.3** (Single-crossing)**:**
(i) $\frac{\partial c/\partial \alpha}{\partial W_w/\partial \alpha}$ is strictly increasing in $\alpha \in (0,1)$.
(ii) $\frac{\partial c/\partial \alpha}{\partial W/\partial \alpha}$ is strictly increasing in $\alpha \in (0,1)$ whenever $\frac{\partial W}{\partial \alpha} > 0$.

For $\mathcal{S}$ being the set of truth-or-noise technologies and $X_i \overset{iid}{\sim} U[0,1]$, Assumption 2.3 is satisfied for convex precision costs.

We now characterize the relation between the worker-optimal and the socially optimal information level if information is costly.

**Theorem 2.5.** *In a precision and matching game, suppose Assumption 1–3 are satisfied.*
*(i) The socially optimal information level of workers is higher than the worker-optimal level,* $\alpha^{wo} \leq \alpha^{so}$*, if*
  *a) workers constitute the long side of the market and* $n$ *is sufficiently large, or*
  *b) the distribution of firms' types* $F_Y$ *has a decreasing hazard rate.*
*(ii) The socially optimal information level of workers is lower than the worker-optimal level,* $\alpha^{so} \leq \alpha^{wo}$*, if workers constitute the short side of the market and the distribution of firms' types has an increasing hazard rate.*

It is not surprising that in a precision and matching game the worker-optimal and socially optimal levels of precision do not coincide. Information of workers imposes an externality on firms. For a higher information level a better allocation can be achieved which increases total match output. However, more information also leads to more differentiation among workers, which increases competition among firms. This results in higher expected investments of firms. The relation between the worker-optimal and the socially optimal level of information depends on whether the overall effect of a higher information level of workers on firms is positive or negative.

Theorem 2.5 illustrates that in a relatively balanced market, with groups of similar sizes on each side of the market, the relation between the worker-optimal and socially optimal information level depends on the distribution of the informed agents' types. However, the socially optimal information level is always higher than the worker optimal information level, if the group of uniformed agents constitutes the long side of the market and is sufficiently large.

# 7 Applications

In this section we discuss implications of the results established in Section 5 and Section 6 for various market design settings, in particular auctions, contests and matching markets. In each of the applications we highlight certain features of our results and contributions.

## 7.1 Auctions

The standard private values auction setting is captured as a special case by our model. It corresponds to the case in which workers represent the bidders of the auction, and there is only one firm of a given, commonly known type, representing for example the auction platform or the object to be sold. The seller is a third party who collects the bids. As summarized in Table 2.4, the expected valuation of the winning bidder is $\frac{1}{2} \cdot O$, his expected information rent is $W_w$, and $T_w$ represents the seller's expected revenue.

| auctions | exp. valuation of the winning bidder | exp. revenue of the seller | exp. information rent of the winning bidder |
|---|---|---|---|
| general model | $\frac{1}{2}O = \mu^S_{1:n}$ | $T_w = \mu^S_{2:n}$ | $W_w = \mu^S_{1:n} - \mu^S_{2:n}$ |

Table 2.4: Translation of our results to the standard auction setting with $n$ bidders.
Reminder: $\mu^S_{i:n}$ denotes the $i^{th}$ mean order statistics of the posterior distribution of bidders' types.

Translating our results of Section 5 to the auction setting yields the following insights: Disclosing information to bidders increases the expected valuation of the winning bidder (Proposition 2.1) as well as his expected information rent (Theorem 2.2).[28] Moreover, the seller's

[28] For the result on expected information rents of the winning bidder a version of *strong precision* is needed. The result follows for both criteria of strong precision, the one used in this essay as well as the concept of *supermodular precision* adopted by Ganuza and Penalva (2010).

expected revenue increases, if there are sufficiently many bidders (Theorem 2.2). If information disclosure is costly, the revenue maximizing level is below the efficient level (Theorem 2.5). These observations correspond to the results established in Ganuza and Penalva (2010) on information disclosure in auctions. Our results therefore include their results on information disclosure in auctions as a special case.

The methods from statistics that we adopt in this essay, provide an alternative and shorter way to prove these results. Using these statistical methods allows furthermore to strengthen the results of Ganuza and Penalva (2010). For example, the result on the expected informational rent of the winning bidder can be strengthened to the statement that providing more information to bidders increases the expected informational rent of the winning bidder in terms of first-order stochastic dominance and not only in expectation.[29]

The statistical methods used in this essay are powerful and should be explored further because they hold the promise to yield interesting results and insights in mechanism design settings with endogenous information. Let us support this point by providing a small, new result that we can establish by exploring these methods further.

Let $\mathcal{S} = \{S_\alpha\}_{\alpha \in \mathcal{A}}$ be a ordered set of information technologies such that $S_{\alpha'}$ is more precise than $S_\alpha$, if $\alpha' > \alpha$. Say that the *information level* of bidder $i$ is $\alpha_i$, if he receives a signal from information technology $S_{\alpha_i}$ in the information stage. Bidders may receive signals from different information technologies, that is, of different precision, resulting in a profile of information levels of bidders $(\alpha_1, \ldots, \alpha_n)$. We say that the information level of bidders *weakly increases* if the information level of at least one bidder strictly increases, and the information levels of all other bidders are non-decreasing.

**Proposition 2.2.** *In a private values auctions setting consider any auction format that implements the efficient allocation. For any weak increase in the information levels of bidders, the expected value of the winning bidder increases.*

To our knowledge, this generalization of the results on information disclosure in auctions is new to the literature. It establishes that any weak increase in the information level of bidders will increase the expected efficiency of the allocation of the auction. For this result, the additional information provided to, or processed by, the individual bidders may be heterogeneous, which is natural feature in many situations.

Consider for example a seller who publicly discloses information that is relevant for bidders to learn about their valuation for the object for sale. Typically, the level of information that individual bidders extract from the publicly available data differs across bidders. Proposition 2.2 establishes that the effect on the expected valuation of the winning bidder does not depend on this detail. Providing more information will always increase the expected valuation of the winning bidder.

[29] The result is a simple corollary to theorem 3.B.31 in Shaked and Shanthikumar (2007). This was pointed out in a footnote in Ganuza and Penalva (2010) but the fact that this alternative proof yields a stronger result was mentioned only recently in Shaked et al. (2012).

## 7.2 Two-sided Matching Markets

When applied to matching markets, our results provide a first theoretical study of the effects of private information and the information level of market participants on the equilibrium outcome in matching markets.

As we have discussed in Section 4, a higher information level of workers leads to more differentiation among workers in the matching game. This effect allows for a better allocation: high-abilitiy workers are more likely to be matched to a firm of similar ranking, which results in an increase of total match output (Proposition 2.1).[30]

Moreover, a higher information level and the resulting higher differentiation among workers, also raises the stakes for the firms of being matched to a better or worse partner. This results in an increase of firms' expected signaling investments (Theorem 2.3).

The effect on workers' expected signaling investments is less clear-cut and depends on certain features of the market (Theorem 2.2). For a higher information level of workers, the (marginal) benefit from obtaining a better match increases for high-ranked workers, whereas it is decreasing for low-ranked workers. This results in high-ranked workers increasing their investments in signaling whereas lower ranked workers may invest less. If workers constitute the long side of the market, only high-ranked workers are matched in equilibrium and workers' expected total investments are increasing. If there are more firms than workers, the effect may be reveres.

These results illustrate, that in finite matching markets, some of the effects of a higher information level of market participants depend on whether information is disclosed to agents on the short or the long side of the market. This new insight is made possible because our comparative statics result in Section 5 apply for arbitrary finite group sizes on the two sides of the market.[31] This observation highlights that it is important to study matching market models with a finite number of agents and not restrict attention solely on the case with a continuum of agents on both sides of the market.[32]

An important insight from our analysis is the following: In a two-sided matching market in which both sides of the market invest in wasteful signaling to compete for match partners, the trade-off between a better allocation and a potential increase in wasteful signaling investments may result in a decrease of expected aggregate welfare when the information level of market participants increases (Example 2.3). This feature is specific to two-sided markets. Welfare of

[30] This effect is observed in empirical studies. For example, in their study Hoxby and Turner (2013) provide a subgroup of high-school seniors with additional information about their college opportunities and find that, for students who received information, the probability to enroll in a college that matches their abilities increases significantly.

[31] This also allows us to consider information disclosure in auctions as a special case of our results. Moreover, the projection of our results to the model of Hoppe et al. (2009) generalizes their results on comparative statics effects of group heterogeneity (they only consider the case $n = k$).

[32] Most models which discuss comparative static settings study models with a continuum of agents on each side of the market. Considering a continuum of agents is often a reasonable and very useful assumption, since it avoids the technicalities of having to deal with order statistics. However, the point we want to make here is that it is also important to study the model with finite sets of agents.

agents on the side of the market receiving more information is always increasing in expectation (Theorem 2.2). Consequently, in settings in which only one side of the market are active agents, providing these agents with additional information would unambiguously increase expected welfare. By contrast, in a two-sided matching market, for an increase in the information level of agents, the amplified competition among agents within their groups and the resulting increase in signaling investments may eat up all additional match surplus.

What can we learn from this? Our results indicate that in a two-sided matching market more information is not necessarily better if the objective is to maximize overall expected welfare. However, there are often different or additional objectives like equalizing the information level across agents, fairness considerations or incentivizing schools or colleges to invest in their quality. Our results yield insights into the last aspect. If parents are provided with more information about school choices, the highest-ranked schools profit most from informed choices of parents, whereas low-ranked schools may be worse off (Theorem 2.3).[33] This may serve as a formal rationale for the claim often raised by the parent empowerment movement, that providing parents with more information results in them making more informed choices, which – in the long run – will increase school quality. From this perspective, it may even be good to let a lobby for one side of the market, determine the level of information of market participants, even though they do not fully interalize the costs and benefits from additional information and thus will not choose the socially optimal information level.[34] Theorem 2.5 suggest that, if the uninformed agents – the workers – constitute the long side of the market, the worker-optimal level of information for this side of the market is higher than the socially optimal level. Given that only high-quality firms profit from the informed choices by workers, this incentivizes firms to compete for the highest-rank among their peers which may induce them to invest in quality-enhancing policies.

## 7.3 Contests

Our model can also includes a rank-order tournament or contest setting as a special case. To see this, interpret agents on one side of the market, say the firms, as representing the prizes in a promotion tournament, with commonly known values $\eta_{1:k} \geq \cdots \geq \eta_{k:k}$. This side of the market is passive. Workers represent the participants of the contest. Workers' types reflect their abilities, and their investments correspond to the effort, which they exert. With this interpretation, workers' investments are not wasteful but they are collected by a third party – the company or organization running the promotion tournament.

In promotion tournaments there are two natural objectives: To promote the best workers and to maximize workers' efforts. Translating our results from Section 5 and Section 6 to

[33] This claim still remains true if schools are not considered to be active agents and therefore do not invest in signaling about their types.

[34] Distributing information among parents, providing more or less detailed information on websites, or determining the precision level of standardized test like the SAT are examples of technologies that can serve to influence the level of information of market participants.

the promotion contest, we obtain the following predictions. More information of competitors increases the probability to promote the best workers (Proposition 2.1). If the ratio of workers to prizes is sufficiently large, then workers' expected overall effort is increasing in their information level. However, if this ratio is too small, providing information to the workers may not be effort enhancing – not even on an aggregate level (Theorem 2.2).

There is a second option to project our model to a contest, interpreting firms as contestants and the mean-order statistics of posterior types of workers, $\mu^S_{1:n} \geq \cdots \geq \mu^S_{n:n}$, as prizes. In this case, the translation of Theorem 2.3 yields the well-known observation that increasing the prize-spread in contests results in an increase in workers' effort.[35]

A typical question in the contest literature is how to design an optimal contest in order to maximize workers' effort. Our results indicate that information management through feedback systems may serve as an useful element of contest design. Let us provide some details for this insight. The standard design element that is usually considered in the contest literature are prizes. The number and distribution of prizes in a contest affect workers' effort, and can therefore be used to design an optimal prize-structure that maximizes workers' efforts. However, in some organizations it may not be feasible to implement the optimal prize-structure suggested by theoretical models, because there are certain constraints on the number or distribution of prizes. For example, in promotion tournaments the prize-structure is determined by the wage schedule and the number of positions on each level of the organization. In situations in which the optimal prize-structure cannot be implemented, a designer could influence workers' effort by implementing a feedback systems to manipulate the information level available to workers. We refer to this design element as *information management.*

Our results shed light on properties of optimal feedback systems in contests. They suggest that we should observe different feedback systems depending on the ratio of workers and prizes in a contest. In organizations with steeper hierarchies or an up-or-out system, we should expect stronger feedback systems to be in action, for example a high frequency of periodical performance reports. By contrast, for organizations with flat hierarchies or *promotion by seniority* practices, our results predict less sophisticated feedback structures. These predictions seem to be in line with common practices. For example, large consulting firms with an up-or-out policy are known to have a very rigorous feedback structure.

# 8 Related literature

This essay is related to various strands of literature. It is connected to the vast matching literature that emerged from the seminal papers by Gale and Shapley (1962), Shapley and Shubik (1971) and Becker (1973). Most of the theoretical analysis of matching markets focuses on complete information models in which agents' preferences, types, and match values are

[35]See for example Lazear and Rosen (1981), Moldovanu et al. (2007), Connelly et al. (2014) and references therein – also of empirical studies supporting these theoretical predictions.

common knowledge. There is an emerging literature studying incomplete information matching models and issues of screening and signaling which arise therein. See for example Hoppe et al. (2009), Hopkins (2012), and Bilancini and Boncinelli (2013). Our second-stage game is based on the models analyzed in Hoppe et al. (2009) and Hopkins (2012). They study two-sided matching markets, in which agents on one or both sides of the market have private information about their characteristics. Agents invest in costly signaling à la Spence (1973) to compete for match partners. As Hoppe et al. (2009) we consider a small market with a finite number of agents, whereas Hopkins (2012) studies a model with a continuum of agents. In a related paper by Bilancini and Boncinelli (2013), agents on both sides of the market have private information about their skills and can choose whether or not to disclose this information. For one side of the market information is not verifiable and disclosing information yields certification costs. All of the aforementioned papers consider matching markets in which agents have private information about their characteristics, and analyze the costs and benefits from disclosing this information. By contrast, our focus is on disclosing information to agents about their types. We study how different information levels of participants in (matching) markets affect the resulting equilibrium properties and welfare.

Related questions are addressed in the literature on information disclosure in auctions. Ganuza and Penalva (2010) discuss the effects of different information levels of buyers in a second-price auction, whereas Bergemann and Pesendorfer (2007), Esö and Szentes (2007), and Ganuza and Penalva (2014) adopt a mechanism design perspective. In a private values environment, these papers discuss the revenue maximizing information structure and selling mechanism for the seller. Similar to Ganuza and Penalva (2010) we focus on a given mechanism and study how different information levels affect the equilibrium.

Our analysis extends the discussion of information disclosure in auctions to two-sided matching markets and can also be applied to contests and rank-order tournaments. A recent survey of the contest literature is provided by Connelly et al. (2014). The focus of most of these papers, for example Moldovanu and Sela (2001, 2006) is on optimal contest design, that is, on the optimal portfolio of prizes or how to split the contestants in subgroups to achieve the designer's objective. There is a growing literature studying the role of feedback and optimal feedback systems in contests. Examples include Aoyagi (2010), Goltsman and Mukherjee (2011), Hansen (2013), and Ederer (2010). All of these models restrict attention to the two-agent case and most of them only consider full or no disclosure policies. By contrast, we allow for arbitrary finite numbers of prizes and workers. A new insight that can be gained from our results is that the optimal feedback policy depends on the ratio of workers to prizes (see discussion in Section 7.3).

This essay also ties to the literature that identifies and explores connections between auctions and matching markets in order to establish new results. See for example Demange and Gale (1985) and Hatfield and Milgrom (2005). The matching market that we study, can be considered as the combination of two multi-object auctions with two sides of active agents (cf. discussion in Section 3). To establish our results we use this connection between matching markets and auctions and identify a relation between the statistical methods used by Hoppe et al. (2009) and the type of precision criterion introduced by Ganuza and Penalva (2010).[36]

Our essay also relates to the growing literature on pre-match investments in matching tournaments. Examples include Cole et al. (2001), Peters and Siow (2002), and Mailath et al. (2013) and Dizdar (2015). Pre-match investments generate first-order effects on agents' types. Similarly, Hopkins (2012) studies such first-order effects, interpreting shifts in agents' type distribution in terms of first-order stochastic dominance as a more competitive environment. By contrast, in our analysis, investments in information yield second-order effects. A higher level of information leads to a more dispersed distribution of workers' posterior types in the second-stage matching game.

# 9 Conclusion

In this essay we studied the impact of the level of information available to market participants in a two-sided matching market. We illustrated that for a higher information level of workers there is a trade-off between the increased match surplus from the better allocation, and the welfare reducing effects of increased competition among agents. It was shown that the increased competition among agents may be so strong that it eats up all additional match surplus. In this case, a higher information level of market participant reduces welfare. Our results not only provide a first study of information disclosure in matching markets, but can also be applied beyond the matching setting. We discussed implications of our results for auctions and contests.

A notable distinction between the effects of information disclosure in auctions and matching markets is the following: In an auction, a seller faces a trade-off between efficiency and having to leave information rents to the buyers. In a matching market information disclosure yields a trade-off between allocative efficiency and the welfare-reducing effects of increased competition among agents on both sides of the market.

The setup of the model and the discussion of applications to different market design settings illustrated how these settings are connected. We used these insights in this essay to identify a relation between the discussions in Hoppe et al. (2009) and Ganuza and Penalva (2010). Establishing a link between the methods adopted in these papers provided us with a new approach to study the impact of information disclosure in two-sided matching markets and related applications. We believe that the connections that we have identified between the different models and concepts will prove to be useful in future research, in particular to study mechanism design problems with endogenous information of agents.

[36] Methodologically, a related paper is Chi (2014), who uses statistical methods to study informational effects in Bayesian decision problems.

# Appendix

## A Technical Prerequisites

In this section we present the main techniques used to prove our results. The methods stem from statistics and reliability theory. Shaked and Shanthikumar (2007) provide a comprehensive treatment of order statistics whereas Marshall et al. (2011) is a good reference for the theory of majorization. If not indicated otherwise, all definitions and theorems stated in this section can be found in these two books.

The single-crossing property of quantile functions that we use in Definition 2.1 is equivalent to the distribution of posterior estimates being ordered in terms of the star order (see Shaked and Shanthikumar (2007), Section 4.B).

**Fact 2.1.** *If $S_2$ is more single-crossing precise than $S_1$, then $E[X|S_2]$ is greater than $E[X|S_1]$ in the* star-order*, $E[X|S_2] \geq_* E[X|S_1]$.*

**Notation.** $X \leq_{MPS} Y$, then $Y$ is a mean-preserving spread of $X$.
$X \leq_* Y$, then $Y$ is greater in the star order than $X$.

**Definition 2.2:** Consider two ordered n-dimensional real-valued vectors $\mathbf{a} = (a_1, \dots, a_n)$, and $\mathbf{b} = (b_1, \dots, b_n) \in \mathbb{R}^n$, with $a_1 \geq \dots \geq a_n$ and $b_1 \geq \dots b_n$. We say that $\mathbf{a}$ *submajorizes* $\mathbf{b}$, ($\mathbf{a} \succ_{sub} \mathbf{b}$), if

$$\sum_{i=1}^{m} a_i \geq \sum_{i=1}^{m} b_i \quad \text{for all } m = 1, \dots, n. \tag{2.2}$$

If in addition (2.2) holds with equality for $m = n$ we say that $\mathbf{a}$ *majorizes* $\mathbf{b}$, ($\mathbf{a} \succ \mathbf{b}$).

A function $\phi : \mathbb{R}^n \supseteq A \longrightarrow \mathbb{R}$ is *Schur-convex* (resp. *Schur-concave*) if, whenever $a$ majorizes $b$, $\mathbf{a} \succ \mathbf{b}$, then $\phi(\mathbf{a}) \geq \phi(\mathbf{b})$ (resp. $\phi(\mathbf{a}) \leq \phi(\mathbf{b})$).
If $\mathbf{a} \succ_{sub} \mathbf{b}$ then $\phi(\mathbf{a}) \geq \phi(\mathbf{b})$ for every Schur-convex and increasing function $\phi$.

To proof our results we repeatedly use the following important results from statistics.[37]

**Theorem 2.6** (Cal and Carcamo 2006). *Let $X$ and $Y$ be integrable random variables with equal means and $F(0) = G(0) = 0$. Then if $X \leq_{MPS} Y$, the vector of mean order statistics of $Y$, $(E[Y_{1:n}], \dots, E[Y_{n:n}])$* majorizes *the vector of mean order statistics of $X$ for all $n \geq 1$. That is,*

$$(E[Y_{1:n}], \dots, E[Y_{n:n}]) \succ (E[X_{1:n}], \dots, E[X_{n:n}]).$$

**Theorem 2.7** (Barlow and Proschan (1966)). *Let $X$ and $Y$ be integrable random variables with equal means and $F(0) = G(0) = 0$. Then if $X \leq_* Y$ this implies $X \leq_{MPS} Y$, and moreover*

[37] In their paper Cal and Carcamo (2006) establish this result for random variables ordered in terms of the convex-order. In our informational setting we always compare random variables with finite and equal means. In this case the convex order is equivalent to a mean-preserving spread.

*(i) For $1 \leq r \leq n$:*

$$\sum_{i=r}^{n} i \cdot \left(E\left[X_{i:n}\right] - E\left[X_{i+1:n}\right]\right) \geq \sum_{i=r}^{n} i \cdot \left(E\left[Y_{i:n}\right] - E\left[Y_{i+1:n}\right]\right).$$

*(ii) For $a_1 \leq \cdots \leq a_n$:*

$$\sum_{i=1}^{n} a_i \cdot i \cdot \left(E\left[X_{i:n}\right] - E\left[X_{i+1:n}\right]\right) \geq \sum_{i=1}^{n} a_i \cdot i \cdot \left(E\left[Y_{i:n}\right] - E\left[Y_{i+1:n}\right]\right).$$

**Lemma 2.2.** *Let $F$ be a distribution function with $F(0) = 0$ and an increasing hazard rate (IHR). Then, for fixed $n$, the normalized spacings of order statistics $i \cdot (X_{i:n} - X_{i+1:n})$ are stochastically increasing in $i = 1, \ldots, n$. That is:*

$$(X_{1:n} - X_{2:n}) \leq_{FOSD} 2 \cdot (X_{2:n} - X_{3:n}) \leq_{FOSD} \cdots \leq_{FOSD} n \cdot (X_{n:n} - X_{n+1:n}).$$

*If $F$ has a decreasing hazard rate (DHR), then the normalized spacings are stochastically decreasing in $i$.*

We will also need the following result

**Lemma 2.3.** *Let $X, Y$ be nonnegative random variables with distribution functions $F$ and $G$, respectively, such that $F(0) = G(0) = 0$. If $X \leq_* Y$ then*
*(i) $\frac{E[Y_{i:n}]}{E[X_{i:n}]}$ is decreasing in $i$, and*
*(ii) $\frac{E[Y_{i:n}]}{E[X_{i:n}]}$ is increasing in $n$.*

# B Proofs

*Proof of Lemma 2.1.* is a direct corollary of theorem 3.6 in Barlow and Proschan (1966). □

*Proof of Proposition 2.1.* If $S_2 \succsim S_1$ then $E[X|S_2] \geq_{MPS} E[X|S_1]$ and, by Theorem 2.6, $(\mu^{S_2}_{1:n}, \ldots, \mu^{S_2}_{n:n}) \succ (\mu^{S_1}_{1:n}, \ldots, \mu^{S_1}_{n:n})$.

For $k \geq n$, $O = \sum_{i=1}^{n} \eta_{i:k}\mu_{i:n}$ is Schur-convex in the vector of mean order statistics of workers' characteristics and consequently if $S_2 \succsim S_1$ then $O(S_2) \geq O(S_1)$.

For $k < n$, $O = \sum_{i=1}^{k} \eta_{i:k}\mu_{i:n}$ is Schur-convex in the *truncated* vector of mean-order statistics of workers' posterior types, $\mu|_{\leq k} = (\mu_{1:n}, \ldots, \mu_{k:n})$. A higher information level of workers only results in (weak) submajorization of the truncated vectors of mean-order statistics, i.e.

$$S_2 \succsim S_1 \quad \Rightarrow \quad \mu^{S_2}\big|_{\leq k} \succ_{sub} \mu^{S_1}\big|_{\leq k}.$$

Since $O$ is increasing and Schur-convex it follows that $O(S_2) \geq O(S_1)$. □

In order to prove Theorem 2.2 we first establish a technical Lemma. To state and prove it we need the following fact.

**Fact 2.2** (Theorem 3.A.5 in Shaked and Shanthikumar 2007)**.** *The following conditions are each sufficient and necessary for* $X \leq_{MPS} Y$

$$\int_0^p \left(G^{-1}(u) - F^{-1}(u)\right)\, du \leq 0 \quad \forall\, p \in [0,\, 1]\,, \quad \textit{and} \tag{2.3}$$

$$\int_p^1 \left(G^{-1}(u) - F^{-1}(u)\right)\, du \geq 0 \quad \forall\, p \in [0,\, 1]\,. \tag{2.4}$$

**Lemma 2.4.** *Let* $X$, $Y$ *be random variables with continuous differentiable distributions* $F$ *and* $G$ *and equal means, such that* $X \leq_{MPS} Y$*. Then, for every* $k \in \mathbb{N}$ *there exists some* $\hat{n}_k$ *such that*

$$E\left[X_{k:n}\right] \leq E\left[Y_{k:n}\right] \quad \forall\, n \geq \hat{n}_k.$$

*Proof.* The methods in this proof are similar to the ideas used to prove theorem 1 in Ganuza and Penalva (2010).
We can apply the probability integral transformation[38] to obtain the following simple formula for the $k^{th}$ order statistics of $X$:

$$E\left[X_{k:n}\right] = \frac{n!}{(k-1)!(n-k)!} \int_0^1 F^{-1}(u) u^{n-k} (1-u)^{k-1}\, \mathrm{d}u$$

Set $\phi(u) := G^{-1}(u) - F^{-1}(u)$. Since $G$ and $F$ are continuously differentiable, by the inverse function theorem $F^{-1}$ and $G^{-1}$ are continuous and so is $\phi$.

Suppose that $\phi(u) \neq 0$ on a subset of $[0,1]$ with nonempty interior.[39] Define $L := \{u \in X\,[0,1] : \phi(u) < 0\}$ and $\overline{u} := \sup\{L\}$. (2.3) and (2.4), continuity of $\phi$ and the assumption that $\phi(u) \neq 0$ on a subset of $[0,1]$ of positive measure imply that $\overline{u} \in (0,\, 1)$. We obtain that there exist $p_1,\, p_2 \in (\overline{u},\, 1]$ such that $\phi(u) > 0$ for all $u \in [p_1,\, p_2]$. Set $c_1 := \min_{u \in [0, p_1]}\{\phi(u)(1-u)^{k-1}\}$ and $c_2 := \min_{u \in [p_2, 1]}\{\phi(u)(1-u)^{k-1}\}$. By construction $c_1 < 0$ and $c_2 > 0$. This yields:

$$\begin{aligned} E\left[Y_{k:n}\right] - E\left[X_{k:n}\right] =& k \binom{n}{k} \int_0^1 \left(G^{-1}(u) - F^{-1}(u)\right) u^{n-k} (1-u)^{k-1}\, \mathrm{d}u \\ \geq& \frac{n!}{(k-1)!(n-k+1)!} p_2^{n-k+1} \left[ \left(\frac{p_1}{p_2}\right)^{n-k+1} (c_1 - c_2) + c_2 \right] \end{aligned}$$

Set $\hat{n} := \lceil k - 1 + \frac{\ln\left(\frac{c_2}{c_2 - c_1}\right)}{\ln\left(\frac{p_1}{p_2}\right)} \rceil$ where $\lceil x \rceil$ denotes the smallest natural number greater or equal than $x$. It follows that:

$$\frac{n!}{(k-1)!(n-k+1)!} p_2^{n-k+1} \left[ \left(\frac{p_1}{p_2}\right)^{n-k+1} (c_1 - c_2) + c_2 \right] \geq 0 \quad \forall\, n \geq \hat{n}$$

[38]For every random variable $X$ with continuous c.d.f. $F$ and density $f$, the transformed random variable $F(X)$ has a standard uniform distribution, $F(X) \sim U\,[0,1]$

[39]The case $\phi(u) = 0$ a.e. is trivial.

□

*Proof of Theorem 2.2.*

(i) *Total expected investments:*

a) By Lemma 2.4 there exists some $\hat{n} > k+1$ such that for every $n \geq \hat{n}$, $\mu^{S_2}_{k+1:n} - \mu^{S_1}_{k+1:n} \geq 0$. Since $\mu^{S_1}_{k+1:n} \geq 0$, it follows that

$$\frac{\mu^{S_2}_{k+1:n}}{\mu^{S_1}_{k+1:n}} \geq 1 \quad \forall n \geq \hat{n}.$$

By Lemma 2.3, $\frac{\mu^{S_2}_{i:n}}{\mu^{S_1}_{i:n}}$ is decreasing in $i$ for every $n$ and it follows that $\mu^{S_2}_{i:n} - \mu^{S_1}_{i:n} \geq 0$ for all $i \leq k+1$. We obtain that for all $n \geq \hat{n} > k$:

$$\begin{aligned} T_w(S_2) - T_w(S_1) &= \sum_{i=1}^{\min\{n,k\}} (\eta_{i:k} - \eta_{i+1:k}) \cdot \left(\mu^{S_2}_{i+1:n} - \mu^{S_1}_{i+1:n}\right) \\ &= \sum_{i=1}^{k} \underbrace{(\eta_{i:k} - \eta_{i+1:k})}_{>0} \cdot \underbrace{\left(\mu^{S_2}_{i+1:n} - \mu^{S_1}_{i+1:n}\right)}_{\geq 0} \geq 0. \end{aligned}$$

b) If $F_Y$ has an increasing hazard rate, then the normalized spacings $i(\eta_{i:k} - \eta_{i+1:k})$ are stochastically increasing in $i$ (Lemma 2.2). Set $\widetilde{T}_w := \sum_{i=0}^{n} i(\eta_{i:k} - \eta_{i+1:k})\mu_{i+1:n}$. Then, for $n \leq k$, $T_w = \widetilde{T}_w$ and $\widetilde{T}_w$ is Schur-concave in the vector of mean order statistics of workers' characteristics. It follows that $T_w$ is decreasing (non-increasing) in the level of information of workers.

(ii) *Workers' expected welfare:*

Set $a_i := -\eta_{i:k}$. Then, applying Theorem 2.7 (ii) yields

$$S_2 \succsim_* S_1 \quad \Rightarrow \quad W_w(S_2) \geq W_w(S_1).$$

□

*Proof of Theorem 2.3.*

(i) *Total expected investments:*

Analogous to the proof of Theorem 2.2 (ii) whereas the case-by-case analysis is now for $k \leq n$ and $k > n$.

(ii) *Firms' expected welfare:*

If $F_Y$ has a decreasing hazard rate, by Lemma 2.2 the normalized spacings $i(\eta_{i:k} - \eta_{i+1:k})$ are stochastically decreasing in $i$. Consequently, $W_f$ is Schur-convex in the vector of conditional mean order statistics of workers. By Theorem 2.6 it follows that $W_f(S_2) \geq W_f(S_1)$, if $S_2$ is more precise than $S_1$. The results for the case when $F_Y$ has an increasing hazard rate, follow from arguments analogous to those used to prove Theorem 2.2 (i). □

*Proof of Theorem 2.4.* $W = W_w + W_f$. Rearranging terms yields:

$$W(S) = \underbrace{\left[ \sum_{i=1}^{\min\{n,k\}} i \cdot \left( \mu_{i:n}^S - \mu_{i+1:n}^S \right) \left( \eta_{i:k} - \eta_{i+1:k} \right) \right]}_{(W_f - T_w)(S)} + \underbrace{\sum_{i=1}^{\min\{n,k\}} \mu_{i:n}^S \eta_{i:k}}_{0.5O}$$

By Proposition 2.1 we know that total match output is increasing in precision. Whether aggregate welfare is increasing or decreasing in the level of workers' information depends on the effect on $W_f - T_w$, and, for the case that $W_f - T_w$ is decreasing, on which of these effects dominates. Throughout the proof, let $S_2$ be more single-crossing precise than $S_1$.

(i) If $F_Y$ has a decreasing hazard rate, both $W_w$ and $W_f$ are increasing in precision (cf. Theorem 2.2 and Theorem 2.3) and so is $W = W_w + W_f$.

(ii) Suppose $n < k$ and $f_Y$ is monotone decreasing. In this case

$$W = \sum_{i=1}^{n} \mu_{i:n}^S \eta_{i:k} + \left[ \sum_{i=1}^{n} i \cdot \left( \mu_{i:n}^S - \mu_{i+1:n}^S \right) \left( \eta_{i:k} - \eta_{i+1:k} \right) \right].$$

We use the following result which establishes that the spacings of order statistics from random variables with monotone density functions can be ordered in terms of stochastic dominance.[40]

**Lemma 2.5.** *Let $Y_1, \ldots, Y_n$ be independently, identically distributed random variables with finite support and density function $f_Y$. Then,*

*(i) if $f_Y$ is monotone increasing (non-decreasing)*

$$Y_{i:n} - Y_{i+1:n} \leq_{FOSD} Y_{i+1:n} - Y_{i+2:n} \quad \forall\, i = 1, \ldots, n-2$$

*(ii) if $f_Y$ is monotone decreasing (non-increasing)*

$$Y_{i:n} - Y_{i+1:n} \geq_{FOSD} Y_{i+1:n} - Y_{i+2:n} \quad \forall\, i = 1, \ldots, n-2$$

It follows directly that if $f_Y$ is monotone decreasing, the expected spacings of mean order statistics $(\eta_{i:k} - \eta_{i+1:k})$, $i = 1, \ldots, k-1$ are decreasing in $i$. Setting $a_i := -(\eta_{i:k} - \eta_{i+1:k})$, by Theorem 2.7 (ii) we obtain $(W_f - T_w)(S_2) \geq (W_f - T_w)(S_1)$, for $S_2 \succsim_* S_1$. It follows that aggregate welfare is increasing in workers' information level. □

**Lemma 2.6.** *For linear information technologies, under Assumption 2.2, for all $\alpha \in (0,1)$, $H^\alpha$ and $h^\alpha$ are continuously differentiable in the precision level $\alpha$. Moreover, $O$, $T_w$, $T_f$, $W_w$, $W_f$ are continuously differentiable in $\alpha \in (0,1)$.*

[40] This result can be found in Shaked and Shanthikumar (2007).

*Proof.* For $\alpha \neq 0$, set $\phi(\alpha, w) := \frac{w-(1-\alpha)E(X)}{\alpha}$. Then, for linear information technologies and $\alpha \neq 0$, $H^\alpha(w) = G(\phi(\alpha, w))$, and $h^\alpha(w) = \frac{1}{\alpha} g(\phi(\alpha, w))$, for $\alpha = 0$, $H = G$. By Assumption 2.2 and since $\phi(w, \alpha)$ is continuously differentiable in $\alpha \in (0, 1)$, $H^\alpha(w)$ and $h^\alpha(w)$ are continuously differentiable in $\alpha$. Moreover, if $H^\alpha$ and $h^\alpha$ are continuously differentiable in $\alpha$ then so are the distributions of order statistics $H^\alpha_{i:n}$. The densities $h^\alpha_{i:n}$ are continuous in $\alpha$ for all $i = 1, \ldots, n$. This implies that the conditional mean order statistics $E\left[\widehat{X}^\alpha_{i:n}\right]$ are continuously differentiable in $\alpha$. It follows that $W$, $W_w$, $T_f$, $T_w$, $O$ are continuously differentiable in $\alpha$. □

*Proof of Theorem 2.5.* The marginal value of information for workers is $\frac{\partial W_w}{\partial \alpha}$ and the socially marginal value is $\frac{\partial W}{\partial \alpha} = \frac{\partial W_w}{\partial \alpha} + \frac{\partial W_f}{\partial \alpha}$. By Theorem 2.3, if $F_Y$ is DHR or if $n \geq \hat{n} > k$ then $\frac{\partial W_f}{\partial \alpha} > 0$ and it follows that, at any information level $\alpha \in (0, 1)$, $\frac{\partial W}{\partial \alpha} > \frac{\partial W_w}{\partial \alpha}$. However, if $F_Y$ is IHR and $n \leq k$, then social marginal gains from higher precision are lower than the marginal gains for workers, $\frac{\partial W}{\partial \alpha} < \frac{\partial W_w}{\partial \alpha}$.

In the *precision and matching game* with costly precision, the optimization problem for workers is:

$$\max_{\alpha \in [0,1]} \{U_c(\alpha) = W_w(S_\alpha) - n \cdot c(\alpha)\}$$

and for the social planner:

$$\max_{\alpha \in [0,1]} \{U_{SP}(\alpha) = W(S_\alpha) - n \cdot c(\alpha)\}$$

Given our assumptions on the cost function and Assumption 2.2, $U_c$ and $U_{SP}$ are continuously differentiable in $\alpha$. By the extreme value theorem this guarantees the existence of a solution to the optimization problem of workers, respectively the social planner. The single-crossing conditions, **(SC)** and **($\mathbf{SC_C}$)**, establish uniqueness.

If $U_c$ is increasing on $[0, 1]$, then the optimal level of precision for workers is $\alpha^{wo} = 1$, otherwise it is characterized by:

$$\left.\frac{\partial W_w}{\partial \alpha}\right|_{\alpha=\alpha^{wo}} = n \left.\frac{\partial c}{\partial \alpha}\right|_{\alpha=\alpha^{wo}} \tag{I}$$

Analogous reasoning shows that the unique socially optimal level of precision is either $\alpha^{so} \in \{0, 1\}$ or an interior solution exists which is characterized by:

$$\left.\frac{\partial W}{\partial \alpha}\right|_{\alpha=\alpha^{so}} = n \cdot \left.\frac{\partial c}{\partial \alpha}\right|_{\alpha=\alpha^{so}} \tag{II}$$

Suppose that $F_Y$ is DHR or that $n \geq \hat{n} > k$. In this case, at any information level $\widetilde{\alpha}$, the marginal gains for workers from higher precision are lower than the social marginal gains, $\left.\frac{\partial W_w}{\partial \alpha}\right|_{\alpha=\widetilde{\alpha}} < \left.\frac{\partial W}{\partial \alpha}\right|_{\alpha=\widetilde{\alpha}}$. Given uniqueness of the worker-optimal and the socially optimal level of precision we obtain $\alpha^{wo} \leq \alpha^{so}$.

The result for $F_Y$ being IHR and $n < k$ follows by analogous reasoning. □

*Proof of Proposition 2.2.* This result is a direct corollary of Theorem 7.6 in Chapter 4, Barlow and Proschan (1981). □

## C Discussion and relation to other informativeness criteria

Given our assumption that signals are monotone the natural informativeness criterion to use is the concept of *effectiveness* introduced by Lehmann (1988).[41] The basic idea behind this concept is that for a given state space $X$, information technology $S_2$ is more informative about $X$ than $S_1$, if the conditional distribution of $S_2$ is more dependent on $X$ than that of $S_1$. Formally,

**Definition 2.3** (Effectiveness, Lehmann (1988))**:** Given $X$, let $S_1$ and $S_2$ be two signals which satisfy the Assumption 2.1. Then $S_2$ is said to be *more effective* than $S_1$ if for all $s$

$$G_{S_2}^{-1}(G_{S_1}(s|x)|x) \quad \text{is nondecreasing in } x.$$

Mizuno (2006) shows that for a more effective signal about $X$ the resulting distribution of conditional expectations is more dispersed.

**Theorem 2.8** (Mizuno 2006)**.** *If signals are monotone, then if $S_2$ is more effective than $S_1$, it follows that $S_2$ is more integral precise than $S_1$ for all priors.*

Our definition of precision is similar to the notion of *integral* and *supermodular precision* in Ganuza and Penalva (2010) but the stochastic orders used to define these concepts differ. Our *precision* criterion in Definition 2.1 is based on the star order whereas *integral precision* is based on the *convex order* and *supermodular precision* is based on the *dispersive order*.[42] We briefly discuss the relation of these criteria which amounts to analyzing the relation of the stochastic orders.[43]

Let $X$ and $Y$ be two random variables with interval support and distribution functions $F$ and $G$, respectively. We write $X \leq_* Y$ for $X$ being smaller than $Y$ in the star order, and $X \leq_{cx} Y$ and $X \leq_{disp} Y$ for $X$ and $Y$ ordered in terms of the convex, respectively dispersive order.

For a random variable $X$ and signals $S_1$ and $S_2$, by the law of iterated expectations $E\left[E\left[X|S_1\right]\right] = E\left[E\left[X|S_1\right]\right] = \mu$. Consequently, in our informational setting we always compare random variables with finite and equal means. In this case the *convex order* is equivalent to

[41] Persico (2000) refers to this concept as *accuracy*. Effectiveness applies to monotone decision problems and requires less restrictive conditions than *sufficiency* (Blackwell, 1951) to compare signals in terms of their informativeness.

[42] For a formal definition of these concepts, see Shaked and Shanthikumar (2007) or Ganuza and Penalva (2010).

[43] For further insights on the relation to other informativeness criteria, like *sufficiency* (Blackwell, 1951) or *accuracy*, respectively *effectiveness* (Lehmann, 1988; Persico, 2000) we refer the reader to the discussion in Ganuza and Penalva (2010).

the concept of second order stochastic dominance. Moreover, the dispersive order and the star order are both stronger than the convex order, that is

$$X \leq_{disp} Y \quad \Rightarrow \quad X \leq_{cx} Y \quad \text{and}$$
$$X \leq_{*} Y \quad \Rightarrow \quad X \leq_{cx} Y.$$

For the star order and the dispersive order the following relation holds:

$$X \leq_{*} Y \quad \Leftrightarrow \quad \log X \leq_{disp} \log Y. \tag{2.5}$$

Thus, the star order and the dispersive order are in general not nested. However, under some conditions they are.

**Lemma 2.7.** *For nonnegative random variables, $X$ and $Y$ with distribution functions $F$ and $G$, respectively,*

*(i) if $X \leq_{FOSD} Y$, then $X \leq_{*} Y$ implies $X \leq_{disp} Y$.*

*(ii) if $F$ and $G$ are absolutely continuous with $F(0) = G(0) = 0$ and $f(0) \geq g(0) > 0$, then $X \leq_{*} Y$ implies $X \leq_{disp} Y$.*

Figure 2.4 summarizes the relation of the three precision criteria and the sets of signals that are ordered in terms of any of these criteria.

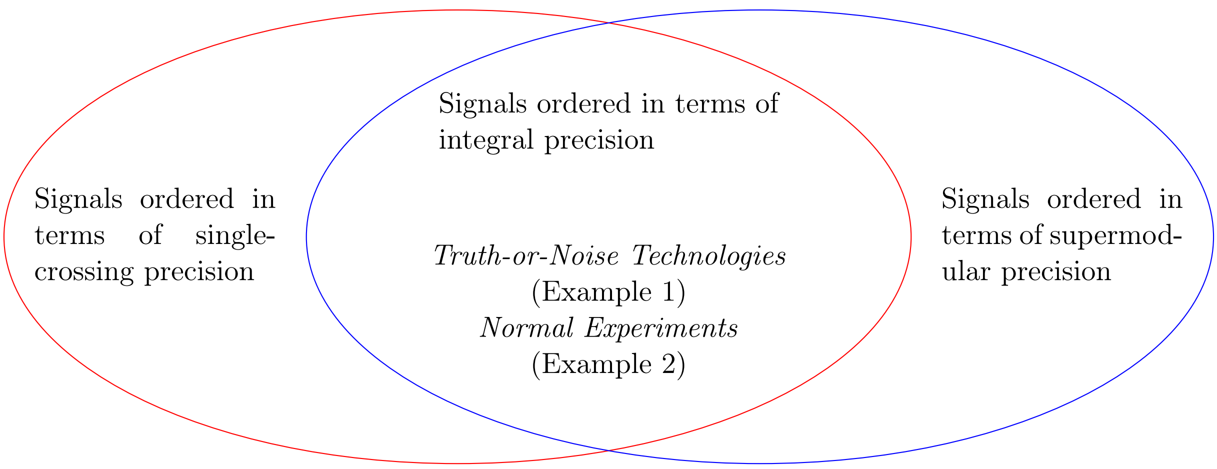



Figure 2.4: Illustration of the relation between the concepts of single-crossing precision, supermodular precision and integral precision.

# Chapter 3

# Mechanism Design with Endogenous Information

In mechanism design problems with endogenous information, regularity properties of the distribution of posterior estimates (types) are essential for tractability. Important properties are a monotone hazard rate, increasing virtual valuations or costs. Difficulties arise since these properties are not preserved under mixtures, and regularity of the prior distribution may not translate to the distribution of posterior types. We identify sufficient conditions on the primitives of an information structure, which guarantee that the distribution of posterior types has a monotone hazard rate, increasing virtual valuations or costs. These characterization results make it possible to study mechanism design problems with endogenous information, without imposing regularity conditions on the interim stage or restricting attention to specific information structures. Applications to information acquisition and disclosure in optimal auctions, and to allocation problems without money are discussed.

## 1 Introduction

Consider a setting with endogenous information, in which the distribution of posterior types of agents emerges from the information acquisition or disclosure choices of the agents. In the process of Bayesian updating, mixtures over distributions are formed, an operation under which the increasing hazard rate property is not generally preserved. That is, for a prior distribution $F$ with support $\mathcal{X} \subseteq \mathbb{R}$ and a family of distributions $\{G(\cdot|x)\}_{x\in\mathcal{X}}$ with support $S \subseteq \mathbb{R}$, even if all of these distributions have an increasing hazard rate this is not generally the case for the mixture distribution

$$G(s) = \int_{\mathcal{X}} G(s|x)\,\mathrm{d}F(x).$$

Consequently, even if the prior distribution of types has an increasing hazard rate, the distribution of posterior types induced by the endogenous choices of agents may not have this property.

In mechanism design settings, in which agents' information is endogenous, conditions that guarantee regularity of the distribution of posterior types are essential for tractability.[1] Without

[1]By contrast, in settings with exogenous information the role of regularity conditions is to simplify the

this assumption a circular effect could arise: small changes in the information level of agents could result in significant changes of the structure of the optimal mechanism, which would change the incentives to acquire or disclose information. This effect would render the model fragile, complicate the analysis tremendously, and make the model untractable. Under what conditions can we guarantee that all feasible choices of agents lead to regularity of the distribution of posterior types?

The main objective of this essay is to identify sufficient conditions on the primitives of an information structure that guarantee that the distribution of posterior estimates has an increasing hazard rate, increasing virtual valuations or costs.[2] This characterization result is important for the emerging literature on mechanism design with endogenous information of agents. This literature dispenses with the common assumption that the distribution of types, and the private information held by agents, is exogenously given. It includes an information stage into the analysis, in which information is either acquired by market participants or disclosed to them.[3]

In our analysis, we focus on the standard setting for mechanism design problems, in which agents are risk-neutral and have quasi-linear preferences. In such a framework, all payoff-relevant information of agents that is necessary to characterize the optimal mechanism in the second stage is captured in the *posterior estimates (types)* of the agents. It is therefore not necessary to know the full posterior distribution, but it suffices to know its mean.

Our first result is an "impossibility result". We identify a class of signal structures for which the resulting distribution of posterior types will always have a decreasing hazard rate, irrespective of the prior distribution of types.[4] If a signal structure from this class is contained in the set of feasible signal structures that agents can choose from, it is impossible to guarantee that the induced distribution of posterior types has an increasing hazard rate for all feasible choices of agents.

Our second result is a "possibility result". We identify sufficient conditions on the signal structure that guarantee that certain regularity properties of the prior distribution – an increasing hazard rate, increasing virtual valuations or costs – translate to the distribution of posterior estimates.[5]

---

analysis and avoid technicalities, specifically ironing-out procedures.

[2] It is a well-known problem in the economic literature that certain properties are not generally preserved under aggregation or mixtures. A prominent example is the single-crossing property introduced by Milgrom and Shannon (1994), which is not preserved under aggregation. Quah and Strulovici (2012) provide sufficient conditions that guarantee that the single crossing condition is preserved under aggregation. We provide a similar result: sufficient conditions for the increasing hazard rate property to be preserved under mixtures.

[3]Examples include Bergemann and Välimäki (2002) and Shi (2012) who study information acquisition, whereas the focus in Bergemann and Pesendorfer (2007), Esö and Szentes (2007), Ganuza and Penalva (2010), Li and Shi (2015) and Ganuza and Penalva (2014) is on information disclosure. Bergemann and Välimäki (2007) provide a good survey of the topic.

[4]This is the case for signal structures that are characterized by a family of conditional distributions that all have a decreasing hazard rate.

[5]Formally, signals must be characterized by a family of survival functions that is log-concave. This property

Straightforward applications of our results are the auction design problems analyzed in Shi (2012) and Ganuza and Penalva (2014). Shi (2012) studies optimal auctions with information acquisition by the bidders, whereas the focus of Ganuza and Penalva (2014) is on information disclosure in optimal auctions. The authors of these papers choose different approaches to circumvent the tractability problems that arise in their models. Shi (2012) imposes the regularity assumption directly on the distribution of posterior estimates, assuming that it has increasing virtual valuations. Ganuza and Penalva (2014) restrict attention to a specific information structure to make their model tractable. The results presented in this essay make it possible to identify classes of information structures to which the results in Shi (2012) and Ganuza and Penalva (2014) apply. These applications are discussed in Section 4.

To further illustrate how our results can be applied, we discuss information disclosure in allocation problems without monetary transfers.[6] We find that, by choosing an appropriate information technology, the designer can guarantee that the optimal mechanism is a full screening mechanism. This result is robust in the sense that the designer does not need to know the prior distribution of agents' types.

The rest of the essay is organized as follows. In Section 2, we introduce the formal model of the informational environment. Section 3 contains our theoretical results, with the main results presented in Section 3.3. Applications are discussed in Section 4. We conclude with some further discussion and remarks in Section 5. All proofs are relegated to the appendix.

# 2 The Informational Setting

Consider the following model. There exists an unknown state, represented by a real-valued random variable $X$. The common, initial beliefs about the state are captured by an absolutely continuous prior distribution $F$ with interval support $\mathcal{X} \subseteq \mathbb{R}$. We assume that $X$ has finite expectation, $\mu := E(X) < \infty$, under $F$.

A signal is characterized by a real-valued random variable $S$ with typical realizations $s \in [\underline{s}, \overline{s}] \subseteq \mathbb{R}$, and a family of conditional distributions $\{G(\cdot|x)\}_{x \in \mathcal{X}}$, where

$$G(s|x) := Pr(S \leq s|X = x)$$

is the probability to observe a signal $s' \leq s$ if the state is $x$.[7] We assume that for every $x \in \mathcal{X}$, $G(\cdot|x)$, is absolutely continuous in $s$, that is, it admits a density function, and $g(s|x) > 0$ almost everywhere.[8] Together with the prior distribution $F$, a signal induces a joint distribution on

is a generalization of the increasing hazard rate property to multivariate distributions.

[6]For the case of exogenous private information of agents this problem has been studied for example in Condorelli (2012) and Chakravarty and Kaplan (2013).

[7]We allow for the supports of $X$ and $S$ to be the real line.

[8]This assumption implies that there is some noise in the signal. That is, upon observing a signal realization, agents cannot exclude any states. The set of states to which an agent attaches a positive probability is the same

$(X, S)$, a so-called *information structure*. We denote the marginal distribution of $S$ by $G$.

Agents update their beliefs according to Bayes' rule. The posterior distribution of $X$ conditional on observing $s$ is $G(\cdot|s)$, and the resulting conditional expectation is

$$\widehat{X}(s) = E\left[X|S = s\right] = \int_{\mathcal{X}} x \,\mathrm{d}G(x|s). \tag{3.1}$$

We call $\widehat{X}(s)$ the *posterior estimate*. Without loss of generality we can assume that $\widehat{X}$ is increasing in $s$, which implies that an inverse function $\widehat{X}^{-1}$ exists.[9] For a given prior distribution $F$, every signal $S$ results in a distribution of posterior estimates, represented by a random variable $\widehat{X} = E\left[X|S\right]$ with distribution function

$$H(\hat{x}) := G\left(\widehat{X}^{-1}(\hat{x})\right) = \int_{\mathcal{X}} G(\widehat{X}^{-1}(\hat{x})|x)\,\mathrm{d}F(x),$$

and quantile function $H^{-1}(p) = \inf\{\hat{x}|H(\hat{x}) \geq p\}$ for $p \in [0, 1]$.

We assume that signals are monotone, that is, that high signal realizations are *more favorable* than low signal realizations in the sense of Milgrom (1981). This condition implies that it is more likely to observe a high signal realization $s$ if the underlying state $x$ is high, than if it is low.

**Assumption 3.1** (Monotone Signals)**:** For all signal realizations $s, s' \in S$ with $s' > s$, signal realization $s'$ is *more favorable* than $s$. That is, for every non-degenerate prior distribution $F$ on $X$, if $s' > s$, then the posterior distribution $G(\cdot|s')$ dominates $G(\cdot|s)$ in terms of first-order stochastic dominance, $G(\cdot|s') \geq_{FOSD} G(\cdot|s)$.

If signal $S$ is characterized by conditional densities $\{g(\cdot|x)\}_{x\in\mathcal{X}}$, then Assumption 3.1 is equivalent to the monotone likelihood ratio property.[10]

## Examples

The model captures many information technologies, among them the following examples that are frequently used in the literature.

**Example 3.1** (Normal Experiments)**:** Suppose that the states are normally distributed $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$, and signal $S$ is given by $S = X + \varepsilon$, where $\varepsilon$ is a normally distributed noise term, $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. In this case, signals are also normally distributed, $S \sim \mathcal{N}(\mu_X, \sigma_X^2 + \sigma_\varepsilon^2)$, and the posterior estimate after observing signal realization $s$ is

$$\widehat{X}(s) = \frac{\sigma_\varepsilon^2}{\sigma_X^2 + \sigma_\varepsilon^2}\mu + \frac{\sigma_X^2}{\sigma_X^2 + \sigma_\varepsilon^2}s.$$

for all signal realizations. This assumption is sometimes called the “non-shifting support” assumption in the literature.

[9] For a formal justification see Shaked et al. (2012).

[10] Signal $S$ has the (strict) *monotone likelihood ratio property* (MLRP), if for every $x > x'$, $\frac{g(s|x)}{g(s|x')}$ is strictly increasing in $s$.

The posterior estimates are linear in $S$ and normally distributed. △

**Example 3.2** (Truth-or-noise technology)**:** For a given state space $X$, let the prior belief be represented by $F$, a continuously differentiable distribution with finite mean $\mu$. A *truth-or-noise technology* provides with some probability $\alpha \in [0,1]$ a perfectly informative signal $s = x$, and with probability $(1-\alpha)$ pure noise, independently drawn from prior distribution $F$. The receiver cannot distinguish which kind of signal he observes. For signal realization $s$, the posterior estimate is $\widehat{X}(s) = \alpha s + (1-\alpha)\mu$. △

**Example 3.3:** Suppose $X \sim U[0,1]$. If the state is $x$, the resulting signal realizations $s$ are normally distributed with mean $x$ and variance 1, that is, $G(\cdot|x) \sim \mathcal{N}(x,1)$. The joint density which characterizes this information structure is

$$f(x,s) = g(s|x)f(x) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-\frac{(s-x)^2}{2}} & \text{if } 0 \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Upon observing signal realization $s$, the resulting posterior estimate is:

$$\widehat{X}(s) = s + \phi(0) \cdot (1-2s) = s \cdot (1 - 2\phi(0)) + \phi(0),$$

where $\phi(s) := \sqrt{\frac{2}{\pi}} \cdot \frac{\left[e^{-\frac{s^2}{s}} - e^{-\frac{(s-1)^2}{2}}\right]}{\operatorname{erf}\left(\frac{s}{\sqrt{2}}\right) - \operatorname{erf}\left(\frac{s-1}{\sqrt{2}}\right)}$, and erf is the error function.[11] Note that, as in the previous examples, the posterior estimate is linear in the signal realizations. △

# 3 Sufficient Conditions

In this section, we first discuss the implications of properties of information structures for the distribution of posterior estimates. Then, sufficient conditions on the primitives of information structures for the distribution of posterior estimates to have a monotone hazard rate, increasing virtual valuations or costs are identified.

**Definition 3.1:** The random variable $X$ with distribution $F$ and density $f$ has an *increasing hazard rate*, if the *hazard rate function*

$$\lambda(x) = \frac{f(x)}{1 - F(x)}$$

is increasing in $x$.
The random variable $X$ has a *decreasing hazard rate*, if $\lambda(x)$ is decreasing in $x$.

An equivalent condition to $X$ having an increasing (decreasing) hazard rate is that the *survival function* $\overline{F}(x) = 1 - F(x)$ is log-concave (log-convex).[12]

[11] $\operatorname{erf}(s) = \frac{2}{\sqrt{\pi}} \int_0^s e^{-t^2}\, \mathrm{d}t$

[12] The natural definition of an increasing hazard rate for random variables without densities is, to say that $X$ has an increasing hazard rate if the survival function is log-concave.

**Remark 3.1.** Interpreting the state $x$ as time, the hazard rate $\lambda(x) = \frac{f(x)}{1-F(x)}$ has a natural interpretation as the *failure rate* of a component: It represents the probability of an instantaneous failure of a component conditional on the component still being intact at time $x$.

To establish our results we proceed in two steps. First, we identify sufficient conditions on the prior and signal distribution for the marginal distribution of signals to have an increasing hazard rate, respectively log-concave density (Lemma 3.1). We then show that these properties transfer to the distribution of posterior estimates (Proposition 3.1 and Proposition 3.2).

## 3.1 Induced Properties of the Marginal Distribution of Signals

The marginal distribution of signals is the mixture distribution over the conditional distributions characterizing the signal, with the prior being the mixing distribution

$$G(s) = \int_{\mathcal{X}} G(s|x)\,\mathrm{d}F(x).$$

It is a well-known result in statistics that the decreasing hazard rate property is preserved under mixtures.[13] For the increasing hazard rate property – the more important property for economics – the result is less clear-cut since the class of increasing hazard rate distribution is not closed under mixtures.

To develop some intuition about why the increasing hazard rate property is not necessarily preserved under mixtures, think about the hazard rate function as representing the failure rate of a component (cf. Remark 3.1). A basic insight is that for mixtures of distributions early failures are likely to arise from distributions with high hazard rates. As Finkelstein and Cha (2013) put it "the weakest items are dying out first". More precisely, for a given prior distribution, consider the hazard rate of a mixture of a family of increasing (respectively decreasing) hazard rate distributions. For the mixture, early failures are more likely to arise from distributions with high hazard rates (at that time) whereas late failures are more likely to originate from low hazard rate distributions. This effect amplifies the features of decreasing hazard rate distributions but may offset the increasing hazard rate properties of the distributions when they are mixed. Consequently the increasing hazard rate property is not necessarily preserved under mixtures. Figure 3.1 illustrates an example of two distributions with increasing hazard rate, whose mixture does not have this property.[14]

The following lemma identifies a set of sufficient conditions for the primitives of an information structure that guarantee that the marginal distribution of signals has an increasing hazard rate.[15]

[13]See Barlow and Proschan (1981). This result is also formally stated in the appendix (Lemma 3.2).

[14]For further examples, see Finkelstein and Cha (2013) and Gurland and Sethuraman (1994).

[15]This set is the least restrictive set of sufficient conditions we are aware of. The lemma is based on a theorem by Lynch (1999). For sufficient conditions for the case of a discrete state space see Block et al. (2003).

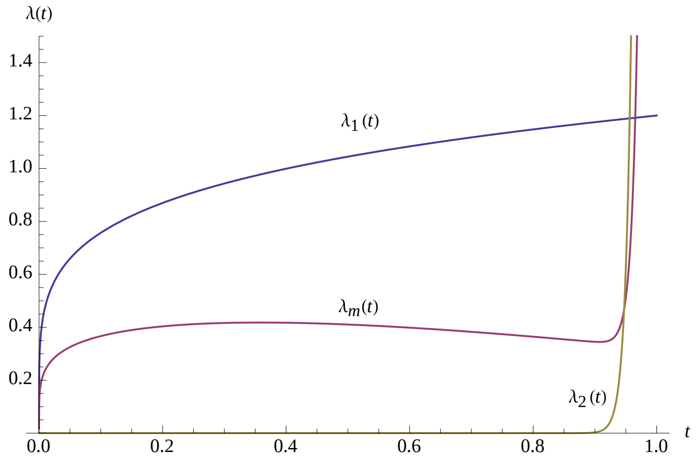



Figure 3.1: Hazard rate functions $\lambda_1, \lambda_2$ of Weibull distributions $\mathcal{W}(1.2, 1)$ and $\mathcal{W}(100, 1)$, with scale parameter 1 and shape parameters $k_1 = 1.2$ and $k_2 = 100$; and mixture hazard rate $\lambda_m$ of their equal-weight mixture.

**Lemma 3.1.** *Suppose the information structure $(X, S)$ satisfies Assumption 3.1.*
*If the family of survival functions $\{\overline{G}(\cdot|x)\}_{x\in\mathcal{X}}$ is log-concave in $(s, x)$, then if the prior distribution $F$ has an increasing hazard rate, so has the marginal distribution of signals $G$.*
*Moreover, if the family of densities $\{g(\cdot|x)\}_{x\in\mathcal{X}}$ is log-concave in $(s, x)$, then if the prior density $f$ is log-concave, so is the marginal density of the signal $g$.*

## 3.2 Link to the Distribution of Posterior Estimates

In order to obtain a general characterization result we still need to establish a relation between the marginal distribution of signals and the distribution of posterior estimates. This is possible for information structures that satisfy one of the following two assumptions, *linearity* or *smoothness*.

**Assumption 3.2** (Linearity)**:** Posterior estimates are a positive linear transformation of the signal:

$$\widehat{X} = aS + b,\ a, b \in \mathbb{R},\ a > 0.$$

Many information structures that are commonly used in the literature satisfy Assumption 3.2, among them those of Example 1 – 3.

For posterior estimates that do not satisfy this linearity condition, we impose the following smoothness condition.

**Assumption 3.3** (Smoothness)**:** The distributions $F$ and $\{G(\cdot|x)\}_{x\in\mathcal{X}}$ are twice continuously differentiable with strictly positive and bounded densities, $0 < f < \overline{f}$ and $0 < g(\cdot|x) < \overline{g}$ $\forall x \in \mathcal{X}$.

This assumption has the flavor of the smoothness assumptions in Mechanism Design that are typically imposed on the prior distribution. Assumption 3.3 requires a certain smoothness on the distributions defining the information structures. Often, information structures satisfy both Assumption 3.2 and Assumption 3.3.

We can now state our result, that for information structures satisfying at least one of these assumptions, the regularity properties of the marginal distribution of signals translate to the distribution of posterior estimates.

**Proposition 3.1.** *Suppose that the information structure $(X, S)$ satisfies Assumption 3.1, and either Assumption 3.2 or Assumption 3.3, and that the posterior estimate is a concave (convex) function of the signal. Then, if the marginal distribution of signals $G$ has an increasing (decreasing) hazard rate, so has the distribution of posterior estimates $H$.*

**Proposition 3.2.** *For information structures satisfying Assumption 3.1 and Assumption 3.2, if the marginal density of signals $g$ is log-concave, so is the density of posterior estimates, $h$.*

**Remark.** To avoid introducing new concepts and notation, the result of Proposition 3.2 is stated for log-concave densities. It should be noted, however, that the result extends to $\rho$-concave densities with $\rho$-concavity defined as in Caplin and Nalebuff (1991a,b) and Ewerhart (2013).

## 3.3 Main Results

Combining the results from Section 3.1 and Section 3.2, we can finally present our main result.

**Theorem 3.1.** *Suppose that the information structure $(X, S)$ satisfies Assumption 3.1, and either Assumption 3.2 or Assumption 3.3.*

*(i) If $\{G(\cdot|x)\}_{x\in\mathcal{X}}$ is a family of decreasing hazard rate distributions and the posterior estimate is a convex function of the signal, then for any prior $F$, the distribution of posterior estimates $H$ has a decreasing hazard rate.*

*(ii) If the family of survival functions $\{\overline{G}(\cdot|x)\}_{x\in\mathcal{X}}$ is log-concave in $(s, x)$ and the posterior estimate is a concave function of the signal, then if the prior distribution $F$ has an increasing hazard rate, so does the distribution of posterior estimates $H$.*

**Remark.** It should be noted that the conditions required to obtain the result in (ii) are significantly stronger than those in (i). Under the assumptions in the theorem, to guarantee that the distribution of posterior estimates has a decreasing hazard rate it suffices that for every $x \in \mathcal{X}$, the conditional distribution $G(\cdot|x)$ has a decreasing hazard rate. In particular, the result holds for any prior distribution. One can also think of this result as an "impossibility result": If signals are characterized by conditional distributions with a decreasing hazard rate, it is impossible that the increasing hazard rate property of the prior distribution translates to the distribution of posterior estimates.

By contrast, the possibility result for the increasing hazard rate in (ii), requires that the prior distribution has an increasing hazard rate and that the family of conditional survival distributions $\{\overline{G}(\cdot|x)\}_{x\in\mathcal{X}}$ is log-concave in $(s, x)$.[16]

[16] A function $\psi : \mathbb{R}^2 \to \mathbb{R}$ is log-concave, if its domain dom $\psi$ is convex and

$$\psi(\alpha x + (1-\alpha)y) \geq \psi(x)^\alpha \psi(y)^{1-\alpha} \quad \forall x, y \in \text{dom } \psi,\ \alpha \in (0, 1)\,.$$

We can also establish sufficient conditions for the virtual valuations and costs of the posterior estimates to be increasing.

**Definition 3.2:** A random variable $X$ with distribution $F$ and density $f$ has *increasing virtual valuations* if

$$J_v(x) = x - \frac{1 - F(x)}{f(x)},$$

is increasing in $x$.
It has *increasing virtual costs* if

$$J_c(x) = x + \frac{F(x)}{f(x)},$$

is increasing in $x$.

The following result is a direct corollary to Theorem 3.1.

**Corollary 3.1.** *Let $(X, S)$ be an information structure that satisfies Assumption 3.1, and either Assumption 3.2 or Assumption 3.3.*
*If the family of survival functions $\{\overline{G}(\cdot|x)\}_{x\in\mathcal{X}}$ is log-concave in $(s, x)$ and the posterior estimate is a concave function of the signal, then if the prior distribution $F$ has an increasing hazard rate, the distribution of posterior estimates has increasing virtual valuations.*

In order to guarantee that the distribution of posterior estimates has increasing virtual costs, slightly stronger conditions are required.

**Theorem 3.2.** *Let $(X, S)$ be an information structure that satisfies Assumption 3.1 and Assumption 3.2.*
*If the family of densities $\{g(\cdot|x)\}_{x\in\mathcal{X}}$ is log-concave in $(s, x)$, then if the prior density $f$ is log-concave, the distribution of posterior estimates has increasing virtual costs.*[17]

# 4 Applications

## 4.1 Auctions

A straightforward starting point for the discussion of applications of our results, is to connect them to the existing research on auction design with endogenous information. The results are useful in settings in which regularity of the posterior estimates does not arise from equilibrium considerations. This is for example the case in Shi (2012), who studies information acquisition in optimal auctions, as well as in Ganuza and Penalva (2014) who analyze information disclosure in optimal auctions.[18] In both settings, the implemented mechanism affects agents' incentives

[17]The statement of the theorem can be strengthened. The conditions stated in the theorem imply that the generalized virtual valuation and cost functions $J_v(x) = x - \gamma\frac{1-F(x)}{f(x)}$ and $J_c(x) = x + \gamma\frac{F(x)}{f(x)}$ are increasing in $x$ for every $\gamma > 0$.

[18] Bergemann and Pesendorfer (2007) study a setting in which the designer can choose what information to provide to bidders as well as the selling mechanism. In this setting the non-decreasing virtual valuations property follows from equilibrium considerations: If providing information would result in non-increasing virtual valuations, the seller would rather not differentiate between buyers, but move the "ironing out" procedure to

to acquire or disclose costly information and these informational effects have to be taken into account when designing the optimal mechanism. In both papers, the set of information technologies from which agents can choose is restricted, such that all information technologies in the feasible set can be compared in terms of their informational content.

**Information Acquisition in Optimal Auctions**

Shi (2012) characterizes an optimal, that is revenue maximizing, selling mechanism in a setting in which buyers do not know their private valuations ex-ante, but can acquire costly information prior to participating in the mechanism. The timing is as follows:

1. The seller announces a mechanism (and suggests an information acquisition profile).
2. Bidders acquire costly information: they choose the precision level of the signal that they will obtain about their valuation of the object.
3. Based on their chosen precision levels of information, bidders obtain a (noisy) signal about their valuation and update their beliefs accordingly.
4. Bidders submit their bids, and the object is sold according to the mechanism previously announced by the seller.

In this environment, when choosing the optimal mechanism, the seller has to take into account that his choice of a mechanism will affect the incentives of bidders' to acquire information. For the symmetric case, in which all bidders acquire the same level of information, Shi (2012) shows that, if the distribution of posterior estimates is regular and the number of bidders is sufficiently large, the optimal mechanism is a standard auction with a reserve price. The optimal reserve price in the case with endogenous information acquisition is closer to the prior mean than the standard reserve price if the equilibrium information level were exogenously given.

The results of Section 3 allow us to identify a class of information structures to which the results of Shi (2012) apply. It is hence not necessary anymore to impose the regularity conditions on the interim stage. Instead it is possible to identify sufficient conditions on the primitives of the information structures such that the results of Shi apply.[19] This class includes all information structures satisfying the conditions in Theorem 3.1 (*ii*) or Corollary 3.1. Examples include truth-or-noise technologies with increasing hazard rate prior distributions; and information structures with $S = X + \epsilon$, an increasing hazard rate prior distribution and noise with log-concave density.[20]

the information stage by not providing information to the agents. The result relies on the richness of the set of feasible information technologies that the designer may choose from. The information technologies in this set are not ordered in terms of informativeness and no statement about the optimal precision-level of disclosed information can be made.

[19]The result in Shi (2012) is based on further assumptions on the distribution of posterior estimates. A sufficient condition for these assumptions to hold is that, when switching from one signal to a more precise (and thus more costly) signal, the resulting distributions of posterior estimates are ordered in terms of the dispersive order. This is the case for information structures satisfying Assumption 3.2 (cf. Proposition 4 in Ganuza and Penalva, 2010).

[20]For classes of functions which have an increasing hazard rate or log-concave densities see Bagnoli and Bergstrom (2005).

**Information Disclosure in Optimal Auctions**

Ganuza and Penalva (2014) study information disclosure in optimal auctions. In their setting, the seller chooses the selling mechanism, as well as the precision level of the information disclosed to bidders before the auction. Prior to the auction, each bidder observes a private, partially informative signal about his valuation for the object and update his beliefs accordingly. The informational content of the signal is determined by the precision level chosen by the seller, which is known to all bidders. More precise information allows bidders to better assess their preferences or tastes for the object, which leads to more heterogeneity among bidders.

In their analysis, Ganuza and Penalva (2014) assume that signals have the structure of a truth-or-noise technology (cf. Example 3.2). The authors state in their conclusion that "the model is standard (and general) in all dimensions but the choice of the set of available signals". This simplifying assumption that signals have the structure of a truth-or-noise-technology has the following convenient implications:

1. For truth-or-noise technologies, the regularity properties, like increasing virtual valuations of the prior distribution translate to the distribution of posterior estimates.[21]
2. Linearity of the information structure keeps the model tractable.
3. Information technologies are naturally ordered in terms of their informational content (precision) and it is straightforward to define a cost-function which captures the idea that information disclosure is costly.

If we allow for a larger set of information structures that satisfy Assumption 3.1 and Assumption 3.2 (linearity), the last two of the aspects mentioned above (2. and 3.) are preserved.[22] However, for these more general information structures the marginal distribution of signals is usually not the same as the prior distribution. Consequently, the increasing virtual valuation property will in general not translate from the prior distribution to the distribution of posterior estimates. In this case, our results of Section 3 can be applied to characterize sufficient conditions on the primitives of information structures for the distributions of posterior estimates to satisfy the increasing virtual valuations property.

The results in Ganuza and Penalva (2014) generalize to the class of information structures satisfying Assumption 3.1, Assumption 3.2 and the conditions in Theorem 3.1 (*ii*) or Corollary 3.1.[23] The main insights are:

1. In an optimal auction, the auctioneer discloses more information than in a standard auction in which the object is always sold. Here, an optimal auctions is a standard auction

[21]For truth-or-noise technologies, the marginal distribution of signals is the same as the prior distribution. Moreover, due to the linearity of the posterior estimates in signals, the regularity properties translate from the marginal distribution of signals to the distribution of posterior estimates.

[22]The class of information structures that satisfy Assumption 3.2, are naturally ordered in terms of *supermodular precision* (cf. Ganuza and Penalva, 2010, Proposition 4).

[23]It is straightforward to replicate the proofs in Ganuza and Penalva (2014) for these more general information structures, using our results in Section 3 and the linearity of posterior estimates (Assumption 3.2). We refer the reader to the discussion in Ganuza and Penalva (2014).

with a reserve price that is optimal given the precision level of the information disclosed to bidders.

2. The level of information disclosed to bidders in an optimal auction is weakly increasing in the number of bidders.

The intuition behind these results is the following. In a standard auction (without reserve price), if a seller discloses information, he has to leave informational rents to the bidders. If information is costly, the auctioneer will therefore not reveal all information. An optimal reserve price reduces the informational rents of bidders and thus increases the seller's incentives to disclose information.

## 4.2 Optimal Mechanisms without Money

Another interesting application of our results are allocation problems without monetary transfers as studied for example in Condorelli (2012) and Chakravarty and Kaplan (2013). These model consider the allocation of $k$ indivisible heterogeneous objects to $n$ agents when monetary transfers or charging personalized prices is infeasible or undesirable, but a benevolent designer can screen the agents.[24] Screening yields non-monetary costs which are wasted, that is, screening generates a deadweigh loss. More specifically, if the seller chooses to screen agents, he implements a mechanism that requires the agents to invest in some costly non-productive action (e.g. exerting effort, spend time in waiting lines, etc.) in order to signal their private types. Incentive compatibility requires that the non-monetary costs incurred by the agents correspond to Vickrey-payments. That is, the expected (wasteful) costs of an agent capture the externalities that he imposes on the other agents.

In this setting, Condorelli (2012) characterizes the optimal mechanism within the class of incentive compatible direct allocation mechanisms, that is, the mechanisms that maximizes ex-ante welfare. Condorelli shows that, if buyers' valuations have a decreasing hazard rate, a *full screening* mechanism is optimal whereas in any other case, due to the trade-off between a more efficient allocation and screening costs, only partial or no screening is optimal.

The results of Section 3 can be applied to extend the model studied in Condorelli (2012) to a setting in which agents do not know their private valuations or tastes ex-ante and the seller can provide information through a noisy channel, for example by advertising a concert or sport event. We assume that the designer has to provide some information to make market participants aware of the availability of his products, but cannot provide perfectly informative private signals. For example, an event manager has to advertise a concert to attract interested customers but cannot perfectly control how interested parties perceive the information provided to them through the advertisement. Formally, this means that we exclude perfectly informative signals and pure noise from the set of feasible information technologies available to the designer.

[24]Typical examples mentioned in the literature are the allocation of donor organs, or ticket sales for concerts or sport events. Waiting lines can serve as costly screening instruments.

The setting studied in Condorelli (2012) is linear, hence the distribution of posterior estimates captures all relevant information to determine the optimal mechanism. Applying our insights from Section 3 yields the following result.

**Proposition 3.3.** *Consider information disclosure in an allocation problem without monetary transfers. Suppose that the signal is characterized by conditional distributions with decreasing hazard rate. Then for any prior distribution of agents' types, it is optimal for the seller to implement a full screening mechanism.*

We want to emphasize the following remarkable robustness feature of this result: As long as the designer can implement a signal characterized by decreasing hazard rate distributions, he knows that a full screening mechanism is optimal, irrespective of the prior type distribution of agents. To implement the optimal mechanism the designer therefore does not need to know the prior distribution.[25]

# 5 Discussion and Concluding Remarks

In this note we discussed properties of information structures and their implications for the distribution of posterior estimates. We specifically focused on identifying conditions such that the induced distribution of posterior estimates satisfies certain regularity properties that are commonly used in the mechanism design literature.

An important insight of the discussion is, that the increasing hazard rate property may not be preserved under mixtures of distribution functions, an operation which occurs during the updating process. For certain signal structures it is impossible that the distribution of posterior estimates has an increasing hazard rate. However, we identified sufficient conditions on the signal structure that guarantee that the increasing hazard rate property translates from the prior distribution to the distribution of posterior estimates.[26]

The theoretical results were used to identify classes of information environments to which the results on information acquisitions and disclosure in optimal auctions of Shi (2012) and Ganuza and Penalva (2014) apply. Moreover, we discussed information disclosure in allocation problems without money as for example studied in Condorelli (2012) or Chakravarty and Kaplan (2013). It was shown that whenever signals are characterized by a family of decreasing hazard rate distributions, a full screening mechanisms is optimal.

[25]The result of Proposition 3.3 extends to models of two-sided matching markets as analyzed in Hoppe et al. (2009) and Roesler (2014), respectively Chapter 2 of this thesis. In these settings, if the designer implements an information technology characterized by decreasing hazard rate distributions, the welfare optimal mechanism is to screen agents, and to implement the positive assortative matching.

[26]As a corollary we obtain sufficient conditions for the distribution of posterior estimates to have increasing virtual valuations. However, these conditions are not tight and could probably be relaxed, using the insight from Ewerhart (2013) that $\left(-\frac{1}{2}\right)$-concavity is a tight sufficient condition for increasing virtual valuations, which is a weaker condition than an increasing hazard rate. The mathematical methods that we use to obtain our results do not extend to the case that would be needed to pursue this question systematically. We therefore second the statement of Hardy et al. (1952) that "the complications introduced by zero or negative values [are] hardly worth pursuing systematically".

We think that our results will be valuable beyond the applications discussed in this note, specifically for research on mechanism design problems with endogenous information. For such problems, the insights of Section 3 can be used to restrict attention to a set of feasible information structures for which the optimal mechanism is of a particular form. This allows to keep mechanism design problems with endogenous information tractable, an important first step to address new questions and develop new insights on this topic.

# Appendix

*Proof of Lemma 3.1.* Suppose the information structure $(X, S)$ satisfies the slightly stronger condition that $f$ is log-concave. In this case, $\overline{G}(x, s) := \overline{G}(s|x)f(x)$ is log-concave, since the product of two log-concave functions is log-concave. The survival function of the marginal distribution $G$ is given by:

$$\overline{G}(s) = \int_X \overline{G}(s|x)f(x)\,\mathrm{d}x.$$

By Prékopa's Theorem (1973), log-concavity is preserved by integration and it follows that $\overline{G}$ is log-concave. Consequently, $G$ has an increasing hazard rate (cf. Definition 3.1).

The same line of reasoning can be used to prove that log-concavity of $f$ and $g(\cdot|x)$ implies log-concavity of $g$.

For the proof of the general case, which only requires that $F$ has an increasing hazard rate, observe that by Assumption 3.1, for every $s > s'$, $G(\cdot|s) \geq_{FOSD} G(\cdot|s')$. This implies, $G(\cdot|x) \geq_{FOSD} G(\cdot|x')$ for all $x > x'$, which is equivalent to $\overline{G}(s|x) = 1 - G(s|x)$ being increasing in $x$ for every $s \in S$. The result then follows by Theorem 2.1 in Lynch (1999). □

*Proof of Proposition 3.1.*

**Case 1:** Suppose the information structure $(X, S)$ satisfies Assumption 3.1 and Assumption 3.2. That is, suppose $\widehat{X} = aS + b$ with $a > 0$. This is equivalent to $S = \frac{\widehat{X}-b}{a}$. Given monotonicity of signals, for every $\hat{x} \in \hat{\mathcal{X}}$, $\widehat{X}(s) \leq \hat{x} \Leftrightarrow s \leq \frac{\hat{x}-b}{a}$. This implies $H(\hat{x}) = G\left(\frac{\hat{x}-b}{a}\right)$.

Let $\eta(\hat{x}) := \frac{\hat{x}-b}{a}$. For $a > 0$, $\eta(\hat{x})$ is increasing in $\hat{x}$. Moreover,

$$h(\hat{x}) = \frac{\mathrm{d}G}{\mathrm{d}\eta}\frac{\mathrm{d}\eta}{\mathrm{d}\hat{x}} = \frac{1}{a} \cdot g\left(\eta(\hat{x})\right). \tag{3.2}$$

It follows that

$$\lambda_H(\hat{x}) = \frac{1}{a} \cdot \frac{g\left(\eta(\hat{x})\right)}{1 - G\left(\eta(\hat{x})\right)}.$$

Given that $a > 0$ and $\eta(\hat{x})$ is increasing in $\hat{x}$, it follows that, if $G$ has an increasing (decreasing) hazard rate then $\lambda_H(\hat{x})$ is increasing (decreasing) in $\hat{x}$ which implies that $H$ (resp. $\widehat{X}$) has an increasing (decreasing) hazard rate.

**Case 2:** Suppose the information structures $(X, S)$ satisfies Assumption 3.1 and Assumption 3.3. These conditions imply that $\widehat{X}(s)$ is continuously differentiable and strictly increasing in $s$. By the inverse function theorem, $\widehat{X}$ is invertible. That is, there exists a twice continuously differentiable function $\eta := \widehat{X}^{-1}$, and the first and second derivative of $\eta$ are given by

$$\eta'(\hat{x}) = \frac{1}{\widehat{X}'\left(\eta(\hat{x})\right)} \quad \text{and} \quad \eta''(\hat{x}) = -\frac{\widehat{X}''(\eta(\hat{x}))}{\left(\widehat{X}'(\eta(\hat{x}))\right)^3}.$$

By Assumption 1, $\eta(\hat{x})$ is strictly increasing in $\hat{x}$. This implies that $\eta'(\hat{x}) > 0$ for every $\hat{x}$. Moreover, if $\widehat{X}$ is concave, then $\eta''(\hat{x}) > 0$, that is, $\eta$ is convex. Similarly, if $\widehat{X}$ is convex, then $\eta''(\hat{x}) < 0$ and $\eta$ is concave.

For every $\hat{x}$, $\eta(\hat{x})$ determines the signal realization $s$ that results in the conditional expectation $\hat{x}$. With these specifications, $S = \eta(\widehat{X})$ and $H(\hat{x}) = G(\eta(\hat{x}))$. Since $G$ and $\eta$ are both continuously differentiable so is $H$, and it follows that

$$h(\hat{x}) = \frac{\mathrm{d}G}{\mathrm{d}\eta}\frac{\mathrm{d}\eta}{\mathrm{d}\hat{x}} = g(\eta(\hat{x})) \cdot \eta'(\hat{x}).$$

It follows that the hazard rate function of $\widehat{X}$ is given by

$$\lambda_H(\hat{x}) = \eta'(\hat{x})\frac{g(\eta(\hat{x}))}{1 - G(\eta(\hat{x}))}.$$

Its derivative is

$$\lambda'_H(\hat{x}) = \eta''(\hat{x})\psi(\hat{x}) + \eta'(\hat{x}) \cdot \psi'(\hat{x}),$$

with $\psi(\hat{x}) := \frac{g(\eta(\hat{x}))}{1-G(\eta(\hat{x}))}$.

If $G$ has an increasing (decreasing) hazard rate, then $\psi(\hat{x}) = \frac{g(\eta(\hat{x}))}{1-G(\eta(\hat{x}))}$ is increasing (decreasing). Given that $\eta'(\hat{x}) > 0$ for every $\hat{x}$, it follows that the second summand of $\lambda'_H(\hat{x})$ is increasing (decreasing) in $\hat{x}$. Moreover, if $\widehat{X}$ is concave (convex) then $\eta''(\hat{x}) > 0$ ($\eta''(\hat{x}) < 0$), which implies that the first summand is increasing (decreasing).

It follows that the distribution of posterior estimates $H$ has an increasing hazard rate, if the posterior estimate is a concave function of the signal and the distribution of signals has an increasing hazard rate. Similarly, the distribution of posterior estimates has a decreasing hazard rate, if the distribution of signals has a decreasing hazard rate and the posterior estimate is a convex function of the signal. □

*Proof of Proposition 3.2.* The proof follows along the same lines as *Case 1* of the proof of Proposition 3.1. The result for log-concavity is a direct consequence of (3.2) and the fact that log-concavity is preserved under affine transformations. □

*Proof of Theorem 3.1.*
(i) As mentioned in Section 3.1, it is a well-known results that the decreasing hazard rate property is preserved under mixtures. Formally

**Lemma 3.2** (Barlow and Proschan 1981)**.** *Consider a family of distributions $\{G(\cdot|x)\}_{x\in\mathcal{X}}$ with support $\mathcal{S}$ that all have a decreasing hazard rate. Then, for any mixing distribution $F$, the mixture distribution*

$$G(s) = \int_{\mathcal{X}} G(s|x)\ dF(x)$$

*has a decreasing hazard rate.*

It is straightforward to apply this result to information structures: If for every $x \in \mathcal{X}$, the distribution $G(\cdot|x)$ has a decreasing hazard rate, then by Lemma 3.2 the marginal distribution of signals $G$ has a decreasing hazard rate. By Proposition 3.1 it follows that the distribution of posterior estimates, $H$, has a decreasing hazard rate.

(ii) The result follows directly by combining Lemma 3.1 and Proposition 3.1. Given the assumptions, Lemma 3.1 implies that the marginal distribution of signals $G$ has an increasing hazard rate. By Proposition 3.1 the distribution of posterior estimates $H$ has the same property. □

*Proof of Corollary 3.1.* The result follows directly from Theorem 3.1 and the fact, that an increasing hazard rate implies increasing virtual valuations. □

*Proof of Theorem 3.2.* Under the assumptions of the theorem, Lemma 3.1 implies that the marginal density function of signals $g$ is log-concave. Applying Proposition 3.1 yields that the distribution of posterior estimates $h$ is log-concave. It is shown in Ewerhart (2013) that log-concavity is equivalent to $\rho-$concavity for $\rho = -\frac{1}{2}$, and that this is a sufficient condition for the virtual cost function $J_c(\hat{x})$ to be increasing in $\hat{x}$. □

*Proof of Proposition 3.3.* Suppose the seller implements a signal that is characterized by conditional distributions with decreasing hazard rate. That is, suppose that for every $x \in \mathcal{X}$, the distribution $G(\cdot|x)$ has a decreasing hazard rate. Then, by Theorem 3.1 (i), the distribution of posterior estimates $H$ has a decreasing hazard rate. By Theorem 1 and Corollary 3 in Condorelli (2012) it follows that the optimal mechanism is a full screening mechanism. □

Chapter 4

# Preference Uncertainty and Conflict of Interest in Committees

We propose and analyze a committee model, in which agents with interdependent values vote on whether to accept an alternative or to stick to the status quo. Agents hold two-dimensional private information: About a dimension of the payoff state, and about their individual preference type, which reflects an agent's level of partisanship.

In equilibrium, committee members adopt cutoff-strategies, and an agent's preference type is reflected in his acceptance standard. We identify how the composition of a committee affects its decisions: As the population of committee members becomes more partisan, agents' adjust their acceptance standards more, and equilibrium cutoffs move away from the sincere voting threshold. By contrast, an agent who finds himself in a committee with more heterogeneity of preference types, is more uncertain about the preferences of others, and hence bases his vote more on his own, privately observed signal.

## 1 Introduction

In most organizations, it is the prevailing practice that complex decisions are made by committees, not by individuals. Corporate boards decide how to invest, whom to hire and whether or not to adopt a new technology. Further examples include the allocation of research grants, the approval of new drugs by the FDA, and academic hiring. Typically, decisions are reached by a voting procedure.

By the complexity of matters that are put to vote, committee members often cannot assess all information about the alternatives. Rather, they pay attention to the aspect of the matter that is most important for them, but are aware that the signals held by the other committee members also contain relevant information. The aforementioned examples are such situations, in which committee members typically have interdependent, but not purely common preferences.[1] Agents differ in how they aggregate available information about the state into preferences.

[1]This is in contrast to the common assumption in the voting literature that agents share a common interest.

Two dimensions about which committee members usually have private information are the following: First, they possess private information about the state, which is payoff relevant to all agents. Second, agents also have private information about their *preference type*, which determines how they aggregate available information about the state into preferences. The private preference type of an agent reflects his level of partisanship, that is, the extent to which he favors his own private signal over the average signal held by other agents. This means that, even if all private signals about the state were publicly revealed, it would still remain private information to the agents how they aggregate these signals into preferences – there is *preference uncertainty*.

In this essay, we introduce a committee model that captures the idea that agents have interdependent preferences and individual private preference types. Our goal is to understand how private preference types, preference uncertainty, and the composition of the committee affect the equilibrium acceptance standards and acceptance sets of committee decisions. We focus on two questions: How does the level of partisanship of the population affect acceptance standards? And what are the effects of more or less heterogeneous distribution of preference types for committee decisions?

The analysis is motivated by the observation that committees differ significantly in their composition. One dimension is the level of partisanship of committee members. The distribution of partisanship levels depends on aspects such as the cultural background of committee members, and the culture within an organization. Agents from individualistic cultures are more likely to show a high level of partisanship compared to their colleagues with a more collectivist, socially-minded (cultural) background.[2] Hence, partisanship levels of members naturally vary across committees.

A second dimension of the composition of committees is the heterogeneity of preference types of its members. Irrespective of whether committee members display a high or a low average level of partisanship, the partisanship levels of committee members may be similar, or show a high level of variation.[3] For example, if committee members stem from a group with diverse cultural backgrounds, one would expect this population to display a high level of preference heterogeneity. Committees or parliaments from small states like Belgium or the Netherlands, are less likely to display strong preference heterogeneity than parliaments of larger states like the US or Canada, or multinational assemblies such as the EU, or the UN general assembly.

However, even in committees whose members stem from a population with a high level of preference heterogeneity, committee members may learn about the other agents' preference types, for example in standing committees. The uncertainty about the preference types of fellow committee members decreases, which corresponds to a decrease in the level of preference heterogeneity in the population. Hence, such committees resemble those with members from a

[2]Hofstede (1991) identifies individualism vs. collectivism as one dimension along which cultural differences can be analyzed; Triandis (2001) links this cultural dimension to differences in personality and behavior.

[3]The special case, in which all committee members have the same level of partisanship is studied in Yildirim (2012) and Moldovanu and Shi (2013).

relatively homogeneous population.

Our results provide insights how the composition of committees regarding the distribution and the heterogeneity of preference types of its members affects the acceptance standards in collective decisions.

We study a model, in which a group of agents faces a binary decision: whether to accept a proposal or to stay with the status quo. The decision is made by generalized majority voting. Each agent indicates whether he wants to accept the proposal. The majority rule determines the minimal number of affirmative votes that are required to adopt the proposal.

Prior to voting, each agent obtains a private signal about the state (quality of the proposal). The payoff of the status quo is assumed to be zero for all agents. For each agent, the payoff from adopting the alternative is determined by an additively separable function of the private signals of all agents, which satisfies a single-crossing condition. Hence, committee members have interdependent values and each agent puts the most weight on his own signal. The extent to which an agent favors his own signal – his partisanship level – is represented by his *preference type*. It is assumed to be private information, capturing the idea that the partisanship level of an individual is intrinsic in nature, and can be regarded as part of his personality. Consequently, there is conflict of interest among committee members but uncertainty about the extent of the conflict. Moreover, agents hold two-dimensional private information.

We begin our analysis by establishing the existence of a Nash equilibrium in undominated strategies, and by characterizing some fundamental properties. In equilibrium, agents adopt cutoff strategies,[4] that is, an agent accepts an alternative whenever his private signal is above a certain threshold.

We find that an agent's private preference type is reflected in the cutoff that he adopts: Strongly partisan agents base their votes mostly on their own observed signal. More socially-oriented types take into account the information about the other agents' signals that they can derive from the event of being pivotal. They adjust their acceptance thresholds accordingly. For example, under unanimity voting, partisan agents adopt higher acceptance standards than the ones adopted by their more socially-oriented colleagues.

Next, we address the questions how private preference types and the composition of a committee affects its decisions. For the comparative statics analysis, we consider the iid case. We establish the existence of a unique symmetric voting equilibrium, on which we focus throughout the analysis.

As a first step, we provide a result that simplifies the analysis: In the present model, identifying the symmetric equilibrium corresponds to finding a fixed point of a mapping between function spaces. We show that it is possible to re-express this problem as a one-dimensional one: There exists a one-dimensional mapping and a one-to-one correspondence between fixed

[4]This is typical for voting models with continuous signals, see e.g. Feddersen and Pesendorfer (1997), Duggan and Martinelli (2001), and Li and Suen (2009).

points of this mapping and symmetric equilibria of the original problem. This reformulation of the problem simplifies the comparative statics analysis. The problem reduces to identifying the direction in which the fixed-point of the one-dimensional mapping moves.

We show that agents, who find themselves among more partisan committee members, adjust their acceptance standards more than if they were in a committee from a more socially-oriented population. For the unanimity rule, this means that acceptance standards decrease in the partisanship level of the population of the committee members. The driving effect for this result is the following: For a fixed profile of strategies, shifting to a more partisan population implies that acceptance standards are on average higher, and hence the event of being pivotal is more informative about the signals of others. Consequently, agents adjust their acceptance standards more, which counteracts the initial effect.

Furthermore, we find that in committees with much heterogeneity of preference types, agents adjust their acceptance standards less than in committees with more homogeneous preference types. Here, more heterogeneity of preference types leads to more uncertainty about why an agent votes affirmatively. He may have observed a high signal, or simply have a socially-oriented preference type. It follows, that the event of being pivotal is less informative about the other agents' signals, and agents adjust their acceptance standards less. For the unanimity rule, one may suspect that more preference uncertainty would lead an agent to adopt a more lenient acceptance standard. It may be somewhat surprising that the opposite occurs: more preference uncertainty causes an agent to focus more on his own private signal, and increase his acceptance standard.

The rest of the essay is structured as follows. The model is introduced in Section 2. In Section 3, we discuss related literature and special features of the present model. An illustrative example is presented in Section 4. Section 5 and Section 6 contain the main results. In Section 5, equilibrium existence is established and fundamental properties of equilibrium strategies are discussed. The comparative static results regarding the partisanship level and the heterogeneity of preference types in the population of committee members are presented in Section 6. Section 7 concludes. All proofs are relegated to the appendix.

# 2 The Model

Consider a committee of $n$ agents, $\mathcal{I} = \{1, \ldots, n\}$, who take a binary decision: whether to accept a proposal (P) or to stay with the status quo (Q). A proposal is characterized by an $n$-dimensional vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. We refer to $x_1, \ldots, x_n$ as the *attribute values* of $x$. Each agent has an individual, private *preference type*, $\theta_i \in \Theta_i \subseteq [0, 1]$, which determines how the agent aggregates the $n$-dimensional payoff state $x$ into preferences.

**Payoffs.**

For agent $i$ with preference type $\theta_i$, the payoff of proposal $x$ is:[5]

$$v_i(\theta_i, x) = \theta_i x_i + (1-\theta_i)\,\frac{1}{n-1}\sum_{j\neq i} x_j. \tag{4.1}$$

The payoff of the status quo is 0 for all members.

The parametric form (4.1) captures that agents have interdependent preferences. If $\theta_i = 1$, agent $i$ has private values, whereas the pure common values case corresponds to $\theta_i = \frac{1}{n}$ for all $i = 1, \ldots, n$. We take the private and the pure common values case as the two boundary benchmarks, and hence focus on $\Theta_i \subseteq \left[\frac{1}{n}, 1\right]$.

**Information Environment.** Attribute values $x_i$ are determined by independent random draws with distributions $F_i$ from interval $\mathcal{X}_i = [\underline{x}, \overline{x}] \subset \mathbb{R}$. The set of proposals, $\mathcal{X} = \times_{i=1}^{n} \mathcal{X}_i$, is a closed convex set in $\mathbb{R}^n$. Let $\mathcal{X}_{-i} := \times_{j\neq i} \mathcal{X}_j$ and $x_{-i} := (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$. The distribution functions $F_i$ are are twice continuously differentiable. The joint distribution of proposals is $F = \prod_{i=1}^{n} F_i$. The realization $x_i$ is private information to agent $i$; the distributions $F_i$ of the respective random variables $X_i$ are common knowledge.

Agents' preference types are independently distributed on $\Theta_i$ with distributions $G_i$ and densities $g_i > 0$. Preference type $\theta_i$ is private information to agent $i$, the distribution of types is common knowledge. Given these model specifications, agents hold two-dimensional private information: $(x_i, \theta_i)$ is private information to agent $i$.

It is common knowledge that the payoff structure has the form of (4.1). We assume that $E(X_i) = 0$ for every $i \in \mathcal{I}$. This implies that, before agents observe their private signal $x_i$ about the proposal, they neither favor the status quo nor the proposal.

**Decision Rule.** The committee decision is made by generalized majority voting. The majority rule, which is characterized by an integer $k \in \{1, \ldots, n\}$, is publicly announced. Agents indicate whether they want to accept or reject the proposal. The proposal is adopted if and only if there are at least $k$ affirmative votes.

**Strategies.** A pure strategy for agent $i$ is a measurable function:

$$\begin{aligned} \sigma_i:\ \Theta_i \times \mathcal{X}_i &\to \{0,\, 1\} \\ (\theta_i,\, x_i) &\mapsto \sigma_i(\theta_i,\, x_i), \end{aligned}$$

where $\sigma_i(\theta_i,\, x_i) = 1$ represents the case in which agent $i$ votes affirmatively (i.e. in favor of the

[5] The specific parametric form of (4.1) is not crucial for the results. The equilibrium characterization results of Section 5 extend to additively separable utility functions that are continuously increasing in $x_i$ for all $i \in \mathcal{I}$, and that satisfy the following single-crossing property with respect to agents' private signals $x_i$:

**Assumption 4.1** (SC)**:** For all $i, j \in \mathcal{I}$, $j \neq i$:

$$\frac{\partial v_i}{\partial x_i}(\theta_i, x) \geq \frac{\partial v_j}{\partial x_i}(\theta_j, x) \quad \forall x \in \mathcal{X}.$$

For the parametric form of (4.1) this is equivalent to $\theta_i \geq \frac{1}{n-1}(1-\theta_j)$, for all $(\theta_i, \theta_j) \in \Theta_i \times \Theta_j$, $j \neq i$. It follows that $\Theta_i \subseteq \left[\frac{1}{n}, 1\right]$ for every $i \in \mathcal{I}$.

proposal) if his type is $\theta_i$ and his private signal is $x_i$. It is sometimes convenient to consider the strategy of a given type, $\theta_i$, of agent $i$. With a slight abuse of notation we denote the strategy of type $\theta_i$ by $\sigma_{\theta_i}$, where $\sigma_{\theta_i}(x_i) := \sigma_i(\theta_i, x_i)$.

**Remark 4.1.** One could allow for mixed strategies here. However, it is easy to show that agents adopt pure strategies in equilibrium. Hence, we directly restrict attention to pure strategies, in order to avoid unnecessary technicalities.

In the binary decision problem that we consider, a pure strategy for an agent characterizes for each of his preference types $\theta_i$ a corresponding *acceptance set* $A_i^+(\theta_i) := \sigma_{\theta_i}^{-1}(1) \subseteq \mathcal{X}_i$. This is the set of private signals $x_i \in \mathcal{X}_i$ that will induce the agent to vote affirmatively. A strategy of agent $i$ thus corresponds to a set of acceptance sets $\{A_i^+(\theta_i)\}_{\theta_i \in \Theta_i}$.

Let $\overline{\mathcal{X}}_i = \mathcal{X}_i \cup \{\tilde{x}\}$ be the space that is obtained by adjoining a point $\tilde{x}$ to $\mathcal{X}_i$ at the upper boundary of $\mathcal{X}_i$.[6] A strategy $\sigma_i$ of agent $i$ is a *cutoff-strategy*, if for every $\theta_i$ there exists some $\chi_i(\theta_i) \in \overline{\mathcal{X}}_i$ such that type $\theta_i$ votes affirmatively, if and only if he observes a signal $x_i \geq \chi_i(\theta_i)$. Here $\chi_i(\theta_i) = \tilde{x}$ represents the case in which agent $i$ always rejects the proposal, irrespective of his private signal. If agents adopt cutoff strategies, the acceptance sets $A_i^+(\theta_i)$ are intervals of the form $[\chi_i(\theta_i), \overline{x}]$.[7]

**Equilibrium Concept.** We employ the concept of undominated Nash equilibrium, that is, we restrict attention to Nash equilibria in which no agent plays a weakly dominated strategy.[8] This is standard in the voting literature and, as in Feddersen and Pesendorfer (1997), we refer to it as a *voting equilibrium.*

# 3 Related Literature and Discussion of the Model

**Related Literature** To the best of our knowledge, the pool of literature addressing the question how the composition of a committee (i.e. the distribution of preference types of its members) affects collective decisions is very small. Two recent, related papers are Yildirim (2012) and Moldovanu and Shi (2013). They analyze voting models in which committee members have interdependent, additively separable utilities. The preference structure adopted in both papers is hence similar to the one in this essay. However, in their models all committee members possess the same preference type, which can be interpreted as a parameter that determines the conflict of interest among committee members. Our model departs in two aspects: First, agents have individual preference types. Second, these individual preference types are private information to the agents, that is, there is preference uncertainty. The new model, that we introduce – in which agents hold two-dimensional private information – makes it possible

[6] Formally, consider $\tilde{x}$ as a duplicate of $\overline{x}$ and $\overline{\mathcal{X}}_i = \mathcal{X}_i \cup \{\tilde{x}\}$ as the space equipped with the following topology: Let the set of open sets $\mathcal{O}$ consist of all subsets $O \subseteq \overline{\mathcal{X}}_i$ such that for each $x \in O$ there is an interval $I_x \in \{(a,b), (a,\overline{x}]\}$, with $x \in I_x \subseteq O$. And let $int\,[a,\tilde{x}] := (a,\overline{x}]$, where $int$ denotes the interior.

[7] with $[\tilde{x}, \overline{x}] := \emptyset$.

[8] This eliminates trivial equilibria where all agents play extreme strategies, i.e., always accept respectively, always reject the project.

to study new question, such as: How do individual preference types and the composition of a committee affect collective decisions?

Given the differences of the models, the analyses in Yildirim (2012) and Moldovanu and Shi (2013) focus on different questions. Moldovanu and Shi (2013) consider a search model, in which the decision to stop is made by a committee by unanimity voting. They characterize a stationary equilibrium in cutoff-strategies. In this equilibrium, acceptance standards increase, and welfare decreases in the level of conflict among committee members.[9] In a static model, Yildirim (2012) identifies time-consistent majority rules, that is, majority rules that a designer can implement, if he cannot commit to a rule prior to observing the votes.

Further interdependent values models that study collective decision making are Grüner and Kiel (2004) and Rosar (2015). They consider a different (quadratic-losses) functional form of utilities, and continuous collective decisions, whereas we focus on a binary decision problem. Grüner and Kiel (2004) show that, with an unrestricted report space, from a utilitarian perspective, the average mechanism performs better in the common values case, whereas for private values the median mechanism is preferable. By contrast, Rosar (2015) finds that with an optimally designed report space, for uniformly distributed information or large electorates, the average mechanism performs better for any degree of interdependence.

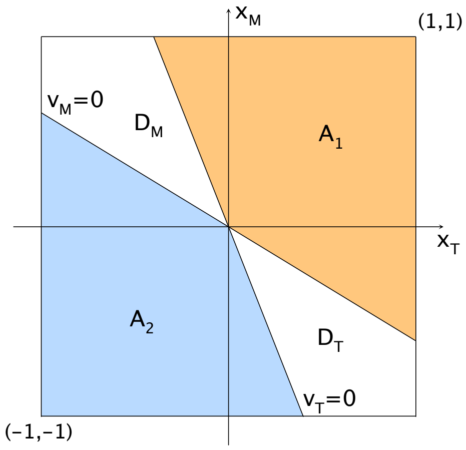
More broadly, this essay contributes to the voting literature. This strand of literature goes back to an early observation made by Condorcet (1785) that, by pooling the information of their members, groups may take better decisions than individuals. This statement, known as the *Condorcet Jury Theorem*, was initially formulated for non-strategic voters and thus a purely statistical result. Starting from this insight, there is an extensive literature on collective decision making, now typically focusing on strategic voters who update their beliefs about the information held by other agents conditional on the event of being pivotal. Li and Suen (2009) provide a good survey.

Most of the theoretical voting models study settings in which individuals share a common interest. That is, committee members would agree on the best outcome if they knew the state. Often an even stronger assumption is made, namely, that agents have perfectly aligned preferences. This assumption implies that there is an underlying consensus: agents would agree on the best action if there were no asymmetric information, that is, if all private information were publicly available.

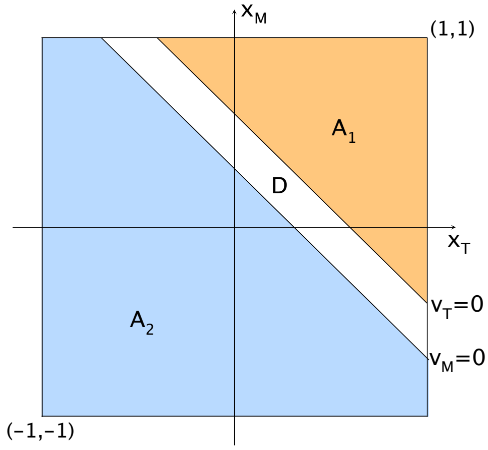
There are different ways to introduce heterogeneity among agents' preferences in the voting model. One option is to introduce a private values component in agents' preferences. The model in this essay, and those of Moldovanu and Shi (2013), Yildirim (2012), Grüner and Kiel (2004), and Rosar (2015) fall into this category. The analyzed interdependent values environments incorporate both the private values case, as well as the common values case.

Li et al. (2001) choose a different approach to introduce heterogeneity among voters. They relax the assumption that agents' preferences are perfectly aligned, but still assume that agents

[9] Meyer and Strulovici (2015) extend some of the results of Moldovanu and Shi (2013) to more general preference structures.




(a) Present model, $(\theta_A, \theta_B) = \left(\frac{1}{2}, \frac{1}{3}\right)$.




(b) Model in Li et al. (2001).

Figure 4.1: Agreement and disagreement sets in the present model, and in Li et al. (2001), if $x$ is common knownledge.
Agreement sets: $A_1$ – agents prefer the proposal; $A_2$ – agents prefer the status quo.
Disagreement sets: $D_i$ – agent $i$ prefers the proposal, agent $j \neq i$ favors status quo; $i \in \{T,\, M\}$.

share a common objective. If there is uncertainty about the state there may be conflict of interest, but disagreement vanishes if all uncertainty is resolved. The authors discuss how the level of conflict among committee members affects their incentives to strategically misrepresent their information and thus may hinder information aggregation.

In the model by Li et al. (2001) and related papers,[10] agents are heterogeneous in the sense that they have different preferences for type-$I$ and type-$II$ errors, and hence require different levels of evidence to prefer the alternative over the status quo. This specification of heterogeneity implies that between any pair of agents there is only one direction of disagreement: If a pair of agents disagrees, it is always the same agent who supports the proposal whilst the other favors staying with the status quo.

By contrast, in the present model, the conflict of interest arises from different preferences of committee members. Even if the payoff state were known, agents may not agree on the best outcome. Moreover, the heterogeneity among agents is such that the direction of conflict depends on the payoff state: If agents disagree, then it is not always the same agent who favors the proposal. For any two agents, $T$ and $M$, with different preference types,[11] there are payoff states in which agent $T$ wants to accept the proposal, while agent $M$ prefers the status quo, and vice versa. The differences in modeling conflict of interest and heterogeneity among agents are illustrated in Figure 4.1.

[10] e.g. Austen-Smith and Feddersen (2006) and Li and Suen (2009).
[11] In the example discussed in Section 4, they represent a technology and a marketing expert.

# 4 Illustrative Example

In order to fix ideas, let us discuss an illustrative, stylized example. Consider a technology company that has to decide whether to accept a proposal for a product update or to keep producing the current version. For simplicity, assume that this decision is taken by a two-member committee, constituted of a technology expert (T), and a marketing expert (M). Acceptance of the proposal requires unanimity.

The attribute values of the proposal represent the quality of its technical features $x_T$ and its design features $x_M$, where the technology expert can asses the former, and the marketing expert can assess the latter. Both experts think that for the success of the product update, the quality of the proposal in their own dimension is more important, but they acknowledge that both aspects are relevant. For agent $i \in \{T, M\}$ with preference type $\theta_i \in \left[\frac{1}{2}, 1\right]$, the payoff of implementing alternative $x$ is

$$u_i(x, \theta_i) = \theta_i x_i + (1 - \theta_i)\, x_j.$$

Formally, attribute values are uniformly distributed, $X_i \overset{iid}{\sim} U\,[-1, 1]$, and preference types are independently distributed on $\left[\frac{1}{2}, 1\right]$ with distributions $G_i$, $i \in \{T, M\}$.

Suppose the marketing expert adopts the cutoff-function $\chi_M \in \mathcal{X}_M{}^{\Theta_M}$, that is, if his type is $\theta_M$, then he votes affirmatively if and only if the design quality of the proposal is at least $\chi_M(\theta_M)$. The vote of the technology expert only matters, if he is pivotal, that is, if the marketing expert votes affirmatively. Taking this into account, conditional on being pivotal, the expected payoff of the proposal for the technology expert with private information $(\theta_T, x_T)$ is:

$$\begin{aligned} V_T(x_T, \theta_T) =& \theta_T x_T + (1 - \theta_T)\mathbb{E}_{\Theta_M}\left[\mathbb{E}_{\mathcal{X}_M}\left[x_M | x_M \geq \chi_M(\theta_M)\right]\right] \\ =& \theta_T x_T + (1 - \theta_T)\int_{1/2}^{1} \frac{\chi_M(\theta_M) + 1}{2}\,\mathrm{d}G_M(\theta_M). \end{aligned}$$

For the technology expert it is optimal to vote for the proposal if and only if his expected payoff from the proposal is higher than that of the status quo, $V_T(x_T, \theta_T) \geq 0$. It is easy to see, that the expert's best response is to adopt a cutoff-strategy. Analogous arguments can be used to determine the best-response function of the marketing expert. Equilibrium cutoff-functions are determined by the equations:

$$\chi_i^*(\theta_i) = -\frac{1 - \theta_i}{\theta_i}\mathbb{E}_{\Theta_j}\left[\mathbb{E}_{\mathcal{X}_j}\left[x_j | x_j \geq \chi_j^*(\theta_j)\right]\right] \quad i \neq j \in \{T, M\}. \tag{4.2}$$

Notice, that the case in which agents' preference types are common knowledge corresponds to $G_i$ being degenerate distributions. In this case, the equilibrium would be described by the cutoffs $x_T^*$, $x_M^*$ that satisfy (4.2) for the realizations of $\theta_T$, $\theta_M$.

Under unanimity voting, being pivotal is good news. It means that the other committee member has received a sufficiently high signal to vote affirmatively. A private value type $\theta_i = 1$, does not attach any value to the signal of the other expert and adopts the sincere voting threshold 0. The lower an agent's preference type, the more weight he attaches to the signal of the other agent. Consequently, equilibrium cutoff-functions $\chi^*(\cdot)$ are non-positive, and increasing in $\theta_i$. More partisan agents adopt higher acceptance standards.[12]

# 5 Equilibrium Characterization

In this section, we establish equilibrium existence and characterize fundamental properties of the equilibrium strategies.

In a voting game, a rational agent conditions his decision on the event of being pivotal. Hence, when choosing the optimal action, each committee members takes into account the information that he can extract from the event of being pivotal. For any majority rule $k \in \{1, \dots, n\}$, this is the event in which exactly $k-1$ of the other agents vote affirmatively, and hence the agent's vote determines the outcome.

Every strategy profile $\sigma_{-i}$, corresponds to a set of acceptance sets, $\left\{A^+_{-i}(\theta_{-i})\right\}$. For any majority rule $k$ and type profile $\theta_{-i}$, the *pivotal set*,

$$A_i^{piv}(\theta_{-i}) := \{x_{-i} : |\{j \in \mathcal{I}\backslash\{i\} : \pi_j(x_{-i}) \in A^+(\theta_j)\}| = k-1\},$$

is the set of signal profiles $x_{-i}$ for which agent $i$ is pivotal. Here, $\pi_j$ denotes the $j^{\text{th}}$ projection map, which maps vector $x_{-i}$ to coordinate $x_j$.

Conditional on being pivotal, the expectation of agent $i$ about the signals of the other agents is:

$$\begin{aligned}\mathbb{E}_{\sigma_{-i}}\left[\overline{x}_{-i} \mid piv\right] =& \mathbb{E}_{\Theta_{-i},\mathcal{X}_{-i}}\left[\overline{x}_{-i} \mid x_{-i} \in A_i^{piv}(\theta_{-i})\right] \\ =& \frac{1}{\mathbb{P}(\sigma_{-i})} \int_{\Theta_{-i}} \int_{\mathcal{X}_{-i}} \overline{x}_{-i} \cdot \mathbb{1}_{A_i^{piv}(\theta_{-i})} \, \mathrm{d}F_{-i}(x_{-i}) \, \mathrm{d}G_{-i}(\theta_{-i}),\end{aligned} \tag{4.3}$$

where

$$\mathbb{P}(\sigma_{-i}) := \int_{\Theta_{-i}} \int_{\mathcal{X}_{-i}} \mathbb{1}_{A_i^{piv}(\theta_{-i})} \, \mathrm{d}F_{-i}(x_{-i}) \, \mathrm{d}G_{-i}(\theta_{-i}), \tag{4.4}$$

is the probability that agent $i$ is pivotal, and

$$\overline{x}_{-i} := \frac{1}{n-1} \sum_{j \neq i} x_j, \tag{4.5}$$

[12] Equilibrium cutoffs are:

$$\chi_i^*(\theta_i) = -\frac{1-\theta_i}{\theta_i} \cdot \frac{1}{1+\ln 4} \quad \text{for } i \in \{T, M\},$$

if preference types are private information; and

$$x_i^* = -\frac{(3\theta_j - 1)(1-\theta_i)}{3\theta_i\theta_j + \theta_i + \theta_j - 1} \quad \text{for } i \in \{T, M\},$$

if preference types $\theta_T$, $\theta_M$ are common knowledge.

is the average signal of the other agents.

It follows that, given strategy profile $\sigma_{-i}$, for agent $i$ with private information $(\theta_i,\, x_i)$, the expected payoff from implementing the alternative, conditional on being pivotal is:

$$V_i\left((\theta_i,\, x_i);\sigma_{-i}\right) = \theta_i x_i + (1-\theta_i)\cdot \mathbb{E}_{\sigma_{-i}}\left[\overline{x}_{-i}\,|\, piv\right]. \tag{4.6}$$

We establish equilibrium existence by using the Tychonoff-Schauder fixed-point theorem. In equilibrium, agents adopt cutoff-strategies.

**Theorem 4.1** (Equilibrium Existence)**.**
*In a committee of $n$ members with interdependent values and private preference types, for any generalized majority rule, $k \in \{1,\ldots,n\}$, there exists a voting equilibrium.*
*In every voting equilibrium agents adopt cutoff strategies, given by:*

$$\chi_i^*(\theta_i) = \begin{cases} \underline{x} & \textit{if } V_i\left((\theta_i,\, x_i);\sigma^*_{-i}\right) \geq 0\;\forall\, x_i \in [\underline{x},\overline{x}] \\ \tilde{x} & \textit{if } V_i\left((\theta_i,\, x_i);\sigma^*_{-i}\right) < 0\;\forall\, x_i \in [\underline{x},\overline{x}] \\ -\frac{1-\theta_i}{\theta_i}\mathbb{E}_{\sigma^*_{-i}}\left[\overline{x}_{-i}\,|\, piv\right] & \textit{otherwise.} \end{cases} \quad \forall\, i \in \mathcal{I}. \tag{4.7}$$

If attribute values and preference types are identically distributed, it makes sense to focus on symmetric equilibria, in which all committee members adopt the same strategies. In the iid case, there exists a unique symmetric voting equilibrium.

**Proposition 4.1** (Symmetric Equilibrium – Uniqueness)**.**
*If attributes values and preference types are iid ($F_i = F_j$, and $G_i = G_j$ for all $i, j \in \mathcal{I}$), then there exists a unique symmetric voting equilibrium.*

The rest of the section is devoted to characterizing and understanding some fundamental properties of equilibrium strategies.

**Lemma 4.1.** *In any voting equilibrium, for all $i \in \mathcal{I}$,*

*(i) agents with private value type $\theta_i = 1$ vote sincerely, $\chi_i^*(1) = 0$.*
*All preference types of an agent adjust their cutoffs in the same direction; that is, agent $i$'s cutoffs $\chi_i^*(\theta_i)$ have the same sign for all types $\theta_i \in \Theta_i\backslash\{1\}$.*

*(ii) the cutoff functions $\chi_i^*$ are continuous in the agent's preference type $\theta_i$, and twice continuously differentiable a.e..*

*(iii) the more partisan an agent is, the less he adjusts his acceptance standard:*
*$|\chi_i^*(\theta_i)|$ is non-increasing in $\theta_i$.*

If an agent has private values ($\theta_i = 1$), the signals of the other agents do not influence his preferences, and the information that he derives from the event of being pivotal does not affect his decision. Hence, a private values type will vote *sincerely*, that is, solely based on his own private signal. He adopts the cutoff 0.

Every other agent with interdependent values, takes into account the expected value that he attaches to the information held by other committee members. In his decision of choosing an optimal cutoff, each agent updates his beliefs conditional on the event of being pivotal. In particular, being pivotal is either good news (if $E_{\sigma_{-i}}\left[\overline{x}_{-i}|piv\right] > 0$) or bad news ($E_{\sigma_{-i}}\left[\overline{x}_{-i}|piv\right] < 0$). In the first case, agent $i$ attaches a positive expectation to the information held by other agents. He will therefore require weaker evidence himself to accept an alternative, adjust his own acceptance standard and adopt a negative cutoff. The expected information derived from the event of being pivotal is the same for all types of agent $i$. Consequently, all types $\theta_i \neq 1$ will adjust their acceptance standards in the same direction. In particular, under unanimity voting, all cutoffs are non-positive.

**Corollary 4.1** (Unanimity Voting). *For the unanimity rule $k = n$, equilibrium cutoffs are non-positive:*

$$\chi_i^*(\theta_i) \in [\underline{x},\, 0] \quad \forall \theta_i \in \Theta_i,\ i \in \mathcal{I},$$

*and $\chi_i^*(\theta_i)$ is increasing in $\theta_i$.*

Are agents' equilibrium strategies *responsive*, that is, do agents condition their voting decision on the private signal that they observe?

**Definition 4.1:** Type $\theta_i$ of agent $i \in \mathcal{I}$ is *responsive*, if he conditions his decision whether or not to vote affirmatively on his observed signal. For cutoff-strategies this is equivalent to adopting an (interior) cutoff, $\chi_i(\theta_i) \in (\underline{x}, \overline{x}]$.
Agent $i$'s strategy is *responsive* if there exists a set of types $\Theta_i^R \subseteq \Theta_i$ with non-empty interior, such that all types $\theta_i \in \Theta_i^R$ are responsive.

Responsiveness of agents' equilibrium strategies is a necessary condition for information aggregation. The next result shows that there is always a set of responsive preference types with non-empty interior. In any voting equilibrium a positive measure of preference types (and profiles) condition their decision on the signal that they observe – some information aggregation occurs.

**Lemma 4.2.** *In any voting equilibrium, agents' cutoff strategies are responsive. There exists some type $\hat{\theta}_i$ such that all types $\theta_i \geq \hat{\theta}_i$ are responsive, whereas $\chi_i^*$ is constant on $\left[0, \hat{\theta}_i\right)$. All of these types adopt the same extreme cutoff, either $\underline{x}$ or $\tilde{x}$. That is,*

$$\chi_i^*(\theta_i) \in (\underline{x}, \overline{x}]\,, \forall\, \theta_i \in \left[\hat{\theta}_i,\, 1\right],\ \textit{and}$$
$$\chi_i^*(\theta_i) \equiv \textit{constant} \in \{\underline{x},\, \tilde{x}\}, \forall\, \theta_i \in \left[0,\, \hat{\theta}_i\right).$$

The following lemma characterizes the shape of the equilibrium cutoff-functions. If equilibrium cutoffs are non-positive (non-negative), then the corresponding cutoff-function is concave (convex) on the set of responsive types.

**Lemma 4.3.** *In any voting equilibrium, every equilibrium cutoff function with non-positive cutoffs $\chi^*(\theta_i) \leq 0$ (non-negative, $\chi^*(\theta_i) \geq 0$), is concave (convex) on the set of responsive types $\left[\hat{\theta}_i, 1\right]$. Equilibrium cutoff functions satisfy:*

$$(\chi_i^*)' \cdot (\chi_i^*)'' \leq 0 \quad \forall\, i \in \mathcal{I}. \tag{4.8}$$

The identified properties determine the shape of equilibrium strategies. Some properties depend on the quorum rule. There are four possible shapes of equilibrium cutoff-functions, which are illustrated in Figure 4.2. An explicit example for a two-member committee was discussed in Section 4.

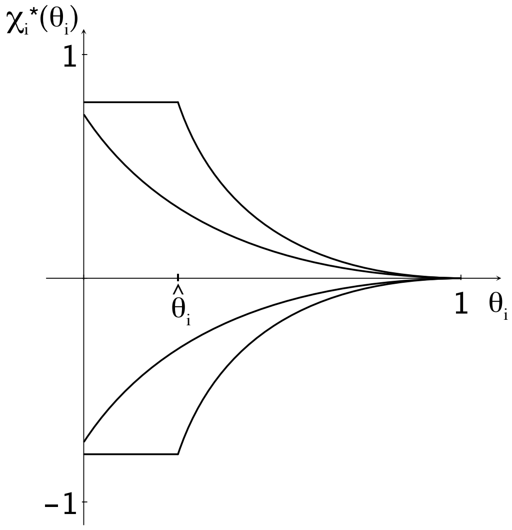



Figure 4.2: Possible shapes of equilibrium cutoff-functions.

# 6 Comparative Static Effects of Preference Uncertainty

One goal of this essay is to understand how the distribution of preference types, and preference uncertainty affect the outcomes of committee decisions. Should we expect the composition of committees to matter for the outcome of collective decisions – and if so how? The model establishes a framework that allows us to study these questions.

Specifically, we answer the following questions: Should we expect committee members who originate from more partisan populations to adopt more or less stringent acceptance standards? And how does this affect the set of alternatives that are accepted by such committees? The same questions are analyzed regarding the effects of more or less heterogeneity of the population of committee members.

For the comparative statics analysis, we consider the symmetric case with iid attribute values and preference types. We focus on symmetric equilibria, which are unique in this case (Theorem 4.1 and Proposition 4.1). For simplicity, we consider decisions that require unanimity.

Hence, equilibrium cutoffs are non-positive (Corollary 4.1). A brief discussion of the case of generalized majority rules is provided at the end of this section.

## 6.1 Extent of Partisanship in the Population

How are committee decisions affected by the extent of partisanship of its members? As the following result shows, agents adopt lower acceptance standards if they find themselves among fellow committee members from a more partisan population.

**Proposition 4.2** (More Partisan Populations)**.**
*In any symmetric voting equilibrium,*[13] *in a committee whose members stem from a more partisan population, each preference type will adopt less stringent acceptance standards than in a committee with less conflict among members' preferences:*

$$\text{If } H \geq_{FOSD} G \ \text{ then } \ \chi^*_H(\theta_i) \leq \chi^*_G(\theta_i) \quad \forall\, \theta_i \in \Theta_i,\ \forall\, i \in \mathcal{I}.$$

A first-order shift in the distribution of preference types of the committee members population represents a shift in the population towards more extreme or partisan types. From Lemma 4.1 we know that those types adjust their thresholds less, and vote more based on their own signal. Hence, for any individual agent who faces fellow committee members from such a more partisan population, being pivotal is more informative about other members' signals than if his fellow committee members are more socially oriented. It follows that any given preference type will adjust his acceptance standard more, and equilibrium cutoffs move away from 0, the sincere voting threshold. In the case of unanimity voting, this means that the event of being pivotal is better news, and equilibrium cutoffs-functions shift downwards.

Let me sketch the idea of the proof of Proposition 4.2, the formal details are relegated to the appendix. A difficulty in the analysis of the current model is that agents hold two-dimensional private information. The multi-dimensionality of the model hence requires working with function spaces and to identify fixed points therein (cf Theorem 4.1). The trick that we adopt is to re-formulate the characterizing fixed-point property for symmetric equilibria. The original problem is mapped to a related problem, which only requires to identify fixed-points of a one-dimensional mapping, and hence is more tractable. Moreover, it is shown that there is a one-to-one correspondence (bijection) between fixed points of this one-dimensional mapping, and symmetric equilibria of the original problem. This trick to re-formulate the problem is illustrated in Figure 4.3.

In the alternative problem, we determine the effects of distributional shifts on the mapping $\Phi$ in order to then derive the induced effects for the equilibrium cutoff-functions of the original problem. In the proof it is shown that the function $\Phi$ is decreasing. Moreover, a shift in the distribution of preferences types only affects the mapping $\Lambda$ whereas the mapping $\beta$ remains

[13] for decisions requiring unanimity

$$\Psi : \quad X^{\Theta} \xrightarrow{\Lambda} \mathbb{R} \xrightarrow{\beta} X^{\Theta} \qquad (\Psi = \beta \circ \Lambda)$$

$$\chi(\theta) \longmapsto c(\chi) := \mathbb{E}_{\Theta}\left[\mathbb{E}_{X}\left[\chi(\theta)|piv\right]\right] \longmapsto \chi_c(\theta) := \chi^{BR}(c, \theta)$$

(a) Best-response correspondence – symmetric case.

$$\Phi : \quad \mathbb{R}_0^+ \xrightarrow{\beta} \mathcal{X}_i^{\Theta_i} \xrightarrow{\Lambda} \mathbb{R} \qquad (\Phi = \Lambda \circ \beta)$$

$$c \longmapsto \max\left\{\underline{x}, \ -\frac{1-\theta_i}{\theta_i} \cdot c\right\} \longmapsto \mathbb{E}_{\Theta_i}\left[\mathbb{E}_{\mathcal{X}}\left[X_i | X_i \geq \chi_i(\theta_i)\right]\right].$$

(b) Related, one-dimensional mapping.

Figure 4.3: Re-formulation of the problem. There is a one-to-one correspondence between fixed points of the mappings Ψ and Φ.

unchanged. Specifically, it is shown that a first-order shift of the distribution of preference types results in a upward shift of the function Φ. It is then easy to see that the fixed-point of the new mapping is to the right of the fixed point of the original mapping Φ. This is illustrated in Figure 4.4, where the fixed points are determined by the intersection points of the graph of Φ with the 45°-line. This fixed-point captures the information that an agent derives in equilibrium from the event of being pivotal – for the unanimity rule it is positive, and higher for more partisan populations. Finally, it follows that agents adjust their acceptance standards more, resulting in a downward shift of the equilibrium cutoff-function.

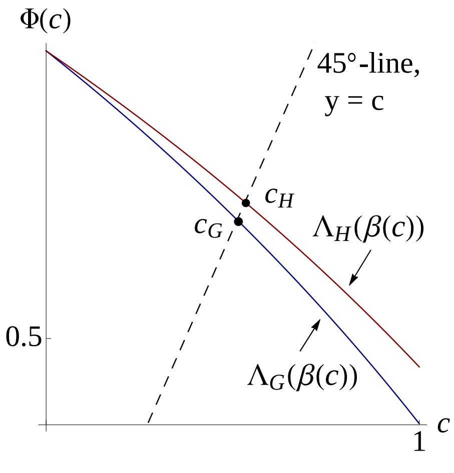



Figure 4.4: Effects of a FOSD-shift of the distribution of preference types on the function Φ and its fixed points.

## 6.2 Level of Preference Heterogeneity of the Populations

How should we expect acceptance standards and sets to change depending on whether the committee consists of members from a population with more or less preference heterogeneity? In the latter case, there is less uncertainty about other agents' preferences. As the following result shows, if agents find themselves among committee members from a more partisan population, they adjust their acceptance standards less and vote more based on their own signal.

**Proposition 4.3** (More Heterogeneous Populations).
*For the unanimity rule, in any symmetric voting equilibrium, in a committee that is constituted of members from a more heterogeneous population, each preference type will adopt more stringent acceptance standards than in a committee with members from a more homogeneous population:*

$$\textit{If } H \geq_{MPS} G \ \ \textit{then} \ \ \chi^*_H(\theta_i) \geq \chi^*_G(\theta_i) \quad \forall\, \theta_i \in \Theta_i,\ \forall\, i \in \mathcal{I}.$$

The proof of the result again uses the re-formulation of the problem discussed in Section 6.1 and is relegated to the appendix.

Let us provide some intuition for the result. An agent, who finds himself among committee members from a population with a high level of preference heterogeneity, faces a lot of uncertainty about the preference type of his fellow committee members. It is hence hard for him to predict whether a fellow committee member voted affirmatively because of a high signal, or because he has a low preference type. In the latter case, an agent may vote affirmatively even though he observes a relatively low signal (cf. Lemma 4.1). Hence, the event of being pivotal is not very informative for an agent, he only weakly adjusts his acceptance standard and votes based more on his own private signal than an agent who finds himself in a committee of members from a more homogeneous population. For the unanimity rule, this implies that members of committees constituted of individuals from a more heterogeneous population adopt higher acceptance standards than committees with members from a more homogeneous population.

The results allow to make predictions, regarding the acceptance standards and acceptance sets of the examples of committees that we mentioned in the introduction. One should expect that committees from large states or assemblies of international organizations, only accept alternatives that are of sufficiently high quality in every dimension. By contrast, the minimum requirements in each dimension for an alternative to pass, may be lower for a committee whose members stem from a more homogeneous population. It should be noticed that this observation does not imply that one or the other committee performs better in term of maximizing utilitarian welfare.

**Generalized Majority Rules**

The comparative static results in this section were established for the unanimity rule. However, the results do not seem to depend on this assumption and should extend to generalized majority rules. Let us provide some intuition about which results one could expect.

For any generalized majority rule, the number of affirmative votes required to adopt the alternative determines whether in the symmetric equilibrium, cutoff-functions are non-positive (as for the unanimity rule), or non-negative. The latter is the case if only a small number of affirmative votes are necessary in order to adopt the alternative.

As discussed in Section 6.1, a first-order shift in the distribution of preference types of the population of committee members, results in the event of being pivotal being more informative about the other agents' signals than before. Hence, each individual preference type will adjust his acceptance standard more and equilibrium cutoffs move away from 0: The equilibrium cutoff-function moves upwards if it is non-negative, and downwards if it is non-positive.

For the case of a second-order shift – meaning that committee members stem from a population with more heterogeneous preference types – the event of being pivotal becomes less informative about the other agents' preferences. As a result, each agent adjusts his acceptance threshold less, and votes more based on his private signal. For each individual preference type, equilibrium cutoffs move closer to 0: The equilibrium cutoff-function moves downwards if it is non-negative, and upwards if it is non-positive.

# 7 Conclusion

In this essay, we have proposed and analyzed a committee voting model in which agents have interdependent preferences and individual, private preference types. It was shown that, in equilibrium, agents choose optimal acceptance standards conditional on being pivotal, and an agent's preference type is reflected in his acceptance standard. We identified how acceptance standards react to changes in the distribution of preferences-types of committee members:

1. Agents adjust their acceptance standards more (compared to the sincere voting threshold), if they believe to find themselves among committee members from a more partisan population.
2. More heterogeneity of preference types – and hence more uncertainty about the preference types of the other committee members – leads an agent to act cautious, adjust his acceptance standard less, and base his decision more on his own, privately observed signal.

The prediction of the model could be tested, either in the lab, or in the field by comparing decisions across committees. Here, cultural dimensions, e.g. on the scale of individualism-collectivism, could be taken as predictors of preference types of committee members.

The analysis can be extended in several dimensions and there are various open questions and directions for future research. It is of interest to (1) allow for sequential voting procedures, (2) study more general mechanisms than generalized majority rules, (3) allow for pre-voting communication; either about the observed signal, or the preference types, or a combination of both, (4) study a dynamic situation in which a standing committee votes on a sequence of alternatives, and may accept multiple of them.

We briefly want to comment on two of these topics, on which we have some basic insights. First, it would be interesting to consider the collective decision problem for a group of agents with interdependent preferences from a mechanism design perspective (not allowing for monetary transfers).

For the case with one-dimensional private information, in which preference types are common knowledge, it is straightforward to show that a necessary condition for implementability of a social choice function is that it is a step function. Such social choice functions can be implemented by some kind of generalized voting procedure, where agents may use finer reports than simple yes/no votes (e.g. yes, maybe, rather not, no). It follows that, in this setting, the socially efficient (utilitarian) policy hence cannot be implemented. The question remains, which social choice functions can be implemented, and how well the utilitarian policy can be approximated.

For the model with two-dimensional private information studied in this essay, an interesting observation is that it falls into the class of "non-generic settings" for which the impossibility result of Jehiel and Moldovanu (2001) does not apply. This raises the question what can be done beyond implementing constant social choice functions.

Second, a straightforward question is whether information sharing through pre-vote communication is possible. Much of the voting literature consider straw polls as communication devices.[14] In the model studied in this essay, information sharing is not possible through straw polls. This is most easily seen for unanimity voting. Suppose agents observe their private signal about an attribute of the proposal. In the straw poll, agents may reveal some of their private information by indicating support for the proposal or not. However, the only interest of an agent under unanimity is to get other agents to vote for the alternative in case that his own private signal is high enough. Hence, an agent will always exaggerate his private signal in the straw poll. This still leaves him with the option to vote against the proposal in the actual vote. It would be of interest to allow for more general communication devices/procedures, or to give agents the option to communicate about their private preference types.

[14] See e.g. Coughlan (2000) and subsequent papers

# Appendix

*Proof of Theorem 4.1.* Consider some agent $i \in \mathcal{I}$. Suppose $\sigma_{-i}$ denotes the strategies adopted by all agents but $i$, and let $\{A^{+}_{-i}(\theta_{-i})\}_{\theta_{-i}\in\Theta_{-i}}$ be the corresponding set of acceptance sets. For agent $i$ with type $\theta_i$, the expected payoff of the alternative, conditional on being pivotal, and observing signal $x_i$ is:[15]

$$V_i\left((\theta_i,\, x_i);\sigma_{-i}\right) = \theta_i x_i + (1-\theta_i)\cdot \mathbb{E}_{\sigma_{-i}}\left[\overline{x}_{-i}\mid piv\right] \quad \text{with} \tag{4.9}$$
$$\mathbb{E}_{\sigma_{-i}}\left[\overline{x}_{-i}|piv\right] := \frac{1}{\mathbb{P}(\sigma_{-i})}\int_{\Theta_{-i}}\int_{\mathcal{X}_{-i}} \overline{x}_{-i}\cdot \mathbb{1}_{A^{piv}_{-i}(\theta_{-i})}\,\mathrm{d}F_{-i}(x_{-i})\,\mathrm{d}G_{-i}(\theta_{-i}).$$

It is agent $i$'s best response to vote in favor of the alternative if and only if the expected payoff of the alternative, conditional on him being pivotal, is greater than the payoff of the status quo, which is 0. Consequently, it is agent $i$'s best response to vote affirmatively if and only if $V_i\left((\theta_i,\, x_i);\sigma_{-i}\right)\geq 0$.

It is easy to see from (4.9) that $V_i\left((\theta_i,\, x_i);\sigma_{-i}\right)$ is continuous and strictly increasing in $x_i$. This readily establishes that agent $i$'s best response is to follow a cutoff-strategy. That is, for every preference type $\theta_i$, there exists some $\chi_i(\theta_i)\in\overline{\mathcal{X}}_i$ such that type $\theta_i$ votes affirmatively if and only if he observes a signal $x_i\geq\chi_i(\theta_i)$.[16] In particular, suppose all other agents adopt strategy profile $\sigma_{-i}$, then agent $i$'s best response is to adopt cutoff-function $\phi_i^{BR}$ characterized by:

$$\phi_i^{BR}(\theta_i) = \begin{cases} \underline{x} & \text{if } V_i\left((\theta_i,\, x_i);\sigma_{-i}\right)\geq 0 \;\forall x_i\in[\underline{x},\overline{x}] \\ \tilde{x} & \text{if } V_i\left((\theta_i,\, x_i);\sigma_{-i}\right)< 0 \;\forall x_i\in[\underline{x},\overline{x}] \\ -\frac{1-\theta_i}{\theta_i}\mathbb{E}_{\sigma_{-i}}\left[\overline{x}_{-i}|piv\right] & \text{otherwise.}\end{cases} \tag{4.10}$$

From now on we assume that agents adopt cutoff-strategies. By a slight abuse of notation, we denote agent $i$'s cutoff strategy by a function $\chi_i:\Theta_i\to\overline{\mathcal{X}}_i$, where $\chi_i(\theta_i)$ is the cutoff that agent $i$ adopts if his type is $\theta_i$.

Let $\overline{\mathcal{X}}_i^{\Theta_i}$ be the space of functions $f:\Theta_i\to\overline{\mathcal{X}}_i$ endowed with the product topology (here: the topology of pointwise convergence). We denote agent $i$'s best response function by

[15] Suppose agent $i$ updates his beliefs, assuming that he is pivotal and all other agents play according to $\sigma_{-i}$. Then his expected payoff if he votes affirmatively is:

$$\begin{aligned}\widetilde{V}_i\left((\theta_i,\, x_i);\sigma_{-i}\right) =& \int_{\Theta_{-i}}\int_{\mathcal{X}_{-i}}\left[\theta_i x_i + \frac{1-\theta_i}{n-1}\sum_{j\neq i} x_j\right]\cdot \mathbb{1}_{A^{piv}_{-i}(\theta_{-i})}\,\mathrm{d}F_{-i}(x_{-i})\,\mathrm{d}G_{-i}(\theta_{-i}) \\ =&\mathbb{P}(\sigma_{-i})\theta_i x_i + \frac{1-\theta_i}{n-1}\int_{\Theta_{-i}}\int_{\mathcal{X}_{-i}}\left(\sum_{j\neq i} x_j\right)\cdot \mathbb{1}_{A^{piv}_{-i}(\theta_{-i})}\,\mathrm{d}F_{-i}(x_{-i})\,\mathrm{d}G_{-i}(\theta_{-i}),\end{aligned}$$

where $\mathbb{P}(\sigma_{-i}) = \int_{\Theta_{-i}}\int_{\mathcal{X}_{-i}}\mathbb{1}_{A^{piv}_{-i}(\theta_{-i})}\,\mathrm{d}F_{-i}(x_{-i})\,\mathrm{d}G_{-i}(\theta_{-i})$ is the probability that agent $i$ is pivotal. Using that $\overline{x}_i := \frac{1}{n-1}\sum_{j\neq i}x_j$ and conditioning on the event of being pivotal yields (4.9).

[16] Reminder: $\overline{\mathcal{X}}_i := \mathcal{X}_i\cup\{\tilde{x}\}$, and cutoff $\tilde{x}$ represents the case in which the agent rejects all proposals.

$\phi_i^{BR} : \overline{\mathcal{X}}^{\Theta} \to \overline{\mathcal{X}}_i^{\Theta_i}$.[17] This is well-defined since we have shown that for any strategy profile $\sigma_{-i}$ agent $i$'s best response is a cutoff function.

The discussion shows that, for every strategy profile $\sigma_{-i}$ a unique best response for agent $i$ exists, and that best responses take the form of cutoff functions. It follows that the best response correspondence is a function, characterized by:

$$\begin{aligned}\Phi : \overline{\mathcal{X}}_1^{\Theta_1} \times \cdots \times \overline{\mathcal{X}}_n^{\Theta_n} &\longrightarrow \overline{\mathcal{X}}_1^{\Theta_1} \times \cdots \times \overline{\mathcal{X}}_n^{\Theta_n} \\ \boldsymbol{\chi} = (\chi_1, \ldots, \chi_n) &\longmapsto \left(\phi_i^{BR}(\boldsymbol{\chi}), \ldots, \phi_n^{BR}(\boldsymbol{\chi})\right).\end{aligned}$$

We use the *Schauder-Tychonoff fixed-point theorem*[18] to establish equilibrium existence.

First, notice that for every $i \in \mathcal{I}$, $\Theta_i$ is compact. Moreover, $\overline{\mathcal{X}}_i$ is compact for the given topology that we have chosen. It is possible to interpret $\overline{\mathcal{X}}_i^{\Theta_i}$ as an infinite product of $\overline{\mathcal{X}}_i$, and hence by Tychonoff's theorem $\overline{\mathcal{X}}_i^{\Theta_i}$ is compact. Applying Tychonoff's theorem again, yields that $K := \overline{\mathcal{X}}_1^{\Theta_1} \times \cdots \times \overline{\mathcal{X}}_n^{\Theta_n}$ is compact. It is easily verified that $K$ is non-empty and convex.

We also have to verify that the best response function $\Phi$ is continuous for which it suffices to show continuity for each of the coordinate functions. Consider the coordinate function

$$\begin{aligned}\Phi_i : \overline{\mathcal{X}}_i^{\Theta_i} \times \overline{\mathcal{X}}_{-i}^{\Theta_{-i}} &\to \overline{\mathcal{X}}_i^{\Theta_i} \\ \boldsymbol{\chi} = (\chi_i, \chi_{-i}) &\mapsto \phi^{BR}(\boldsymbol{\chi})\end{aligned}$$

It is easily seen that $\Phi_i$ is constant in $\chi_i$. Moreover, since the expectation operator is linear, and in the given setting bounded, it follows that $\Phi_i$ is continuous in $\chi_{-i}$ (cf. (4.10)). This shows that every coordinate function $\Phi_i$ is continuous, and so is $\Phi$.

Finally, the existence of a fixed point of $\phi$ follows by the Schauder-Tychonoff fixed-point theorem, which completes the proof of equilibrium existence. □

*Proof of Proposition 4.1.* Consider the symmetric case, in which attributes values and preference types are iid: $F_i = F_j$, and $G_i = G_j$ for all $i, j \in \mathcal{I}$.

Given the cutoff-function profile $\chi = (\chi_1, \ldots, \chi_n)$, agent $i$'s expected payoff of the proposal, conditional on being pivotal, is

$$\begin{aligned}V_i(\theta_i, x_i) =& \theta_i x_i + (1-\theta_i)\mathbb{E}_{\Theta_{-i}}\left[\mathbb{E}_{\mathcal{X}}\left[\frac{1}{n-1}\sum_{j \neq i} X_j | X_j \geq \chi_j(\theta_j)\right]\right] \\ =& \theta_i x_i + (1-\theta_i)\frac{1}{n-1}\sum_{j\neq i}\mathbb{E}_{\Theta_j}\left[\mathbb{E}_{\mathcal{X}}\left[X_j | X_j \geq \chi_j(\theta_j)\right]\right]\end{aligned}$$

[17] $\phi^{BR}$ identifies for every $\sigma_{-i}$ a corresponding cutoff-function $\phi^{BR}(\sigma_i, \sigma_{-i}) = \chi_i \in \overline{\mathcal{X}}_i^{\Theta_i}$. Notice that $\phi^{BR}$ is constant in $\sigma_i$.

[18] cf. Aliprantis and Border (2006) p. 583

On an individual basis, we thus have

$$\begin{aligned} \Lambda_j :\ \overline{\mathcal{X}}_i^{\Theta_i} &\longrightarrow \mathbb{R} \\ \chi_i &\longmapsto \mathbb{E}_{\Theta_i}\left[\mathbb{E}_{\mathcal{X}}\left[X_i | X_i \geq \chi_i(\theta_i)\right]\right] \end{aligned} \tag{4.11}$$

Define the function

$$\begin{aligned} \widetilde{\beta} :\ \mathbb{R}_0^{+\,n-1} &\longrightarrow \overline{\mathcal{X}}_i^{\Theta_i} \\ (c_1, c_2, \ldots, \widehat{c_i}, c_{i+1}, \ldots c_n) &\longmapsto \chi_i(\theta_i) = \max\Big\{\underline{x},\ -\frac{1-\theta_i}{\theta_i} \cdot \frac{1}{n-1} \sum_{j \neq i} c_j\Big\}. \end{aligned}$$

We focus on symmetric equilibria, in which all agents adopt the same strategy. In a symmetric equilibrium $\chi_i^* \equiv \chi_j^*$, and hence $c_j^* = c_i^* := \mathbb{E}_{\Theta_i}\left[\mathbb{E}_{\mathcal{X}}\left[X_i | X_i \geq \chi_i^*(\theta_i)\right]\right]$ for all $i, j \in \mathcal{I}$. Consequently, it suffices to consider the following reduced version of the function $\widetilde{\beta}$:

$$\begin{aligned} \beta :\ \mathbb{R}_0^{+} &\longrightarrow \overline{\mathcal{X}}_i^{\Theta_i} \\ c &\longmapsto \chi_i(\theta_i) = \max\Big\{\underline{x},\ -\frac{1-\theta_i}{\theta_i} \cdot c\Big\}. \end{aligned}$$

The discussion shows that, for the unanimity rule, there exists a constant, $c \in \mathbb{R}_0^+$ that determines the equilibrium cutoff-functions.

Consider the following composite function

$$\begin{aligned} \Phi :\ \mathbb{R}_0^{+} \xrightarrow{\beta} \quad & \overline{\mathcal{X}}_i^{\Theta_i} \quad \xrightarrow{\Lambda} \mathbb{R} \\ c \longmapsto \max\Big\{\underline{x},\ -\frac{1-\theta_i}{\theta_i} \cdot c\Big\} & \longmapsto \mathbb{E}_{\Theta_i}\left[\mathbb{E}_{\mathcal{X}}\left[X_i | X_i \geq \chi_i(\theta_i)\right]\right]. \end{aligned}$$

The problem to find a symmetric equilibrium is equivalent to finding a fixed-point, $c^*$, of this mapping $\Phi$. Every fixed-point $c^*$ uniquely determines the equilibrium cutoff-functions and vice versa.

For $c \geq 0$, $\beta$ is decreasing in $c$ (in a "set-value" sense), meaning that if $c' \geq c$, then $\chi_{c'}(\theta_i) \leq \chi_c(\theta_i)$ for every $\theta_i \in \Theta_i$.

The left-truncated expectation operator is non-decreasing in the cutoff; that is, $\mathbb{E}\left[X | X \geq \hat{x}\right]$ is non-decreasing in $\hat{x}$. Consequently, if $\widetilde{\chi}(\theta_i) \leq \chi(\theta_i)$ for all $\theta_i \in \Theta_i$, then $\mathbb{E}_{\Theta_i}\left[\mathbb{E}_{\mathcal{X}}\left[X_i | X_i \geq \widetilde{\chi}_i(\theta_i)\right]\right] \leq \mathbb{E}_{\Theta_i}\left[\mathbb{E}_{\mathcal{X}}\left[X_i | X_i \geq \chi_i(\theta_i)\right]\right]$. It follows that $(\Lambda \circ \beta)(c)$ is decreasing in $c$ (for $c \in \mathbb{R}_0^+$).

We wish to determine a fixed point of $\Phi$, that is, some $c^* \in \mathbb{R}_0^+$ such that

$$c^* = (\Lambda \circ \beta)(c^*) \tag{4.12}$$

We know that for (4.12):

(i) For $c = 0$: $lhs < rhs$.
Indeed, $(\Lambda \circ \beta)(0) = \mathbb{E}\left[X | X \geq 0\right] > \mathbb{E}\left[X\right] = 0$.

(ii) For $c = \overline{x}$: $lhs \geq rhs$.
Indeed, $\mathbb{E}[X|X \geq \hat{x}] \leq \overline{x}$, with strict inequality for $\hat{x} < \overline{x}$. Moreover, for every $c \in \mathbb{R}_0^+$, $\chi_i(\theta_i)$ is continuous in $\theta_i$, with $\chi_i(\theta_i) \leq 0$ for all $\theta_i \in \Theta_i$, and $\chi_i(1) = 0$. It follows that $(\Lambda \circ \beta)(c) = \mathbb{E}_{\Theta_i}[\mathbb{E}_{\mathcal{X}}[X_i|X_i \geq \chi_c(\theta_i)]] < \overline{x}$.

The lhs of (4.12) is strictly increasing in $c$, whereas the rhs of (4.12) is strictly decreasing in $c$. The functions on both sides of (4.12) are continuous. It follows that there exists a fixed point $c^*$ of $\Phi = \Lambda \circ \beta$. Moreover, given that the functions on both sides are strictly decreasing, respectively increasing, there exists a unique fixed point of $\Phi$. □

*Proof of Lemma 4.1.*
(i): An agent with preference type $\theta_i = 1$ has private values, and hence $V_i\left((1, x_i); \sigma_{-i}^*\right) = x_i$, $\forall x_i \in X_i$. Given that the payoff of the status quo is 0 it follows that $\chi_i^*(1) = 0$, that is, in equilibrium, preference type $\theta_i = 1$ always votes sincerely.

Consider any equilibrium strategy profile $\sigma^*$ with corresponding acceptance sets $\{A_i^*(\theta_i)\}_{i \in \mathcal{I}, \theta_i \in \Theta_i}$. Notice that, for every agent $i$, $\mathbb{E}_{\sigma_{-i}}[\overline{x}_{-i}|piv]$ is constant in $\theta_i$. Moreover, $-\frac{1-\theta_i}{\theta_i} < 0$ for all $\theta_i \in (0,1)$. It follows that equilibrium cutoffs $\chi_i^*(\theta_i)$ have the same sign for all types $\theta_i \in \Theta_i \backslash \{0\}$:

$$sign\left[\chi_i^*(\theta_i)\right] = -sign\, \mathbb{E}_{\sigma_{-i}}[\overline{x}_{-i}|piv].$$

(ii): *Continuity of equilibrium cutoff-functions.*
Consider a committee member with type $\theta_i$ who adopts and interior equilibrium cutoff, that is, $\chi_i^*(\theta_i) \in (\underline{x}, \overline{x}]$. Given the equilibrium characterizing conditions in Theorem 4.1, it must hold that $V_i\left((\theta_i, x_i); \sigma_{-i}^*\right) \underset{(<)}{>} 0$ for $x_i \underset{(<)}{>} \chi_i^*(\theta_i)$. Since $V_i\left((\theta_i, x_i); \sigma_{-i}^*\right)$ is continuous in $\theta_i$ and $x_i$, this implies that there exist an $\epsilon > 0$ s.t. $V_i\left((\theta_i', \overline{x}); \sigma_{-i}^*\right) > 0$ and $V_i\left((\theta_i', \underline{x}); \sigma_{-i}^*\right) < 0$ for all $\theta_i' \in \mathcal{B}_\epsilon(\theta_i)$, where $\mathcal{B}_\epsilon(\theta_i)$ is the open $\epsilon-$ball about $\theta_i$. It then follows from the equilibrium characterizing conditions of Theorem 4.1, that all preference types $\theta_i' \in \mathcal{B}_\epsilon$ adopt interior cutoffs. In this case equilibrium cutoffs are characterized by $\chi_i^*(\theta_i) = -\frac{1-\theta_i}{\theta_i}\mathbb{E}_{\sigma_{-i}}[\overline{x}_{-i}|piv]$. Since $\mathbb{E}_{\sigma_{-i}}[\overline{x}_{-i}|piv]$ is constant and $-\frac{1-\theta_i}{\theta_i}$ is twice continuously differentiable in $\theta_i$, it follows that $\chi_i^*(\theta_i)$ is continuously differentiable in $\theta_i$.

Now, consider a preference type $\theta_i \in \Theta_i$ who adopts a boundary cutoff, $\chi_i^*(\theta_i) \in \{\underline{x}, \tilde{x}\}$, that is for type $\theta_i$ either $V_i\left((\theta_i, x_i); \sigma_{-i}^*\right) \geq 0$ for all $x_i \in X_i$, or $V_i\left((\theta_i, x_i); \sigma_{-i}^*\right) < 0$ for all $x_i \in X_i$. Consider the first case, that is, $\chi_i^*(\theta_i) = \underline{x}$ and $V_i\left((\theta_i, x_i); \sigma_{-i}^*\right) \geq 0$, $\forall\, x_i \in X_i$.[19] Given that $V_i\left((\theta_i, x_i); \sigma_{-i}^*\right)$ is monotone increasing in $x_i$, a necessary and sufficient condition for $V\left((\theta_i, x_i); \sigma_{-i}^*\right) \geq 0$ is $V\left((\theta_i, \underline{x}_i); \sigma_{-i}^*\right) \geq 0$. Since $V_i\left((\theta_i, x_i); \sigma_{-i}^*\right)$ is continuous in $\theta_i$, the set $\{\theta_i \in \Theta_i | V_i\left((\theta_i, \underline{x}_i), \sigma_{-i}^*\right) \geq 0\}$ is closed. Since $\frac{\partial V_i}{\partial \theta_i}\big|_{x_i = \underline{x}_i} \geq 0$, it follows that there exists some $\hat{\theta} \in \Theta_i$ such that $\{\theta_i \in \Theta_i | V_i\left((\theta_i, x_i), \sigma_{-i}^*\right) \geq 0\} = \left[\underline{\theta}, \hat{\theta}\right]$. By Theorem 4.1, $\chi^*(\theta_i) = \underline{x}$ for $\theta_i \in \left[0, \hat{\theta}\right]$. The equilibrium cutoff function is constant on this set and thus twice continuously

[19] The second case can be easily verified using analogous arguments.

differentiable in $\theta_i$. It is easy to check that $\lim_{\theta_i \downarrow \hat{\theta}} \chi^*(\theta_i) = \underline{x}$, which establishes continuity of the equilibrium cutoff function. However, $\chi^*(\theta_i)$ is not differentiable at $\hat{\theta}$, hence equilibrium cutoff-functions are only differentiable almost everywhere and the same holds true for higher order differentiability.

(iii): To prove the last statement we use again that $\mathbb{E}_{\sigma_{-i}}\left[\overline{x}_{-i}|piv\right]$ is constant in $\theta_i$. For all $\theta_i$ such that $\chi_i^*(\theta) \in (\underline{x}, \overline{x}]$, we have shown that the cutoff-function is continuously differentiable. We obtain:

$$\left|\frac{\partial}{\partial\theta_i}\chi_i^*\right| = -\frac{1}{\theta_i^2}\cdot\left|\mathbb{E}_{\sigma_{-i}}\left[\overline{x}_{-i}|piv\right]\right|,$$

from which it follows that $|\chi_i^*(\theta_i)|$ is non-increasing in $\theta_i$ whenever cutoffs are interior. Moreover, all types $\theta_i \in \left[0, \hat{\theta}_i\right]$ of agent $i$ adopt extreme cutoffs in $\{\underline{x}, \tilde{x}\}$. Since cutoff-functions are continuous, it follows that $|\chi_i^*(\theta_i)|$ is non-increasing in $\theta_i$ for all types $\theta_i \in \Theta_i$. □

*Proof of Corollary 4.1.* For unanimity voting, agent $i$ is pivotal if and only if all other committee members vote affirmatively. This implies

$$\begin{aligned}\mathbb{E}_{\sigma_{-i}}\left[\overline{x}_{-i}|piv\right] =& \frac{1}{n-1}\sum_{j\neq i}\mathbb{E}_{\Theta_j}\left[\mathbb{E}_{\mathcal{X}_j}\left[X_j \mid X_j \geq \chi_j^*(\theta_j)\right]\right]\\ >& \frac{1}{n-1}\sum_{j\neq i}\mathbb{E}_{\Theta_j}\left[\mathbb{E}_{\mathcal{X}_j}\left[X_j \mid X_j \geq \underline{x}\right]\right] = 0\end{aligned}$$

We obtain the last equality because $\mathbb{E}\left[X_j\right] = 0$ for every $j \in \mathcal{I}$. The inequality is strict because $\chi_j^*(1) = 0$ and $\chi_j^*(\theta_j)$ is continuous in $\theta_j$, which implies that in a neighborhood of $\boldsymbol{\theta_{-i}} = \mathbf{1}$, $\chi_j^*(\theta_j) \neq \underline{x}$ for all $j \in \mathcal{I}\backslash\{i\}$.

It follows that

$$\chi_i^*(\theta_i) = -\frac{1-\theta_i}{\theta_i}\mathbb{E}_{\Theta_{-i},\mathcal{X}_{-i}}\left[\frac{1}{n-1}\sum_{j\neq i}X_j \mid X_j \geq \chi_j^*(\theta_j)\right] \leq 0,$$

and hence by Lemma 4.1 (iii), $\chi_i^*(\theta_i)$ is increasing in $\theta_i$ for every $i \in \mathcal{I}$. □

*Proof of Lemma 4.2.* In Lemma 4.1 it was shown that $\chi_i^*$ is continuous in $\theta_i$, $\chi_i^*(1) = 0$ and $|\chi_i^*(\theta_i)|$ is non-increasing in $\theta_i$. Moreover, if $\chi_i^*(\theta_i) \in \{\underline{x}, \tilde{x}\}$, then $\chi_i^*(\theta_i') = \chi_i^*(\theta_i) \in \{\underline{x}, \tilde{x}\}$ for all $\theta_i' \leq \theta_i$. Continuity of $\chi_i^*$ for interior cutoffs and $\chi_i^*(1) = 0$ yield that if there are types which adopt extreme cutoffs in $\{\underline{x}, \tilde{x}\}$, then there exists some type $\hat{\theta}_i$ such that all types $\left[0, \hat{\theta}_i\right]$ adopt extreme cutoffs in $\{\underline{x}, \tilde{x}\}$ whereas all types $\left(\hat{\theta}_i, 1\right]$ adopt interior cutoffs. □

*Proof of Lemma 4.3.* By Lemma 4.1 $\chi_i^*(\theta_i)$ is constant on the set of non-responsive types. That is, $(\chi_i^*)'(\theta_i) = 0$ for all $\theta_i \in \left[0, \hat{\theta}_i\right)$ and (4.8) is trivially satisfied.

Now consider any responsive type $\theta_i \in \left(\hat{\theta}, 1\right]$. By Lemma 4.1 we know that $\chi_i^*$ is twice continuously differentiable at $\theta_i$. It is easily verified that for $\theta_i \in \left(\hat{\theta}, 1\right]$ we obtain

$$(\chi_i^*)' \cdot (\chi_i^*)''(\theta_i) = -\frac{2}{\theta_i^5}\mathbb{E}_{\sigma^*_{-i}}\left[\overline{x}_{-i}|piv\right]^2 \leq 0.$$

The result about concavity/convexity of equilibrium cutoff functions follows by combining this with the result of Lemma 4.1 (iii). □

*Proof of Proposition 4.2.* As shown in the proof of Proposition 4.1, there is a one-to-one correspondence between symmetric equilibria, and fixed-points of the composite function:

$$\begin{aligned} \Phi : \mathbb{R}_0^+ \xrightarrow{\beta} \quad & \mathcal{X}_i^{\Theta_i} \xrightarrow{\Lambda} \mathbb{R} \\ c \longmapsto \max\left\{\underline{x},\ -\frac{1-\theta_i}{\theta_i}\cdot c\right\} & \longmapsto \mathbb{E}_{\Theta_i}\left[\mathbb{E}_{\mathcal{X}}\left[X_i|X_i \geq \chi_i(\theta_i)\right]\right]. \end{aligned}$$

with

$$\begin{aligned} \beta(c) &:= \max\left\{\underline{x},\ -\frac{1-\theta_i}{\theta_i}\cdot c\right\}, \quad \text{and} \\ \Lambda(\chi_i) &:= \mathbb{E}_{\Theta_i}\left[\mathbb{E}_{\mathcal{X}}\left[X_i|X_i \geq \chi_i(\theta_i)\right]\right]. \end{aligned} \tag{4.13}$$

Based on this observation, we will first study the effects of changes in the distribution of preference types on the fixed-point(s) of the one-dimensional mapping $\Phi$. We then derive the induced effects on the equilibrium cutoff-functions of the original problem.

Every distribution function on $\Theta$ defines an operator $\Lambda$ as in (4.13), with the expectation taken with respect to the distribution of preferences types. The operators corresponding to the distributions $H$ and $G$ are denoted $\Lambda_H$ and $\Lambda_G$.

Notice that a change in the distribution function of agents' preference types only affects the operator $\Lambda$, whereas the mapping $\beta$ remains unchanged.

We know from Corollary 4.1 that, under unanimity voting, any equilibrium cutoff-function is non-positive and increasing in $\theta_i$. Moreover, if $H \geq_{FOSD} G$, then for every increasing function $\chi \in \mathcal{X}_i^{\Theta_i}$, it holds that

$$\Lambda_H(\chi) \geq \Lambda_G(\chi).$$

Since $\beta$ remains unchanged, and $\beta(c)$ is increasing in $\theta_i$ whenever $c \geq 0$, it follows that:

$$\Phi_H(c) = \Lambda_H(\beta(c)) \geq \Lambda_G(\beta(c)) = \Phi_G(c) \quad \forall\, c \geq 0. \tag{4.14}$$

Now, notice that $\Phi_G$ and $\Phi_H$ are functions from the unit interval to $\mathbb{R}$. Consequently, it is easy to graphically determine a fixed point of these mappings. It is simply an intersection point of the graph $\Phi_G(c)$, respectively $\Phi_H(c)$ with the 45°-line. We denote these points by $c_G$, respectively $c_H$. Figure 4.4 (page 113) illustrates the graphs and fixed points.

From Theorem 4.1, we know that fixed points of the mappings $\Phi_H$ and $\Phi_G$ exist. Moreover, by (4.14), $\Phi_H(c) \geq \Phi_G(c)$ for all $c \geq 0$, and hence the graph of $\Phi_H$ always lies above the graph of $\Phi_G$. As shown in the proof of Theorem 4.1, $\Phi_H$ and $\Phi_G$ are decreasing in $c$. Hence, for the intersection points of the graphs with the 45°-line it follows that $\Phi_H(c^*_H)$ lies above $\Phi_G(c^*_G)$, that is, $\Phi_H(c^*_H) \geq \Phi_G(c^*_G)$. This implies $c^*_H \geq c^*_G$.

For the equilibrium cutoff-functions it follows that

$$\chi^*_H(\theta_i) = \chi_{c^*_H}(\theta_i) \leq \chi_{c^*_G}(\theta_i) = \chi^*_G(\theta_i). \tag{4.15}$$

In the symmetric equilibrium, on an individual basis, acceptance standards are lower for more partisan populations than for more socially-oriented populations. □

*Proof of Proposition 4.3.* We again use the trick that it is possible to decompose the map $\Psi$ into $\Psi = \beta \circ \Lambda$, and that there is a one-to-one correspondence between fixed-points of the composition map $\Phi = \Lambda \circ \beta$ and symmetric voting equilibria. For two distribution functions $H \geq_{MPS} G$, we determine and compare the fixed points of $\Phi_H$ and $\Phi_G$, in order to then derive the implied relation of the equilibrium cutoff-functions for populations distributed according to $H$ and $G$, respectively.

As pointed out in the proof of Proposition 4.2, a shift in the distribution function of preference types only affects the second component, the function $\Lambda$, of the decomposition of $\Phi$.

By Lemma 4.3 and Corollary 4.1, under unanimity, equilibrium cutoff functions are increasing and concave, and $c^* = \mathbb{E}_{\Theta_i}\left[\mathbb{E}_{\mathcal{X}}\left[X_i | X_i \geq \chi^*_i(\theta_i)\right]\right] \geq 0$. Moreover, if $H \geq_{MPS} G$, then for every increasing and concave function $\chi \in \mathcal{X}_i^{\Theta_i}$, it holds that

$$\Lambda_H(\chi) \leq \Lambda_G(\chi).$$

Since $\beta$ remains unchanged, it follows that for every $c \geq 0$:

$$\Phi_H(c) = \Lambda_H(\beta(c)) \leq \Lambda_G(\beta(c)) = \Phi_G(c). \tag{4.16}$$

As in the proof of Proposition 4.2, it is easy to determine and compare the fixed points of the functions $\Phi_G$ and $\Phi_H$ graphically. They are simply the intersection points of the graph $\Phi_G(c)$, respectively $\Phi_H(c)$, with the 45°-line. Figure 4.5 illustrates the graphs and fixed points.

If $H \geq_{MPS} G$, by (4.16), the graph of $\Phi_H$ lies below $\Phi_G$ for all $c \geq 0$. Since $\Phi_G$ and $\Phi_H$ are decreasing in $c$, it follows that the intersection point of $\Phi_H$ with the 45°-line lies below the intersection point of $\Phi_G$ with that line. Consequently, $c^*_H \leq c^*_G$, and for the equilibrium

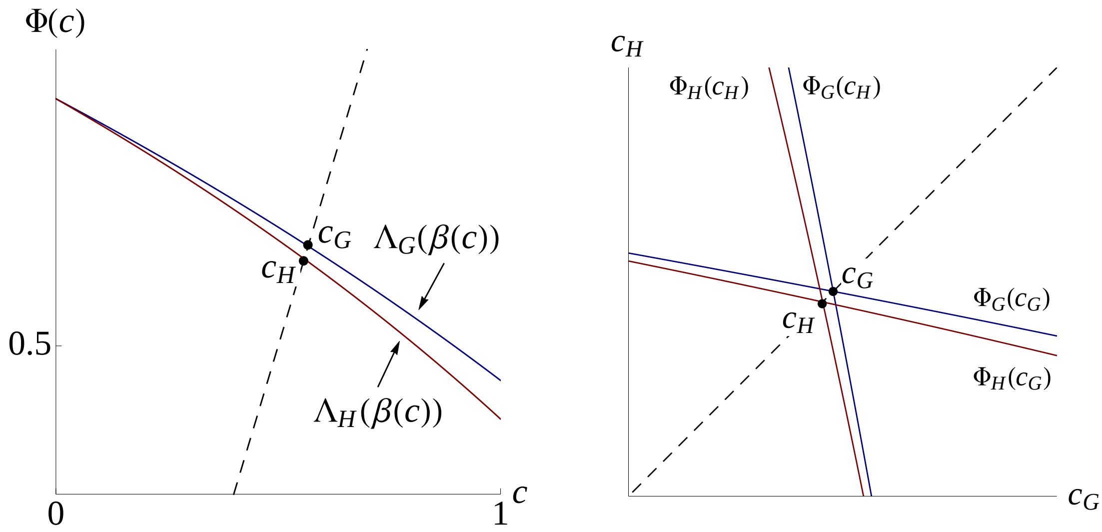



Figure 4.5: Illustration of the effect of a mean-preserving spread of the distribution of the populations' preference types on the expected value of the information of the other agent conditional on the event of being pivotal.

cutoff-functions it follows that:

$$\chi_H^*(\theta_i) = \chi_{c_H^*}(\theta_i) \geq \chi_{c_G^*}(\theta_i) = \chi_G^*(\theta_i) \quad \forall\, \theta_i \in \Theta_i,$$

which proves the result. □

# Bibliography


Acharya, Viral V., Peter DeMarzo, and Ilan Kremer (2011), "Endogenous Information Flows and the Clustering of Announcements." *American Economic Review*, 101, 2955–2979.

Aliprantis, Charalambos D. and Kim C. Border (2006), *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer, Berlin.

Alles, Lakshman A and John Kling (1994), "Regularities in the Variation of Skewness in Asset Returns." *Journal of Financial Research*, 17, 427–438.

Aoyagi, Masaki (2010), "Information Feedback in a Dynamic Tournament." *Games and Economic Behavior*, 70, 242–260.

Arnold, Barry C., Narayanaswamy Balakrishnan, and Haikady N. Nagaraja (1992), *A First Course in Order Statistics*. John Wiley & Sons, Inc., New York.

Aumann, Robert J. and Michael B. Maschler (1995), *Repeated Games with Incomplete Information*. MIT Press, Cambridge, MA.

Austen-Smith, David and Timothy J. Feddersen (2006), "Deliberation, Preference Uncertainty, and Voting Rules." *American Political Science Review*, 100, 209–217.

Bagnoli, Mark and Ted Bergstrom (2005), "Log-Concave Probability and its Applications." *Economic Theory*, 26, 445–469.

Barlow, Richard E. and Frank Proschan (1966), "Inequalities for Linear Combinations of Order Statics from Restricted Families." *The Annals of Mathematical Statistics*, 37, 1574–1592.

Barlow, Richard E. and Frank Proschan (1981), *Statistical Theory of Reliability and Life Testing. Probability Models*. Holt, Rinehart and Winston Inc., New York.

Becker, Gary S. (1973), "A Theory of Marriage: Part I." *Journal of Political Economy*, 91, 813–846.

Beedles, William L. (1979), "Return, Dispersion and Skewness: Synthesis and Investment Strategy." *Journal of Financial Research*, 2, 71–80.

Bergemann, Dirk, Benjamin A. Brooks, and Stephen Morris (2015), "The Limits of Price Discrimination." *American Economic Review*, 105, 921–957.

Bergemann, Dirk and Martin Pesendorfer (2007), "Information Structures in Optimal Auctions." *Journal of Economic Theory*, 137, 580–609.

Bergemann, Dirk and Juuso Välimäki (2002), "Information Acquisition and Efficient Mechanism Design." *Econometrica*, 70, 1007–1033.

Bergemann, Dirk and Juuso Välimäki (2007), "Information in Mechanism Design." In *Proceedings of the 9th World Congress of the Econometric Society*, 186–221.

Bilancini, Ennio and Leonardo Boncinelli (2013), "Disclosure of Information in Matching Markets with Non-Transferable Utility." *Games and Economic Behavior*, 82, 143–156.

Blackwell, D. (1951), "Comparison of Experiments." In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (J. Neyman, ed.), 93–102, University of California Press, Berkeley, CA.

Block, Henry W., Thomas H. Savits, and Eshetu T. Wondmagegnehu (2003), "Mixtures of Distributions with Increasing Linear Failure Rates." *Jounal of Applied Probability*, 40, 485–504.

Busse, Meghan R., Nicola Lacetera, Devin G. Pope, Jorge Silva-Risso, and Justin R. Sydnor (2013), "Estimating the Effect of Salience in Wholesale and Retail Car Markets." *American Economic Review, Papers & Proceedings*, 103, 575–579.

Cal, Jesus de la and Jarvier Carcamo (2006), "Stochastic Orders and Majorization of Mean Order Statistics." *Jounal of Applied Probability*, 43, 704–712.

Caplin, Andrew, Mark Dean, and John Leahy (2014), "The Behavioral Implications of Rational Inattention with Shannon Entropy." Working Paper, New York University.

Caplin, Andrew and Barry Nalebuff (1991a), "Aggregation and Social Choice: A Mean Voter Theorem." *Econometrica*, 59, 1–23.

Caplin, Andrew and Barry Nalebuff (1991b), "Aggregation and Imperfect Competition: On the Existence of Equilibrium." *Econometrica*, 59, 25–59.

Chakravarty, Surajeet and Todd R. Kaplan (2013), "Optimal Allocation without Transfer Payments." *Games and Economic Behavior*, 77, 1–20.

Chi, Chang-Koo (2014), "The Value of Information and Dispersion." Dissertation, University of Wisconsin - Madison.

Clarke, Edward H. (1971), "Multipart Pricing of Public Goods." *Public Choice*, 11, 17–33.

Cole, Harold L., George J. Mailath, and Andrew Postlewaite (2001), "Efficient Non-Contractibel Investments in Large Economies." *Journal of Economic Theory*, 101, 333–373.

Condorcet, Marquis de. (1785), "Essai sur l'Application de l'Analyse a la Probabilite des Decisions Rendues a la Pluralite des Voix." In *Classics of Social Choice. (transl.)*, 91–112, University of Michigan Press (1995), Ann Arbor.

Condorelli, Daniele (2012), "What Money Can't Buy: Efficient Mechanism Design with Costly Signals." *Games and Economic Behavior*, 75, 613–624.

Connelly, Brian L., Laszlo Tihanyi, T. Russell Crook, and K. Ashley Gangloff (2014), "Tournament Theory: Thirty Years of Contests and Competitions." *Journal of Management*, 40, 16–47.

Coughlan, P. (2000), "In Defense of Unanimous Jury Verdicts: Mistrials, Communication, and Sincerity." *American Political Science Review*, 94, 375–394.

Crawford, Vincent P. and Joel Sobel (1982), "Strategic Information Transmission." *Econometrica*, 50, 1431–1451.

Damodaran, Aswath (1985), "Economic Events, Information Structure, and the Return-Generating Process." *The Journal of Financial and Quantitative Analysis*, 20, 423–434.

de Clippel, Geoffroy, Kfir Eliaz, and Kareen Rozen (2014), "Competing for Consumer Inattention." *Journal of Political Economy*, 122, 1203–1234.

DellaVigna, Stefano (2009), "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature*, 47, 315–72.

Demange, Gabrielle and David Gale (1985), "The Strategy Structure of Two-Sided Matching Markets." *Econometrica*, 53, 873–888.

Dizdar, Deniz (2015), "Two-sided Investments and Matching with Multi-dimensional Types and Attributes." Working Paper, available at SSRN: http://ssrn.com/abstract=2357087.

Duggan, John and Cesar Martinelli (2001), "A Bayesian Model of Voting in Juries." *Games and Economic Behavior*, 37, 259 – 294.

Ederer, Florian (2010), "Feedback and Motivation in Dynamic Tournaments." *Journal of Economics & Management Strategy*, 19, 733–769.

Ely, Jeffrey, Alexander Frankel, and Emir Kamenica (2014), "Suspense and Surprise." *Journal of Political Economy*, 123, 1215–260.

Englmaier, Florian, Arno Schmöller, and Till Stowasser (2013), "Price Discontinuities in an Online Used Car Market." In *Annual Conference 2013: Competition Policy and Regulation in a Global Economic Order.*

Esö, Peter and Balazs Szentes (2007), "Optimal Information Disclosure in Auctions and the Handicap Auction." *Review of Economic Studies*, 74, 705–731.

Ewerhart, Christian (2013), "Regular Type Distributions in Mechanism Design and $\rho$-concavity." *Economic Theory*, 53, 591–603.

Feddersen, Timothy and Wolfgang Pesendorfer (1997), "Voting Behavior and Information Aggregation in Elections with Private Information." *Econometrica*, 65, 1029–1058.

Finkelstein, Maxim and Ji Hwan Cha (2013), *Stochastic Modeling for Reliability: Shocks, Burn-in and Heterogeneous populations*. Springer Series in Reliability Engineering, Springer, London.

Gabaix, Xavier and David Laibson (2003), "A New Challenge for Economics: The Frame Problem." In *Collected Essays in Psychology and Economics* (I. Brocas and J. Carillo, eds.), Oxford University Press, London.

Gabaix, Xavier, David Laibson, Guillermo Moloche, and Stephen Weinberg (2006), "Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model." *American Economic Review*, 96, 1043–1068.

Gale, David and Lloyd S. Shapley (1962), "College Admissions and the Stability of Marriage." *The American Mathematical Monthly*, 69, 9–15.

Ganuza, Juan-Jose and Jose S. Penalva (2010), "Signal Orderings Based on Dispersion and the Supply of Private Information in Auctions." *Econometrica*, 78, 1007–1030.

Ganuza, Juan-Jose and Jose S. Penalva (2014), "Information Disclosure in Optimal Auctions." Working Paper 14-02, Universidad Carlos III de Madrid.

Gentzkow, Matthew and Emir Kamenica (2014), "Costly Persuasion." *American Economic Review, Papers & Proceedings*, 104, 457–62.

Goltsman, Maria and Arijit Mukherjee (2011), "Interim Performance Feedback in Multistage Tournaments: The Optimality of Partial Disclosure." *Journal of Labor Economics*, 29, 229–265.

Groves, Theodore (1973), "Incentives in Teams." *Econometrica*, 41, 617–631.

Grüner, Hans Peter and Alexandra Kiel (2004), "Collective Decisions with Interdependent Valuations." *European Economic Review*, 48, 1147–1168.

Gul, Faruk, Wolfgang Pesendorfer, and Tomasz Strzalecki (2015), "Coarse Competitive Equilibrium and Extreme Prices." Working Paper, Harvard and Princeton University.

Gurland, John and Jayaram Sethuraman (1994), "Reversal of Increasing Failure Rates When Pooling Failure Data." *Technometrics*, 36, 416–418.

Hansen, Stephen E. (2013), "Performance Feedback with Career Concerns." *Journal of Law, Economics, and Organization*, 29, 1279–1316.

Hardy, Godfrey H., John E. Littlewood, and George Polya (1952), *Inequalities.* Cambridge University Press, Cambridge.

Hatfield, John William and Paul R. Milgrom (2005), "Matching with Contracts." *American Economic Review*, 95, 913–935.

Hofstede, Geert (1991), *Cultures and Organizations: Software of the Mind.* McGraw-Hill UK, London.

Hopkins, Ed (2012), "Job Market Signalling of Relative Position, or Becker Married to Spence." *Jounal of the European Economic Association*, 10, 290–322.

Hoppe, Heidrun, Benny Moldovanu, and Aner Sela (2009), "The Theory of Assortative Matching Based on Costly Signals." *The Review of Economic Studies*, 76, 253–281.

Hoxby, Caroline and Sarah Turner (2013), "Expanding College Opportunities for High-Achieving, Low Income Students." Technical report, SIEPR Discussion Paper 12-014.

Jehiel, Philippe and Benny Moldovanu (2001), "Efficient Design with Interdependent Valuations." *Econometrica*, 69, 1237–1259.

Kamenica, Emir and Matthew Gentzkow (2011), "Bayesian Persuasion." *American Economic Review*, 101, 2590–2615.

Kessler, Anke S. (1998), "The Value of Ignorance." *The RAND Journal of Economics*, 29, 339–354.

Lacetera, Nicola, Devin G. Pope, and Justin R. Sydnor (2012), "Heuristic Thinking and Limited Attention in the Car Market." *American Economic Review*, 102, 2206–2236.

Lazear, Edward P. and Sherwin Rosen (1981), "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy*, 89, 841–864.

Lehmann, E. (1988), "Comparing Location Experiments." *The Annals of Statistics*, 16, 521–533.

Li, Hao, Sherwin Rosen, and Wing Suen (2001), "Conflicts and Common Interests in Committees." *American Economic Review*, 91, 1478–1497.

Li, Hao and Xianwen Shi (2015), "Discriminatory Information Disclosure." Working Paper, University of Toronto.

Li, Hao and Wing Suen (2009), "Decision-Making in Committees." *Canadian Journal of Economics*, 42, 359–392.

Lynch, James D. (1999), "On Conditions for Mixtures of Increasing Failure Rate Distributions to have an Increasing Failure Rate." *Probability in the Engineering and Informational Sciences*, 13, 33–36.

Mailath, George J, Andrew Postlewaite, and Larry Samuelson (2013), "Pricing and Investments in Matching Markets." *Theoretical Economics*, 8, 535–590.

Marshall, Albert W., Ingram Olkin, and Barry C. Arnold (2011), *Inequalities: Theory of Majorization and its Applications*, 2nd edition. Springer Series in Statistics, New York.

Matejka, Filip (2015), "Rigid Pricing and Rationally Inattentive Consumer." *Journal of Economic Theory*, 158, Part B, 656–678. Symposium on Information, Coordination, and Market Frictions.

Meyer, Margaret and Bruno Strulovici (2015), "Beyond Correlation: Measuring Interdependence Through Correlation." Working Paper, University of Oxford.

Milgrom, Paul R. (1981), "Good News and Bad News: Representation Theorems and Applications." *The Bell Journal of Economics*, 12, 380–391.

Milgrom, Paul R. and Chris Shannon (1994), "Monotone Comparative Statics." *Econometrica*, 62, 157–180.

Mizuno, Toshihide (2006), "A Relation between Positive Dependence of Signal and the Variability of Conditional Expectation Given Signal." *Journal of Applied Probability*, 43, 1181–1185.

Moldovanu, Benny and Aner Sela (2001), "The Optimal Allocation of Prizes in Contests." *American Economic Review*, 91, 542–558.

Moldovanu, Benny and Aner Sela (2006), "Contest Architecture." *Journal of Economic Theory*, 126, 70–96.

Moldovanu, Benny, Aner Sela, and Xianwen Shi (2007), "Contests for Status." *Journal of Political Economy*, 115, 338–363.

Moldovanu, Benny and Xianwen Shi (2013), "Specialization and Partisanship in Committee Search." *Theoretical Economics*, 8, 751–774.

Myerson, Roger B. (1997), *Game Theory*. Harvard University Press, Cambridge.

Persico, Nicola (2000), "Information Acquisition in Auctions." *Econometrica*, 68, 135–148.

Peters, Michael and Aloysius Siow (2002), "Competing Premarital Investments." *Journal of Political Economy*, 110, 592–608.

Prekopa, Andras (1973), "On Logarithmic Concave Measures and Functions." *Acta Scientiarum Mathematicarum*, 34, 335–343.

Quah, John K.-H. and Bruno Strulovici (2012), "Aggregating the Single Crossing Property." *Econometrica*, 80, 2333–2348.

Riley, J. G. and R. Zeckhauser (1983), "Optimal Selling Strategies: When to Haggle, When to Hold Firm." *Quarterly Journal of Economics*, 98, 267–290.

Roesler, Anne-Katrin (2014), "Information Disclosure in Markets: Auctions, Contests, and Matching Markets." Working Paper, University of Bonn.

Rosar, Frank (2015), "Continuous Decisions by a Committee: Median versus Average Mechanisms." *Journal of Economic Theory*, 159, Part A, 15–65.

Shah, Anuj K. and Daniel M. Oppenheimer (2008), "Heuristics Made Easy: An Effort-Reduction Framework." *Psychological Bulletin*, 207–222.

Shaked, Moshe and J. George Shanthikumar (2007), *Stochastic Orders*. Springer Series in Statistics, Springer, New York, NY.

Shaked, Moshe, Miguel A. Sordo, and Alfonso Suarez-Llorens (2012), "Global Dependence Stochastic Orders." *Methodology and Computing in Applied Probability*, 14, 617–648.

Shapley, Lloyd S. and Martin Shubik (1971), "The Assignment Game I: The Core." *International Journal of Game Theory*, 1, 111–130.

Shi, Xianwen (2012), "Optimal Auctions with Information Acquisition." *Games and Economic Behavior*, 74, 666–686.

Sims, Christopher A. (1998), "Stickiness." *Carnegie-rochester Conference Series On Public Policy*, 49, 317 – 356.

Sims, Christopher A. (2003), "Implications of Rational Inattention." *Journal of Monetary Economics*, 50, 665–690.

Sims, Christopher A. (2006), "Rational Inattention: Beyond the Linear-Quadratic Case." *American Economic Review*, 96, 158–163.

Sobel, Joel (2012), "Complexity versus Conflict in Communication." In *Proceedings of 46th Annual CISS (Conference on Information)*.

Spence, Michael (1973), "Job Market Signaling." *Quarterly Journal of Economics*, 87, 296–332.

Triandis, Harry C. (2001), "Individualism-Collectivism and Personality." *Journal of Personality*, 69, 907–924.

Vickrey, William (1961), "Counterspeculation, Auctions, and Competitive Sealed Tenders." *The Journal of Finance*, 16, 8–37.

Wilson, Andrea (2014), "Bounded Memory and Biases in Information Processing." *Econometrica*, 82, 2257–2294.

Woodford, Michael (2012), "Inattentive Valuation and Reference-Dependent Choice." Working Paper, Columbia University.

Yang, Ming (2015a), "Coordination with Flexible Information Acquisition." *Journal of Economic Theory*, 158, Part B, 721–738. Symposium on Information, Coordination, and Market Frictions.

Yang, Ming (2015b), "Optimality of Debt under Flexible Information Acquisition." Working Paper, Duke University.

Yee, Michael, Ely Dahan, John R. Hauser, and James Orlin (2007), "Greedoid-Based Noncompensatory Inference." *Marketing Science*, 26, 532–549.

Yildirim, Huseyin (2012), "Time-consistent majority rules and heterogeneous preferences in group decision-making." Working paper, Duke University.