UNIVERSITY OF CAMBRIDGE
Faculty of Economics

Institute for
New Economic Thinking

Cambridge-INET Institute

Cambridge-INET Working Paper Series No: 2019/19

Cambridge Working Papers in Economics: 1982

CHECKING CHEAP TALK

Ian Ball
(Yale University)

Xin Gao
(University of Cambridge)

We study a sender-receiver game. The sender observes the state and costlessly transmits a message to the receiver, who selects one component of the state to check and then chooses a binary action. The receiver's preferred action depends on the state. The sender has a state-independent preference for one action over the other. Nevertheless, communication can strictly benefit both players. We characterize the symmetric equilibria. In each one, the sender tells the receiver which components of the state are highest. The same equilibria exist in an extension where the receiver can check multiple components. We also find that with commitment power, the sender can extract more rents from the receiver by randomizing between signals that induce different actions. However, the receiver's ability to partially verify the state has an ambiguous effect on the sender's utility unless the sender has commitment power, in which case verification can only restrict the set of posteriors the sender can induce.

Checking Cheap Talk*

Ian Ball[†] Xin Gao[‡]

September 11, 2019

Abstract

We study a sender-receiver game. The sender observes the state and costlessly transmits a message to the receiver, who selects one component of the state to check and then chooses a binary action. The receiver's preferred action depends on the state. The sender has a state-independent preference for one action over the other. Nevertheless, communication can strictly benefit both players. We characterize the symmetric equilibria. In each one, the sender tells the receiver which components of the state are highest. The same equilibria exist in an extension where the receiver can check multiple components. We also find that with commitment power, the sender can extract more rents from the receiver by randomizing between signals that induce different actions. However, the receiver's ability to partially verify the state has an ambiguous effect on the sender's utility unless the sender has commitment power, in which case verification can only restrict the set of posteriors the sender can induce.

Keywords: cheap talk, partial verification, Bayesian persuasion
JEL Classification Codes: D82, D83

---

*We are grateful to Dirk Bergemann, V. Bhaskar, Tilman Börgers, Gabriel Carroll, Laura Doval, Mira Frick, Hans Peter Grüner, Takakazu Honryo, Ryota Iijima, Navin Kartik, Dmitriy Knyazev, Volker Nocke, Martin Peitz, Larry Samuelson, Johannes Schneider, Nicolas Schutz, Andrzej Skrzypacz, Konrad Stahl, Takuo Sugaya, Emanuele Tarantino, Juuso Välimäki as well as the audiences at Yale and Mannheim for insightful comments and discussions. Xin Gao is deeply indebted to his advisor, Thomas Tröger, for his continuous support and guidance. Part of the work was conducted during Xin Gao's visit to the Department of Economics at Yale University, for which he is thankful.

[†]Department of Economics, Yale University, 28 Hillhouse Avenue, New Haven, CT 06511, USA. Email: ian.ball@yale.edu

[‡]Cambridge INET-Institute, Faculty of Economics, University of Cambridge, Sidgwick Avenue, Cambridge CB3 9DD, United Kingdom. Email: xg264@cam.ac.uk

Contents

1 Introduction

Less informed agents often turn to biased experts for guidance. Even when the expert's private information is important for the agent's preferences, the expert's own preferences may be independent of this private information if the bias is sufficiently strong. For example, a salesperson wants a shopper to buy her product, no matter its true quality; a prosecutor wants a judge to convict the defendant, without regard to the defendant's actual guilt or innocence; a politician wants a voter's support, regardless of whether her proposal would actually benefit the voter. This level of bias makes credible cheap talk communication impossible. Indeed, the sender will send whichever message that induces the highest probability of the receiver taking the sender-preferred action.

However, in practice, the receiver can often gather additional information after receiving the sender's message. For instance, the shopper can inspect some attributes of the product, the judge can check some evidence submitted by the prosecutor, and the voter can do some research on the proposal. Then a natural question to ask is whether and why does an agent still solicit information from an expert who is known to be biased if the agent can *partially* observe the state of the world? In other words, (why) does an agent bother to listen to the biased expert for soft information (cheap talk messages) if the agent can get access to some hard information (through verification)?

In this paper, we show that with strategic partial verification, cheap talk messages by an extremely biased sender can be credible and strictly benefit both the sender and the receiver. The vital channel of influence is that the sender's message can influence *which* information the receiver seeks to acquire. The salesperson can point the shopper to the best attributes of the good; the prosecutor can guide the judge to inspect the strongest evidence against the defendant; the politician can highlight the merits of her proposal to the public.

It is well known in economic theory that if the sender is too biased, no information can be credibly communicated. To facilitate informative communication, our novel idea is to use partial verification as a remedy to the extreme conflict of interests between the two parties. In other words, the verification serves as a substitute for the missing preference alignment so that with the help of hard information, soft information from the sender becomes credible and strictly benefits both the sender and the receiver.

We obtain this result in a simple sender-receiver game. The receiver chooses between two actions and the sender has a strict, state-independent preference for one action over the other. The state is a vector of N binary attributes (i.e., either good or bad), drawn from a symmetric common prior. The sender observes the state and then sends a message to the receiver. After seeing the message, the receiver chooses some (but not all) components of the state to costlessly verify and then decides whether to buy the product at an exogenous price or not. We interpret the restriction on verification as a time or cognitive constraint, as in Glazer and Rubinstein (2004). For instance, a shopper may not bother to become an expert on all the tech specs of a smartphone; the judge may have to complete the case in a certain amount of time; the voter may only glance at a news article before voting.

The main insights can be captured by the case where the receiver can only check one attribute, which is our baseline model in Section 2 and 3. We construct a natural family of symmetric equilibria, called *top equilibria*. Under a top-k equilibrium, the sender points to the k best components of the state, without indicating their relative positions. Ties are broken by uniform randomization. The receiver then uniformly selects one of these recommended attributes to check, and buys if and only if that attribute is good. For any $1 \leq k \leq N$, we provide the sufficient and necessary condition for the existence of the top-k equilibrium, which requires the price below an upper bound. In Section 4, we extend the setting to the case where the receiver can check multiple attributes and we show that the same equilibria exist. Naturally, "checking one attribute" is simply a special case of "checking multiple attributes." We discuss other equilibria in Appendix D.

This family of equilibria has an intuitive structure. The probability of purchase is strictly decreasing in k because for smaller k, the relative quality of the checked attributes is higher and the probability of good attributes being checked is decreasing as the sender points to more and more attributes. However, the range of prices at which the equilibria can be sustained is increasing in k. For larger k, the checked attributes are more representative of the overall quality, which makes the receiver more optimistic and therefore more willing to pay a higher price. Moreover, the expected utility of the receiver upon buying is also increasing in k. As k increases, seeing a good attribute becomes a rarer but stronger signal about the quality of the product. This features an interesting trade-off between how often a good signal is observed and how strong that signal can be.

Since the sender tries to persuade an initially uninformed receiver, we compare the sender's utility under these equilibria with her utility under alternative communication structures. Specifically, we consider two information structures for the receiver—no verification and partial verification—and three communication protocols for the sender—no communication, cheap talk, and Bayesian persuasion. This gives six pairs of an information structure and a communication protocol.

We find that, even though the sender has state-independent preferences, for a range of prices, the sender strictly benefits from the ability to communicate. To illustrate this, suppose that without any information the receiver will not buy. With state-independent preferences, the sender's only objective is to maximize the probability of buying. If the receiver can randomly pick one attribute to check without talking to the sender, he will buy *only if* he sees a good attribute, which may increase the probability of buying to some extent. However, in our top-1 equilibrium, in which the receiver checks the recommended attribute after receiving the sender's message, the probability of buying significantly increases to a higher level so that all the non-zero types can sell the good with probability one. Intuitively, under partial verification without communication, seeing a good signal is purely random. But if the sender can communicate with the receiver (although the message is cheap talk), she can guide the receiver to the good attributes. Hence, we can think of the cheap talk message as a belief-coordinating device that guides the receiver to the good signals and this is incentive compatible for the sender due to the verification by the receiver.

We also find that if the sender has commitment power, as in Kamenica and Gentzkow (2011), then she can further increase her utility by committing to randomize between various messages that induce different buying probabilities. This essentially smooths the discreteness in the equilibria arising from the sender's incentive constraints. In light of the Bayesian persuasion literature, it is not surprising that the sender can benefit from commitment. However, we show that cheap talk can do as well as Bayesian persuasion in some cases. This happens with prices such that the receiver is exactly indifferent between checking the recommended attributes and checking the unrecommended ones. Since the commitment power substantially enlarges the sender's set of communication strategies, if the receiver strictly prefers to follow the sender's recommendation when the message is cheap talk, then with commitment the sender can gradually adjust the signal structure to increase her utility until the receiver is indifferent between obedience and disobedience. However, if the receiver is already indifferent when the sender cannot commit, then the sender can only replicate the cheap talk equilibrium when she can commit.

Finally, we observe that the receiver's ability to partially verify the state has an ambiguous effect on the sender's utility unless the sender has commitment power, in which case verification can only restrict the set of posteriors the sender can induce. If the sender cannot commit so that the message is cheap talk, depending on the prior, the sender may or may not benefit from verification. Specifically, when the receiver decides to buy with his prior, then verification hurts the sender since the receiver will buy with probability less than one, as indicated in our top equilibria. However, if the receiver does not buy without any information a priori, then the sender benefits from verification since seeing a good signal makes the receiver more optimistic about the product. On the other hand, if the sender can commit to a signal structure, verification can only hurt the sender since it imposes restrictions on the sender's strategy set, whereas without verification the sender has full control over the receiver's information structure by Bayesian persuasion.

Related Literature

Crawford and Sobel (1982) introduced cheap talk games where the sender's messages are costless, unrestricted, and unverifiable. They showed that informative equilibria exist as long as the sender is not too biased. In many settings of interest, however, the sender prefers that the receiver take a particular action, no matter the state. With this level of bias, the sender cannot use cheap talk messages to credibly communicate information about the state, and hence the unique equilibrium is "babbling." If various assumptions about the cheap talk setting are relaxed, then there can be an informative equilibrium. For instance, messages can be credible if their cost depends on the state, as in Spence's (1973) classical signaling model or, more recently, in Kartik's (2009) model of cheap talk with lying costs.

Another approach, commonly referred to as "persuasion games" beginning with Milgrom (1981) and Grossman (1981), is to restrict the sender's strategy set to allow for information to be concealed but not misreported. This can be thought of as a reduced form approach to incorporate verification or evidence.

We take a different approach. The sender's strategy set is unrestricted and there is no exogenous lying cost. Instead, the receiver, after seeing the sender's message, can verify part (but not all) of the state. Instead of exogenously restricting the sender's messages, the receiver's actions, through verification, discipline the messages that the sender chooses to send in equilibrium, much in the spirit of Crawford and Sobel (1982). Indeed, one interpretation of verification is that it adds another dimension to the receiver's action space in a way that preferences are sufficiently aligned to support an informative equilibrium.

The advantage, relative to persuasion games, is that we can study the receiver's strategic decision about which information to verify. In each of the top-k equilibria, the binding deviation often involves checking one of the attributes that is not recommended. This captures an important strategic consideration that is absent in the models of persuasion. If the relevant attributes for the receiver's decision are the highest attributes, then an equilibrium can be sustained. Depending on the parameters of the environment, in other cases, the receiver would rather make his decision on the basis of the worst attributes. This introduces a new role for cheap talk. It is not about telling the receiver whether the product is of high quality, but rather *which* attributes are of high quality, so that the receiver can make his decision on the basis of those attributes.

There have been numerous extensions of cheap talk in various dimensions. The most relevant strand is the sequence of papers on multi-dimensional cheap talk (e.g., Battaglini, 2002; Levy and Razin, 2007; Ambrus and Takahashi, 2008). It is worth stressing the comparison with Chakraborty and Harbaugh (2007, 2010). What is important here is to distinguish the information structure from the preference structure. Our paper features a multi-dimensional information structure, but the preference structure is one-dimensional. Their papers can be seen as a complementary way to generate influential equilibria when the sender has state-independent preferences.

Our baseline model is rooted in Glazer and Rubinstein (2004) who study a mechanism design problem that minimizes the probability of the decision maker taking wrong action given that the information provider is biased in favor of one alternative of the decision maker. We depart from them in the following ways: On one hand, we assume away the commitment of the receiver, i.e., we study the existence and properties of a class of equilibria of the communication game rather than an optimal mechanism design problem. Although Glazer and Rubinstein (2004) show that the resulting optimal mechanism can be supported as an equilibrium outcome, i.e., commitment is not needed for the optimality of the mechanism. However, the converse is not true. In other words, in other equilibria of the game without commitment, the optimal mechanism with commitment may not be obtained. And what we are investigating belongs to the general group of equilibria which does not necessarily correspond to the optimal mechanism considered in Glazer and Rubinstein (2004). On the other hand, they take a mechanism design approach and focus on receiver-preferred equilibrium. Consequently, their focus is on extracting information from an informed sender rather than persuading an uninformed receiver. However, we are focused on a class of equilibria, and particularly the equilibrium that is sender-optimal. And we gain novel insights on whether and how the sender benefits from cheap talk, verification and commitment.

A recent extension of Glazer and Rubinstein (2004) is Carroll and Egorov (2017) who consider a similar setting (i.e., a receiver can verify only one dimension of a sender's multi-dimensional information) but focus on the range of sender's payoff functions that can support full information extraction, which is also a receiver-optimal mechanism design problem.

Another recent paper by Lipnowski and Ravid (2017) complements our work on how the sender benefits from commitment and cheap talk when the sender's preferences are state-independent. However, in their setting the receiver has no access to hard information, i.e., there is no verification.

The remainder of this paper is organized as follows. In Section 2, we present the baseline model in which the receiver can only check one attribute. In Section 3, we analyze the equilibrium and convey the main insights of this paper in this simple setting. We extend the environment to the general case where the receiver can check any number of attributes in Section 4. Some remarks about our model are made in Section 5. We conclude in Section 6. The proofs are relegated to Appendices A, B and C. We discuss other extensions in Appendices D and E.

2 Model

There are two players, a sender (she) and a receiver (he). The state has $N \geq 2$ binary attributes. Formally, the state $\theta \in \Theta := \{0,1\}^N$ is drawn from a symmetric prior $\pi \in \Delta(\Theta)$ with full support. Specifically, we assume that $\theta_1, \ldots, \theta_N$ are exchangeable. The natural interpretation is that the state captures the (binary) quality of a product along N dimensions. Alternatively, the components can be interpreted as the outcomes of binary product tests, which are independently and identically distributed conditional on the (unobserved) product quality. The i-th component θ_i equals 1 or 0 according to whether the i-th attribute is good or bad. The sender's utility is 1 if the receiver buys and zero otherwise. Let $|\theta| := \theta_1 + \cdots + \theta_N$.[1] The receiver's utility from buying when the state is θ is $v(\theta) - P$ where $v : \Theta \to \mathbb{R}$ is a symmetric function such that $v(\theta) = v(\theta')$ if $|\theta| = |\theta'|$ and $v(\theta) > v(\theta')$ if $|\theta| > |\theta'|$, and P is the price of the good, which is exogenously given.[2] If the receiver does not buy, his utility is normalized to 0.

The timing is as follows. The sender observes the state realization and sends a message to the receiver. The receiver sees the message, updates his beliefs by Bayes' rule, and then costlessly checks one attribute of the state.[3] The receiver then decides whether to buy the product.

To complete the description of the model, we define strategies for both players. A *message strategy* for the sender is a function $m\colon \Theta \to \Delta(\mathcal{M})$ that maps each state realization to a distribution over a finite message space $\mathcal{M}$. For any message $A \in \mathcal{M}$, let $m(A|\theta)$ denote the probability that $m(\theta)$ assigns to A. Let $[N] := \{1, \ldots, N\}$. A strategy for the receiver is a pair (c, b) specifying a *checking strategy* $c\colon \mathcal{M} \to \Delta([N])$ with c_i being the probability that the receiver checks attribute i and a *buying strategy* $b = (b^0, b^1)\colon \mathcal{M} \to [0,1]^N \times [0,1]^N$ specifying the probability of buying upon seeing a bad attribute and a good attribute. Specifically, b_i^0 (respectively, b_i^1) denotes the probability that the receiver buys after checking attribute i and seeing that it is bad (respectively, good).

The payoffs of the sender and the receiver from this combination of strategies (m, c, b)

[1] We use $|\cdot|$ to denote the sum of the components of a vector. Therefore, $|\theta_B| = \sum_{i \in B} \theta_i$ for any $B \subset \{1, \ldots, N\}$. And we will refer to $|\theta|$ as the quality of the good below.

[2] We believe the exogenous price is an innocuous assumption since our main focus is on the strategic interaction between the sender and the receiver rather than how the price arises. We can think of it as the prevailing price in a perfectly competitive market.

[3] We assume the receiver has limited time or cognitive capacity for processing information as in Glazer and Rubinstein (2004) so that the receiver can check the attributes without any cost within his checking capacity but it is prohibitively costly to check attributes beyond his checking capacity. Our results still hold if the verification cost is sufficiently small, which will be discussed in Section 5.2.

are

$$u_S(m;c,b) = \sum_{\theta,A,i} c_i(A)(\theta_i b_i^1(A) + (1-\theta_i)b_i^0(A))m(A|\theta)\pi(\theta),$$
$$u_R(m;c,b) = \sum_{\theta,A,i} (v(\theta)-P)c_i(A)(\theta_i b_i^1(A) + (1-\theta_i)b_i^0(A))m(A|\theta)\pi(\theta)$$

respectively, where each sum is taken over all $\theta \in \Theta, A \in \mathcal{M}$, and $i \in \{1,\dots,N\}$.

The solution concept is perfect Bayesian equilibrium.

3 Equilibrium Analysis: Checking One Attribute

3.1 Existence

First notice that, in a game of cheap talk without verification, no communication could be sustained in equilibrium. As all the cheap talk games, the babbling equilibrium always exists in which the sender sends messages that are independent of her type and the receiver just ignores the sender's message. With the state-independent preferences of the sender, this would be the only equilibrium. The reason is that no matter how coarse the messages are, the sender would always find it advantageous to send the message that induces the highest probability of buying. Thus in equilibrium, the receiver's behavior would be independent of the messages.

However, by introducing partial state verification, there can be a lot of non-babbling equilibria. We would focus on a particular family of equilibria in which the sender is playing a specific message strategy, called the *top-k* strategy, where k is a parameter in $[N]$. Under this strategy, the sender effectively "points" to the k highest attributes of the realized state vector, without indicating any ordering among those k attributes. If there is a tie, the sender will break it uniformly. For example, suppose that $N=3$, $k=2$, and $\theta=(1,0,0)$. Then the sender claims "Attribute 1 and 2 are my two highest attributes" with probability 1/2 and "Attribute 1 and 3 are my two highest attributes" with probability 1/2. So the message space is $\mathcal{M}=\mathcal{P}_k$, where $\mathcal{P}_k$ is the set of k-element subsets of $[N]$. The top-k message strategy, denoted $m^k : \{0,1\}^N \to \Delta(\mathcal{P}_k)$, is formally defined as follows. First let

$$T_k(\theta) = \arg\max_{I\in\mathcal{P}_k} |\theta_I|.$$

For any $\theta \in \{0,1\}^N$,

$$m^k(A \mid \theta) = \begin{cases} 1/|T_k(\theta)| & \text{if } A \in T_k(\theta), \\ 0 & \text{otherwise.} \end{cases}$$

It will be convenient here and below to use the notation $a \wedge b := \min\{a, b\}$.[4] Notice that $T_k(\theta) = \{I \in \mathcal{P}_k \mid |\theta_I| = k \wedge |\theta|\}$. Therefore, for any $A \in \mathcal{P}_k$, we have

$$m^k(A \mid \theta) = \begin{cases} 1/\binom{N-|\theta|}{k-|\theta|} & \text{if } |\theta_A| = k \wedge |\theta| = |\theta|, \\ 1/\binom{|\theta|}{k} & \text{if } |\theta_A| = k \wedge |\theta| = k, \\ 0 & \text{otherwise.} \end{cases}$$

where $\binom{b}{a}$ is the number of combinations without repetitions of a elements from b elements.

We would like to exclude trivial equilibria in which no trade occurs. The following definition is in order:

Definition. *A perfect Bayesian equilibrium is called top-k equilibrium if the sender is playing the top-k strategy and the probability of trading is positive.*

Given the sender's top-k strategy, it is not hard to conjecture that in a top-k equilibrium, the receiver's strategy is evenly pick one attribute from those that are recommended to check and buy if and only if the result of verification is 1. Clearly, the sender is incentive compatible to honestly communicate since she is maximizing the probability of buying by playing top-k strategy given the receiver's strategy. But we need to make sure the strategy specified is a best response for the receiver. The following theorem gives the sufficient and necessary condition for the existence of a top-k equilibrium. Let $\bar{S}_a^b$ ($\underline{S}_a^b$) denote the sum of a attributes chosen uniformly from b of the top (bottom) attributes.[5]

Theorem 1. *Let $k \in [N]$. There exists a top-k equilibrium if and only if*

$$P \leq \mathbb{E}[v(\theta) \mid (\bar{S}_1^k, \underline{S}_1^{N-k}) = (1, 0)].$$

The theorem states that there exists a top-k equilibrium if and only if the price is sufficiently low. In that case, the checking and buying strategies specified above together with the top-k strategy constitute an equilibrium. Intuitively, top-k equilibria can be sustained if the price is low enough such that the receiver will actually check the recommended attributes. If the price is too high, the receiver would rather deviate to the unrecommended attributes. But then the sender would have pointed to the worst attributes in the first place, knowing the receiver will not check them, and the equilibrium breaks down.

---

[4] To simplify the notations below, we will assume the wedge operator "$\wedge$" has higher precedence than addition, so, e.g., $a \wedge b + c = \min\{a, b\} + c$.

[5] To determine these top and bottom attributes, ties are broken uniformly. To simplify the notations below, set $\bar{S}_a^b = \bar{S}_b^b$ and $\underline{S}_a^b = \underline{S}_b^b$ for $a > b$.

To better understand this theorem, we start with top-1 equilibrium to elaborate. It can be shown that when $k = 1$, the sufficient and necessary condition above can equivalently be expressed as

$$\mathbb{E}[(v(\theta) - P)\mathbb{1}(\theta \neq \mathbf{0})] \geq \mathbb{E}[(v(\theta) - P)\theta_1].$$

For necessity, suppose there exists an equilibrium of the desired form at the given price P. That is, the sender uniformly points to her highest attribute and the receiver checks the recommended attribute and buys if and only if the result of verification is 1. Then, on the equilibrium path, the receiver will buy if and only if $\theta \neq \mathbf{0}$. Moreover, the receiver must get a weakly higher payoff from obedience than from the following deviation: check attribute 1 and buy if and only if $\theta_1 = 1$. But this requirement is precisely the inequality above since the receiver's payoff from buying is $v(\theta) - P$.

For sufficiency, suppose the price P satisfies the condition. Given the receiver's strategy, the sender clearly has no profitable deviation. It remains to check that the receiver has no profitable deviation. Specifically, we need to show two kinds of deviation are not profitable for the receiver: first, the receiver ignores the sender's message and checks a random attribute and then buys if and only if that attribute is good; second, the receiver checks an unrecommended attribute based on the sender's message and buys if and only if it is good. By the necessity above, it is clear that the first deviation is not profitable. The difficulty lies in the second deviation since checking different unrecommended attributes induces different beliefs of the receiver. When the number of attributes is large, the second deviation can be very overwhelming to deal with.

To solve this problem, we will prove something slightly stronger than needed: even if the receiver could check the recommended attribute i and one other attribute $j \neq i$ of his choosing, his optimal strategy would still be to buy if and only if $\theta_i = 1$. Formally, we are proving optimality in the larger class of strategies that are measurable with respect to (θ_i, θ_j) for some j. In other words, we consider a hypothetical game in which the receiver can check not only one recommended attribute but also one unrecommended attribute, which gives the receiver a more general information partition structure than any information partition structure in the original game. And we show the optimality under this finer information partition structure for the receiver.

Now consider a general top-k equilibrium. Recall that $\bar{S}_1^k$ ($\underline{S}_1^{N-k}$) records the result of verification from checking the recommended (unrecommended) attributes. Notice that when the sender uniformly points to k highest attributes, the support of $(\bar{S}_1^k, \underline{S}_1^{N-k})$ is contained in the set

$$\mathbf{S} = \{(0, 0), (1, 0), (1, 1)\}.$$

Then there is a threshold $(a, b) \in \mathbf{S}$ such that it is a best response to buy if and only if $(\bar{S}_1^k, \underline{S}_1^{N-k}) \geq (a, b)$. If $b = 0$, then this buying strategy is measurable with respect to $\bar{S}_1^k$ and hence the receiver need only check recommended attributes. If $b = 1$, then this buying strategy is measurable with respect to $\underline{S}_1^{N-k}$ and hence the receiver need only check the unrecommended attributes. Therefore, Theorem 1 essentially indicates that top-k equilibria can be supported if the buying threshold is low enough so that the receiver will actually only check the recommended attributes. Otherwise, the receiver would rather deviate to the unrecommended attributes.[6]

It is worth noting that among the family of top-k equilibria, top-1 equilibrium is the sender's most preferred equilibrium since the highest probability of buying is achieved in this equilibrium. In effect, all the non-zero types manage to sell the good with probability one if the top-1 equilibrium is sustainable.

3.2 Parametrization

In order to bring about the novel insights, we parametrize the model by taking $v(\theta) = |\theta|$ and imposing a specific strictly positive prior $\pi = (\pi_0, \ldots, \pi_N)$ over $|\theta|$, i.e., $\Pr(|\theta| = i) = \pi_i$ for $i = 0, \ldots, N$. Because $\theta_1, \ldots, \theta_N$ are assumed exchangeable, the probability distribution of the vector θ is pinned down by π. Specifically, $\Pr(\theta = \hat{\theta}) = \pi_{|\hat{\theta}|} / \binom{N}{|\hat{\theta}|}$ for $\hat{\theta} \in \{0, 1\}^N$. Then Theorem 1 reduces to the following proposition:

Proposition 1. *Let $v(\theta) = |\theta|$. There exists a top-1 equilibrium if and only if*

$$P \leq \frac{N\mathbb{E}(|\theta|) - \mathbb{E}(|\theta|^2)}{N(1 - \pi_0) - \mathbb{E}(|\theta|)}.$$

For any given strictly positive prior π, we can find the upper bound of the price by Proposition 1 to sustain a top-1 equilibrium when $v(\theta) = |\theta|$. To make things more interesting, we further assume that $\mathbb{E}(|\theta|) < P$. Therefore, the receiver will not buy in the ex ante stage with the prior belief. However, the price is not necessarily below the upper bound dictated by Proposition 1 when it is above the ex ante expectation of $|\theta|$. The requirement $\mathbb{E}(|\theta|) < \frac{N\mathbb{E}(|\theta|) - \mathbb{E}(|\theta|^2)}{N(1 - \pi_0) - \mathbb{E}(|\theta|)}$ implies $\pi_0 > \frac{\mathrm{Var}(|\theta|)}{N\mathbb{E}(|\theta|)}$.[7]

We consider two benchmarks to which we compare top-1 equilibrium with respect to the probability of buying.

Benchmark 1. *There is only cheap talk communication but no verification.*

[6] See more discussions in Section 4 after we present the more general result Theorem 2.

[7] Note that neither Theorem 1 nor Proposition 1 relies on this requirement.

Benchmark 2. *There is no communication but the receiver can randomly pick one attribute to verify.*

Under Benchmark 1, as we discussed before, there is only the babbling equilibrium and the receiver will ignore the sender's message and take an action based on his prior belief, which implies he will not buy by the assumption $\mathbb{E}(|\theta|) < P$.

Under Benchmark 2, note that the receiver would be determined to purchase the good *only if* the result of verification is 1. To see why, suppose the receiver checks attribute i with some probability $q_i \in [0,1]$ such that $\sum_{i=1}^{N} q_i = 1$. If he finds the result of verification is 0, his expectation over all the attributes is $\mathbb{E}[|\theta||\ \theta_i = 0] < P$.[8] Thus the receiver will not buy. Therefore, the probability of buying is *at most* $\sum_{i=1}^{N} q_i \cdot \left(\sum_{j=1}^{N} \frac{\pi_j \binom{N-1}{j-1}}{\binom{N}{j}} \right) = \frac{\mathbb{E}(|\theta|)}{N}$.[9]

Compared to these two benchmarks, the probability of buying in top-1 equilibrium is strictly higher than either of them.[10] This comparison is summarized in Figure 1. It is clear

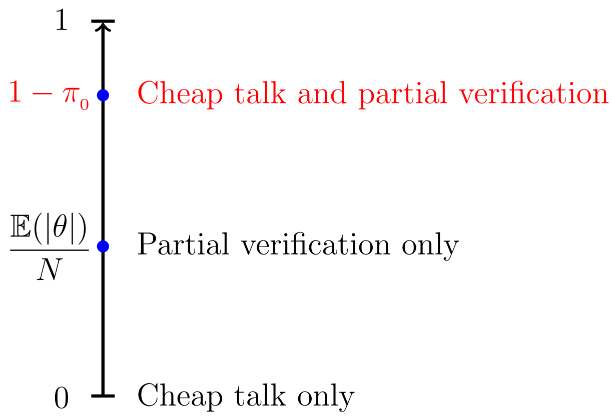


Figure 1: Probability of Trading

that even though the sender has state-independent preferences, for a range of prices, the sender strictly benefits from the ability to communicate. The underlying idea is simple. We can think of the cheap talk messages as a belief-coordinating device which guides the receiver to the good attributes. The receiver finds it optimal to follow the sender's recommendation

[8]To see this, note that $\mathbb{E}[|\theta||\ \theta_i = 0] \leq \mathbb{E}(|\theta|)$ (see Lemma 5 in Appendix C.1 for the formal proof) and by the assumption that $\mathbb{E}(|\theta|) < P$, the desired inequality is attained.

[9]Since $\mathbb{E}(|\theta|) \leq \mathbb{E}[|\theta||\ \theta_i = 1]$ (see Lemma 5 in Appendix C.1 for the formal proof), under the assumption $\mathbb{E}(|\theta|) < P$, the price P can be higher than $\mathbb{E}[|\theta||\ \theta_i = 1]$ so that the probability of buying can be zero even if the receiver finds the result of verification is 1.

[10]It is straightforward to see that $\mathbb{E}(|\theta|) = \sum_{i=0}^{N} i \cdot \pi_i = \sum_{i=1}^{N} i \cdot \pi_i < \sum_{i=1}^{N} N \cdot \pi_i = N(1 - \pi_0)$. Therefore, we have $\frac{\mathbb{E}(|\theta|)}{N} < 1 - \pi_0$.

in equilibrium and therefore he sees a good attribute more often than if he randomly picks one attribute to check and sees a "1."

3.3 Comparative Statics

For the following comparative statics results, we will keep N and π fixed.

3.3.1 Equilibrium Parameter k

First, we study how the equilibrium existence depends on the equilibrium parameter k. Define the price threshold by

$$\bar{P}_k = \mathbb{E}[v(\theta) \mid (\bar{S}_1^k, \underline{S}_1^{N-k}) = (1, 0)].$$

This is the highest price at which there exists a top-k equilibrium when the receiver can only check one attribute.

Proposition 2. *The price threshold $\bar{P}_k$ is strictly increasing in k.*

This means that the set of prices at which the top-k equilibrium can be sustained is increasing in k. The intuition is that as k increases, the sample of the top-k attributes is moderated so that seeing ones becomes a better signal of quality. Therefore, a positive signal from this sample can induce purchase at higher prices.

Next, we study the comparative statics of the equilibrium buying probability and the expected utility of the receiver upon buying. Let Z_k be an indicator function for whether the receiver buys. Clearly, the equilibrium buying probability is

$$\mathbb{E}[Z_k] = \frac{\mathbb{E}[k \wedge |\theta|]}{k}.$$

Proposition 3. *The equilibrium buying probability $\mathbb{E}[Z_k]$ is decreasing in k but the conditional expected utility of the receiver $\mathbb{E}[v(\theta) \mid Z_k = 1]$ is increasing in k.*

Concerning the equilibrium buying probability, as k increases, it is more and more difficult to see a good attribute for the receiver when he uniformly picks one recommended attribute to check. For example, the type (1,0,0) can sell the good with probability one in the top-1 equilibrium whereas the receiver will check her first attribute with probability 1/2 in the top-2 equilibrium. Concerning the conditional expected utility of the receiver, intuitively, if the sender points to one attribute and the receiver indeed sees a "1," he will not take it as a big deal since he suspects that the sender may only have a single "1." However, if the sender points to ten attributes and claims they are good, when the receiver sees a "1," he would reasonably believe the sender may have quite a few ones and therefore he can easily see a "1" when he randomly checks a recommended attribute. In other words, seeing a "1" becomes a rarer but stronger signal as k increases.

Proposition 3 implies that if the sender can select which top-k equilibrium is played, she will choose the one with the smallest k. However, she can achieve equilibria with smaller and smaller k only as the price goes down by Proposition 2. At some point, the price is low enough so that the receiver will buy without communication. The relative position of this threshold $\mathbb{E}[v(\theta)]$ and the thresholds $\bar{P}_k$ depends on the distribution.

3.3.2 Full Taxonomy

We have already seen that in terms of the probability of trading, the top-1 equilibrium is the sender's most preferred equilibrium. And the sender can strictly benefit from cheap talk even though the preferences of the two parties are misaligned. Then a natural question arises: what if a top-1 equilibrium is not sustainable because the price is too high? For example, the price is just between the price threshold of top-1 equilibrium and top-2 equilibrium so that the top-1 strategy fails to be an equilibrium. Then could the sender do better than in a top-2 equilibrium? The answer is yes. The idea is to give the sender commitment power. Here we relate to the Bayesian persuasion literature where the sender can commit to a signal structure. We will show that the sender can strictly benefit from commitment under partial verification. This may not be surprising given the insights of the Bayesian persuasion literature, e.g., Kamenica and Gentzkow (2011). However, we will also show that in some cases, the sender does not benefit from commitment and cheap talk can do as well as Bayesian persuasion.

Altogether, we consider two information structures for the receiver: no verification and partial verification (V) and three communication protocols for the sender: no communication (NC), cheap talk (CT), and Bayesian persuasion (BP). This gives six ways to pair an information structure with a communication protocol, which we label as NC, NCV, CT, CTV, BP and BPV, respectively. It is natural to compare the sender's payoff in her corresponding most-preferred equilibrium in each of these six settings. We have completed the full taxonomy. It is summarized in the following proposition.

Proposition 4. *The ranking of trading probability in the sender-optimal equilibrium of different settings is as follows:*

$$NC \leq CT \leq BP \text{ and } NCV \leq CTV \leq BPV \leq BP.$$

This proposition shows that no communication is weakly worse than cheap talk which is in turn weakly worse than Bayesian persuasion. This is not surprising because the sender has more and more control over the receiver's information structure and therefore she can induce higher and higher probability of buying. This is also true when we introduce verification. Since under Bayesian persuasion, the sender has full control over the receiver's information structure, she always achieves the highest probability of trading. These are the only relationships that hold in general. Any two pairs that are not related by the inequalities above have ambiguous relationships. For example, there are examples where the sender's payoff is higher under NC than NCV and also examples where the sender's payoff is higher under NCV than NC.

Numerical Example. Let $N = 3$, $v(\theta) = |\theta|$ and $|\theta|$ follows a Binomial distribution with Bernoulli parameter q where $0 < q < 1$, i.e., $|\theta| \sim \text{Binomial}(N, q)$. Note that the (ex ante) expected quality $\mathbb{E}(|\theta|)$ is simply $Nq = 3q$. By Theorem 1, we can pin down $\bar{P}_1$ and $\bar{P}_2$ for each prior indexed by q, which are $\dfrac{2}{2-q}$ and $\dfrac{1+3q}{1+q}$, respectively. If there is no communication but the receiver can randomly pick one attribute to check, $\mathbb{E}[|\theta| \mid \theta_i = 0] = 2q$ and $\mathbb{E}[|\theta| \mid \theta_i = 1] = 2q + 1$ for $i = 1, 2, 3$ are the expected quality upon seeing "0" and "1," respectively. We plot $\bar{P}_1$, $\bar{P}_2$ and the expectations in Figure 2.

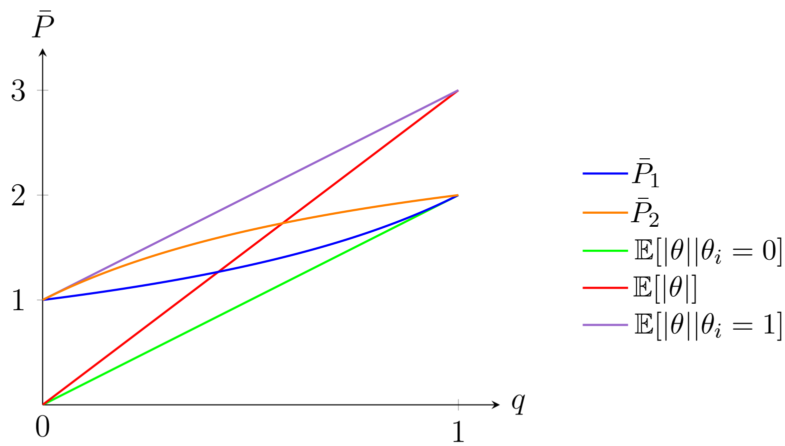


Figure 2: Price Cutoffs

To illustrate the potential gains from commitment power, we will restrict attention to prices $P \in (\bar{P}_1, \bar{P}_2]$. For this range of prices, we see that without communication, if the receiver randomly selects an attribute to check, he will buy if and only if it is good. For prices southeast of the red line, the receiver would buy even if he could not verify, and hence verification makes the sender worse off under no communication or cheap talk. For prices northwest of the red line, the receiver would not buy if he could not verify, and hence verification makes the sender better off under no communication or cheap talk.

Since we are more interested in the effect of the communication structure, taking verification as given, we will see an example where we have the strict inequalities NCV<CTV<BPV<BP. See Figure 3.

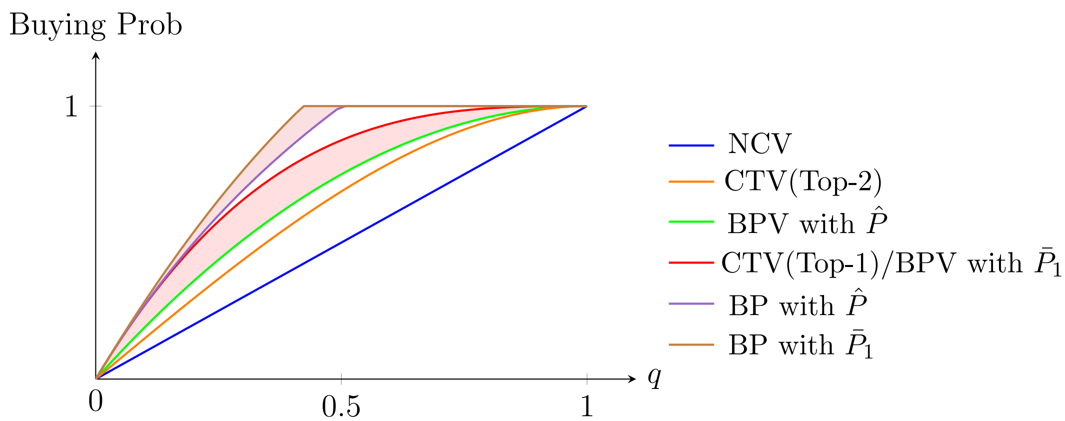


Figure 3: Communication Structures under Verification

We fix the price at $\hat{P}$ which is strictly between $\bar{P}_1$ and $\bar{P}_2$, specifically, $\hat{P} = 1 + q$. With this price, in the verification without communication setting (NCV), the buying probability is simply q corresponding to the blue line since the receiver will buy if and only if the result of verification is 1. In the cheap talk with verification setting (CTV), it jumps to the orange line since now only top-2 equilibrium is sustainable. When the sender has commitment power, namely, in the Bayesian persuasion with verification setting (BPV), the buying probability jumps to the green line. How? The commitment power substantially enlarges the sender's set of communication strategies. Specifically, the sender will always point to exactly one attribute, but she will mix between pointing to the highest attribute and the second highest attribute. The mixing probability will be computed by the receiver's indifference condition: the receiver will be indifferent between (i) checking the indicated attribute and buying if and only if it is good; and (ii) checking a random attribute and buying if and only if it is good. Note that at the price $\hat{P}$, the receiver strictly prefers obedience to disobedience in the top-2 equilibrium. However, under BPV, the sender can gradually adjust the signal structure to increase the buying probability and hence decrease the receiver's utility until the receiver is exactly indifferent.

Without surprise, in pure Bayesian persuasion setting, the sender's hands are not tied any more. She can induce an even higher buying probability. Specifically, the equilibrium of the BP setting will have a monotonicity property: if $|\theta|$ is larger than some marginal number, the sender will recommend the receiver to buy; if it is lower than the marginal number, the sender will recommend the receiver not to buy; and if it is equal to the marginal number, the sender will mix between the two recommendations. The marginal number and mixing probability will be pinned down by the condition that the receiver is indifferent between buying and not buying when the sender recommends she buy. It turns out that at $\hat{P}$ the marginal number is 0, which implies, now the zero type gets some chance to sell. Therefore, the buying probability is higher than any communication protocol with verification setting.

Another interesting thing is that, if we reduce the price from $\hat{P}$ to $\bar{P}_1$, the buying probability curves under BPV trace out the pink region and the buying probability curves under BP trace out the dark pink region, respectively. Under BPV, when the price varies from $\hat{P}$ to $\bar{P}_1$, the sender puts more and more weight on the highest attribute and hence the buying probability rises. When the price drops to $\bar{P}_1$, the sender cannot benefit from commitment any more under verification (CTV does as well as BPV). Under BP, when the price tends to $\bar{P}_1$, the sender will tell the receiver to buy if her type is not $\mathbf{0}$ and also more and more often when her type is $\mathbf{0}$ according to the outcome of her randomization device. And hence the buying probability rises.

4 Extension: Checking Multiple Attributes

Assuming that the receiver can only check one attribute provides a simple illustration of the main insights we want to bring up. However, the analysis can be generalized. In this section, we will allow the receiver to costlessly and simultaneously check any number of attributes after seeing the sender's message.[11] Of course, it is still partial verification, i.e., $n < N$ where n denotes the receiver's information (or checking) capacity.

4.1 Existence

To define the receiver's strategy in this general setting, let $\mathcal{P}_n$ denote the family of subsets of $\{1, \dots, N\}$ of size at most n. Here $\mathcal{P}_n$ consists of all attribute sets that the receiver can check. A strategy for the receiver is a pair (c, b) consisting of a *checking strategy* $c : \mathcal{M} \to \Delta(\mathcal{P}_n)$ and a *buying strategy*

$$b : \mathcal{M} \times \bigcup_{B \in \mathcal{P}_n} \{0,1\}^B \to [0,1].$$

A checking strategy specifies the set of attributes the receiver checks after each message. It maps each message from the sender to a distribution over subsets of attributes. Let $c(B|A)$ denote the probability that, upon receiving the message A, the receiver checks exactly the attributes in B. The checking strategy specifies the possibly random set of attributes the receiver checks upon receiving the sender's message. The realizations of these checked

[11] It turns out that the results are not affected if the receiver sequentially checks the attributes that he chose. See Corollary 1.

attributes are summarized by a function $y : B \to \{0,1\}$. And $b(A, y)$ is the probability that the receiver buys the good upon receiving message A, checking exactly the attributes $i \in B$, and observing $\theta_i = y(i)$ for each $i \in B$.[12]

Since $n = 1$ is taken as a special case of what we will discuss in this section, the family of equilibria we will focus on is still the top-k equilibria. We reproduce the sender's top-k strategy and the definition of top-k equilibrium here:

Definition. *The top-k message strategy, denoted $m^k : \{0,1\}^N \to \Delta(\mathcal{P}_k)$, is formally defined as follows: For any $A \in \mathcal{P}_k$,*

$$m^k(A \mid \theta) = \begin{cases} 1/\binom{N-|\theta|}{k-|\theta|} & \text{if } |\theta_A| = k \wedge |\theta| = |\theta|, \\ 1/\binom{|\theta|}{k} & \text{if } |\theta_A| = k \wedge |\theta| = k, \\ 0 & \text{otherwise.} \end{cases}$$

Definition. *A perfect Bayesian equilibrium is called top-k equilibrium if the sender is playing the top-k strategy and the probability of trading is positive.*

Now we define a strategy for the receiver. If he can check all the attributes recommended by the sender, he does so (and does not check any others). If he cannot check all the recommended attributes, then he checks n of them. He buys if the number of ones he sees exceeds some threshold. We give the sufficient and necessary condition of the existence of top-k equilibrium that holds in general.

Theorem 2. *Suppose $n \leq N - 2$. For each $k = 1, \ldots, N$, there exists a top-k equilibrium if and only if*

$$P \leq \mathbb{E}[v(\theta) \mid (\bar{S}^k_n, \underline{S}^{N-k}_n) = (n \wedge k, 0)].$$

If $n = N - 1$, it goes through except the "only if" statement holds for all but finitely many prices.[13]

First note that by taking $n = 1$ Theorem 2 reduces to Theorem 1.

The theorem states that there exists a top-k equilibrium if and only if the price is sufficiently low. In that case, the checking and buying strategies specified above together with the top-k strategy constitute an equilibrium. In partibular, if the top-k equilibrium can be sustained for some $k < n$, then the receiver need not even check n of the attributes. He can

[12] We will use the function notation $y(i)$ and the vector notation y_i interchangeably throughout.

[13] Specifically, when $n = N - 1$ and $k \leq N - 2$, we can also construct top-k equilibrium if $P = \mathbb{E}[v(\theta) \mid (\bar{S}^k_n, \underline{S}^{N-k}_n) = (n \wedge k, j)]$ for some $j = 1, \ldots, N - k - 1$.

base his decision on only k of them. Of course this is not a best response for the receiver in general. The intuition is that if the receiver would gain from checking unrecommended attributes, then he would check as many of the unrecommended attributes as possible. But then the sender would have an incentive to deviate and point to lower attributes, and the equilibrium would break down.

The formal proof of this theorem is relegated to the Appendix. Here we would like to sketch the basic idea of the proof. For simplicity and intuition, let $n < k$ and $n < N - k$ so that the receiver's information capacity is lower than the number of (un)recommended attributes. It seems quite overwhelming because we need to take care of quite a lot of possible deviations and different deviations induce different beliefs of the receiver. Here the idea we are using is that, since the receiver's buying strategy is based on an information partition structure induced from the checking decision and message received, we construct a hypothetical game in which the receiver can check not only n recommended attributes but also n unrecommended attributes. That is, the receiver is given a more general information partition structure and it is finer than any information partition structure in the original game. Then if we can show the receiver's buying strategy is optimal in this hypothetical game, we can conclude the receiver's buying strategy is also optimal in the original game, if it is still attainable. In other words, we enlarge the strategy space of the receiver, and if we can show the optimality of the receiver's buying strategy in this larger set, we can conclude it is still optimal in the smaller set if it is still attainable. Therefore, we will effectively prove something stronger than needed (we allow for more possible deviations in the hypothetical game than we need to take care of in the original game).

The key observation is that, given the sender's top-k strategy, the support of $(\bar{S}_n^k, \underline{S}_n^{N-k})$ is contained in the set

$$\mathbf{S} = \{(0,0), (1,0), \ldots, (n,0), (n,1), \ldots, (n,n)\}$$

on which the product order is total.[14] If the receiver could check n of the recommended attributes and n of the unrecommended attributes, then there is a threshold $(a, b) \in \mathbf{S}$ such that it is a best response to buy if and only if $(\bar{S}_n^k, \underline{S}_n^{N-k}) \geq (a, b)$. If $b = 0$, then this buying strategy is measurable with respect to $\bar{S}_n^k$ and hence the receiver need only check recommended attributes. Alternatively, if $b > 0$, then this buying strategy is measurable with respect to $\underline{S}_n^{N-k}$ and hence the receiver need only check the unrecommended attributes. This is the heart of Theorem 2.

Essentially, top-k equilibria can be sustained if the buying threshold is low enough so that the receiver will actually only check the recommended attributes. If the buying threshold is

[14] Recall that the product order means $(a, b) \geq (a', b')$ if and only if $a \geq a'$ and $b \geq b'$.

too high, the receiver would rather deviate to the unrecommended attributes. But then the sender would have pointed to the worst attributes in the first place, knowing the receiver will not check them, and the equilibrium breaks down.

It is not hard to observe that the receiver does not benefit from sequential verification.

Corollary 1. *Theorem 2 still holds if the receiver can check the attributes sequentially rather than simultaneously.*

4.2 Comparative Statics

Compared to $n = 1$, not all the previous comparative statics results can be carried over to $1 < n < N$. We will discuss the reasons and give counter-examples. We still keep N and π fixed.

First, recall that the receiver's buying strategy is a cutoff strategy. That is, he will buy if the number of ones he sees exceeds some threshold. The threshold $\bar{s}^*$ for recommended attributes is defined by

$$\bar{s}^*(P, k, n) = (n \wedge k + 1) \wedge \min \Big\{ s \in \{0, \ldots, n \wedge k\} \mid \mathbb{E}[v(\theta) \mid \bar{S}_n^k = s] \geq P \Big\}.$$

Analogously, the threshold $\underline{s}^*$ for unrecommended attributes is defined by

$$\underline{s}^*(P, k, n) = (n \wedge (N-k) + 1) \wedge \min \Big\{ s \in \{0, \ldots, n \wedge (N-k)\} \mid \mathbb{E}[v(\theta) \mid \underline{S}_n^{N-k} = s] \geq P \Big\}.$$

The following proposition gives how the buying thresholds depend on the price P, the equilibrium parameter k, and the receiver's information capacity n.

Proposition 5. *Both thresholds $\bar{s}^*(P, k, n)$ and $\underline{s}^*(P, k, n)$ are weakly increasing in P and n but weakly decreasing in k.*

The comparative statics in P and n are straightforward. If the price is higher, the receiver demands a more favorable signal of quality in order to buy the good. If the sample is larger, more successes are required to provide the same signal of quality. Increasing k pushes the sample of recommended attributes closer to random but also pushes the sample of unrecommended attributes away from random. In each case, the sample is made lower relative to the population as a whole, and hence fewer successes are needed to induce purchase.

Next, we study the comparative statics of equilibrium existence. Define the price threshold by

$$P^*(k, n) = \mathbb{E}[v(\theta) \mid (\bar{S}_n^k, \underline{S}_n^{N-k}) = (n \wedge k, 0)].$$

This is the highest price at which there exists a top-k equilibrium when the receiver's information capacity is n.

Proposition 6. *For any fixed* n*, the price threshold* $P^*(k,n)$ *is strictly increasing in* k*.*

The intuition is similar to the previous result. As the sender points to more and more high attributes, the receiver's sample of recommended attributes becomes more and more representative and he is more optimistic when seeing ones and therefore would like to pay a higher price.

This result is a generalization of Proposition 2. However, we do not necessarily have monotonicity in n. Following the equilibrium is more informative as n increases, but so is deviating. We can construct a counter-example by taking $N = 3$, $v(\theta) = |\theta|$ and $|\theta| \sim \text{Binomial}(3, 1/3)$. By Theorem 2, we can calculate the corresponding price thresholds. See Table 1. When fixing k and varying n, there is no unambiguous monotonicity in n.

$P^*(k,n)$	$n = 1$		$n = 2$
$k = 1$	1.2	$\searrow$	1
$k = 2$	1.5	$\nearrow$	2

Table 1: Price Threshold $P^*(k,n)$

At last, with respect to equilibrium buying probability and the receiver's expected utility upon buying, we lose the preceding nice comparative statics results in Proposition 3 for the more general case $1 < n < N$. The fundamental reason is that the buying threshold can change when the receiver can check more than one attribute. When $n = 1$, the only equilibrium buying threshold is one.

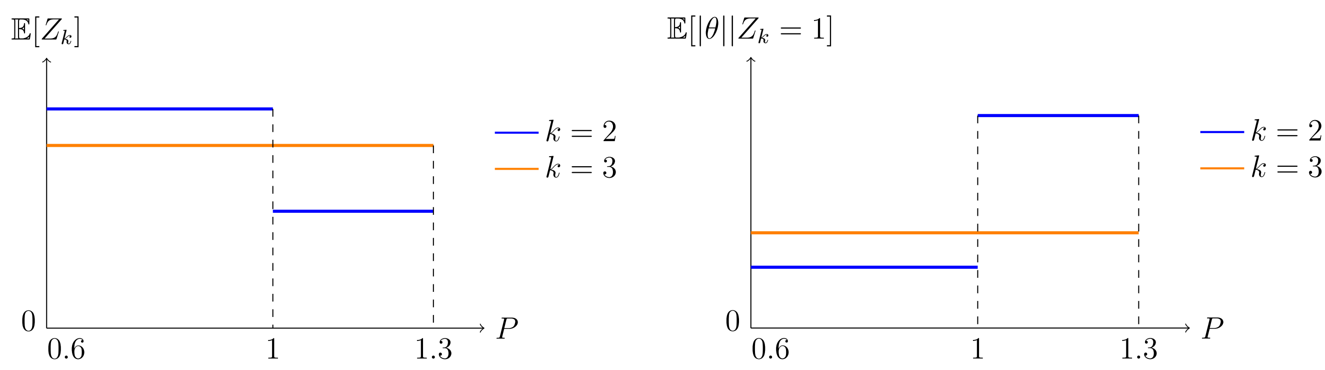
For intuition, let's focus on the case where $n < k$ and $v(\theta) = |\theta|$. So after the sender recommends the k highest attributes, the receiver randomly and uniformly selects n of these k attributes to check. He will buy if and only if at least $\bar{s}^*$ of these n attributes are good. If $n = 1$, then we must have $\bar{s}^* = 1$. That is, the receiver will check a single attribute and buy if and only if it is good. When $n > 1$, however, we may have $\bar{s}^* > 1$. For example, if $n = 3$ and $\bar{s}^* = 2$, the receiver will check three attributes and buy if and only if at least two of them are good. Recall that $\bar{s}^*(P,k)$ is weakly decreasing in k (see Proposition 5).[15] If k is lower, then the receiver knows that the sample of attributes he is observing is more upwardly biased, so he demands to see more high attributes in order to be willing to buy. Also notice that no matter the value of n, if the sender points to the k highest attributes, then the average quality of these recommended attributes is weakly decreasing in k. For

[15] In general, the value of $\bar{s}^*$ is also determined by the receiver's information capacity n. For fixed n, we denote it by $\bar{s}^*(P,k)$ as a shorthand.

example, the average quality of the two highest attributes is weakly greater than the average quality of the four highest attributes.

Let's look at a simple case where $N = 4$, $n = 2$ and we can compare $k = 2$ with $k = 3$. We could get a simple failure of monotonicity as follows. For P low enough, we might have $\bar{s}^*(P, 2) = \bar{s}^*(P, 3) = 1$. That is, regardless of whether the sender points to the two or three highest attributes, the receiver will check $n = 2$ of them and buy if and only if at least one of those attributes is good. It follows that for this range of prices, the top-2 equilibrium induces a higher buying probability and a lower expected quality conditional on buying. But when P increases beyond some threshold, say to some P', we might have $\bar{s}^*(P', 2) = 2$ and $\bar{s}^*(P', 3) = 1$. This means that in the top-2 equilibrium, the receiver will only buy if both attributes he checks are good, but in the top-3 equilibrium, the receiver will buy if at least one of the two attributes he checks is good. In this region, where $\bar{s}^*(P', 3) < \bar{s}^*(P', 2)$, the buying probability may be higher and expected quality conditional on buying lower in the top-3 equilibrium than in the top-2 equilibrium. This phenomenon cannot happen when $n = 1$ because in that case either $\bar{s}^* = n = 1$ or else the receiver always buys.

More concretely, consider the following numerical example: Let $N = 4$, $n = 2$, $v(\theta) = |\theta|$, $|\theta| \sim \text{Binomial}(N, q)$ where $q = 0.3$ and $P \in [0.6, 1.3]$. We plot the change of the equilibrium buying probability $\mathbb{E}[Z_k]$ and the expected quality upon buying $\mathbb{E}[|\theta||\ Z_k = 1]$ as the price P varies between 0.6 and 1.3 in Figure 4a and Figure 4b, respectively. Specifically, when $0.6 < P < 1$ the buying probability is higher and expected quality conditional on buying lower in the top-2 equilibrium than in the top-3 equilibrium but the relationships are reversed when $1 < P < 1.3$.



(a) Equilibrium Buying Probability **(b)** Expected Quality upon Buying

Figure 4: Equilibrium Buying Probability and Conditional Expected Quality

5 Discussion

5.1 Sender- and Receiver-Optimal Equilibrium

Even though the state space Θ is finite and the payoff-relevant component of the receiver's action is binary, allowing for (payoff-irrelevant) partial verification means that the analysis must be performed on functions from Θ to $\mathcal{A} := \Delta([N]) \times [0,1]^N \times [0,1]^N$. Such functions form quite a large space, and this fact, together with the structure of the incentive constraints, makes it so difficult to characterize all equilibria.

In the core of the paper, we have identified a natural class of equilibria, the *top equilibria*, which capture what is often observed in practice. And our comparison across different combinations of information structure and communication protocol is also confined within this family of symmetric equilibria.

Depending on the price and the payoff functions, not all top strategy profiles will be equilibria. It is easy to check that the values of k for which the top-k strategy profile is an equilibrium will take the form of an interval of consecutive integers. Among these values of k, let k_S denote the sender's favorite top equilibrium and let k_R denote the receiver's favorite top equilibrium. Clearly, k_S will simply be the smallest value of k, so $k_S \leq k_R$. It can be shown that the top-k_R equilibrium is the receiver's most-preferred equilibrium among all possible equilibria (see Appendix D.1).

The top-k_S equilibrium is by definition the sender's most-preferred equilibrium among all *top equilibria*, but it is not necessarily the sender's most-preferred equilibrium among *all* equilibria. In other words, there are other, less natural equilibria, that can increase the buying probability further, and therefore make the sender even better off. Indeed, we can give an example of an *asymmetric* equilibrium that strictly increases the buying probability above the buying probability under the top-k_S equilibrium (see Appendix D.2).

Then one natural question to ask is whether the sender-optimal *symmetric* equilibrium is a top equilibrium. However, this is a hard question to answer at the moment and we leave it to future work. On one hand, there are other symmetric equilibria (we include the construction in Appendix D.3), but we have not found another symmetric equilibrium that increases the buying probability above the sender-preferred top equilibrium. On the other hand, at the current stage we are not able to construct a formal proof to show that such an increase is impossible.

5.2 Costly Verification

In the main model we have assumed the verification is costless within the receiver's checking capacity but becomes prohibitively costly beyond this capacity to capture the receiver's time or cognitive constraint as in Glazer and Rubinstein (2004). However, we can relax this assumption and allow for a sufficiently small verification cost and our top-k equilibria can still be sustained. We illustrate this by the following simple setting.

Consider $N = 2$, $v(\theta) = |\theta|$, $|\theta| \sim \text{Binomial}(N, q)$ where $q = 1/2$ and $P = 1$. In the ex ante stage without any verification, under the prior the receiver is indifferent between buying and not buying (note that $\mathbb{E}(|\theta|) = P$) and assume he will not buy which yields payoff 0. Suppose the receiver will incur a cost c for each attribute he checks. Also note that if the receiver checks both attributes, he will not buy since the expected payoff from buying is $\mathbb{E}(|\theta|) - 2c - P < 0$.

To see we still have the top-1 equilibrium, let m_1 denote the message "Attribute 1 is my highest attribute" and m_2 denote the message "Attribute 2 is my highest attribute." In the top-1 equilibrium, the sender with type $(1, 0)$ sends m_1 with probability 1, type $(0, 1)$ sends m_2 with probability 1 and the type $(0, 0)$ and $(1, 1)$ uniformly randomize over these two messages while the receiver checks the recommended attribute and buys if and only if it is good. Clearly, the sender is incentive compatible since she already maximizes the probability of buying. To see the receiver's strategy is a best response, note that upon receiving m_1, the receiver's expected payoff from checking attribute 1 is

$$\frac{1}{4} \times 0 + \frac{3}{4}(1 + \frac{1}{3} \times 1 - P) - c = 1 - \frac{3}{4}P - c,$$

and his expected payoff from checking attribute 2 is

$$\frac{3}{4} \times 0 + \frac{1}{4}(1 + 1 - P) - c = \frac{1}{2} - \frac{1}{4}P - c.$$

When $P = 1$, these two expected payoffs are both equal to $1/4 - c$. Therefore, if $c < 1/4$, the receiver strictly prefers checking exactly one attribute over checking no attribute or both attributes. Similar argument for the case where the receiver obtains m_2. Hence, if the cost of verification is sufficiently small ($c < 1/4$ in this setting), top-1 equilibrium still exists.

In general, when the verification is costly, new issues will arise. For example, whether the receiver checks at all and if he does how many attributes he checks may also depend on the cost, etc. Clearly, this makes the analysis more involved and we leave it to future work. Here, we would like to mention several papers that touch upon related issues. Ben-Porath et al. (2014) characterize a favored-agent mechanism to allocate an indivisible good among a group of senders where monetary transfers are not allowed and the receiver can learn each sender's type at a given cost. In a similar model, Erlanson and Kleiner (2017) study optimal mechanism for the principal in collective choice problems. They show that this mechanism can be implemented as a weighted majority voting rule. In contrast, Mylovanov and Zapechelnyuk (2017) study the allocation of an indivisible prize among multiple agents in a setting where the principal learns the true value from allocating the prize ex post, namely, after the allocation decision has been made.

5.3 Equilibrium Refinement

As we have already seen, the top equilibria are not even unique among the symmetric equilibria, although we still think the top equilibria are the most natural equilibria. Therefore, we suffer from a plethora of equilibria as other cheap talk games.

As for further equilibrium refinements, since θ has full support, each node at which the sender selects a message is on the equilibrium path. Since for all messages $A \in \mathcal{M}$, we have $m(A|\theta) > 0$ for some θ, it follows that each information set at which the receiver chooses a set of attributes B to check is on the equilibrium path. The subtlety arises at the information sets indexed by (y, B, A) where the receiver chooses whether to buy the good. Such a node may be off the equilibrium path either because $c(B|A) = 0$ and hence it is inconsistent with the receiver's strategy, or because $m(A|\theta) = 0$ for all θ such that $\theta_B = y$, and hence it is inconsistent with the sender's strategy. As long as a node is consistent with the sender's strategy, there is a natural way to update the receiver's beliefs if he totally mixes. So it is natural to impose either sequential equilibrium or trembling-hand perfect equilibrium. However, even these refinements will not place any restrictions on the receiver's beliefs at nodes where the sender has trembled.

To be clear on this, let's see an example. Suppose $N = 3$ and $n = 2$. For simplicity, suppose the sender is playing the top-1 strategy. Suppose the realized type is $(0, 1, 0)$ and the sender deviates by pointing to attribute 1. At first, the receiver does not know that the sender has deviated because he believes that attribute 1 could be one of the highest attributes. Suppose the receiver's strategy prescribes that, following this message (i.e., "Attribute 1 is my highest attribute"), he checks the recommended attribute 1 and also the unrecommended attribute 2 (recall $n = 2$). Upon seeing that the first attribute is 0 and the second attribute is 1, the receiver now knows that the sender deviated. The receiver also knows the realizations of the first two attributes, but he must form a belief about the third. To form his belief, as a rational Bayesian, the receiver compares the probabilities of (i) $\theta = (0, 1, 0)$ and the sender deviated by pointing to attribute 1; and (ii) $\theta = (0, 1, 1)$ and the sender deviated by pointing to attribute 1. The relative likelihood of these two events is not pinned down because either of them is a "tremble." To pin down the belief, we have to take a position on the relative likelihood of different trembles. All we can say for sure is that the receiver knows $|\theta|$ is either 1 or 2, but the relative probability she assigns to 1 and 2 is undetermined. The conditional expectation of $|\theta|$ can be any number in $[1, 2]$.

5.4 Real Cheap Talk?

Although there is no intrinsic content of the "cheap talk" message, we take the intuitive interpretation that the sender announces her k highest attributes in a top-k equilibrium. The information disclosure seems verifiable. However, this is not true since the sender's message does not indicate any ordering among those k attributes and the receiver cannot outright observe the true state of the world. That is, in our model, the sender has no evidence to present. This clarifies the difference with the evidence game defined by Hart et al. (2017) where each type of the sender is characterized by a set of verifiable statements from which the sender chooses. As different messages are available for different types of the sender, the messages amount to evidence. Nevertheless, in our model all the messages are available for all types of the sender and hence the messages are "cheap talk."

6 Conclusion

As new technologies revolutionize the collection, storage, and analysis of data, information is becoming an increasingly important commodity. In this paper, we study the scope for persuasion in a static environment where a sender who is perfectly informed about the state costlessly transmits a message to a receiver, who then chooses which aspect of the state to verify. Specifically, we consider a sender-receiver game with a multi-dimensional state. The sender observes the state and costlessly sends a message to the receiver, who selects *some* components of the state to check and then chooses a binary action. Even when the sender always prefers one action independent of the state, we show that for a range of state-dependent preferences for the receiver, there exists a natural family of informative equilibria. In these equilibria, the sender indicates which attributes are highest; the receiver checks some of those attributes and then chooses his action based on their realizations. Across the family of equilibria we construct, the receiver faces a trade-off between the frequency of seeing a good signal and the strength of that signal: a good signal is not representative of good quality when it is observed too often.

We find that compared to alternative communication structures, in the equilibria we characterize even though the sender has state-independent preferences, she can strictly benefit from the ability to communicate. If the sender has commitment power, as in Kamenica and Gentzkow (2011), then she can further increase her utility by committing to randomize between various messages. This is not so surprising by the Bayesian persuasion literature. However, we find that in some non-generic cases, the sender cannot benefit from commitment any more and costless message (cheap talk) can do as well as committing to a signal structure (Bayesian persuasion). Finally, we observe that the receiver's ability to partially verify the state has an ambiguous effect on the sender's utility unless the sender has commitment power, in which case verification can only restrict the set of posteriors the sender can induce.

References

Ambrus, A. and S. Takahashi (2008). Multi-sender cheap talk with restricted state spaces. *Theoretical Economics 3*, 1–27.

Battaglini, M. (2002). Multiple referrals and multidimensional cheap talk. *Econometrica 70*(4), 1379–1401.

Ben-Porath, E., E. Dekel, and B. L. Lipman (2014). Optimal allocation with costly verification. *The American Economic Review 104*(12), 3779–3813.

Carroll, G. and G. Egorov (2017). Strategic communication with minimal verification. Working paper.

Chakraborty, A. and R. Harbaugh (2007). Comparative cheap talk. *Journal of Economic Theory 132*(1), 70–94.

Chakraborty, A. and R. Harbaugh (2010). Persuasion by cheap talk. *The American Economic Review 100*(5), 2361–2382.

Crawford, V. P. and J. Sobel (1982). Strategic information transmission. *Econometrica: Journal of the Econometric Society*, 1431–1451.

Erlanson, A. and A. Kleiner (2017). Costly verification in collective decisions. Working paper.

Glazer, J. and A. Rubinstein (2004). On optimal rules of persuasion. *Econometrica 72*(6), 1715–1736.

Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *The Journal of Law and Economics 24*(3), 461–483.

Hart, S., I. Kremer, and M. Perry (2017). Evidence games: truth and commitment. *The American Economic Review 107*(3), 690–713.

Kamenica, E. and M. Gentzkow (2011). Bayesian persuasion. *The American Economic Review 101*(6), 2590–2615.

Kartik, N. (2009). Strategic communication with lying costs. *The Review of Economic Studies 76*(4), 1359–1395.

Levy, G. and R. Razin (2007). On the limits of communication in multidimensional cheap talk: a comment. *Econometrica 75*(3), 885–893.

Lipnowski, E. and D. Ravid (2017). Cheap talk with transparent motives. Working paper.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, 380–391.

Mylovanov, T. and A. Zapechelnyuk (2017). Optimal allocation with ex post verification and limited penalties. *The American Economic Review 107*(9), 2666–94.

Spence, M. (1973). Job market signalling. *Quarterly Journal of Economics 87*(3), 355–374.

Appendix

A Proofs for Section 3

Theorem 1. *Let $k \in [N]$. There exists a top-k equilibrium if and only if*

$$P \leq \mathbb{E}[v(\theta) \mid (\bar{S}_1^k, \underline{S}_1^{N-k}) = (1, 0)].$$

Proof. Let $n = 1$ and apply Theorem 2. □

Proposition 1. *Let $v(\theta) = |\theta|$. There exists a top-1 equilibrium if and only if*

$$P \leq \frac{N\mathbb{E}(|\theta|) - \mathbb{E}(|\theta|^2)}{N(1 - \pi_0) - \mathbb{E}(|\theta|)}.$$

Proof. Let $v(\theta) = |\theta|$ and $n = 1$. By routine calculation, we have $\mathbb{E}[v(\theta) \mid (\bar{S}_1^k, \underline{S}_1^{N-k}) = (1, 0)] = \frac{N\mathbb{E}(|\theta|) - \mathbb{E}(|\theta|^2)}{N(1 - \pi_0) - \mathbb{E}(|\theta|)}$. Apply Theorem 1. □

Proposition 2. *The price threshold $\bar{P}_k$ is strictly increasing in k.*

Proof. Let $n = 1$ and apply Proposition 6. □

Proposition 3. *The equilibrium buying probability $\mathbb{E}[Z_k]$ is decreasing in k but the conditional expected utility of the receiver $\mathbb{E}[v(\theta) \mid Z_k = 1]$ is increasing in k.*

Proof. For the first statement, note that

$$\mathbb{E}[Z_k] = \frac{\mathbb{E}[k \wedge |\theta|]}{k} = \sum_{i=1}^{k} \pi_i \cdot \frac{i}{k} + \pi_{k+1} + \cdots + \pi_N.$$

Then we have

$$\begin{aligned}
\mathbb{E}[Z_{k+1}] &= \sum_{i=1}^{k+1} \pi_i \cdot \frac{i}{k+1} + \pi_{k+2} + \cdots + \pi_N \\
&= \sum_{i=1}^{k} \pi_i \cdot \frac{i}{k+1} + \pi_{k+1} + \pi_{k+2} + \cdots + \pi_N \\
&< \sum_{i=1}^{k} \pi_i \cdot \frac{i}{k} + \pi_{k+1} + \cdots + \pi_N = \mathbb{E}[Z_k].
\end{aligned}$$

The inequality holds for any $k \in \{1, \ldots, N\}$. Therefore, $\mathbb{E}[Z_k]$ is decreasing in k.

For the second statement, let $\tilde{v} : \{0, \dots, N\} \to \mathbb{R}$ such that $\tilde{v}(i) = v(\theta)$ if $|\theta| = i$.

$$\mathbb{E}[v(\theta)|Z_k = 1] = \frac{\mathbb{E}[v(\theta)\mathbb{1}(Z_k = 1)]}{\Pr(Z_k = 1)} = \frac{\sum_{i=1}^{N} \tilde{v}(i) \cdot \Pr(|\theta| = i) \cdot \Pr(Z_k = 1 \mid |\theta| = i)}{\mathbb{E}[Z_k]}$$
$$= \frac{\sum_{i=1}^{k} \tilde{v}(i) \cdot \pi_i \cdot \frac{i}{k} + \sum_{i=k+1}^{N} \tilde{v}(i) \cdot \pi_i \cdot 1}{\sum_{i=1}^{k} \pi_i \cdot \frac{i}{k} + \pi_{k+1} + \cdots + \pi_N}$$
$$= \frac{\frac{k}{k+1} \sum_{i=1}^{k} \tilde{v}(i) \cdot \pi_i \cdot \frac{i}{k} + \frac{k}{k+1} \sum_{i=k+1}^{N} \tilde{v}(i) \cdot \pi_i}{\frac{k}{k+1} \sum_{i=1}^{k} \pi_i \cdot \frac{i}{k} + \frac{k}{k+1} \sum_{i=k+1}^{N} \pi_i}.$$

Then we have

$$\mathbb{E}[v(\theta)|Z_{k+1} = 1] = \frac{\sum_{i=1}^{k+1} \tilde{v}(i) \cdot \pi_i \cdot \frac{i}{k+1} + \sum_{i=k+2}^{N} \tilde{v}(i) \cdot \pi_i \cdot 1}{\mathbb{E}[Z_{k+1}]}$$
$$= \frac{\sum_{i=1}^{k} \tilde{v}(i) \cdot \pi_i \cdot \frac{i}{k+1} + \sum_{i=k+1}^{N} \tilde{v}(i) \cdot \pi_i}{\sum_{i=1}^{k} \pi_i \cdot \frac{i}{k+1} + \pi_{k+1} + \pi_{k+2} + \cdots + \pi_N}$$
$$= \frac{\frac{k}{k+1} \sum_{i=1}^{k} \tilde{v}(i) \cdot \pi_i \cdot \frac{i}{k} + \frac{k}{k+1} \sum_{i=k+1}^{N} \tilde{v}(i) \cdot \pi_i + \frac{1}{k+1} \sum_{i=k+1}^{N} \tilde{v}(i) \cdot \pi_i}{\frac{k}{k+1} \sum_{i=1}^{k} \pi_i \cdot \frac{i}{k} + \frac{k}{k+1} \sum_{i=k+1}^{N} \pi_i + \frac{1}{k+1} \sum_{i=k+1}^{N} \pi_i}.$$

Let

$$a := \frac{k}{k+1} \sum_{i=1}^{k} \tilde{v}(i) \cdot \pi_i \cdot \frac{i}{k} + \frac{k}{k+1} \sum_{i=k+1}^{N} \tilde{v}(i) \cdot \pi_i, \quad b := \frac{k}{k+1} \sum_{i=1}^{k} \pi_i \cdot \frac{i}{k} + \frac{k}{k+1} \sum_{i=k+1}^{N} \pi_i,$$
$$c := \frac{1}{k+1} \sum_{i=k+1}^{N} \tilde{v}(i) \cdot \pi_i, \quad d := \frac{1}{k+1} \sum_{i=k+1}^{N} \pi_i,$$
$$e := \frac{k}{k+1} \sum_{i=1}^{k} \tilde{v}(i) \cdot \pi_i \cdot \frac{i}{k}, \quad f := \frac{k}{k+1} \sum_{i=1}^{k} \pi_i \cdot \frac{i}{k}.$$

Then,

$$a = e + k \cdot c, \quad b = f + k \cdot d,$$
$$\mathbb{E}[v(\theta)|Z_k = 1] = \frac{a}{b}, \quad \mathbb{E}[v(\theta)|Z_{k+1} = 1] = \frac{a+c}{b+d}.$$

So we have $a \cdot (f + k \cdot d) = b \cdot (e + k \cdot c) \Leftrightarrow (ad - bc)k = be - af$.
Note that

$$\begin{aligned}
be - af = &\Big(\frac{k}{k+1}\sum_{i=1}^{k}\pi_i \cdot \frac{i}{k} + \frac{k}{k+1}\sum_{i=k+1}^{N}\pi_i\Big)\Big(\frac{k}{k+1}\sum_{i=1}^{k}\tilde{v}(i)\cdot\pi_i\cdot\frac{i}{k}\Big) - \\
&\Big(\frac{k}{k+1}\sum_{i=1}^{k}\tilde{v}(i)\cdot\pi_i\cdot\frac{i}{k} + \frac{k}{k+1}\sum_{i=k+1}^{N}\tilde{v}(i)\cdot\pi_i\Big)\Big(\frac{k}{k+1}\sum_{i=1}^{k}\pi_i\cdot\frac{i}{k}\Big) \\
&= (\frac{k}{k+1})^2\Big[\Big(\sum_{i=k+1}^{N}\pi_i\Big)\Big(\sum_{i=1}^{k}\tilde{v}(i)\cdot\pi_i\cdot\frac{i}{k}\Big) - \Big(\sum_{i=k+1}^{N}\tilde{v}(i)\cdot\pi_i\Big)\Big(\sum_{i=1}^{k}\pi_i\cdot\frac{i}{k}\Big)\Big] \\
&< (\frac{k}{k+1})^2\Big[\Big(\sum_{i=k+1}^{N}\pi_i\Big)\Big(\sum_{i=1}^{k}\tilde{v}(k+1)\cdot\pi_i\cdot\frac{i}{k}\Big) - \Big(\sum_{i=k+1}^{N}\tilde{v}(i)\cdot\pi_i\Big)\Big(\sum_{i=1}^{k}\pi_i\cdot\frac{i}{k}\Big)\Big] \\
&= (\frac{k}{k+1})^2\Big[\Big(\sum_{i=k+1}^{N}\tilde{v}(k+1)\cdot\pi_i\Big)\Big(\sum_{i=1}^{k}\pi_i\cdot\frac{i}{k}\Big) - \Big(\sum_{i=k+1}^{N}\tilde{v}(i)\cdot\pi_i\Big)\Big(\sum_{i=1}^{k}\pi_i\cdot\frac{i}{k}\Big)\Big] \\
&< 0.
\end{aligned}$$

The first inequality follows from the fact that $\tilde{v}(k+1) > \tilde{v}(i)$ for all $i \leq k$ and the last inequality follows from the fact that $\tilde{v}(k+1) < \tilde{v}(i)$ for all $i \geq k+2$. Therefore, we have

$$(ad - bc)k = be - af < 0 \Rightarrow ad < bc \Rightarrow \frac{a}{b} < \frac{a+c}{b+d} \Leftrightarrow \mathbb{E}[v(\theta)|Z_k = 1] < \mathbb{E}[v(\theta)|Z_{k+1} = 1].$$

The inequality holds for any $k \in \{1, \ldots, N\}$. Therefore, $\mathbb{E}[v(\theta)|Z_k = 1]$ is increasing in k.

□

Proposition 4. *The ranking of trading probability in the sender-optimal equilibrium of different settings is as follows:*

$$NC \leq CT \leq BP \text{ and } NCV \leq CTV \leq BPV \leq BP.$$

Proof. First, we show NC≤CT≤BP. If there is no verification, as we discussed in the main text, the only equilibrium in the cheap talk game is the babbling equilibrium in which the receiver ignores the sender's message and makes a purchase decision based on his prior belief about the sender's type. Given the receiver's strategy, the sender finds it not worthwhile to convey any information in the message she sends. This is due to the sender's state-independent preferences in the sense that she only aims to maximize the probability of buying. If there are two messages one of which induces a higher probability of buying, then all types of the sender will be pooling on that message. Therefore, cheap talk (CT) is equivalent to no communication (NC) between the two parties. And the probability of buying in NC is equal to that in CT.

If the sender can commit to a statistical experiment (BP), one option for the sender which is always available is an experiment that generates a single signal about the sender's type, i.e., there is actually no experiment. The receiver then makes a purchase decision after observing this signal. Hence, the probability of buying is determined by the two parties' common prior about the sender's type. Thus, the probability of buying induced by this experiment is equal to that in NC or CT. Clearly, the seller can be weakly better off by designing a non-degenerate experiment. Therefore, the probability of buying in BP is weakly higher than NC or CT.

Second, we show NCV≤CTV≤BPV≤BP. We start with showing NCV≤CTV. Clearly, $\mathbb{E}[v(\theta) \mid \theta_i = 0] \le \mathbb{E}[v(\theta)] \le \mathbb{E}[v(\theta) \mid \theta_i = 1]$ for any $i \in \{1, \ldots, N\}$. If $P \le \mathbb{E}[v(\theta) \mid \theta_i = 0]$, NCV is trivially equal to CTV since the price is too low so that the receiver always buys. Then consider $\mathbb{E}[v(\theta) \mid \theta_i = 0] < P$. We further restrict attention to prices that can sustain top-k equilibria, i.e., $P \le \bar{P}_k$. Then (i) if $\mathbb{E}[v(\theta) \mid \theta_i = 1] < P \le \bar{P}_k$, then the probability of buying in NCV is 0 but positive in CTV; (ii) if $P \le \bar{P}_k \le \mathbb{E}[v(\theta) \mid \theta_i = 1]$ or $P \le \mathbb{E}[v(\theta) \mid \theta_i = 1] \le \bar{P}_k$, then the receiver will buy only if she sees a "1" in NCV and the corresponding probability of buying is $\mathbb{E}(|\theta|)/N$. Recall that in CTV, the probability of buying of a top-k equilibrium is

$$\begin{aligned}\mathbb{E}[Z_k] &= \sum_{i=1}^{k} \pi_i \cdot \frac{i}{k} + \sum_{i=k+1}^{N} \pi_i \\ &> \sum_{i=1}^{k} \pi_i \cdot \frac{i}{N} + \sum_{i=k+1}^{N} \pi_i \cdot \frac{i}{N} = \frac{E(|\theta|)}{N}.\end{aligned}$$

Next, we show CTV≤BPV≤BP. If the sender can commit to a signal structure, let $b^{\mathrm{BP}} = (b_0, \ldots, b_N)$ be the buying vector under the optimal persuasion mechanism, where b_j denotes the probability of buying conditional on $|\theta|= j$. It can be computed by maximizing the sender's utility subject to the receiver's utility constraint. And it will be of the form $b^{\mathrm{BP}} = (0, \ldots, 0, \alpha, 1, \ldots, 1)$, where $\alpha \in (0, 1]$. Then the optimal persuasion mechanism must be split into two cases. If $b_1^{\mathrm{BP}} = 1$, then the sender can simply tell the receiver whether to buy or not. If $b_1^{\mathrm{BP}} < 1$, however, this strategy will not work, and the sender must instead point to an attribute and tell the receiver to buy if and only this attribute is one. The attribute should be chosen so that it is distributed uniformly and this behavior results in the desired buying vector.

Why this difference? The key is that the sender must find a way to achieve his optimal buying vector without giving the receiver any valuable information. If $b_1^{\mathrm{BP}} < 1$, then the optimal buying vector does not always require buying when there is some good attribute.

This just comes down to whether the top-1 strategy profile is an equilibrium. If it is, then under the optimal persuasion mechanism, the sender will always recommend buying when there is at least one attribute. If the top-1 strategy profile is not an equilibrium, then the sender will not always recommend buying even when there is some positive attributes. In this case, telling the agent to buy could lead into trouble because the receiver may be better off deviating and checking a random attribute, secure in the knowledge that he has dodged some very bad states.

In other words, with the extra instrument of commitment, the sender can always replicate the sender-optimal top equilibrium in CTV, and by manipulating her randomization device, the sender can achieve a (weakly) better buying vector without giving the receiver any more valuable information.

□

B Proofs for Section 4

Theorem 2. *Suppose $n \leq N-2$. For each $k = 1, \ldots, N$, there exists a top-k equilibrium if and only if*

$$P \leq \mathbb{E}[v(\theta) \mid (\bar{S}_n^k, \underline{S}_n^{N-k}) = (n \wedge k, 0)].$$

If $n = N-1$, it goes through except the "only if" statement holds for all but finitely many prices.[16]

To make the exposition more clear, we relabel the state θ as X and therefore $v(\theta) = v(|X|)$, and denote a specific realization of the random variable X by x.

In order to prove Theorem 2, a number of statistical lemmas are needed. To simplify the arguments below, we first construct a random variable as follows. Given a message strategy m, define an $\mathcal{M}$-valued random variable $M = M(m)$ on the same probability space as X, with the joint distribution of (X, M) determined by π and m. That is, the joint probability

[16]Specifically, when $n = N-1$ and $k \leq N-2$, we can also construct top-k equilibrium if $P = \mathbb{E}[v(\theta) \mid (\bar{S}_n^k, \underline{S}_n^{N-k}) = (n \wedge k, j)]$ for some $j = 1, \ldots, N-k-1$.

mass function is given by $f(x, A) = m(A|x)\pi(x)$. We will refer to M as the *random message induced by m*. Note, however, that the sender's strategy is still a function.

Next, we introduce some notations for the hypergeometric distribution. Let $\mathrm{HG}(n, K, N)$ denote the hypergeometric distribution when the sample size is n, the number of successes is K, and the total population size is N. Denote the corresponding probability mass function by

$$p(x; n, K, N) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}, \quad 1 \leq x \leq n \wedge K. \tag{1}$$

Lemma 1. *Let M be the random message induced by the strategy m^k. Conditional on M and $|X|$,*

(a) The random vectors X_M and $X_{[N]\setminus M}$ are independent.

(b) The components X_i for $i \in M$ are exchangeable, and for any $I \subset M$, we have $|X_I|\sim$ $\mathrm{HG}(|I|, k \wedge |X|, k)$.

(c) The components X_j for $j \in [N] \setminus M$ are exchangeable, and for any $J \subset [N] \setminus M$, we have $|X_J|\sim$ $\mathrm{HG}(|J|, |X| - k \wedge |X|, N - k)$.

Proof. By exchangeability, $\pi(x) = \pi(x')$ whenever $|x| = |x'|$. By construction, for any $A \in \mathcal{P}_k$,

$$\Pr((X, M) = (x, A)) = \begin{cases} \pi(x) / \binom{N - |x|}{k - |x|} & \text{if } |x_A| = k \wedge |x| = |x|, \\ \pi(x) / \binom{|x|}{k} & \text{if } |x_A| = k \wedge |x| = k, \\ 0 & \text{otherwise.} \end{cases}$$

Notice that this expression depends only on $|x|$ and $|x_A|$. Therefore, conditional on the event that $|X| = s$ and $M = A$, the vector $(X_A, X_{[N]\setminus A})$ is uniformly distributed over the set

$$\{x \in \{0,1\}^A \mid |x| = k \wedge s\} \times \{y \in \{0,1\}^{[N]\setminus A} \mid |y| = s - k \wedge s\}.$$

This proves (a), and parts (b) and (c) now follow from standard computations, which are omitted. $\square$

Next, we define random variables that are equal in distribution to $|X_I|$ and $|X_J|$ from Lemma 1. For $0 \leq a \leq b \leq N$, define the random variables $\bar{S}_a^b$ and $\underline{S}_a^b$ on the same probability space as (X, M) as follows. Conditional on X, these random variables are mutually independent of each other and M, with

$$\bar{S}_a^b \sim \mathrm{HG}(a, b \wedge |X|, b) \quad \text{and} \quad \underline{S}_a^b \sim \mathrm{HG}(a, |X| - (N - b) \wedge |X|, b).$$

The idea is that $\bar{S}^b_a$ ($\underline{S}^b_a$) is equal in distribution to the sum of a attributes chosen uniformly from b of the top (bottom) attributes.[17] To simplify the notations below, set $\bar{S}^b_a = \bar{S}^b_b$ and $\underline{S}^b_a = \underline{S}^b_b$ for $a > b$.

After checking the attributes in some set B, the receiver's information is summarized by the pair (X_B, M). We introduce a pair of simpler statistics. Let

$$T(X_B, M) = (|X_{B\cap M}|, |X_{B\setminus M}|) \quad \text{and} \quad U(M) = (|B \cap M|, |B \setminus M|). \tag{2}$$

In particular, T and U depend on the set B of attributes that the receiver checks. Note that the symbol $|\cdot|$ denotes the sum of components of vectors in T and the cardinality of sets in U.

Lemma 2. *Let M be the random message induced by the strategy m^k. Fix a nonempty subset B of $[N]$. Given the sample (X_B, M), the statistic (T, U) defined in (2) is sufficient for $|X|$.*

Proof. With p denoting the hypergeometric probability mass function (see (1)), Lemma 1 gives

$$\begin{aligned}&\Pr((X_B, M) = (x_B, A) \mid |X| = s)\\ &= \binom{s}{k \wedge s}^{-1} p(|x_{B\cap A}|; |B \cap A|, k \wedge s, k)p(|x_{B\setminus A}|; |B \setminus A|, s - k \wedge s, N - k).\end{aligned}$$

Since this expression depends on (x_B, A) only through

$$(T, U) = (|x_{B\cap A}|, |x_{B\setminus A}|, |B \cap A|, |B \setminus A|),$$

sufficiency follows. □

This statistic (T, U) satisfies an intuitive and useful monotonicity property. For any $u = (u_1, u_2) \in \operatorname{supp} U$, define the conditional support of T, given $U = u$, by

$$\mathbf{S}(u) = \{t \mid (t, u) \in \operatorname{supp}(T, U)\}.$$

Note that

$$\mathbf{S}(u) = \{(0, 0), (1, 0), \ldots, (u_1, 0), (u_1, 1) \ldots, (u_1, u_2)\},$$

which is obviously totally ordered by the product order on $\mathbb{Z}^2$. We will use the usual symbol $\geq$ to denote the product order; the meaning should be clear from the context.[18]

[17] To determine these top and bottom attributes, ties are broken uniformly.

[18] Recall that the product order means $(a, b) \geq (a', b')$ if and only if $a \geq a'$ and $b \geq b'$.

We are interested in whether one signal realization is "more favorable" than another in the spirit of Milgrom (1981).

Definition. *Given realizations* (t, u) *and* (t', u') *in* $\text{supp}(T, U)$, *we say* (t, u) *is* more favorable *than* (t', u'), *denoted* $(t, u) \succ_{\text{fav}} (t', u')$, *if the conditional distribution of* $|X|$ *given* $(T, U) = (t, u)$ *first-order stochastically dominates the conditional distribution of* $|X|$ *given* $(T, U) = (t', u')$, *for any strictly positive distribution of* $|X|$.[19]

Lemma 3. *Fix a nonempty subset* B *of* $[N]$. *For any* $(t, u), (t', u) \in \text{supp}(T, U)$, $t > t'$ *implies* $(t, u) \succ_{\text{fav}} (t', u)$.

Proof. Fix $(t, u), (t', u) \in \text{supp}(T, U)$ with $t > t'$. Let $f(\cdot|\cdot)$ denote the conditional probability mass function of (T, U) given $|X|$. Following the argument of Milgrom (1981), it suffices to show that

$$f(t, u|s)f(t', u|s') \geq f(t, u|s')f(t', u|s),$$

for all $s > s'$ with strict inequality for some $s > s'$.[20] Divide each side by $f(u|s)f(u|s') > 0$ and then plug in the expressions for f to obtain

$$\begin{aligned} &p(t_1; u_1, k \wedge s, k)p(t'_1; u_1, k \wedge s', k) \\ &\qquad \cdot p(t_2; u_2, s - k \wedge s, N - k)p(t'_2; u_2, s' - k \wedge s', N - k) \\ &\geq p(t'_1; u_1, k \wedge s, k)p(t_1; u_1, k \wedge s', k) \\ &\qquad \cdot p(t'_2; u_2, s - k \wedge s, N - k)p(t_2; u_2, s' - k \wedge s', N - k). \end{aligned}$$

We will use the fact that for the hypergeometric family of distributions (parametrized by the number of successes) has a strictly monotone likelihood ratio. Suppose $s > s'$. Then $k \wedge s \geq k \wedge s'$ with strict inequality if $s' < k$; likewise, $s - k \wedge s \geq s' - k \wedge s'$ with strict inequality if $s > k$. Comparing the first two terms on each side of the inequality and then the last two terms on each side gives the weak inequality. We claim that the inequality is strict if $s = t_1 + t_2$ and $s' = t'_1 + t'_2$, which satisfies $s > s'$ since $t > t'$. Clearly, all terms on the LHS are strictly positive. We separate into cases. If $t_1 > t'_1$ then $t'_1 < u_1 \wedge k$ and hence $t'_2 = 0$ so $s' < k$. Therefore, the product of the first two terms is strictly greater on the left than on the right. If $t_2 > t'_2$, then $t_2 > 0$ so $s > k$. Therefore, the product of the last two terms is strictly greater on the left than on the right. In either case, the inequality is strict. □

[19] Milgrom (1981) defines a weak and a stronger notion of favorability. Our definition lies between the two in strength.

[20] Milgrom (1981) uses a slightly different notion of favorability, but his argument can be easily modified to apply to our definitions.

With these statistical lemmas established, we now define the receiver's strategy as follows. If he can check all the attributes recommended by the sender, he does so (and does not check any others). If he cannot check all the recommended attributes, then he checks n of them. Then he buys if the number of ones he sees exceeds some threshold. Formally, let $c^*(A)$ be the uniform distribution over $\mathcal{P}_{n\wedge k}(A)$, the family of $(n \wedge k)$-element subsets of A, and let $b^*(y, B, A) = \mathbb{1}\{|y| \geq \bar{s}^*\}$ where the threshold $\bar{s}^*$ is defined by

$$\bar{s}^*(P, k, n) = (n \wedge k + 1) \wedge \min\Big\{s \in \{0, \dots, n \wedge k\} \mid \mathbb{E}[v(|X|) \mid \bar{S}_n^k = s] \geq P\Big\}.$$

Notice that under this strategy, the receiver buys the product when indifferent. We also define the analogous threshold $\underline{s}^*$ for unrecommended attributes by

$$\underline{s}^*(P, k, n) = (n \wedge (N-k)+1) \wedge \min\Big\{s \in \{0, \dots, n \wedge (N-k)\} \mid \mathbb{E}[v(|X|) \mid \underline{S}_n^{N-k} = s] \geq P\Big\}.$$

Now we turn to the proof proper.

Proof. First suppose the price inequality in the theorem statement holds, i.e., $P \leq \mathbb{E}[v(|X|) \mid (\bar{S}_n^k, \underline{S}_n^{N-k}) = (n \wedge k, 0)]$. We show that (m^k, c^*, b^*) is a top-k equilibrium. Suppose counterfactually that the receiver were allowed to check $n \wedge k$ recommended attributes and also $n \wedge (N-k)$ unrecommended attributes. By Lemmas 2 and 3, there exists a best response for the receiver that is a threshold strategy in T, i.e., buy if and only if $T \geq t^*$ for some $t^* \in \mathbf{S}(n \wedge k, n \wedge (N-k))$. Since $P \leq \mathbb{E}[v(|X|) \mid (\bar{S}_n^k, \underline{S}_n^{N-k}) = (n \wedge k, 0)]$, the threshold t^* can be chosen so that $t^* \geq (n \wedge k, 0)$. Therefore the threshold strategy is measurable with respect to $n \wedge k$ of the recommended attributes, and hence (c^*, b^*) achieves the same payoff as the best response in the counterfactual game. We conclude that (c^*, b^*) remains a best response in the actual game in which the receiver is only allowed to check n attributes in total, and that the receiver is playing a best response to sender's strategy.

Given the receiver's strategy (c^*, b^*), the sender's strategy is clearly a best response because it maximizes the probability of purchase. Finally, (m^k, c^*, b^*) is a top-k equilibrium because, in particular, the receiver buys with probability one whenever $|X| \geq n \wedge k$, which has positive probability since $n < N$ and X has full support.

Next, suppose the price inequality in the theorem statement does not hold. Suppose for a contradiction that there is a top-k equilibrium (m, c, b) with $m = m^k$. First assume

$$P \neq \mathbb{E}[v(|X|) \mid (\bar{S}_n^k, \underline{S}_n^{N-k}) = (n \wedge k, j)] \tag{3}$$

for all $j = 1, \dots, n \wedge (N-k)$. Following the same reasoning as above, it can be shown that under any best response, the receiver must check $n \wedge (N-k)$ of the unrecommended attributes and must buy if and only if $t_2 \geq t_2^*$ for some $t_2^* > 1$. Since this is a top-k equilibrium, we must have $t_2^* \leq n \wedge (N-k)$. But then the sender can profitably deviate when $|X| = k + \bar{s}^* - 1$. If the sender follows the top-k strategy and chooses A with $|X_A| = k$, then $|X_{[N] \setminus B}| = \bar{s}^* - 1$, and the receiver will never buy the good. If instead, the sender sends a message A such that $|X_A| = k - 1$, then $|X_{[N] \setminus A}| = \bar{s}^*$ and with positive probability, $|X_{B \setminus A}| = \bar{s}^*$ and the receiver buys. This contradiction completes the proof under the generic assumption (3).

Lastly, suppose

$$P = \mathbb{E}[v(|X|) \mid (\bar{S}_n^k, \underline{S}_n^{N-k}) = (n \wedge k, \hat{j})]$$

for some $\hat{j} \geq 1$. Now the receiver's optimal strategy is not uniquely pinned down because the buying probability is completely free when he sees j. However, the argument still goes through as before. If $n \geq N - k$, there is a special case to consider because there may exist (A, B) on the equilibrium path with $|B \setminus A| = N - k - 1$. Then the receiver's strategy is pinned down. He must buy if and only if $|y| \geq \underline{s}^* - 1$, at least along the equilibrium path. Suppose $|X| = k + s^* - 2$ and $|X_A| = k - 1$ and $X_i = 0$ for the unique attribute $i \notin A \cup B$. On the equilibrium path, the receiver will never buy. By sending message A instead, the receiver will buy with positive probability. This completes the proof.

$\square$

Corollary 1. *Theorem 2 still holds if the receiver can check the attributes sequentially rather than simultaneously.*

Proof. Clearly if the inequality is satisfied, the buyer is doing as well as she can. If it is violated, sequential checking must have the same result in the particular cases shown, and hence there is no benefit. $\square$

Proposition 5. *Both thresholds $\bar{s}^*(P, k, n)$ and $\underline{s}^*(P, k, n)$ are weakly increasing in P and n but weakly decreasing in k.*

Proof. Recall

$$\bar{s}^*(P, k, n) = (n \wedge k + 1) \wedge \min \Big\{ s \in \{0, \dots, n \wedge k\} \mid \mathbb{E}[v(|X|) \mid \bar{S}_n^k = s] \geq P \Big\},$$

and

$$\underline{s}^*(P, k, n) = (n \wedge (N-k) + 1) \wedge \min \Big\{ s \in \{0, \dots, n \wedge (N-k)\} \mid \mathbb{E}[v(|X|) \mid \underline{S}_n^{N-k} = s] \geq P \Big\}.$$

As P increases, the minima are taken over smaller sets and hence increase. For the comparative statics in n and k, it suffices to check that both conditional expectations in these definitions are weakly decreasing in n and weakly increasing in k. To simplify notation, let $\bar{f}_{n,k}(\cdot|\cdot)$ denote the conditional probability mass function of $\bar{S}_n^k$ given $|X|$. To show that $\mathbb{E}[v(|X|)|\bar{S}_n^k = s]$ is decreasing in n, it suffices to check that for all t and $s > s'$, we have

$$\bar{f}_{n,k}(t|s)\bar{f}_{n',k}(t|s') \leq \bar{f}_{n,k}(t|s')\bar{f}_{n',k}(t|s)$$

for all k and $n > n'$. To show that $\mathbb{E}[v(|X|) \mid \bar{S}_n^k = s]$ is decreasing in k, it suffices to check that for all t and $s > s'$, we have

$$\bar{f}_{n,k}(t|s)\bar{f}_{n,k'}(t|s') \geq \bar{f}_{n,k}(t|s')\bar{f}_{n,k'}(t|s)$$

for all n and $k > k'$. To establish the comparative statics in $\underline{s}^*(P, k, n)$, we simply prove the equivalent inequalities with $\underline{f}$ in place of $\bar{f}$, where $\underline{f}(\cdot|\cdot)$ denotes the conditional probability mass function of $\underline{S}_n^{N-k}$ given $|X|$. These inequalities can all be verified from the standard properties of the hypergeometric distribution listed in Appendix C.2.

□

Proposition 6. *For any fixed n, the price threshold $P^*(k, n)$ is strictly increasing in k.*

Proof. Fix n. For any k, let

$$\bar{f}(k|s) = \Pr(\bar{S}_n^k = n \wedge k \mid |X| = s), \quad \underline{f}(k|s) = \Pr(\underline{S}_n^{N-k} = 0 \mid |X| = s).$$

By conditional independence,

$$\bar{f}(k|s)\underline{f}(k|s) = \Pr((\bar{S}_n^k, \underline{S}_n^{N-k}) = (n \wedge k, 0) \mid |X| = s).$$

We need to show that for all $k > k'$, we have

$$\bar{f}(k|s)\underline{f}(k'|s') \geq \bar{f}(k'|s)\underline{f}(k|s'),$$

for all $s > s'$, with strict inequality for some $s > s'$ (which may depend on k). For strictness simply take $s = k$ and $s' = k'$. Then the LHS is unity but the RHS is strictly less than unity since $\underline{f}(k|s') < 1$. For the weak property, it suffices to check the monotone likelihood ratio property (MLRP) separately for $\bar{f}$ and $\underline{f}$. We begin with $\bar{f}$. By definition,

$$\bar{f}(k|s) = p(n \wedge k; n \wedge k, k \wedge s, k) = \frac{\binom{k \wedge s}{n \wedge k}}{\binom{k}{n \wedge k}}.$$

For convenience, we will work with s and $s - 1$ and k and $k - 1$. It suffices to prove the result when the terms are positive.

If $s \geq k$, then the result is trivial, so we may assume $s < k$ and and hence $s-1 < k-1$. Hence,

$$\begin{aligned}
\frac{\bar{f}(k|s)\bar{f}(k-1|s-1)}{\bar{f}(k|s-1)\bar{f}(k-1|s)} &= \frac{\binom{s}{n\wedge k}\binom{s-1}{n\wedge(n-1)}}{\binom{s-1}{n\wedge k}\binom{s}{n\wedge(k-1)}} \\
&= \frac{(s-1-n\wedge k)!\,(w-n\wedge(k-1))!}{(s-n\wedge k)!\,(s-1-n\wedge(k-1))!}, \\
&= \frac{s-n\wedge(k-1)}{s-n\wedge k} \\
&> 1.
\end{aligned}$$

Now we give the similar proof for $\underline{f}$. By definition,

$$\underline{f}(k|s) = p(0; n\wedge(N-k), s-k\wedge s, N-k) = \frac{\binom{N-k-(s-k\wedge s)}{n\wedge(N-k)}}{\binom{N-k}{n\wedge(N-k)}}.$$

The result is trivial if $s \leq k$, so we may assume $s > k$. Hence,

$$\begin{aligned}
\frac{\underline{f}(k|s)\underline{f}(k-1|s-1)}{\underline{f}(k|s-1)\underline{f}(k-1|s)} &= \frac{\binom{N-s}{n\wedge(N-k)}\binom{N-s+1}{n\wedge(N-k+1)}}{\binom{N-s+1}{n\wedge(N-k)}\binom{N-s}{n\wedge(N-k+1)}} \\
&= \frac{(N-s+1-n\wedge(N-k))!\,(N-s-n\wedge(N-k+1))!}{(N-s-n\wedge(N-k))!\,(N-s+1-n\wedge(N-k+1))!} \\
&= \frac{N-s+1-n\wedge(N-k)}{N-s+1-n\wedge(N-k+1)} \\
&> 1.
\end{aligned}$$

This completes the proof.

□

C Statistical Background

C.1 Exchangeability

A vector of random variables $X = (X_i)_{i\in I}$ (with finite index set I) is called exchangeable if

$$(X_i)_{i\in I} =_d (X_{\tau(i)})_{i\in I}$$

for any permutation τ on I.

Lemma 4. *Suppose* $X = (X_i)_{i \in I}$ *is exchangeable. Then for any* $k \in I$*,*

$$Cov(X_k, \sum_{i \in I} X_i) \geq 0.$$

Proof. By exchangeability, $\text{Cov}(X_k, \sum_{i \in I} X_i)$ is independent of k. Therefore,

$$0 \leq \text{Var}(\sum_{i \in I} X_i) = \sum_{j \in I} \text{Cov}(X_j, \sum_{i \in I} X_i) = |I| \text{Cov}(X_k, \sum_{i \in I} X_i)$$

for any $k \in I$, as needed. □

When the components of X are binary, Lemma 4 implies that observing a high realization of one component can only increase the expected value of the sum $\sum_{i \in I} X_i$.

Lemma 5. *Suppose* $X = (X_i)_{i \in I} \in \{0, 1\}^I$ *is exchangeable and*

$$0 < Pr(X_1 = 1) < 1.$$

Then for any $k \in I$*,*

$$\mathbb{E}[\sum_{i \in I} X_i \mid X_k = 0] \leq \mathbb{E}[\sum_{i \in I} X_i] \leq \mathbb{E}[\sum_{i \in I} X_i \mid X_k = 1].$$

Proof. By exchangeability, $\Pr(X_k = 1)$ does not depend on k. Denote this common value by q. By assumption, $q \in (0, 1)$. By Lemma 4,

$$\begin{aligned} q\mathbb{E}[\sum_{i \in I} X_i] &= \mathbb{E}[X_k] \cdot \mathbb{E}[\sum_{i \in I} X_i] \\ &\leq \mathbb{E}[X_k \cdot \sum_{i \in I} X_i] \\ &= q\mathbb{E}[\sum_{i \in I} X_i \mid X_k = 1], \end{aligned}$$

where the outer equalities hold because $X_k \in \{0, 1\}$. Then we have $\mathbb{E}[\sum_{i \in I} X_i] \leq \mathbb{E}[\sum_{i \in I} X_i \mid X_k = 1]$ since $q \in (0, 1)$. Therefore,

$$\begin{aligned} \mathbb{E}[\sum_{i \in I} X_i] &= (1 - q)\mathbb{E}[\sum_{i \in I} X_i \mid X_k = 0] + q\mathbb{E}[\sum_{i \in I} X_i \mid X_k = 1] \\ &\geq (1 - q)\mathbb{E}[\sum_{i \in I} X_i \mid X_k = 0] + q\mathbb{E}[\sum_{i \in I} X_i] \end{aligned}$$

which implies

$$\mathbb{E}[\sum_{i\in I} X_i] - q\mathbb{E}[\sum_{i\in I} X_i] \geq (1-q)\mathbb{E}[\sum_{i\in I} X_i \mid X_k = 0],$$

so

$$(1-q)\mathbb{E}[\sum_{i\in I} X_i] \geq (1-q)\mathbb{E}[\sum_{i\in I} X_i \mid X_k = 0].$$

Then we have $\mathbb{E}[\sum_{i\in I} X_i \mid X_k = 0] \leq \mathbb{E}[\sum_{i\in I} X_i]$ since $q \in (0,1)$. This completes the proof. □

C.2 Hypergeometric Distribution

We now check that the hypergeometric distribution satisfies the needed monotone likelihood ratio properties (MLRP). We are interested in the hypergeometric probability mass function

$$p(x; n, K, N) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}, \quad 1 \leq x \leq n \wedge K.$$

We claim that the hypergeometric distribution satisfies the MLRP with respect to n and K and the reverse MLRP with respect to N. In words, no matter your prior over one of the parameters, observing a larger realization of X will always cause you to update your beliefs in the direction of larger n, larger K, but smaller N. To see this, we will use the shorthand $f(x|n)$ when K and N are to be held fixed, and likewise for $f(x|K)$ and $f(x|N)$.

We begin with K. We need to show that for all $x > x'$ and $K > K'$,

$$f(x|K)f(x'|K') \geq f(x'|K)f(x|K').$$

It suffices to verify this for $x' = x-1$. Suppose $x' = x-j$. For each $i = 0, \dots, j-1$, we have

$$f(x-i|K)f(x-i-1|K') \geq f(x-i-1|K)f(x-i|K').$$

Taking the product of these j inequalities gives a telescoping product. All terms except those involving x and $x' = x-j$ cancel and we are left with the desired inequality

$$f(x|K)f(x'|K') \geq f(x'|K)f(x|K').$$

In each case the support shifts in the desired direction so that if $f(x|K) = 0$, then either x is above the support in which case $f(x|K') = 0$, or x is below the support in which case $f(x'|K) = 0$. Therefore, it suffices to verify the desired inequality when all terms are positive.

Thus, it suffices to check the following three inequalities:

$$\frac{f(x|K)f(x-1|K')}{f(x|K;)f(x-1|K)} \geq 1,$$
$$\frac{f(x|n)f(x-1|n')}{f(x|n')f(x-1|n)} \geq 1,$$
$$\frac{f(x|N)f(x-1|N')}{f(x|N')f(x-1|N)} \leq 1.$$

We now prove these in turn. In each case, we can simplify the ratio by noting that the only terms that do not cancel are the factorials involving both x and the parameter of interest.

For the MLRP in K, we have

$$\begin{aligned}
&\frac{f(x|K)f(x-1|K')}{f(x-1|K)f(x|K')} \\
&= \frac{(K-x+1)!\,(N-K-n+x-1)!\,(K'-x)!\,(N-K'-n+x)!}{(K-x)!\,(N-K-n+x)!\,(K'-x+1)!\,(N-K'-n+x-1)!} \\
&= \frac{K-x+1}{K'-x+1}\frac{N-K'-n+x}{N-K-n+x} \\
&> 1,
\end{aligned}$$

assuming all terms appearing in this expression are positive.

For the MLRP in n, we have

$$\begin{aligned}
&\frac{f(x|n)f(x-1|n')}{f(x-1|n)f(x|n')} \\
&= \frac{(n-x+1)!\,(N-K-n+x-1)!\,(n'-x)!\,(N-K-n'+x)!}{(n-x)!\,(N-K-n+x)!\,(n'-x+1)!\,(N-K-n'+x-1)!} \\
&= \frac{n-x+1}{n'-x+1}\frac{N-K-n'+x}{N-K-n+x} \\
&> 1.
\end{aligned}$$

For the MLRP in N, we have a simpler expression:

$$\begin{aligned}
&\frac{f(x|N)f(x-1|N')}{f(x-1|N)f(x|N')} \\
&= \frac{(N-K-n+x-1)!\,(N'-K-n+x)!}{(N-K-n+x)!\,(N'-K-n+x-1)!} \\
&= \frac{N'-K-n+x}{N-K-n+x} \\
&< 1.
\end{aligned}$$

This completes the proof. But for much of the analysis we do in the paper, we are actually interested in a different relationship between the parameters. Before we showed that seeing a higher outcome is more indicative of a higher value of K. But if we see the same outcome, a lower value of n is more indicative of a higher value of K.

We will use the following MLRP of the hypergeometric distribution.

Lemma 6. *For all* $0 \leq K' < K \leq N$, *the extended-real-valued likelihood ratio*

$$\frac{p(x; M, K, N)}{p(x; M, K', N)} = \frac{\binom{K}{x}\binom{N-K}{M-x}}{\binom{K'}{x}\binom{N-K'}{M-x}}$$

is increasing in x *over the range* $[(K' + M - N)_+, M \wedge K]$, *where either the numerator or denominator is positive.*

Proof. Within this range, the likelihood ratio is zero if $x < (K + M - N)_+$ and infinite if $x > M \wedge K'$, so we restrict attention to the range

$$[(K + M - N)_+, M \wedge K'].$$

For integers x such that $(K + M - N)_+ + 1 \leq M \wedge K'$, some algebra gives

$$\frac{p(x; M, K, N)}{p(x; M, K', N)} \Bigg/ \frac{p(x-1; M, K, N)}{p(x-1; M, K', N)} = \frac{K - x + 1}{K' - x + 1} \cdot \frac{N - K' - M + x}{N - K - M + x} > 1,$$

so the proof is complete. □

D Other Equilibria

D.1 Receiver-Optimal Equilibrium

Since payoffs depend only on the state and whether the product is purchased. Therefore, the payoff-relevant outcome is the buying function $b \colon \Theta \to [0, 1]$, where $b(\theta)$ is the probability of buying given that the realized state is θ. By symmetry, payoffs depend only on $|\theta|$, not θ. Therefore, we can focus on the *buying vector* $b = (b_0, \ldots, b_N)$, where b_j denotes the probability of buying conditional on $|\theta| = j$, which can be computed from the buying function:

$$b_j = \binom{N}{j}^{-1} \sum_{\theta : |\theta| = j} b(\theta). \tag{4}$$

Let $\tilde{v} : \{0, \ldots, N\} \to \mathbb{R}$ such that $\tilde{v}(i) = v(\theta)$ if $|\theta|= i$. The payoffs from a buying vector b, under the state distribution π, are

$$u_S(b, \pi) = \sum_{j=0}^{N} \pi_j b_j, \qquad u_R(b, \pi) = \sum_{j=0}^{N} \pi_j b_j (\tilde{v}(j) - P).$$

Ideally we would like to characterize which buying vectors b can be induced by some equilibrium. A full characterization is not possible with our current methods, but we can obtain bounds on the equilibrium buying vectors by imposing the incentive constraints for the sender and the receiver.

Recall that the receiver can always choose an attribute to check randomly and can then buy if this randomly chosen attribute is good. This results in the buying vector $\bar{b} = (0, 1/N, 2/N, \ldots, 1)$. Let $\bar{u} = u_R(\bar{b}, \pi)$. The receiver can also choose to never buy, yielding a utility of 0. A fundamental constraint on any equilibrium buying vector is that the receiver's utility must be at least $\bar{u}_+$. Of course this is a fairly crude lower bound. The sender's messages in equilibrium may reveal valuable information to the receiver that make new, more profitable deviations available.

The sender's incentive constraints prove much more useful. We can establish the following bound by considering a particular class of symmetric deviations by the sender. Namely, any type can uniformly mimic the types that have the same number of good attributes, plus one more.

Lemma 7. *If a buying vector b in $[0, 1]^{N+1}$ is induced by an equilibrium, then $b_{j-1} \geq (j - 1)b_j/j$ for each $j \in [N]$.*

Proof. Fix an equilibrium f, and let b denote the associated buying function. The buying vector is given in (4). Fix $j \in [N]$, and consider some state θ with $|\theta|= j$.

We can think of the state θ in two equivalent ways. As a vector, we have $\theta = (\theta_1, \ldots, \theta_N)$, where θ_i equals 1 if the i-th attribute is good and 0 if the i-th attribute is bad. Alternatively, we can think of θ as a subset of $[N]$, where $i \in \theta$ if and only if the i-th attribute is good. In the following argument, we take the latter interpretation.

In the main text, we consider an abstract message space and define strategies in the standard way. That is, the sender chooses a map from states to messages, and the receiver chooses a map from messages to actions. This is a useful way to describe particular equilibria. For the following argument, it is helpful to apply some insights from the revelation principle of mechanism design.

Consider an equilibrium. By identifying each equilibrium message with the action it induces the receiver to take, we can represent an equilibrium by a stochastic mapping from states to actions. Intuitively, consider any arbitrary equilibrium. Then we can simply pool all the messages that induce the same action for the receiver, and then relabel each message by its induced action. Therefore, if (c, b^0, b^1) is an action profile of the receiver that is chosen with positive density in equilibrium, then we can regard (c, b^0, b^1) as a message sent by the sender in equilibrium as well.

By the sender's incentive constraint,

$$b(\theta) = \sum_{i \in \theta} c_i b_i^1 + \sum_{i \in [N] \setminus \theta} c_i b_i^0.$$

For each $\ell \in \theta$, the sender's incentive compatibility constraint at $\theta \setminus \{\ell\}$ can send this message (c, b^0, b^1) and the receiver will then buy with probability

$$\sum_{i \in \theta \setminus \{\ell\}} c_i b_i^1 + \sum_{i \in [N] \setminus (\theta \setminus \{\ell\})} c_i b_i^0 = b(\theta) - c_\ell (b_\ell^1 - b_\ell^0).$$

The sender's incentive compatibility constraint at $\theta \setminus \{\ell\}$ implies that

$$b(\theta \setminus \{\ell\}) \geq b(\theta) - c_\ell (b_\ell^1 - b_\ell^0).$$

Summing over $\ell \in \theta$ gives

$$\sum_{\ell \in \theta} b(\theta \setminus \{\ell\}) \geq j b(\theta) - \sum_{\ell \in \theta} c_\ell (b_\ell^1 - b_\ell^0) \geq (j-1) b(\theta),$$

where last inequality follows from comparing the sum here with the expression for $b(\theta)$ above, and using the nonnegativity of b^0 and b^1. Recall that θ was an arbitrary state with j good attributes. Summing over all such states gives

$$\sum_{\theta : |\theta| = j} \sum_{\ell \in \theta} b(\theta \setminus \{\ell\}) \geq (j-1) \sum_{\theta : |\theta| = j} b(\theta).$$

The sum on the left side includes every type with $j-1$ good attributes exactly $N-j+1$ times, so we have

$$(N-j+1) \sum_{\theta' : |\theta'| = j-1} b(\theta') \geq (j-1) \sum_{\theta : |\theta| = j} b(\theta).$$

That is,

$$(N-j+1) \binom{N}{j-1} b_{j-1} \geq (j-1) \binom{N}{j} b_j.$$

Expanding the binomial coefficients and simplifying yields the desired inequality. □

Combining this lemma, which follows from the sender's incentive constraints, with the crude bound derived above from the receiver's constraint, we obtain the following theorem.

Theorem 3. *If a buying vector b in $[0,1]^{N+1}$ is induced by some equilibrium, then*

$$b_{j-1} \geq \frac{j-1}{j} b_j, \quad j = 1, \dots, N,$$
$$\sum_{j=0}^{N} \pi_j b_j (\tilde{v}(j) - P) \geq 0 \vee \sum_{j=0}^{N} \pi_j (j/N)(\tilde{v}(j) - P).$$

Recall again that the first set of inequalities comes from the sender's ability to imitate higher types and the second inequality holds because the receiver can always deviate to the strategy of uniform checking.

Also note that the first set of inequalities is independent of the distribution π, the valuation function v, and the price P. Let B denote the set of vectors $b \in [0,1]^{N+1}$ such that

$$b_{j-1} \geq \frac{j-1}{j} b_j, \quad j = 1, \dots, N.$$

Geometrically, the theorem states that any equilibrium buying vector must lie in B and additionally lie above some (π, v, P)-dependent hyperplane slicing through the set B.

By maximizing $\sum_{i=0}^{N} \pi_i b_i$ over the inequalities above, we get a linear program whose value is an upper bound on the sender's utility in any equilibrium. With a little trick, we can use this bound to show that the top equilibria are not Pareto dominated by any other equilibria.

Theorem 4. *Among all equilibria, the sender-optimal top equilibrium is strongly Pareto optimal.*

Proof. For each $j = 1, \dots, N$, let

$$e_j = (0/j, 1/j, \dots, (j-1)/j, j/j, 0, \dots, 0).$$

Then each vector b in B can be expressed uniquely as

$$b = \sum_{j=1}^{N} \alpha_j e_j$$

for some coefficients $a_j \in [0,1]$ such that

$$\sum_{j'=j}^{N} (j/j') \alpha_j \leq 1.$$

Specifically, $\alpha_j = b_j - jb_{j+1}/(j+1)$ for each j.

With this parametrization, maximizing the sender's utility subject to a constraint on receiver's utility and vice versa is given by a collection, so this can be solved by a greedy algorithm. So we have an exact representation of the Pareto frontier, and this applies even to asymmetric equilibria as well.

Specifically, suppose $\bar{u} \geq 0$, so that the receiver is willing to buy with some positive probability under the uniform strategy. Let k_R be the smallest value of k for which $u_R(e_k, \pi) \geq \bar{u}$. It can be shown that the Pareto frontier of B is precisely given by mixtures of top-k and top-k' equilibria for consecutive k and k' between 0 and k_R, where top-0 is interpreted as always buying. All of these are not incentive compatible though. Let k_S be the smallest value of k such that

$$u_R(e_N + N^{-1}e_{N-1} + \cdots + (k+1)^{-1}e_k, \pi) \geq \bar{u}.$$

By construction $k_S \leq k_R$. The top equilibria are precisely the top-k equilibria for k between k_S and k_R. The top-k_R equilibrium is the receiver-optimal equilibrium (Indeed he could not do any better through commitment, by Glazer and Rubinstein (2004)). The top-k_S equilibrium is the sender's best top equilibrium, but all these equilibria are Pareto efficient among *all* the equilibria.

□

This theorem shows that the receiver-optimal equilibrium (optimal among all equilibria, symmetric or asymmetric) is a top equilibrium. However, we cannot rule out that the sender can do slightly better under some other equilibria, and indeed we can construct asymmetric equilibria where this is the case.

D.2 Asymmetric Equilibrium

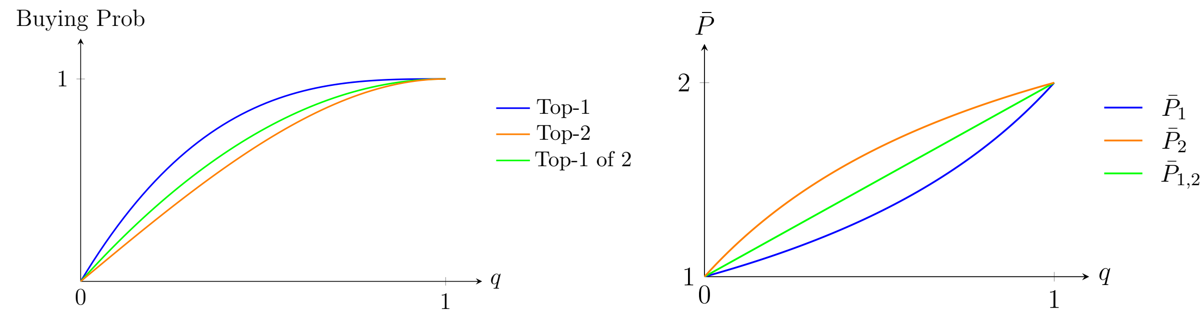
In the top-k equilibrium, the sender is pointing to the k highest attributes out of all the N attributes. Now we let the sender point to the k highest attributes among a prescribed subset consisting of K attributes where $K \leq N$. We could allow the sender to randomly choose K attributes first but for intuition let's say the sender points to the k highest attributes out of the first K attributes and ignores the remaining $N - K$ attributes. Recall that there is a trade-off between the frequency of seeing a good signal and how strong that signal can be. By doing this, the sender makes a signal stronger when $K < N$ than $K = N$. We call this strategy top-k of K strategy and use (k, K) as a shorthand. Then (k, N) corresponds to the regular top-k strategy we characterized before.

To be clear about the intuition, let's see an example. Suppose $N = 4$ and consider two types of the sender $(0,0,1,0)$ and $(0,0,1,1)$ with $\Pr(\theta = (0,0,1,0)) = p$ and $\Pr(\theta = (0,0,1,1)) = q$. We claim the signal in $(1, N-1)$ is stronger than $(1, N)$. To see why, note that in $(1, N)$, sender $(0,0,1,0)$ will point to attribute 3 with probability 1 and sender $(0,0,1,1)$ will point to attribute 3 with probability 1/2. Denote the message "my highest attribute of all attributes is attribute 3" by m_1 and "my highest attribute of the first three attributes is attribute 3" by m_2. And hence, $\frac{\Pr(m_1 \mid \theta=(0,0,1,0))}{\Pr(m_1 \mid \theta=(0,0,1,1))} = \frac{1}{1/2} = 2$. However, under $(1, N-1)$, both types will point to the third attribute with probability 1. We have $\frac{\Pr(m_2 \mid \theta=(0,0,1,0))}{\Pr(m_2 \mid \theta=(0,0,1,1))} = \frac{1}{1} = 1$. Then the receiver's posterior is: in $(1, N)$, $\frac{\Pr(\theta=(0,0,1,0) \mid m_1)}{\Pr(\theta=(0,0,1,1) \mid m_1)} = \frac{\Pr(m_1 \mid \theta=(0,0,1,0))}{\Pr(m_1 \mid \theta=(0,0,1,1))} \times \frac{\Pr(\theta=(0,0,1,0))}{\Pr(\theta=(0,0,1,1))} = 2\frac{p}{q}$ and in $(1, N-1)$, $\frac{\Pr(\theta=(0,0,1,0) \mid m_2)}{\Pr(\theta=(0,0,1,1) \mid m_2)} = \frac{\Pr(m_2 \mid \theta=(0,0,1,0))}{\Pr(m_2 \mid \theta=(0,0,1,1))} \times \frac{\Pr(\theta=(0,0,1,0))}{\Pr(\theta=(0,0,1,1))} = \frac{p}{q}$. So the receiver is more biased to $(0,0,1,1)$ when pointed to the highest attribute out of the first three attributes. It turns out that for fixed k, the signal strength (k, K) is decreasing in K and for fixed K, the signal strength (k, K) is increasing in k, so we have a partial order.

With respect to equilibrium buying probability, top-k of K equilibrium is the intermediate case between top-k equilibrium and top-K equilibrium. If we go back to the numerical example we used in Section 3.3.2, we can clearly see the price threshold and equilibrium buying probability of top-1 of 2 equilibrium are exactly between top-1 equilibrium and top-2 equilibrium.

Recall that in that numerical example, we let $N = 3$, $n = 1$, $v(\theta) = |\theta|$ and $|\theta| \sim \text{Bino}(N, q)$ where $0 < q < 1$. In a top-1 of 2 equilibrium, the sender ignores one attribute and points to the highest attribute of the remaining two. The receiver just checks the recommended attribute and buys unless both of those two attributes are bad. Thus, the equilibrium buying probability is $1-(1-q)^2$. And we know that the receiver's expected utility from equilibrium strategy should be weakly larger than that from checking a random attribute and buying if and only if that attribute is good. Therefore, the condition for the existence of top-1 of 2 equilibrium is $(1-(1-q)^2)(q + \frac{2q}{1-(1-q)^2} - P) \geq q(2q+1-P)$. So the highest price to sustain a top-1 of 2 equilibrium is $\bar{P}_{1,2} = 1+q$. Recall that the price threshold of top-1 and top-2 equilibrium is $\bar{P}_1 = \frac{2}{2-q}$ and $\bar{P}_2 = \frac{1+3q}{1+q}$, respectively. Clearly, $\bar{P}_1 < \bar{P}_{1,2} < \bar{P}_2$ for all $0 < q < 1$. We plot the equilibrium buying probability and price thresholds in Figure 5a and Figure 5b, respectively. It is clear that top-1 of 2 equilibrium is an intermediate case between top-1 and top-2 equilibrium.

Hence, at $\bar{P}_{1,2}$ top-1 equilibrium is not sustainable. The sender's most-preferred top equilibrium is top-2 equilibrium. However, the sender is strictly better off in top-1 of 2 equilibrium which is an asymmetric equilibrium.



(a) Equilibrium buying probability

(b) Price thresholds

Figure 5: Equilibrium Buying Probability and Price Thresholds

It is worth noting that the price threshold $\bar{P}_{1,2}$ in top-1 of 2 equilibrium coincides with the price $\hat{P}$ we used before (see Section 3.3.2). This implies we can replicate the outcome of top-1 of 2 equilibrium by a Bayesian persuasion with verification (BPV) setting. The intuition is as follows. It can be shown that at $\bar{P}_{1,2}$ (or $\hat{P}$), under BPV, the sender points to the highest attribute with probability 2/3 and the second highest attribute with probability 1/3. Under cheap talk, the sender will point to the highest attribute of two preselected attributes. It turns out that they are the same. Why? Suppose exactly one attribute is good (otherwise, the highest and second highest attributes are indistinguishable (either both one or both zero)). Under BPV, the sender will point to the good attribute with probability 2/3. Under cheap talk, the sender will point to the good attribute exactly when it is one of the two preselected attributes, which occurs with probability 2/3. Therefore, the buying probability as a function of $|\theta|$ is the same under both equilibria.

D.3 Other Symmetric Equilibria

This section constructs a parametric family of symmetric equilibria that are not top equilibria. We start with an example. Notice that it is sensitive to the parameters being "just right" in order to maintain indifference, so it is more fragile than the top equilibria. Moreover, in this example, there is a top equilibrium that induces a higher buying probability.

Suppose $N = 4$ and $n = 1$. The attributes are i.i.d. with success probability 1/2. So that the probabilities of $|\theta| = 0, 1, 2, 3, 4$ are

$$1/16, 1/4, 3/8, 1/4, 1/16$$

respectively. Set $P = 7/4$. Consider the equilibrium where the receiver asks the sender to indicate the two highest attributes (unordered) and also the third-highest attribute. The receiver will check one of the three highest attributes and buy if and only if it is one, but he will check each of the two highest attributes with probability $c \in (1/3, 1/2)$ and he will check the third attribute with probability $1 - 2c$.

Next, we verify that this is an equilibrium. When the receiver checks one of the two highest attributes, his utility is

$$(1/4)(1/2)(1 - 7/4) + (3/8)(2 - 7/4) \\ + (1/4)(3 - 7/4) + (1/16)(4 - 7/4) = 0.453$$

When the receiver checks the third highest attribute, his expected utility is

$$(1/4)(3 - 7/9) + (1/16)(4 - 7/4) = 0.453.$$

If the receiver picks randomly, then

$$(1/4)(1/4)(1 - 7/4) + (3/8)(1/2)(2 - 7/4) \\ + (1/4)(3/4)(3 - 7/4) + (1/16)(4 - 7/4) = 0.375.$$

Therefore, this constitutes an equilibrium.

Ultimately, however, this is not of much interest. We know the sender does strictly better in a top-2 equilibrium. However, the top-1 equilibrium

$$2 - (15/16)(7/4) = 0.359.$$

is not available.

Now we construct this parametric family of symmetric equilibria.

Equilibria that only provide comparative information about the attributes can never increase the buying probability above a top equilibrium. Some absolute information is needed. For equilibria in this family, there are two types of messages. The first message indicates the top $\underline{k}_1$ attributes and then the next $\bar{k}_1 - \underline{k}_1$ attributes; the receiver chooses from among the highest attributes with probability c_1, and among the next $\bar{k}_1 - \underline{k}_1$ with complementary probability $1 - c_1$. The second message indicates the top $\underline{k}_2$ attributes and then the next $\bar{k}_2 - \underline{k}_2$ attributes; the receiver choose from among the highest attributes with probability c_2 and among the next $\bar{k}_2 - \underline{k}_2$ with probability $1 - c_2$. No matter the message and the attribute checked, the receiver buys if and only if he sees a one, except in the edge case where $\underline{k}_1 = 0$, in which case the receiver buys upon checking the "zeroth-highest" attribute.

Types $t \geq \bar{k}_2$ are indifferent about which message they send as buying is guaranteed either way. Among types $t < \bar{k}_2$, it turns out that "extreme" types (either below $\underline{t}$ or above $\bar{t}$) will choose to send the first message and "moderate" types (between $\underline{t}$ and $\bar{t}$ inclusive) will choose to send the second message.

The cutoff values $\underline{k}_1, \bar{k}_1$ for the first message, $\underline{k}_2, \bar{k}_2$ for the second message, and $\underline{t}$ and $\bar{t}$ for the types must satisfy the following inequality:

$$0 \leq \underline{k}_1 < \underline{t} < \underline{k}_2 < \bar{t} < \bar{t} + 1 < \bar{k}_1 < \bar{k}_2 \leq N.$$

This implies $N \geq 6$, and in order to make the first and last inequalities strict, we must in fact have $N \geq 8$. The inequalities can be explained as follows. We analyze the inequalities from the outside in. The outermost (weak) inequalities are trivial. The next inequalities $\underline{k}_1 < \underline{t}$ and $\bar{k}_1 < \bar{k}_2$ are needed so that the agents do not just split between the messages as high types and low types. then the inequalities $\underline{t} < \underline{k}_2 < \bar{t}$ ensure that the quality of the attribute the receiver checks after receiving the second message is uncertain. Finally, the inequalities $\bar{t} + 1 < \bar{k}_1$ ensures that seeing a zero among the lower attributes of the first message does not guarantee that the type is below $\underline{t} - 1$.

The continuous parameters are the price P and the interior probabilities

$$c_1, c_2, \pi_0, \ldots, \pi_N \in (0, 1),$$

such that

$$\sum_t \pi_t = 1.$$

Technically, c_1 and c_2 are equilibrium parameters, while P and (π_t) are parameters of the environment. We will see that the incentive constraints imply $c_1 < c_2$ and $\underline{t} < P < \bar{t}$, but these constraints need not be included explicitly.

To simplify notation, let $\mathcal{M} = [\underline{t}, \bar{t}]$ and $\mathcal{E} = [0, \underline{t} - 1] \cup [\bar{t} + 1, \bar{k}_2 - 1]$, so $\mathcal{M} \cup \mathcal{E} = [0, \bar{k}_2 - 1]$. The types $t \geq \bar{k}_2$ will always buy in any equilibrium. Now we impose the equilibrium constraints.

- Attribute ordering:

$$c_1/\underline{k}_1 \geq (1 - c_1)/(\bar{k}_1 - \underline{k}_1), \qquad c_2/\underline{k}_2 \geq (1 - c_2)/(\bar{k}_2 - \underline{k}_2).$$

- Extreme sender:

$$c_1 + (1 - c_1)(\underline{t} - 1 - \underline{k}_1)/(\bar{k}_1 - \underline{k}_1) \geq c_2(\underline{t} - 1)/\underline{k}_2,$$
$$c_1 + (1 - c_1)(\bar{t} + 1 - \underline{k}_1)/(\bar{k}_1 - \underline{k}_1) \geq c_2 + (1 - c_2)(\bar{t} + 1 - \underline{k}_2)/(\bar{k}_2 - \underline{k}_2).$$

- Moderate sender:

$$c_2\underline{t}/\underline{k}_2 \geq c_1 + (1-c_1)(\underline{t}-\underline{k}_1)/(\bar{k}_1-\underline{k}_1),$$
$$c_2 + (1-c_2)(\bar{t}-\underline{k}_2)/(\bar{k}_2-\underline{k}_2) \geq c_1 + (1-c_1)(\bar{t}-\underline{k}_1)/(\bar{k}_1-\underline{k}_1).$$

- Receiver:

$$\sum_{t\in\mathcal{E}}(\tilde{v}(t)-P)\pi_t\left[\frac{t}{\underline{k}_1}\wedge 1 - \frac{(t-\underline{k}_1)_+}{\bar{k}_1-\underline{k}_1}\wedge 1\right] = 0,$$
$$\sum_{t\in\mathcal{M}}(\tilde{v}(t)-P)\pi_t\left[\frac{t}{\underline{k}_2}\wedge 1 - \frac{t-\underline{k}_2}{\bar{k}_2-\underline{k}_2}\wedge 1\right] = 0,$$

where $\tilde{v}(t) = v(\theta)$ such that $|\theta|= t$. A few changes are needed in the special case $\underline{k}_1 = 0$: the initial attribute ordering constraint is trivially satisfied (which is immediate upon expanding the product), and in the receiver's first constraint, $0/0$ must be replaced with 1.

Of course we also want to impose the constraint that the receiver does not want to always buy:

$$\sum_t(\tilde{v}(t)-P)\pi_t t/N > \sum_t(\tilde{v}(t)-P)\pi_t.$$

And we will want to compare to the top-k equilibria for $\underline{k}_1 < k \leq \underline{k}_2$. We will then impose the inequality that receiver cannot achieve the top-$(k-1)$ equilibria:

$$\sum_t(\tilde{v}(t)-P)\pi_t\frac{t}{N} > \sum_t(\tilde{v}(t)-P)\pi_t\frac{t}{k-1}\wedge 1.$$

Finally we will compare the buying probability in the top-k equilibrium

$$\sum_t \pi_t\frac{t}{k}\wedge 1$$

to the probability under the proposed equilibrium:

$$c_1\sum_{t\in\mathcal{E}}\pi_t\frac{t\wedge\underline{k}_1}{\underline{k}_1} + (1-c_1)\sum_{t\in\mathcal{E}}\frac{(t-\underline{k}_1)_+}{\bar{k}_1-\underline{k}_1}\wedge 1$$
$$+ c_2\sum_{t\in\mathcal{M}}\pi_t\frac{t\wedge\underline{k}_2}{\bar{k}_2-\underline{k}_2} + (1-c_2)\sum_{t\in\mathcal{M}}\pi_t\frac{t-\underline{k}_2}{\bar{k}_2-\underline{k}_2}\wedge 1 + \sum_{t\geq\bar{k}_2}\pi_t.$$

Formally, we will maximize this difference subject to the constraint that the top-$(k-1)$ strategy profile is not an equilibrium (which implies that always buying is not an equilibrium).

Suppose $k \leq \underline{k}_2$. The gain over the top-k equilibrium is

$$c_1 \sum_{t \in \mathcal{E}} \pi_t \left[\frac{t \wedge \underline{k}_1}{\underline{k}_1} - \frac{t \wedge k}{k} \right] + (1 - c_1) \sum_{t \in \mathcal{E}} \pi_t \left[\frac{(t - \underline{k}_1)_+}{\bar{k}_1 - \underline{k}_1} \wedge 1 - \frac{t \wedge k}{k} \right]$$
$$+ c_2 \sum_{t \in \mathcal{M}} \pi_t \left[\frac{t \wedge \underline{k}_2}{\bar{k}_2 - \underline{k}_2} - \frac{t \wedge k}{k} \right] + (1 - c_2) \sum_{t \in \mathcal{M}} \pi_t \left[\frac{t - \underline{k}_2}{\bar{k}_2 - \underline{k}_2} \wedge 1 - \frac{t \wedge k}{k} \right].$$

The key observation is that this reduces to a linear programming problem once P is fixed. Therefore, we will iterate over values of P, and for each fixed P solve the maximization problem. To make the inequalities weak, we will replace the strict inequalities with weak inequalities with a small tolerance $\varepsilon > 0$.

For fixed P, there are $N + 2$ parameters. We will group the constraints according to equalities with zero on the right side, inequalities with zero on the right side, inequalities with ε and the RHS and one inequality with 1 on the RHS.

E Correlation Between the Attributes

In this section, we are interested in how the pairwise correlation of the attributes affects the price threshold. Consider a conditional independence model for exchangeable binary random variables. Since each attribute follows the same Bernoulli distribution, suppose the common Bernoulli parameter p is random with cumulative distribution function F whose support is $[0, 1]$. Given p, the binary random variables $\theta_1, \ldots, \theta_N$ are conditionally i.i.d. We have

$$\Pr(\theta = \hat{\theta}) = \int_0^1 p^s (1 - p)^{N-s} dF(p)$$

where $s = |\hat{\theta}| \in \{0, \ldots, N\}$.

Furthermore, assume F is a Beta(α, β) distribution, with density

$$f(p) = [B(\alpha, \beta)]^{-1} p^{\alpha - 1} (1 - p)^{\beta - 1}$$

where $0 < p < 1$, $B(\alpha, \beta)$ is the beta function $(\alpha > 0, \beta > 0)$ to ensure that the total probability integrates to 1. Then we have

$$\Pr(\theta = \hat{\theta}) = \frac{B(\alpha + s, \beta + N - s)}{B(\alpha, \beta)}.$$

where $s = |\hat{\theta}| \in \{0, \ldots, N\}$.

The pairwise correlation coefficient is

$$\rho = \frac{1}{\alpha + \beta + 1}$$

which is a function of α and β. Therefore, we can vary the correlation of the attributes by changing the value of the distribution parameters α and β.

If we assume p follows a Beta(α, β) distribution, then $|\theta|$ has a Beta-binomial(α, β) distribution. The probability mass function is

$$\Pr(|\theta|= s) = \binom{N}{s} \frac{B(\alpha + s, \beta + N - s)}{B(\alpha, \beta)}$$

where $s = 0, \ldots, N$, and $\alpha, \beta > 0$. In particular, we have

$$\pi_0 = \Pr(\theta = \mathbf{0}) = \frac{B(\alpha, \beta + N)}{B(\alpha, \beta)}.$$

Moreover, by standard computation, we have

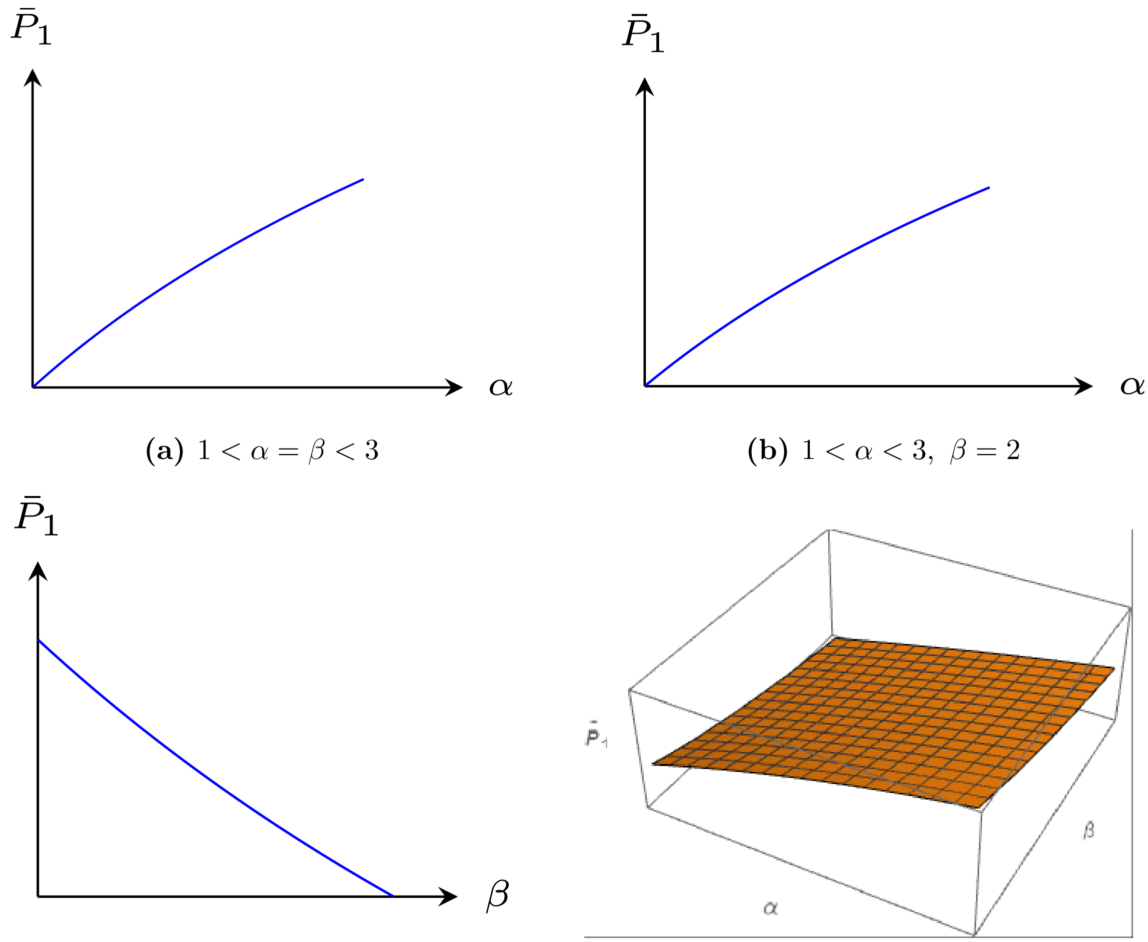
$$\mathbb{E}(|\theta|) = \frac{N\alpha}{\alpha + \beta} \quad \text{and} \quad \mathrm{Var}(|\theta|) = \frac{N\alpha\beta(\alpha + \beta + N)}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

Consider the special case in which $v(\theta) = |\theta|$. By Proposition 1, the price threshold reduces to

$$\bar{P}_1 = \frac{N\mathbb{E}(|\theta|) - \mathbb{E}(|\theta|^2)}{N(1 - \pi_0) - \mathbb{E}(|\theta|)}.$$

Note that $\mathbb{E}(|\theta|^2) = (\mathbb{E}(|\theta|))^2 + \mathrm{Var}(|\theta|)$. Hence, we can rewrite $\bar{P}_1$ as a function of only α and β. However, we are not able to have a closed form of the price threshold as a function of the correlation coefficient ρ because of the Beta functions. We will plot the relationship between $\bar{P}_1$ and α and β numerically.

In the following numerical example, let $N = 10$. Since the values of the two parameters α and β control the shape of the Beta distribution, we start with the case that the Beta density function is symmetric about $1/2$, i.e., $\alpha = \beta$. In particular, we vary their value between 1 and 3 which corresponds to symmetric unimodal Beta density functions. Notice that the pairwise correlation coefficient ρ is decreasing in α and β. But as Figure 6a shows, as α increases, the price threshold is increasing when the attributes become less and less correlated. Next, we fix β at 2 and increase α from 1 to 3. The Beta density function now is left skewed, i.e., a negatively skewed distribution. Figure 6b shows similar effect of the increase of α on the price threshold as the first case. Third, we fix α at 2 and increase β from 1 to 3. But this is just the mirror image (the reverse) of the Beta density function curve in the second case. And this is clear as Figure 6c illustrates. At last, we allow for simultaneous change of α and β. The change of the price threshold in this case is ambiguous since the effects of the increase of α and β offset. See Figure 6d. Of course, ρ also decreases when both α and β go up.



(a) $1 < \alpha = \beta < 3$ **(b)** $1 < \alpha < 3,\ \beta = 2$

(c) $\alpha = 2,\ 1 < \beta < 3$ **(d)** $1 < \alpha < 3,\ 1 < \beta < 3$

Figure 6: Price Threshold and Correlation

We see that when β is fixed, increasing α increases the price threshold, as the correlation decreases. When α is fixed, increasing β decreases the price threshold, as the correlation decreases. It seems a bit counter-intuitive that when the attributes are less correlated, the receiver is more willing to pay a higher price. However, notice that the difference between the effects of α and β arises from the (ex ante) expected quality $\mathbb{E}(|\theta|)$. Increasing α increases the (ex ante) expected quality while increasing β decreases the (ex ante) expected quality. Different pairs of α and β can even result in the same correlation coefficient but different values of the expected quality. So in effect many things are changing at the same time. Intuitively, higher expected quality and higher correlation each tend to increase the price threshold. But in our numerical example the mean-quality and correlation effects are offsetting each other.