Is Ignorance Bliss?
Rational Inattention and Optimal Pricing*

Anne-Katrin Roesler†

JOB MARKET PAPER

December 30, 2014

Abstract

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 consumer-optimal information structure satisfies three properties: It is partitional (*coarse perception*), guarantees *seller indifference*, and induces *efficient trade*. The consumer benefits from having access to information structures with more signal realizations. Every consumer-optimal 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 the resulting equilibrium, trade is efficient and the consumer is strictly better off than under fully informed monopoly pricing. Surprisingly, even in the absence of information processing constraints and costs, the consumer does not want to become perfectly informed.

Keywords: Rational inattention, Information design, Monopoly pricing, Information constraints.

JEL Classifications: C72, D42, D82, D83, L12.

*I am grateful to Dirk Bergemann for guidance and encouragement throughout the project. I thank Johannes Hörner, Daniel Krähmer, and Benny Moldovanu for insightful conversations and valuable suggestions. The paper has also benefited from comments by Ben Brooks, Max Conze, Albin Erlanson, Dennis Gärtner, Eugen Kovac, Stephan Lauermann, Simon Loertscher, Michael Mandler, Filip Matejka, Stephen Morris, Aniko Oery, Larry Samuelson, Philipp Strack, Michael Woodford, and audiences in Bonn and at Yale. Financial support from the DFG through BGSE and from the German Academic Exchange Service is gratefully acknowledged. Part of this research was conducted while visiting Yale University. I thank the Economics Department and the Cowles Foundation for their hospitality.

†Department of Economics, University of Bonn, 53113 Bonn, Germany, aroesler@uni-bonn.de. The latest version of this paper is available here.

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 paper, 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 paper 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 paper 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 paper 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*."

An essential contribution of this paper is to characterize the *consumer-optimal information structure*, that is, the information structure that the consumer chooses in equilibrium. The key features of the optimal information processing structure for a capacity-constrained consumer are as follows:

Coarse Perception: It is optimal for the consumer to obtain a coarse perception of his valuation. The consumer-optimal information structure is a monotone partition. 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.

Efficiency: By choosing an information structure that induces the seller to charge a price that yields efficient trade, the consumer ensures that 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.

Seller-Indifference: The consumer-optimal information structure induces an interim demand curve that yields the same expected revenue for the seller, for each of the potentially optimal prices. Hence, the seller is indifferent between charging the equilibrium price and a price equal to any of the higher value estimates that are induced by the information structure. By adopting such an information structure, the consumer induces the seller to charge a price that yields efficient trade while – at that price – leaving her with just enough revenue in order to guarantee that the seller does not want to deviate and charge a higher price.

The main features of the consumer-optimal information structure also persist in the unconstrained case. The unconstrained consumer-optimal information structure induces efficient trade and satisfies a form of seller-indifference. 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 obtain only a coarse perception about an interval of low values, whereas his perception of higher values is finer and may be perfectly informative. This unconstrained consumer-optimal information structure induces efficient trade.

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 the present paper, 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 paper are mostly constructive.

Related Literature This paper 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. (2014), 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 the present paper resembles the approach in Gul et al. (2014). In the competitive market analyzed in Gul et al. (2014), none of the agents has strategic influence. This is precisely the opposite of what is assumed in this paper; 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. (2014). In their paper, market prices serve to allocate attention, whereas in the present model the allocation of attention determines the induced price.

The present paper 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] That is, 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 the present paper identify the consumer-optimal information structure, if the seller best-responds with a revenue-maximizing mechanism.

This paper analyzes how the information environment of the consumer affects prices and

[1] Some recent papers, such as Yang (2013, 2014), also study flexible information acquisition in other contexts, for instance asset pricing and coordination problems.

the market outcome. A related question is studied in Bergemann et al. (2014). 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. (2014) 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.

The paper is also related 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 (2012). 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 Clippel et al. (2014).

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

[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.

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 μ_0. The seller's valuation for the object is 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 π, 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:

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

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

Capacity constraints. In the present paper, 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. (2014) 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 π, 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 π, 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. The timing of the private signal and the price setting decision can be interchanged, or be simultaneous.

Strategies and Solution Concept. Every information structure π 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 π and a mapping from value estimates and prices to a purchasing decision,

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

[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 an optimal selling mechanism is a posted price 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.

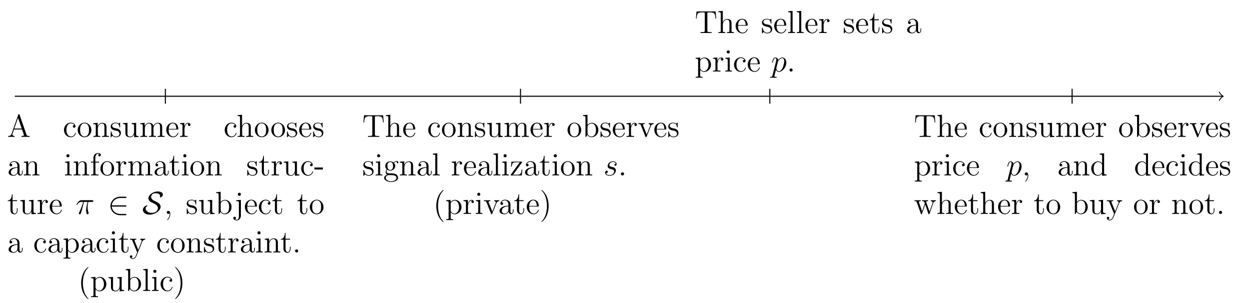


Figure 1: Timing in the monopoly model with a rationally inattentive consumer.

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 that the consumer may choose to the set of price distributions

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

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

The solution concept is perfect Bayesian equilibrium.

The information structure π is said to *induce* the price p, if p is a best-response for the seller to the information structure π. Say that the information structure π *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 π, and the consumer best-responds to this.

3 Illustrative Example: The Uniform Prior Case

In order 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 which he privately knows his true valuation.

[5] In the linear setting with risk-neutral agents considered in this paper, the distribution over value estimates captures all information about π 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 and Dean (2013). They reduce the problem to posterior beliefs. The model in the present paper considers risk-neutral agents and linear utilities. Hence, a reduction of the problem to value estimates is possible.

In the first case, the consumer is uninformed. 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 2.

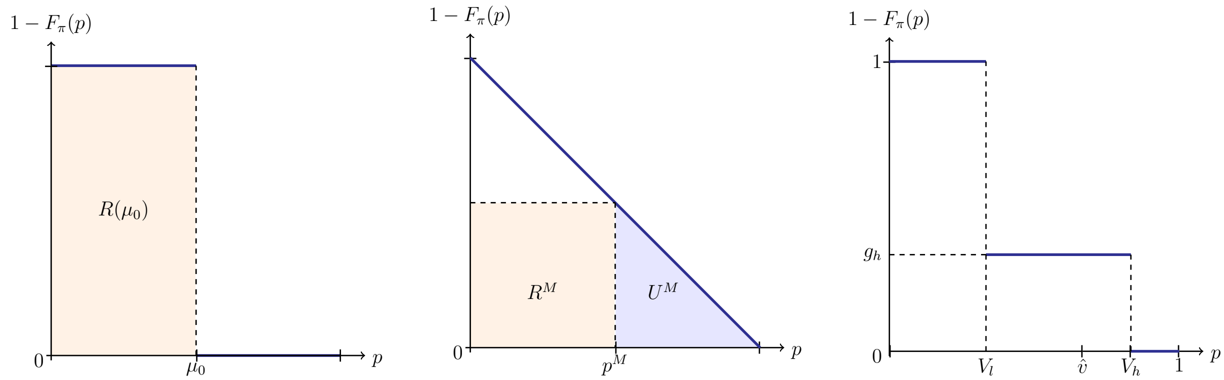
Optimal partitional two-signal information structure. Suppose that information is fully and freely available, and that the consumer has to decide how to process this information. Consider the case in which the consumer can only distinguish two categories, one of which he interprets as "good" and the 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 two 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. This means that the induced value estimate V_h is larger than the prior mean. By contrast, a realization of the low signal s_l yields a decrease of the consumer's willingness to pay, $V_l \leq \mu_0$.

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 willingness to pay of a consumer who observes a high signal. This is illustrated in Figure 2(c).

The seller's objective is to maximize her expected revenue. It is straightforward, that

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

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_l 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 determines both, the possible price realizations V_l 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 π, the seller charges the inclusive price, $V_l(\pi)$, if it yields a weakly higher revenue than the exclusive price, $V_l(\pi) \geq g_h(\pi)V_h(\pi)$.

The consumer's problem reduces to:

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

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 π.

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 $\hat{v}$

[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.

that 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 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.

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 an exclusive price is high if the "good" category is large, and this revenue is decreasing in the threshold that determines the information structure. Moreover, the inclusive price is increasing in the threshold. The seller's revenue from charging the exclusive price increases whereas her revenue from charging the inclusive price decreases. It follows that there exists a critical threshold at which $V_l(\pi_{\hat{v}}) = g_l(\pi_{\hat{v}})V_h(\pi_{\hat{v}})$, and the seller charges the inclusive price for all higher thresholds and the exclusive price otherwise. For a uniform prior, the critical threshold is $\frac{1}{2}\left(\sqrt{5}-1\right)$. That is, the seller charges an inclusive price for any information structure with threshold $\hat{v} \in \left[\frac{1}{2}\left(\sqrt{5}-1\right),\, 1\right]$.

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

$$U\left(\pi_{\hat{v}}\right) = \mu_0 - p^*\left(\pi_{\hat{v}}\right), \tag{3}$$

which is decreasing in the price. It follows that the consumer-optimal information structure is the one that induces the lowest price among all information structures that induce the seller to charge an inclusive price. The consumer-optimal threshold is $v^* = \frac{1}{2}\left(\sqrt{5}-1\right)$. This optimal monotone partitional two-signal information structure is illustrated in Figure 3.

The identification of the consumer-optimal information structure in the example was based on the assumption that the set of accessible information structures consists of the monotone partitional two-signal information structures. Certainly, the question arises whether this information structure is optimal for the consumer among larger classes of accessible infor-

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

[9]Indeed,

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

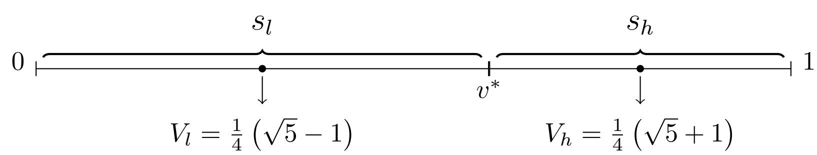


Figure 3: Consumer-optimal two-signal information structure, $\pi^*_{(2)}$, with threshold $v^* = \frac{1}{2}\left(\sqrt{5}-1\right)$. The high signal, s_h, realizes if the consumer's valuation is in the good category, otherwise the low signal, s_l, realizes.

mation structures. Is a partitional information structure optimal, or can the consumer benefit from noise in the 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 consumer-optimal information structure is identified, which determines how a capacity constrained consumer should optimally process or categorize information. In this section, first some preliminary observations are discussed, then properties are derived that any consumer-optimal information structure must satisfy, if it exists. These properties are then used in order to reduce the problem of finding a consumer-optimal information structure, and to prove existence.

For any information structure π, with n signal realization $s_1, \ldots, s_n$, 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, $V_1 \leq \cdots \leq V_n$.

4.1 Basic Observations about Optimal Purchasing, Price-Setting and Information Processing

Consumer's purchasing decision. The information structure π together with the signal realization s yields the value estimate $V_s = \mathbb{E}\left[v|s\right]$ for the consumer. It is immediate that for any given information structure π and price p, a best-response for the consumer is to buy the object if and only if his value estimate is greater or equal to the price,

$$\phi^*(V_s, p) = 1 \;\Leftrightarrow\; V_s \geq p. \tag{4}$$

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

$$\text{supp}\,(F_\pi) = \{V_s : s \in S\}.$$

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 (Selling Mechanism)**.**
For an information structure π that induces the distribution of value estimates F_π, 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_π).

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 π, 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 consumers with a high value estimate buy the good, and demand is low. By contrast, for a fully inclusive price the probability of trade is 1. Given that the valuation of the seller is 0, 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 optimal information structure: His information processing decision does not only determine the distribution over his value estimates, but also the interim demand function, and hence influences the price that the seller will charge.

In the analysis of the consumer-optimal information structure the following complication arises. A change in the information structure that has a seemingly small effect on the distribution and values of the value estimates, may induce the seller to switch from charging an inclusive price to charging an exclusive price. In this case, the resulting effect on the consumer's expected profit is large. 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 profit if the seller charges an inclusive price. It follows that the consumer will always choose an information structure π that induces the seller to charge an inclusive price.

Lemma 2 (Inclusive Prices)**.**
The consumer chooses an information structure that induces the seller to charge a (partially or fully) inclusive price.

4.2 Seller's Indifference

A central result in the characterization of the consumer-optimal information structure is a property that I refer to as *seller-indifference*. This property requires that, if the consumer-optimal 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 (Seller-Indifference)**.**
Suppose that π is a consumer-optimal 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_π:

$$R(V_i) = R(V_j) \quad \forall j \geq i. \tag{5}$$

The seller is indifferent between charging any of these prices.

For any finite information structure π, the optimal price for the seller is equal to one of the value estimates induced by π (Lemma 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_π. Proposition 1 establishes that the consumer-optimal information structure induces a distribution over value estimates that is an equal revenue curve above the equilibrium price charged by the seller. 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.

Let me now sketch the idea of the proof. The result is proven by an indirect argument. For any information structure π that does not satisfy the seller-indifference property (5), a new information structure, $\widetilde{\pi}$, is constructed that makes the consumer better off.

[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*".

Suppose that π is an information structure that does not satisfy (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 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 consumer-optimal information structure is that it lies on the efficient frontier, that is, it induces efficient trade.[11]

Theorem 1 (Efficient Trade)**.**
The consumer-optimal information structure lies on the efficient frontier. Any optimal finite information structure induces efficient trade. The exclusion region is empty.

The intuition for this result is as follows. If trade is not efficient, then some low consumer types are excluded from trade. Hence, the potential gains from trade are not fully realized. 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 consumer-optimal information

[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.

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 conditions 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 now sketch the proof of Theorem 1. It consists of two steps, and the formal proof is in the appendix. First, the problem is reduced by showing that any consumer-optimal 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 π 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 π and $\widetilde{\pi}$ coincide.

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 hence efficient trade – and yields a weakly higher expected surplus for the consumer.

4.4 Optimality of Coarse Perception

The previous findings have established properties (*seller-indifference* and *efficient trade*) for the outcome induced by the consumer-optimal information structure. In this section, the question how the consumer should optimally process information is addressed. It is shown that any consumer-optimal finite information structure is monotone partitional. The consumer only learns in which range of values his true valuation lies, and thus obtains only a coarse perception of his valuation for the good.

Theorem 2 (Optimality of Coarse Perception)**.**
For every $n \in \mathbb{N}$, any consumer-optimal information structure, $\pi^*_{(n)}$ is monotone partitional.

The result shows that, if the consumer is restricted to information structures with a limited number of signal realizations, then he cannot profit from adding noise to the signals. An optimal categorization of the value space for the consumer is to partition it into sub-intervals and to learn in which of the sub-intervals his true valuation falls.[12] That is, in equilibrium, the consumer has a coarse perception of the value space.

4.5 Existence of a consumer-optimal information structure

Building on the necessary properties for a consumer-optimal information structure that were identified in the previous sections, the next result establishes equilibrium existence.

Theorem 3 (Existence of Finite Consumer-Optimal Information Structures).
For every $n \in \mathbb{N}$ there exists a consumer-optimal information structure $\pi^*_{(n)}$ that maximizes the induced expected surplus of the consumer among all information structures with at most n signal realizations.

This result is established by first using that any consumer-optimal information structure must induce seller-indifference, efficient trade and be monotone partitional in order to reduce the problem of identifying a consumer-optimal information structure. Using this simplification of the problem, existence of a consumer-optimal information structure is proven. This result also establishes equilibrium existence.

4.6 More signals are better

As already mentioned, a straightforward question to ask is, whether two signal realizations are enough, or the consumer can profit from having access to information structures with more signal realizations. It is obvious that having access to more information structures increases the choice set of the consumer. Hence, he will be weakly better off. But will he strictly benefit from having access to more signal realizations? If so, is there a maximal number of signal realizations such that the consumer cannot profit from information structures with 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

[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.

[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.

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 (Two signal realizations are not enough)**:**
Reconsider the uniform prior example discussed in Section 3. The consumer-optimal monotone partitional two-signal information structure $\pi^*_{(2)}$ was identified (illustrated in Figure 3). By Theorem 2, this is indeed the consumer-optimal 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 4 illustrates such an information structure, $\widetilde{\pi}_{(3)}$.

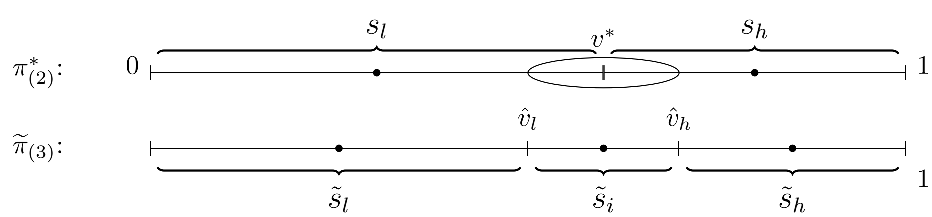


Figure 4: Illustration of the use of an additional signal.

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 and the seller's revenue under the optimal two-signal information structure. Given the seller-indifference property (Proposition 1), for the latter, one can consider the seller's revenue if she charges the exclusive price, induced by the original good signal. Notice that 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 2.

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 consumer-optimal finite information structure induces efficient trade (Theorem 1). Hence, the consumer-optimal n-signal information structure induces the minimal price $p^*_{(n)}$. It follows that 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 2 (More Signals are Better)**.**
The consumer strictly profits from having access to information structures with more signal realizations. The sequence of equilibrium prices $\{p^*_{(n)}\}_{n\in\mathbb{N}}$ is strictly decreasing in n, and the consumer's expected surplus $U^*_{(n)}$ is strictly increasing in n.

Evolution of optimal information structures and thresholds with n

How should a capacity constrained consumer optimally allocate his attention? As discussed above, the consumer will obtain a coarse perception about his valuation for the good. Still,

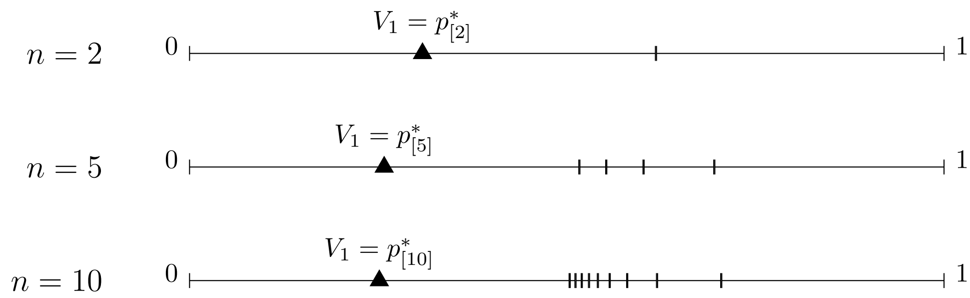


Figure 5: Thresholds and induced prices for the consumer-optimal n-signal information structures.

the question remains, which pieces of information are the most valuable for the consumer. Should the consumer obtain a finer perception about low or high valuations?

For the uniform prior example, the evolution of optimal information structures and prices is illustrated in Figure 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 Optimal Information Structure

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 consumer-optimal information structure pools all values below a certain threshold.

Theorem 4 (Optimal Information Structure)**.**
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* π^**:*

The information structure π^* *pools all values below the threshold* $\underline{v}$*, and the induced distribution of value estimates,* F_{π^*}*, 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 high mass 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 optimal information structures, respectively equilibria, arises. The optimal information structure π^* identified in Theorem 4, is the limit of the consumer-optimal finite 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 consumer-optimal information structures in more detail, I discuss three examples. Each of these examples illustrates a typical consumer-optimal information structure. The discussion shall provide some intuition how specific features of optimal information structures depend on properties of the prior distribution.

Example 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 consumer-optimal information structures. Two examples of consumer-optimal information structures are the following:

a) The information structure π^*_{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 6.

b) The information structure $\pi^*_{(\infty)}$ pools all types in the interval $\left[0, \frac{1}{2}\right)$. On $\left[\frac{1}{2}, 1\right]$, it induces a distribution over value estimates with constant discrete virtual valuation equal to zero.[15] This information structure is the limit of the consumer-optimal finite information

[15] This information structure is characterized by a decreasing sequence of thresholds $\{\hat{v}_i\}_{i\in\mathbb{N}_0}$ with

$$\hat{v}_0 := 1, \quad \hat{v}_1 := \frac{1}{2}\sqrt{2}, \quad \hat{v}_{i+1} := \frac{1}{2} + \frac{1}{2}\left(\sqrt{4\hat{v}_i\hat{v}_{i-1} + 5\hat{v}_i^2 - 4\hat{v}_{i-1} - 2\hat{v}_i + 1} - \hat{v}_i\right), \text{ and } \hat{v}_\infty := \frac{1}{2}.$$

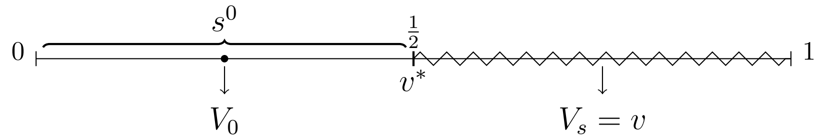


Figure 6: The consumer-optimal information structure π^*_{max} for the uniform prior case, $v \sim U\,[0,\,1]$.

structures,

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

The information structure π^*_{max} is the finest consumer-optimal information structure whereas $\pi^*_{(\infty)}$ is the coarsest consumer-optimal information structure. The distributions of value estimates induced by these two information structures are illustrated in Figure 7.

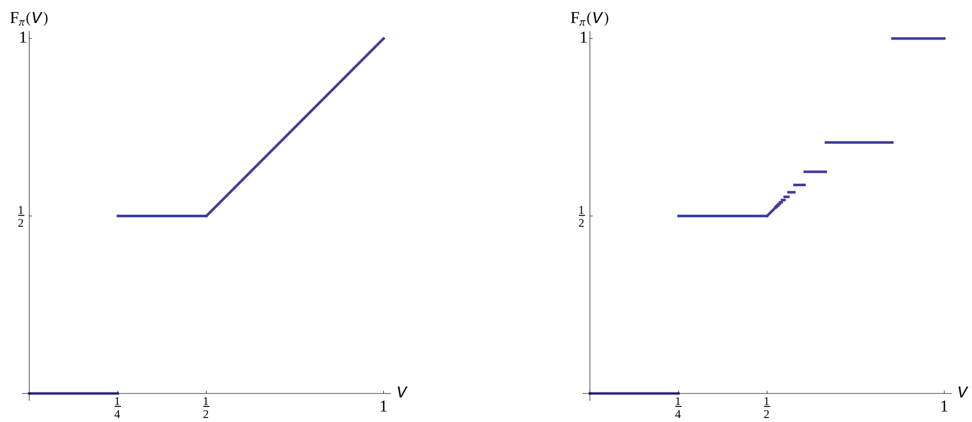


Figure 7: Distributions over value estimates F_{π^*} that are induced by the consumer-optimal information structures π^*_{max} and $\pi^*_{(\infty)}$, for the uniform prior case, $v \sim U\,[0,\,1]$.

In the uniform prior case, $\mathbb{E}\left[v|v \le p^M\right] = \frac{1}{4} = p^M\left(1 - F(p^M)\right)$. 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 $\mathrm{supp}\,(F_{\pi^*}) \setminus \{p^*\}$. In equilibrium, the seller will charge the fully inclusive price.

Example 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, $\mathbb{E}\left[v|v \le p^M\right] > R^M$. 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 consumer-optimal information structure π^* 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 consumer-optimal information structure π^*_{max} is illustrated in Figure 8. It induces efficient trade, $T^* = \frac{1}{3}$, and the induced expected revenue of the seller is below the monopoly revenue $R^* < R^M$.

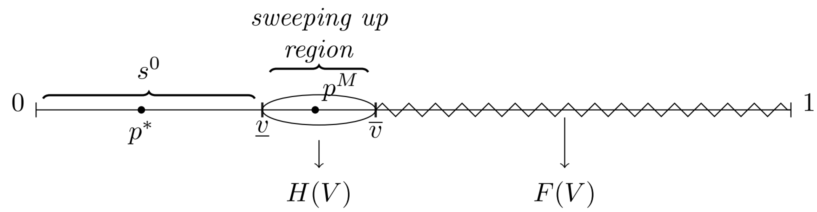


Figure 8: Consumer-optimal information structures π^*_{max} for a decreasing prior density.

Example 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 consumer-optimal information structure. The seller's expected revenue induced by this information structure is smaller than the monopoly revenue. This consumer-optimal information structure is illustrated in Figure 9.

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 sur-

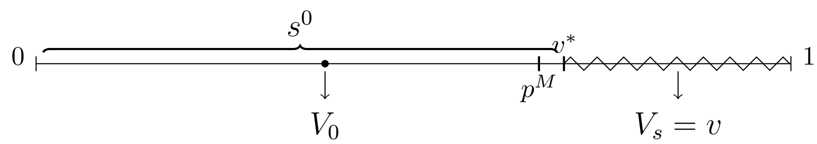


Figure 9: The consumer-optimal information structure π^*_{max} for, $v \sim \beta\left(\frac{1}{2}, 1\right)$.

plus. In order to analyze these effects, the equilibrium outcomes that are induced by the consumer-optimal 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 consumer-optimal 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 consumer-optimal 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 consumer-optimal 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 2 it was shown that the consumer strictly profits from having access to more signal realizations. The following corollary generalizes this result.

Corollary 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 3 and Example 4 it may be strictly smaller.

In each of the above examples, it can be shown that the outcome induced by the consumer-optimal two-signal information structure is a Pareto improvement compared to the benchmark case with a fully informed consumer. Numerical values are provided in Table 4.

It is instructive to illustrate the results on the consumer's expected surplus and the seller's expected revenue in a surplus triangle. By Theorem 1, all equilibrium outcomes lie

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

Table 4: 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.

on the efficient frontier. The locations of the equilibrium outcomes in the surplus triangle are illustrated in Figure 10.

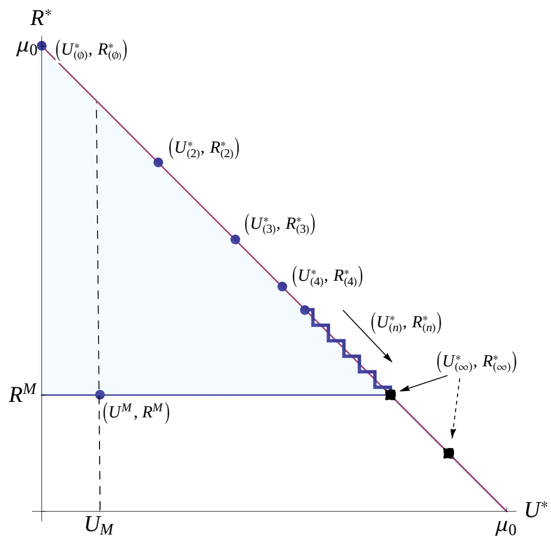


Figure 10: Surplus Triangle

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 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 3 and Example 4, in which the seller's expected revenue is strictly below the monopoly revenue.

7 Discussion and Conclusion

7.1 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 paper, 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 paper highlights the role of rational inattentiveness of the consumer in such a setting. By committing to a coarse information structure, 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. 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[16] is central to the analysis in this paper. 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

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

of the consumer. The effects established in this paper 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 paper, 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 total surplus 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 5. 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.

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

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 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.

7.2 Concluding Remarks

The objective of this paper 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 paper 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 remaining partially uninformed, and to only obtain a coarse perception of his true valuation, 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 consumer-optimal 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 paper, 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. Suppose that the consumer chooses the n-signal information structure π, 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_π. If the seller charges a price $p \in (V_k, V_{k+1}]$ then by (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

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

$\square$

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

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

For a given information structure that induces the distribution F_π 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 (Seller-Indifference). Suppose that the consumer-optimal information structure π 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.[17] Hence, w.l.o.g. we can assume that $\overline{v}_j \in \mathcal{V}_j$, that is, $\overline{v}_j = \sup \mathcal{V}_j = \max \mathcal{V}_j$.

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

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 π:

Construction 1.

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

$$\begin{aligned} &\delta = (\delta_i, \dots, \delta_n) \ \text{ with } \delta_j > 0\ \forall\, j = 1, \dots, n, \ \text{ such that} \\ &\big[\underline{v}_j,\, \underline{v}_j + \delta_j\big) \subseteq \big(\mathcal{V}_j \cup \big[\underline{v}_{j+1},\, \underline{v}_{j+1} + \delta_{j+1}\big]\big)\ \forall\, j > k, \text{ and} \\ &(\overline{v}_j - \delta_j,\, \overline{v}_j] \subseteq (\mathcal{V}_j \cup (\overline{v}_{j-1} - \delta_{j-1},\, \overline{v}_{j-1}])\ \forall\, i \le j < k, \end{aligned} \tag{6}$$

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:

$$\begin{aligned} &\text{For every } j > k: \\ &\widetilde{g}(s_j|v) = \begin{cases} g(s_j|v) + g(s_{j+1}|v) & \text{for } v \in \big[\underline{v}_{j+1},\, \underline{v}_{j+1} + \delta_{j+1}\big) \\ 0 & \text{for } v \in \big[\underline{v}_j, \underline{v}_j + \delta_j\big) \\ g(s_j|v) & \text{otherwise,} \end{cases} \\ &\text{For every } j, \ \text{s.t. } i \le j < k: \\ &\widetilde{g}(s_j|v) = \begin{cases} 0 & \text{for } v \in (\overline{v}_j - \delta_j,\, \overline{v}_j] \\ g(s_j|v) + g(s_{j-1}|v) & \text{for } v \in (\overline{v}_{j-1} - \delta_{j-1},\, \overline{v}_{j-1}] \\ g(s_j|v) & \text{otherwise,} \end{cases} \\ &\widetilde{g}(s_k|v) = \begin{cases} g(s_k|v) + g(s_{k+1}|v) & \text{for } v \in \big[\underline{v}_{k+1},\, \underline{v}_{k+1} + \delta_{k+1}\big] \\ g(s_k|v) + g(s_{k-1}|v) & \text{for } v \in (\overline{v}_{k-1} - \delta_{k-1},\, \overline{v}_{k-1}] \\ g(s_k|v) & \text{otherwise,} \end{cases} \ \text{and} \\ &\widetilde{g}(s_j|v) = g(s_j|v) \quad \forall\, j < i. \end{aligned} \tag{7}$$

For every $v \in [0, 1]$, this construction satisfies $\widetilde{g}(s_j|v) \ge 0$ for all $s_j \in S$, as well as $\sum_{j=1}^{n} \widetilde{g}(s_j|v) = 1$. Hence, $\widetilde{\pi}$ is well-defined as an information structure. This construction is illustrated in Figure 11.

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

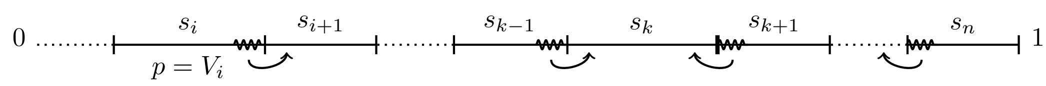


Figure 11: Illustration of Construction 1.

that 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 δ that satisfies (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 δ that satisfies (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 π, 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, $\sum_{j=i}^{n} \widetilde{g}(s_j) = \sum_{j=1}^{n} g(s_j)$, as well as the expected value of types that participate in trade and hence 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 $V_j < \widetilde{V}_i$.

But in this case the consumer is obviously better off than under information structure π.

This contradicts the assumption that there exists a finite consumer-optimal information structure π that does not satisfy the seller-indifference property. $\square$

Proof of Theorem 1: First, two lemmas are established, which are then combined in order to prove Theorem 1.

Lemma 3 (Small Exclusion Region)**.**
Any consumer-optimal information structure π^* 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 3. Suppose that π^* is an optimal information structure. Let $\mathcal{V} := \operatorname{supp}(F_{\pi^*})$ be the support of value estimates that are induced by this information structure. The seller will charge a price equal to an element of $\mathcal{V}$ (Lemma 1), say $p^* = V_p$. Let $\mathcal{V}_E \subseteq \mathcal{V}$ be the subset of types 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. 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 $V < V'$. Moreover, the seller charges a price $p^* = V_p$, 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 form π 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

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

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): In this case, switching from π to $\widetilde{\pi}$ has no effect on the expected realized surplus, the expected revenue of the seller and the expected surplus of the consumer. Hence, π and $\widetilde{\pi}$ are *outcome-equivalent*.

Case (2): Under the new information structure $\widetilde{\pi}$, the seller is induced to charge a price $p = \overline{V} < V' < V_p$. That is, switching from π to $\widetilde{\pi}$, reduces the price charged by the seller. Notice that Construction 2 does not affect the value estimates induced by signal realizations that yield information rents for the consumer under π. For these value estimates, neither their valuation nor the probability that they arise changes under Construction 2. However, by Construction 2 the price charged by the seller reduces and hence the expected information rents increase. Moreover, the probability for the consumer to obtain information rents increases, since now the type V_p is greater than the price $\widetilde{p} = \overline{V}$ that is charged by the seller. Therefore, if the value estimate V_p realizes, the consumer obtains an information rent. Consequently, the consumer is strictly better off under information structure $\widetilde{\pi}$ than under π – a contradiction to the assumption that π is a consumer-optimal information structure. □

Lemma 4 (Implementability by Fully Inclusive Prices)**.**
Suppose that the price p_0 is inducible by the n-signal information structure π, and U_π 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_π for the consumer.

Proof of Lemma 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 π.

Suppose now that the exclusion region is non-empty $\mathcal{V}_E \neq \emptyset$. By Lemma 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 thus 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

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

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 π is a consumer-optimal information structure.

Case (2): Suppose that $R(V_2) > R(V_1)$. For $p_0 = V_2$, expected total surplus from trade is[18]

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

The expected revenue of the seller is:

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

and the expected surplus of the consumer is:

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

It follows that

$$\begin{aligned} p =& \mu_0 - [(1 - g_1)\mathbb{E}[v|s\neg s_1] - (1 - g_1)p_0] \\ =& g_1 V_1 + (1 - g_1)p_0. \end{aligned} \tag{8}$$

Claim 2. The price p, given by (8), can be implemented by an n-signal information structure as a fully inclusive price.

Construct a new information structure $\widetilde{\pi}$ from π 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.[19]

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,

[18] 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.

[19] To be precise, it is of course only possible to reach prices $p_B \leq \mu_0$ with this construction. However, it is never an optimal action for the consumer to induce a price greater than μ_0. If the consumer chooses a purely noisy, that is uninformative, information structure, then the seller would set a price equal to μ_0. The consumer would always buy but obtain zero surplus – any information structure that induces an inclusive price makes the consumer better off.

by Proposition 1, there exists another information structure that makes the consumer better off. $\square$

Proof of Theorem 1 (Efficient Trade).
Suppose not, that is, suppose that for some $n \in \mathbb{N}$, there exists a consumer-optimal 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 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 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 π. Recall that the construction in Lemma 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 implies that there exist an n-signal information structure that yields a strictly higher surplus for the consumer than the surplus induced by $\widetilde{\pi}$, which is also higher than the surplus induced by π. A contradiction. $\square$

Proof of Theorem 2 (Optimality of Coarse Perception).
The result is established by an indirect proof that is again constructive. The argument of the proof proceeds as follows: Suppose that there exists some $n \in \mathbb{N}$ and a consumer-optimal information structure $\pi^*_{(n)}$ that is not partitional. I construct an n-signal information structure that induces a strictly higher surplus for the consumer, contradicting the assumption that the initial non-partitional information structure is consumer-optimal.

Suppose that π is a non-partitional consumer-optimal 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

$$\hat{g}(s_k|v) = \begin{cases} g(s_k|v) + g(s_j|v) & \text{for } v \in \mathcal{V}_k \\ 0 & \text{otherwise,} \end{cases} \tag{9}$$

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

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

This construction adds "missing mass" to the values $v \in \widehat{\mathcal{V}}_k$.[20] The information structure $\widehat{\pi}$ is obtained from π by shifting masses across signal realizations. Hence, by construction, $\widehat{\pi}$ is well-defined as an information structure.

For the construction in (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 $v \in [0,\, 1]$.

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

$$\begin{aligned} &\delta = (\delta_{j+1}, \ldots, \delta_k, 0) \ \text{ with } \delta_l > 0 \ \text{ such that} \\ &[\underline{v}_l,\, \underline{v}_l + \delta_l] \subseteq \big(\mathcal{V}_l \cup \big[\underline{v}_{l+1},\, \underline{v}_{l+1} + \delta_{l+1}\big]\big) \ \forall\, l \in \{j+1, \ldots, k\}, \end{aligned} \tag{10}$$

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

$$\widetilde{g}(s_k|v) = \begin{cases} 0 & \text{for } v \in [\underline{v}_k,\, \underline{v}_k + \delta_k) \\ \hat{g}(s_k|v) & \text{otherwise,} \end{cases} \tag{11}$$

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

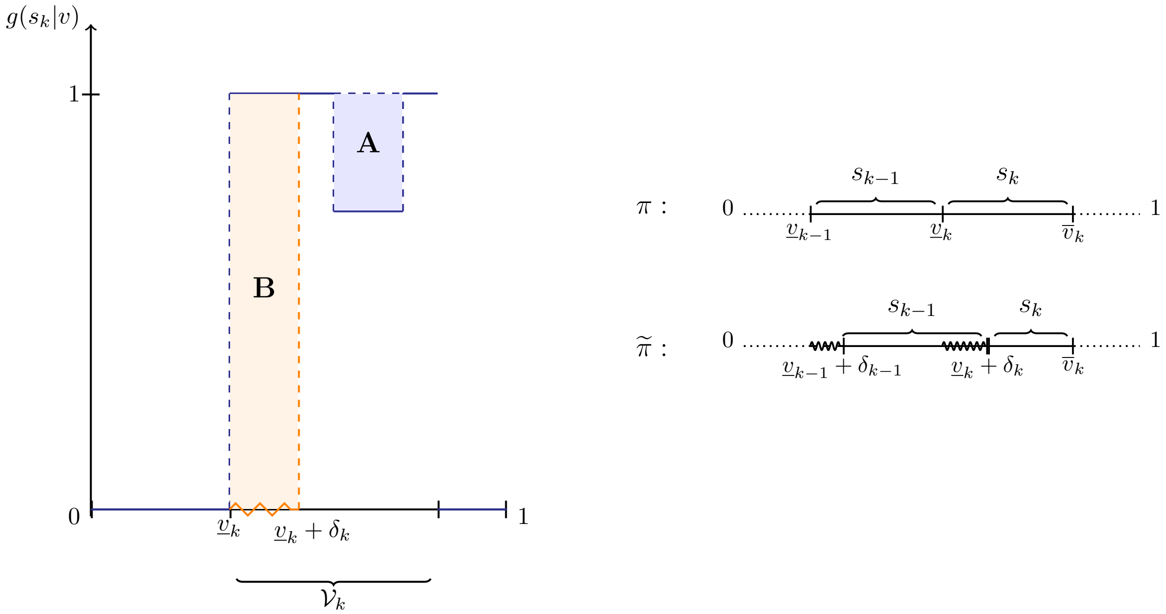
$$\widetilde{g}(s_j|v) = \begin{cases} \hat{g}(s_j|v) + \hat{g}(s_{j+1}|v) & \text{for } v \in \big[\underline{v}_{j+1} + \delta_{j+1},\, \underline{v}_{j+1}\big) \\ \hat{g}(s_j|v) & \text{otherwise,} \end{cases}$$

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

This construction in illustrated in Figure 12.

For every s_l with $l \in \{s_{j+1}, \ldots, s_k\}$, Construction 3 increases the probability that signal s_l realizes for some high values, which increases the revenue that the seller could extract by charging the price $\mathbb{E}\,[v|s_l]$. Construction 3 then takes probability mass off values in the

[20] 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.



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

Figure 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 $F(\cdot|s_k)$.

lower part of the support $\mathcal{V}_l$, which again reduces the revenue that the seller could extract by charging the price $\mathbb{E}\,[v|s_l]$.

For a continuous distribution function, it is possible to find a δ that satisfies (10) such that:

$$R(\widetilde{V}_l) = R\,(V_l) \quad \forall l \in \{j+1, \dots, k\}. \tag{12}$$

Claim 3. For any $l \in \{j+1, \dots, k\}$ such that $\widetilde{g}_l \leq g_l$, the distribution of values in the support of $\widetilde{F}(\cdot|s_l)$ first-order stochastically dominates those in $F(\cdot|s_l)$.

Indeed,

$$\widetilde{f}(v|s_l) = \begin{cases} f(v|s_l) = 0 & \forall\, v \in [0,\, \underline{v}_l) \\ 0 < f(v|s_l) & \forall\, v \in [\underline{v}_l,\, \underline{v}_l + \delta_l) \\ \frac{\widetilde{g}(s_l|v)f(v)}{\widetilde{g}(s_l)} \geq \frac{g(s_l|v)f(v)}{g(s_l)} = f(v|s_l) & \forall\, v \in [\underline{v}_l + \delta_l, 1]\,. \end{cases}$$

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

Claim 4. Construction 3 with δ such that (12) is satisfied yields:

(i) $\sum_{m=l}^{n} \widetilde{g}_m < \sum_{m=l}^{n} g_m$,

(ii) $\widetilde{V}_l > V_l$, and
(iii) $\sum_{m=l}^{n} \widetilde{g}_m \widetilde{V}_m \geq \sum_{m=l}^{n} g_m V_m$,
for all $l \in \{j+1, \ldots, k\}$.

For $\widetilde{g}_k$ this is easy to see. Suppose that $\widetilde{g}_k = g_k$. Then, by Claim 3, $\widetilde{F}(\cdot|s_k) \geq_{FOSD} F(\cdot|s_k)$, which would imply:

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

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 δ_k. It follows that $\widetilde{g}_k < g_k$ and (12) implies $\widetilde{V}_k > V_k$.

The statements of (i) and (ii) of Claim 4 for $l \in \{j+1, \ldots, k-1\}$ follow by induction. They are verified by applying the same arguments as for $\widetilde{g}_k$ and $\widetilde{V}_k$ to $\sum_{m=l}^{n} \widetilde{g}_m$ and $\widetilde{V}_l$ using that by (12)

$$\left(\sum_{m=l}^{n} \widetilde{g}_m\right) \widetilde{V}_l = \left(\sum_{m=l}^{n} g_m\right) V_l.$$

Part (iii) of Claim 4 is still left to show. Notice that the value of $\sum_{m=l}^{n} g_m V_m$ is just the expected value of the values in the supports of $F(\cdot|s_m)$ for signals $s_l, \ldots, s_n$ (cf. Figure 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 (12).

Claim 5. Construction 3 with δ such that (12) is satisfied implies $R(\widetilde{V}_j) < R(V_j)$.

The construction only re-distributes mass among signals s_l with $l \geq j$. 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 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_l with $l \geq j$. Hence, it must hold that:

$$\widetilde{g}_j \widetilde{V}_j + \left(\sum_{k=j+1}^{n} \widetilde{g}_k \widetilde{V}_k \right) = g_j V_j + \left(\sum_{k=j+1}^{n} g_k V_k \right)$$

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

It follows that

$$R(\widetilde{V}_j) = \widetilde{V}_j \left(\sum_{l=j}^{n} \widetilde{g}_l \right) < V_j \left(\sum_{l=j}^{n} g_l \right) = R(V_j),$$

which verifies Claim 5.

The case $j > 1$:
Under information structure $\widetilde{\pi}$, $R(\widetilde{V}_j) < R(V_j)$ and $R(\widetilde{V}_l) = R(V_l)$ for all $l \neq j$. The initial information structure was assumed to be consumer-optimal and hence, by Theorem 1 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, there exists another information structure that yields a strictly higher expected surplus for the consumer. A contradiction to the assumption that π is a non-partitional consumer-optimal 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 $R(\widetilde{V}_l) > R(\widetilde{V}_1)$ for all $l > 1$.

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 $R(\check{V}_1) = R(\widetilde{V}_l)$, $l > 2$, 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 there exists an information structure that yields a strictly higher expected surplus for the consumer, which contradicts the assumption that π is consumer optimal. $\square$

Proof of Theorem 3.
Consider any $n \in \mathbb{N}$. From Theorem 1 it follows that if a consumer-optimal 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 and Theorem 2. $p^* = \min P^*_{(n)}$.

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 consumer-optimal information structure must be partitional (Theorem 2). Hence, the lowest threshold $\hat{v}_1$ is determined by:

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

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

By Proposition 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 Theorem 2, the threshold $\hat{v}_2$ is determined by

$$\mathbb{E}\left[v | \hat{v}_1 \leq v \leq \hat{v}_2\right],$$

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 $V_n := \mathbb{E}\left[v | v \geq \hat{v}_{n-1}\right]$, and $g_n := 1 - F(\hat{v}_{n-1})$.

The above construction is well-defined for every $p \in [0, 1]$, given that 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 for every $i = 1, \ldots, n$, $V_i(p)$ is strictly increasing in p. Moreover, since F is twice continuously differentiable, $V(p)$ is continuous 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, $p = \min P^*_{(n)}$, and the partitional information structure that was constructed in the above construction is the optimal n-signal information structure $\pi^*_{(n)}$.

□

Proof of Proposition 2 (More Signals are Better).
The proof is by induction. For $n \in \mathbb{N}$, let $\pi^*_{[n]}$ be the consumer-optimal 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 consumer-optimal 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 consumer-optimal 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 consumer-optimal n-signal information structure $\pi^*_{[n]}$,and let $V_1^{(n)}, \dots V_n^{(n)}$ 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, the consumer-optimal information structure $\pi^*_{[n+1]}$ is an element of the set $\mathcal{S}^{(n+1)}$, and yields a strictly higher expected surplus than $\pi^*_{[n]}$ for the consumer, $U\left(\pi^*_{[n+1]}\right) > U\left(\pi^*_{[n]}\right)$.

Let $\widehat{\pi}$ be an information structure with $n+1$ signal realizations $\{s_1, \dots, s_{n-1}, \hat{s}_n, \hat{s}_{n+1}\}$ that is constructed 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, let $g(\hat{s}_n|v) + g(\hat{s}_{n+1}|v) = g(s_n|v)$ for all $v \in [0,\, 1]$. The construction can be chosen, such that[21]

$$V_{n-1} < \widehat{V}_n < V_n < \widehat{V}_{n+1}, \tag{13}$$

and $R(V_n) > R(\widehat{V}_{n+1})$. If mass from V_n is split such that (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 there exists an information structure with at most $(n+1)$-signals that yields a strictly higher surplus for the consumer. Consequently, the consumer-optimal information structure $\pi^*_{[n+1]}$ yields a strictly higher surplus than the expected surplus that the consumer would obtain under $\pi^*_{[n]}$. This implies that the surplus $U(\pi^*_{[n+1]})$ is not achievable

[21] Notice that $\pi^*_{[n]}$ is a consumer-optimal 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.

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 consumer-optimal information structure induces efficient trade (Theorem 1). This implies $U^*_n = \mu_0 - p^*_n$, and it follows that $p^*_{n+1} < p^*_n$. $\square$

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

Definition 1: Say that the sequence of finite consumer-optimal 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 4 (Optimal Information Structure).

Claim 6. The sequence of finite consumer-optimal 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 $F^*_{(n_k)} \Rightarrow \widetilde{F}$.

Every finite consumer-optimal information structure $\pi^*_{(n)}$ induces the seller to charge a fully inclusive price. Hence, the sequence $\{\pi^*_{(n)}\}_{n\in\mathbb{N}}$ induces a sequence of prices $\{p^*_{(n)}\}_{n\in\mathbb{N}}$. By Proposition 2, this sequence of prices is strictly decreasing in n. Moreover, every finite consumer-optimal information structure is partitional and satisfies the seller-indifference property (Theorem 2 and Proposition 1). Hence, the thresholds of the interval partition that characterizes a finite consumer-optimal 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, $F^*_{(n)} \Rightarrow \widetilde{F}$. This establishes convergence of the sequence of information structures, $\{\pi^*_{(n)}\}_{n\in\mathbb{N}}$. Define $\widetilde{\pi}^*_{(\infty)} := \lim_{n\to\infty} \pi^*_{(n)}$.

Convergence in the Levy-Prokhorov metric implies convergence of prices, that is, $p^*_n \xrightarrow{n\to\infty} p^*_{(\infty)}$. Every cconsumer-optimal finite information structures satisfies the seller-indifference condition (Proposition 1), which implies

$$p^*_{(n)} = \mathbb{E}\left[v|v \leq \hat{v}_{1,(n)}\right] = \mathbb{E}\left[v|\hat{v}_{1,(n)} \leq v \leq \hat{v}_{2,(n)}\right]\left(1 - F(\hat{v}_{1,(n)})\right), \tag{14}$$

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 $\pi^*_{(n)}$. For a given $\hat{v}_{1,(n)}$, the right-hand side of (14) is increasing in $\hat{v}_{2,(n)}$. This implies that $\lim_{n\to\infty} \hat{v}_{2,(n)} = \hat{v}_{1,(n)}$, and

$$\lim_{n\to\infty} \mathbb{E}\left[v|\hat{v}_{1,(n)} \leq v \leq \hat{v}_{2,(n)}\right]\left(1 - F(\hat{v}_{1,(n)})\right) = \hat{v}_{1,(n)}\left(1 - F(\hat{v}_{1,(n)})\right).$$

It follows that, for the information structure $\widetilde{\pi}^*_{(\infty)}$, the lowest threshold $\underline{v}$ is given by

$$\underline{v} := \min\{\hat{v} \in (0, 1] : (1 - F(\hat{v}))\,\hat{v} = \mathbb{E}\left[v|v \leq \hat{v}\right]\}. \tag{15}$$

The property that there exist a sweeping up region, directly follows from the feature that all finite consumer-optimal information structures satisfy seller-indifference (Proposition 1), and that this property is preserved in the limit.

It is still left to verify that $\pi^*_{(\infty)}$ is indeed an consumer-optimal 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 a consumer-optimal information structure. □

Proof of Corollary 1. It was already established in Proposition 2, that the consumer's expected equilibrium surplus $U^*_{(n)}$ is strictly increasing in the number of signal realizations n. Given that every consumer-optimal information structure induces efficient trade (Theorem 1), 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 4, weak convergence of the sequence of consumer-optimal finite information structures has been established. A direct implication is

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

□

B Evolution of Information Structures and Thresholds with n

Reconsider the uniform prior example. The aim of this section is to identify the thresholds of the consumer-optimal finite information structures and analyze how the optimal information structures evolve with the number of signals.

Any consumer-optimal finite information structure is partitional (Theorem 2). 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 $0 = a_0 < a_1 < \dots < a_n = 1$. 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 consumer-optimal 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 consumer-optimal n-signal information structure $\pi^*_{(n)}$, if the thresholds satisfy the following system of equations:

$$\begin{aligned} a_k =& \frac{a_1}{(1 - a_{k-1})} - a_{k-1}, \text{ for } \ k = 1, \dots, n \\ a_0 =& 0 \ \text{ and } \ a_n = 1. \end{aligned} \tag{16}$$

The thresholds of the consumer-optimal information structure for $n \in \{2, 5, 10\}$ are illustrated in Figure 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, $\lim_{n\to\infty} a_n = \frac{1}{2}\sqrt{2}$.

C Special classes of information structures

In order to fix ideas, I briefly discuss three examples of special classes of information structures, *partitional*, *noisy*, and *finite information structures*.

Example 5 (Partitional Information Structures)**:**
A class of information structures that will be important in our analysis is the class of *partitional information structures*. An information structure $\pi = \big(S,\, \{G(\cdot|v)\}_{v\in[0,\,1]}\big)$ is *partitional* if there exists a function $\overline{s}:\ [0,\,1] \to S$ such that for all $s \in S$ and $v \in [0,\,1]$

$$G(s|v) = \begin{cases} 0 & \text{if } s < \overline{s}(v) \\ 1 & \text{if } s \geq \overline{s}(v). \end{cases} \tag{17}$$

An information structure is *monotone partitional* if Equation 17 is satisfied for an increasing function $\overline{s} : \mathcal{V} \to S$.

Examples:

1. For every $\hat{v} \in (0,\,1)$, the information structure $\pi_{\hat{v}}$ for which signal s_1 realizes if $v \in [0, \hat{v}]$ and s_2 realizes if $v \in [\hat{v},\,1]$, is a monotone partitional information structure.
2. Given $0 < \hat{v}_1 < \hat{v}_2 < 1$, the information structure for which signal s_1 realizes if $v \in [0, \hat{v}_1]$ or $v \in [\hat{v}_2,\,1]$, and s_2 realizes if $v \in [\hat{v}_1,\,\hat{v}_2]$, is partitional but not monotone.

Partitional information structures only provide a *coarse perception* of the true state.

Example 6 (Noisy Information Structures)**:**
A *noisy information structure* provides only a noisy signal of the true state. That is, for a noisy information structure, the random variable that represents the signal is equal to the state variable plus a random noise term, $S = V + \epsilon$. A standard example is normally distributed noise, $\epsilon \sim \mathcal{N}\,(0,\,\sigma^2)$.

Example 7 (Finite Information Structures)**:**
An information structure is *finite*, if it only has a finite number of signal realizations. That is, $\pi = (S,\,\{G(\cdot|v)\}_{v\in\mathcal{V}})$ is finite if $|S| = n < \infty$.

References

Aumann, Robert J. and Michael B. Maschler (1995), *Repeated Games with Incomplete Information.* MIT Press, Cambridge, MA.

Bergemann, Dirk, Benjamin A. Brooks, and Stephen Morris (2014), "The Limits of Price Discrimination." American Economic Review, *forthoming.*

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.

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.

Caplin, Andrew and Mark Dean (2013), "The Behavioral Implications of Rational Inattention with Shannon Entropy." Working Paper.

Clippel, Geoffroy de, Kfir Eliaz, and Kareen Rozen (2014), "Competing for Consumer Inattention." Journal of Political Economy, *forthcoming*.

Crawford, Vincent P. and Joel Sobel (1982), "Strategic Information Transmission." *Econometrica*, 50, 1431–1451.

DellaVigna, Stefano (2009), "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature*, 47, 315–72.

Ely, Jeffrey, Alexander Frankel, and Emir Kamenica (2014), "Suspense and Surprise." Journal of Political Economy, *forthcoming*.

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*.

Gabaix, Xavier and David Laibson (2003), "A New Challenge for Economics: The Frame Problem." In *Essays in Psychology and Economics*, Oxford University Press.

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.

Gentzkow, Matthew and Emir Kamenica (2014), "Costly Persuasion." *American Economic Review, Papers & Proceedings*, 104, 457–62.

Gul, Faruk, Wolfgang Pesendorfer, and Tomasz Strzalecki (2014), "Coarse Competitive Equilibrium and Extreme Prices." Working Paper.

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.

Matejka, Filip (2012), "Rigid Pricing and Rationally Inattentive Consumer." Working paper.

Myerson, Roger B. (1997), *Game Theory*. Harvard University Press.

Riley, J. G. and R. Zeckhauser (1983), "Optimal Selling Strategies: When to Haggle, When to Hold Firm." *Quarterly Journal of Economics*, 98, 267–290.

Shah, Anuj K. and Daniel M. Oppenheimer (2008), "Heuristics Made Easy: An Effort-Reduction Framework." *Psychological Bulletin*, 207–222.

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)*.

Wilson, Andrea (2014), "Bounded Memory and Biases in Information Processing." Econometrica, *forthcoming*.

Woodford, Michael (2012), "Inattentive Valuation and Reference-Dependent Choice." Working paper.

Yang, Ming (2013), "Optimality of Debt under Flexible Information Acquisition." Working paper.

Yang, Ming (2014), "Coordination with Flexible Information Acquisition." Journal of Economic Theory, *forthcoming*.

Yee, Michael, Ely Dahan, John R. Hauser, and James Orlin (2007), "Greedoid-based non-compensatory inference." *Marketing Science*, 26, 532–549.