# Compromise without Continuity

Alexander Clark*

April 27, 2017

**Abstract**

As consumers choose among products, schools seek to admit students, firms hire applicants, or political bodies take advice, they often rely on comparative statements provided by a biased sender. I develop a multidimensional cheap talk model for these situations and provide a necessary and sufficient condition for when communication can be influential and I characterize a natural class of equilibria where the sender is allowed to recommend a fixed number of propositions. While previous literature has relied on the continuity of actions to allow for compromise, I consider a model where actions are binary: salespeople are limited by price maintenance, admissions decisions are binary, etc. I consider the effect of asymmetric sender preferences, a sender favoring one proposition over another. Existing literature shows that, in similar games, a receiver can be made better off when the number of propositions becomes large. However, I show this result depends upon assumptions about receiver preferences and equilibrium selection. The power of commitment is shown to be valuable, as equilibria may be Pareto inefficient.

*Contact: atclark@wisc.edu; Department of Economics, University of Wisconsin–Madison, 1180 Observatory Drive, Madison WI 53706. I would like to thank Antonio Penta, Marzena Rostek, Bill Sandholm, Lones Smith, and Marek Waretka for their many helpful and clarifying comments.

# 1 Introduction

This paper develops a model for games where a sender seeks to advocate for the acceptance of many propositions by a receiver. For example, a salesperson might like to influence a customer to purchase many items. However, a salesperson is constrained by a customer's expectations that not all items can be worth purchasing and by an inability to demonstrate immediately the quality of each item. Instead, we are left with a situation where a salesperson's recommendation must be conservative enough to be useful in equilibrium. One might also imagine similar recommendation games played by high schools and colleges or between universities and firms, as an instructor hopes to influence the admissions or hiring decision of the college or firm, or as a biased lobby might advance several policy proposals to a legislative body. Distinctively, the receiver's action is binary: accept or reject. In application, a salesperson cannot negotiate price due to price maintenance, colleges can cannot find a compromise between acceptance and rejection of a student, and babies ought not be split.

The model is one of multidimensional cheap talk, similar to Chakraborty and Harbaugh (2007, 2010). When a sender cannot influence a receiver's decision regarding one proposition in isolation, they may nonetheless fashion an informative message when many propositions are at play. Chakraborty and Harbaugh (2007) examines comparative cheap talk, where one sender ranks, perhaps only partially, many issues for a decision maker. Chakraborty and Harbaugh (2010) again examines multidimensional cheap talk with comparative statements. Each of these papers rely on continuity of actions and the state space.

I provide a general existence result, showing communicative equilibria are always a Pareto improvement on no information, unlike in Chakraborty and Harbaugh (2010). I also characterize a natural class of equilibria where it is as if the receiver asks for the $k$ best. The setup allows for a deeper understanding of asymptotic properties of equilibria and equilibria when the sender's preferences may be asymmetric, favoring the acceptance of one proposition over that of another. I show that as the gain to either party as the number of propositions becomes large depends on selection—the gains may be concentrated to, or split between, either the sender or the receiver. Asymmetry in preferences can lead to lower equilibrium receiver payoffs even when they expect propositions to be more likely to be worth of acceptance.

In many related games, a receiver instead has inelastic demand. For example, in admissions problems, a school may have room for $k$ students, and seeks to find the $k$ best, as in Frankel (2016). Che et al. (2013) studies a problem of selecting at most one option from a set of two or more options. In this paper, a receiver will want to accept as many propositions as are of good quality. This is not a restrictive assumption, because elasticity of demand can be captured by the receiver's expectations. For instance, if a receiver would like to accept the $k$ best, the distribution will have support only on states where $k$ propositions are worthy of acceptance. Without a definite capacity, a receiver may face a commitment problem when facing a probability distribution that is not exchangeable, so that all equilibria may be Pareto inefficient.

Applications like college admissions and hiring are not games of pure cheap talk in reality, but it is true that the decision-makers are often interested only in verifiable observables to establish some minimum competency. Ultimate decisions are often based on essays, recommendations, or interviews. These are difficult to verify and limited institutional memory prevents reputational effects (e.g. colleges do not track how well a student performs relative to the impression given in essays and recommendations). Similarly, limited interaction between hiring managers and recruits' eventual bosses may also hinder the development of institutional memory. As noted by Rivera (2012, 2015), many firms look for basic competency and hiring decisions are thereafter swayed by cultural matching criteria, paying attention to job-unrelated items like hobbies. A hiring process that emphasizes such criteria could either be modeled with a cheap talk component or by treating the interview as a complex questionnaire. I take the former approach, imagining that all types of applicants can easily imitate other types. Taking the latter approach might rely on a model like in Glazer and Rubinstein (2012, 2014) where a boundedly rational applicant must be able to solve a questionnaire where low types will find it difficult to provide a persuasive answer.

Glazer and Rubinstein (2004) also considers a binary persuasion problem, using a model which differs from mine by featuring a single issue, a single receiver, a single sender, and evidence. Still, the results in this paper apply to the model in Glazer and Rubinstein (2004). Messages are partially verifiable via evidence. After the evidence stage, the probability distribution is updated and the game in Glazer and Rubinstein (2004) resembles one of cheap talk, meaning these models would demonstrate a similar change in welfare with multiple propositions. Though the optimal persuasion rule was shown to be credible (meaning a sequential equilibrium) for these games, these results would be overturned when the model is extended to multiple propositions, again underscoring the value of commitment in many settings.

Section 2 introduces the model. Section 3 presents the results, analyzing the game with a single sender, a single receiver, and $n$ propositions. As in previous work, I show that multi-dimensional cheap talk can lead to a Pareto improvement even with strong biases. An influential cheap talk equilibrium can be found as a partially separating equilibrium in a signaling game, producing a Pareto improvement. Section 4 further discusses potential applications of the model in admissions and policy debates.

# 2 Model

There is a sender and a known finite and nonempty set, $S$, of $n$ binary *propositions*. A receiver must take a binary action with respect to each proposition, choosing to either accept or reject. The *decision* is a vector $d \in D = \{0,1\}^n$, where $d_s = 1$ indicates the acceptance of $s \in S$ and $d_s = 0$ indicates its rejection. Senders have known preferences $\succeq_i$ over the receiver's possible decisions. We sometimes assume that a sender's objective is to maximize the number of accepted propositions so that the sender's preferences can be represented by a utility function $u(d) = v(\sum_S d_s)$ with $v$ increasing. This is similar to the symmetric model of Chakraborty and Harbaugh (2007), and so we refer to such sender preferences as *symmetric.* We also define a broader class of preferences, *partition symmetric*, as those that can be represented

$$u(d) = \sum_{i=1}^{K} v_i \left( \sum_{s \in S_i} d_s \right)$$

for strictly increasing $v_i$ and $K \in \mathbb{N} \cup \{+\infty\}$ so that it is as if the problem can be partitioned into $K$ symmetric problems. However, the existence result makes no such assumptions on sender preferences.

The preferences of the receiver are state dependent. The state space is $\Theta = \{0,1\}^n$, where $\theta_s \in \{0,1\}$ is a random variable giving the truth-value or quality of the proposition and $\theta$ a random sequence. The receiver prefers $d_s = \theta_s$. The receiver minimizes the expected cost-weighted number of mistakes per issue given by a *loss function* $L : D \times \Theta \to \mathbb{R}_+$. Throughout, we make the simplifying assumption that the receiver's costs depend only on the number of type one and type two errors.[1] Let $\varphi : D \times \Theta \to \mathbb{Z}_+^2$ count the number of errors and $\zeta : \mathbb{Z}_+^2 \to \mathbb{R}$ return the cost based on these errors. Thus, $L = \zeta \circ \varphi$. The first and second dimensions of the argument of $\zeta$ are the counts of type one and two errors, respectively, and $\zeta$ is strictly increasing in each error. Conditional on $k$ propositions accepted and $\sum_S \theta_s = j$, the minimum value of $\varphi$ is attained at $\left((j-k)^+, (k-j)^+\right)$ where $x^+ = \max\{x, 0\}$. Expectations are determined by a prior distribution on $\Theta$ and Pr denotes a probability. The receiver's problem is

$$\min_d \frac{1}{n} \mathbb{E} L(d, \theta).$$

Play is modeled as a signaling game where each sender chooses a direct message $m \in M = \Theta$, so that a sender places each issue in either the top or bottom tier. This message space makes the game one of comparative cheap talk. We follow a naming convention so that ranking an issue in the top tier ($m_s = 1$) can be thought of as recommending the corresponding proposition for acceptance, and a low ranking would be a recommendation for rejection. The timing is simple. Nature determines the state of the world, which is observed only by the sender.[2] Then, the sender's type is co-identified with the state as $\theta$. The sender then moves, selecting a message to send to the receiver. The uninformed receiver then takes an action, affecting the payoffs of both players.

A sequential equilibrium consists of

- Receiver beliefs $\mu$ which are consistent given $\sigma$.

[1] To be precise, a type one error is when $d_s = 0$ and $\theta_s = 1$, so that the receiver "convicts the innocent." A type two error identifies the opposite mistake. Implicitly, this defines a null hypothesis that $\theta_s = 1$.

[2] Though the sender need not literally know the true state of the world, but only what decision the receiver would make with the same information.

- Sender strategies $\sigma : \Theta \to \Delta M$, which are sequentially rational given $\mu$, and $\sigma_\theta(m)$ gives the probability that a sender of type $\theta$ sends message $m$.
- Receiver strategy $\sigma_R : M \to \Delta D$, which is sequentially rational given $\mu$.

# 3 Results

A communication equilibrium is *influential* if for two distinct messages $m$ and $m'$ produce different outcomes: $\sigma_R^m \neq \sigma_R^{m'}$ and $\sigma_\theta(m), \sigma_{\theta'}(m') > 0$ for some sender types $\theta, \theta'$. With state-independent preferences for the senders, we will be interested in equilibria where a sender tells the truth when indifferent. In particular, if the sender is a professor recommending multiple students, they will give priority to the students of good quality in their recommendations. A strict preference could be assumed or indirectly attained by adding a suitable evidence structure.

First, we briefly address cases where the probability distribution is such that for some $s \in S$, the optimal decision in a babbling equilibrium features $d_s = 1$. If this is true for some $s \in S$ and $\theta$ is a sequence of independent random variables, we should not expect influential communication regarding proposition $s$.

If $\theta$ is not an independent sequence, then it is plausible that a sender could make a deal with the receiver if preferences are sufficiently aligned. By admitting when $\theta_s = 0$, they might be able to convince the receiver that a $\theta_{s'} = 1$ for a negatively correlated proposition $s'$. This could be maintained when the sender is indifferent between $d_{s,s'} = (0, 1)$ or $(1, 0)$ given that the receiver might never choose $d_{s,s'} = (1, 1)$. However, this paper focuses on, and assumes, the case where the distribution on $\theta$ is such that the receiver's optimal decision is to reject all propositions in a babbling equilibrium. In this case, influential communication will give a Pareto improvement.

The following lemmas allow us to put some structure on costs, regardless of the probability distribution. Throughout the paper, we consider the convexity and related properties of functions with nonconvex domains. For a function $h$, we call it *convex* (concave) when it has nondecreasing (nonincreasing) first forward differences in each component. That is, $\Delta_i h(x) = h(x+e_i) - h(x)$, where $e_i$ is the $i^{\text{th}}$ unit vector, is nondecreasing for any $i$. First, we establish a few properties related to our cost functions.

**Lemma 1.** *If $\theta$ is an exchangeable random sequence and costs $\zeta$ are convex and submodular, then a local minimum of $\mathbb{E}L$ is also a global minimum.*

Submodular costs $\zeta$ require some interpretation. Submodularity requires decreasing differences in the error type levels. For $x < x'$ and $y < y'$,

$$\zeta(x', y') - \zeta(x, y') \leq \zeta(x', y) - \zeta(x, y).$$

This implies that the marginal benefit of avoiding a type 1 error (wrongful rejection) is smaller when more type 2 errors have already been made. Equivalently, we could use a production function that is supermodular in the number of correct rejections and acceptances. In the context of college admissions this amounts to a peer effects story, where the benefit of accepting a high quality student is reduced when more low quality students have been mistakenly admitted, pointing towards complementarities between good students or that it is as if low quality characteristics are contagious. Together, these assumptions make for a decreasing marginal benefit of acceptances. Later, we will consider the reverse assumption of supermodular and concave costs. Such an assumption is appropriate when low quality students, if mistakenly admitted, do not reduce the benefit to a college of admitting good quality students or if the high quality characteristics are contagious.

**Lemma 2.** *If costs $\zeta$ are convex and submodular*[3]*, then a local minimum of $\mathbb{E}L(d, \theta)$ at $d = \mathbf{0}$ is also a global minimum.*

The first question to ask is, when is comparative cheap talk influential in this environment? To begin our analysis, we must introduce $f$, a companion to the loss function $L$. Define $f : \mathbb{Z}_+ \to \mathbb{R}$,

$$f(k) := \frac{1}{n} \sum_{i=0}^{k} \zeta(0, k-i) \Pr\Big(\sum_S \theta_s = i\Big) + \frac{1}{n} \sum_{i=k+1}^{n} \zeta(i-k, 0) \Pr\Big(\sum_S \theta_s = i\Big).$$

[3] This is weaker than a function being L-convex, a concept from discrete convex analysis.

Note, $f$ provides a lower bound on the expected loss, where only one type of error is made. Given a decision $d$ that accepts $k$ propositions, $f(k) \leq \mathbb{E}L(d, \theta)$. The inequality is strict when $d$ will make errors of both types with some probability.

**Lemma 3.** *If costs $\zeta$ are convex, then $f$ is convex.*

Theorem 1 provides a necessary and sufficient condition, allowing for large asymmetries in sender preferences and an arbitrary probability distribution. It is important to note that, for arbitrary preferences, the following existence result is derived even for situations in which not every type of sender will report a correct ranking. That is, a sender may rank a low quality proposition ($\theta_{s'} = 0$) over a high quality proposition ($\theta_s = 1$). Still, the receiver finds it worthwhile to listen to comparisons.

**Theorem 1.** (Existence of an influential equilibrium): *For arbitrary sender preferences, receiver costs $\zeta$ that are submodular and convex, and an arbitrary distribution on $\theta$, an influential cheap talk equilibrium exists if and only if $f(0) \geq f(1)$.*

For $\zeta(x, y) = x + cy$, the condition on $f$ reduces to $\Pr(\theta = \mathbf{0}) \leq \frac{1}{1+c}$. For $\zeta(x, y) = x^2 + y^2$, the inequality is equivalent to $\mathbb{E}\sum_S \theta_s \geq \frac{1}{2}$. Finally, we see this generalizes to a condition on the moments over which a receiver has preferences. For $\zeta(x, y) = \sum_{i=1}^m \alpha_i x^i + \sum_{i=1}^m \beta_i y^i$ and a normalization of $\sum_{i=1}^m \beta_i = 1$, $f(0) \geq f(1)$ requires

$$\Pr(\theta = \mathbf{0}) \leq \sum_{i=1}^{m} -\alpha_i \sum_{j=1}^{i-1} \binom{i}{j} (-1)^j \mathbb{E}\left[\left(\sum_S \theta_s\right)^{i-j}\right] - \sum_{i=1}^{m} \alpha_i (-1)^i \left(1 - \Pr(\theta = \mathbf{0})\right).$$

The existence of an influential cheap talk equilibrium relies on a high enough probability that at least one issue is of high quality ($\theta_s = 1$ for some $s \in S$), and the intuition is easily grasped in the case of linear costs and $\theta$ being drawn from a Binomial$(n, p)$ distribution. Figure 1 demonstrates how $f(0) \geq f(1)$ is reduced to $\Pr(\theta = \mathbf{0}) \leq \frac{1}{1+c}$ in this case.

However, arbitrary sender preferences necessitate randomization by the receiver, eroding the potential gains which might accrue to the receiver in such an equilibrium. In the constructed equilibrium, the receiver does better when

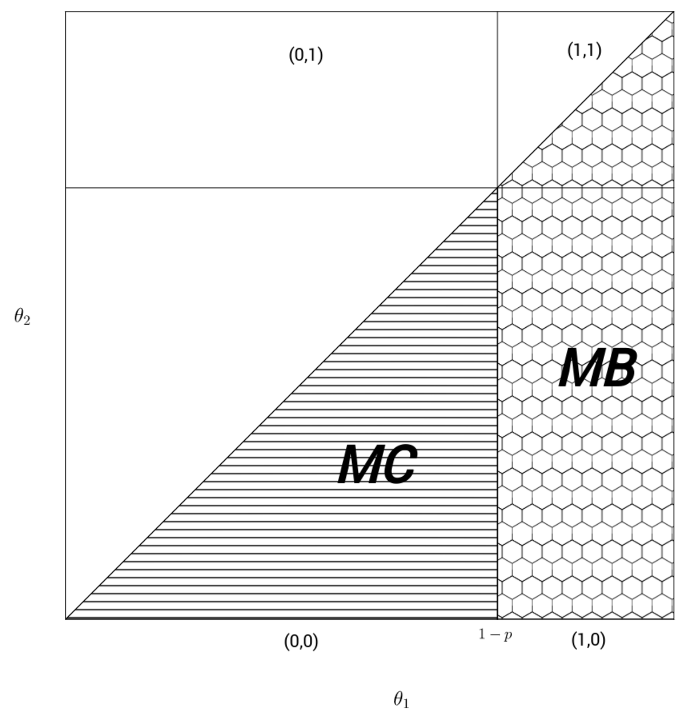


Figure 1: For $\theta \sim \text{Binom}(n, p)$, $n = 2$ and $\zeta(x, y) = x + y$, the necessity of $\Pr(\theta = \mathbf{0}) \leq \frac{1}{2}$ for influential cheap talk can be seen in this figure. The box is divided into regions where $\theta = (0, 0), (1, 0), (1, 1)$, or $(0, 1)$, moving counterclockwise from the bottom left corner and areas represent probabilities. Suppose that when $\theta_{s_1} = \theta_{s_2}$, the sender randomizes uniformly across recommending $s_1$ or $s_2$ for acceptance and the sender tells the truth if $s_1 \neq s_2$. In either of the shaded regions, half the time, the sender recommends the first proposition, $s_1$. When implementing $d = m$ in the horizontally shaded region, the receiver incurs the cost of a type 2 error. The receiver avoids a type 1 error in the honeycomb area. We see that the message is influential if the honeycomb area, $MB$, is larger than the $MC$ region. For general $n$, the latter is a hyperpyramid with volume $\frac{1}{n}(1-p)^n$, which is compared to a marginal benefit of $\frac{1}{n} - \frac{1}{n}(1-p)^n$.

there are several propositions which, when ranked, are towards the bottom and give the same utility from acceptance. This ensures a weak preference for different types to separate to the benefit of the receiver.

**Example 1a.** Let $S = \{s_1, s_2\}$ and $\zeta(x, y) = x+y$. The joint mass function for $\theta \in \Theta$ is given in the table below. Suppose that sender preferences are asymmetric, generating the ordering $u(e_1 + e_2) = \bar{u} \geq u(e_1) = u_1 \geq u(e_2) = u_2 > u(\mathbf{0})$.

| $\theta_{s_1} \backslash \theta_{s_2}$ | 0 | 1 |
|---|---|---|
| 0 | 2/5 | 1/5 |
| 1 | 1/5 | 1/5 |

In equilibrium, all sender types must obtain the same expected utility. In any pooling equilibrium, the receiver will best respond by rejecting both propositions. The following equilibrium separates all types such that $\theta \geq \mathbf{0}$ and thins the lowest type $\theta = \mathbf{0}$ across the three messages sent in equilibrium. Another equilibrium component with only two messages sent also exists, producing a coarser separation.

$$\begin{aligned}
\sigma_\theta(\theta) &= 1 \text{ for } \theta \neq \mathbf{0} \text{ (all but the lowest type report their type truthfully)} \\
\sigma_{\mathbf{0}} &= (\sigma_{\mathbf{0}}(e_1), \sigma_{\mathbf{0}}(e_2), \sigma_{\mathbf{0}}(e_1 + e_2)) = \left(\frac{1}{2}, 0, \frac{1}{2}\right) \\
\sigma_R^{e_1}(e_1) &= \frac{u_2}{u_1} = 1 - \sigma_R^{e_1}(\mathbf{0}),\ \sigma_R^{e_2}(e_2) = 1,\ \sigma_R^{e_1+e_2}(e_1 + e_2) = \frac{u_2}{\bar{u}} = 1 - \sigma_R^{e_1+e_2}(\mathbf{0}) \\
\mu^m(\mathbf{0}) &= 1 \text{ for } m \text{ unsent.}
\end{aligned}$$

**Remark:** "Improving" the distribution can make the receiver worse off in equilibrium. Consider a modification of Example 1a.

| $\theta_{s_1} \backslash \theta_{s_2}$ | 0 | 1 |
|---|---|---|
| 0 | $2/5 - 2\epsilon$ | $1/5 + \epsilon$ |
| 1 | $1/5{+}\epsilon$ | 1/5 |

**Example 1b.** We assume the same sender and receiver preferences as in Example 1a, specifying $u(e_1 + e_2) = 3$, $u(e_1) = 2$, and $u(e_2) = 1$, and modifying only the probability distribution according to the table above. As we increase small $\epsilon$, this improves the distribution in the sense that it becomes more likely that a proposition is of good quality.

First, consider a *$k$-allowance equilibrium* where the sender will recommend $k$ propositions for acceptance. In particular, let $k = 1$. With $s_1$ the most attractive, it is efficient to have types $(1, 1)$ and $(1, 0)$ always recommend $s_1$. The lowest type $(0, 0)$ will imitate these types and send the same recommendation. For $\epsilon > 0$, the sender strictly prefers $d = (1, 0)$ in response. Now, the sender $(0, 1)$ will also send this recommendation part of the time, because $d = (1, 0)$ gives higher utility than always telling the truth and attaining a utility according to $d = (0, 1)$. They will mix between recommending $s_1$ and $s_2$ so that receiver is indifferent between $d = (1, 0)$ and $d = (0, 0)$ upon receiving a recommendation of $s_1$. The receiver then mixes to maintain type $(0, 1)$'s indifference.

As a result, the receiver's expected equilibrium cost is $\frac{3}{5} + 4\epsilon$ given $\zeta(x, y) = x + y$.

We can also construct a three-message equilibrium where types $(1, 1)$ and $(1, 0)$ separate. Here, the receiver's expected equilibrium costs are $\frac{3}{5} + 4\epsilon$.

This example shows that more messages and a finer separation is not always beneficial for the receiver and a better distribution can produce a lower receiver payoff in either case. The two-message equilibrium is quite natural in the recommendation setting: a professor can recommend a single student to a potential employer. In the three-message equilibrium, the professor can recommend a single student, or both as long as that signal is rare enough. One might therefore characterize the three-message equilibrium as being more informative.[4] This differs from work on information disclosure, as in Ostrovsky and Schwarz (2010), where schools may reduce the informativeness of a transcript to improve average job placement, harming the employer. Here we see that in a cheap talk game, more messages and so a finer separation, need not benefit the receiver.

Now, we begin to further characterize $k$-allowance equilibria. Assuming $\theta$

[4] An entropy measure would agree, as entropy is lower for the 2 message equilibrium.

to be a sequence of identical and independent Bernoulli random variables would restrict the usefulness of any model of persuasion. The truth-values of many propositions are naturally correlated. A sender might be a think tank which issues policy recommendations that are founded on normative positions. When the receiver would like to accept recommendations regarding social programs, this likely informs the prudence of fiscal policies. Allowing for dependence, the following results rely on $\theta$ being an exchangeable random sequence ($\Pr(\theta) = \Pr(\theta')$ if $\theta'$ is a permutation of the elements of $\theta$). This is a less restrictive assumption than independence of identical trials, though not unrelated, as de Finetti's theorem shows that an infinite exchangeable sequence of Bernoulli random variables comes from a mixture of sequences which are independent and identically distributed.

**Theorem 2.** (Sequential Equilibria with Symmetric Preferences) *Assume the sender has symmetric preferences and $\theta$ is an exchangeable random sequence.*

1. *Then a sequential equilibrium exists where the sender recommends $k \in \mathbb{Z}_+$ propositions, all of which are accepted, if $k$ minimizes $f$.*
2. *If costs $\zeta$ are concave and supermodular, then a sequential equilibrium exists where the sender recommends $k \in \mathbb{Z}_+$ propositions, all of which are accepted, if and only if $f(k) \leq f(0)$.*

Next, we consider the receiver's optimal strategy.

**Theorem 3.** *Assume the sender has symmetric preferences and that $\zeta(x, y) = x + cy$. In pure strategies, the receiver minimizes the expected loss per issue by committing to accept $k^* = \max\{k \in \mathbb{Z}_+ : \Pr(\sum_{s\in S} \theta_s < k) \leq \frac{1}{1+c}\}$ of the propositions from $S$.*

If $\theta$ is an exchangeable sequence, this cost minimization can be achieved in a sequential equilibrium by Theorem 2.

**Example 2.** This example demonstrates why Theorem 2 need not hold when $\theta$ is not an exchangeable sequence. Let $c = 1$ and $S = \{s_1, s_2, s_3\}$ with a simple joint mass function as given in the following table.

Then, the receiver would prefer to commit to always accepting $k^* = 1$ propositions and rely on senders giving priority to propositions of high quality. An ideal separation would feature a sender recommending $s_2$ or $s_3$ when the

| $\theta_{s_1} \backslash \theta_{s_2} = \theta_{s_3}$ | 0 | 1 |
|---|---|---|
| 0 | 1/4 | $1/2 - \epsilon$ |
| 1 | $1/4 + \epsilon$ | 0 |

state is $(0, 1, 1)$ and the sender would recommend $s_1$ when the state is $(1, 0, 0)$. Suppose $s_2$ or $s_3$ are recommended when the state is $(0, 0, 0)$. Yet, Bayesian beliefs would require the receiver to infer the state is most likely $(0, 1, 1)$ if a recommendation for $s_2$ or $s_3$ is received. Accordingly, a sequentially rational receiver would accept all but the first proposition. Understanding this, the sender deviates from recommending $s_1$ when $\theta = (1, 0, 0)$ and the separation collapses.

Now, we turn our attention to asymptotics. The following proposition shows that the receiver makes zero mistakes per issue as the number of issues becomes arbitrarily large. This result is particular to one of many sequential equilibria however.

**Theorem 4.** (The receiver makes zero mistakes per issue in the limit.) *If sender preferences are symmetric and*

1. *$\theta$ is a sequence of independent and identically distributed random variables and $\zeta(0, n), \zeta(n, 0) \in \mathcal{O}(n)$*

*or*

2. *$\theta$ is a sequence of independent random variables and $\zeta(n, n) \in \mathcal{O}(n)$*

*then as $|S| \to \infty$, $L(d^*, \theta) = 0$ almost surely in the receiver-optimal sequential equilibrium.*

Theorem 4 will not always extend when the state is a sequence of correlated random variables. A stark example would be when either all propositions are of high quality, so $\theta$ is a vector of all ones, or all propositions are of low quality, so that $\theta$ is a vector of all zeroes. For this exchangeable random sequence, the traditional law of large numbers will not apply.

Theorem 4 also does not apply when $\zeta$ grows too fast with the number of mistakes. Take for example, $\zeta(x, y) = x^2 + y^2$. Then $f$ is minimized by $k^\star = \text{round}(\mathbb{E} \sum_S \theta_s)$ and $f(k^\star; n) \approx \frac{1}{n}\text{Var}(\sum_S \theta_s)$.

However, Theorem 4 can extend to preferences that are not symmetric when the biases are known and separable so that the problem can be partitioned into subproblems where the sender preferences are symmetric within.

**Corollary 1.** *Theorem 4 also applies in the case of partition symmetric preferences if all but a finite number of subsets of the partition on S are of unbounded cardinality when* $|S| \to \infty$.

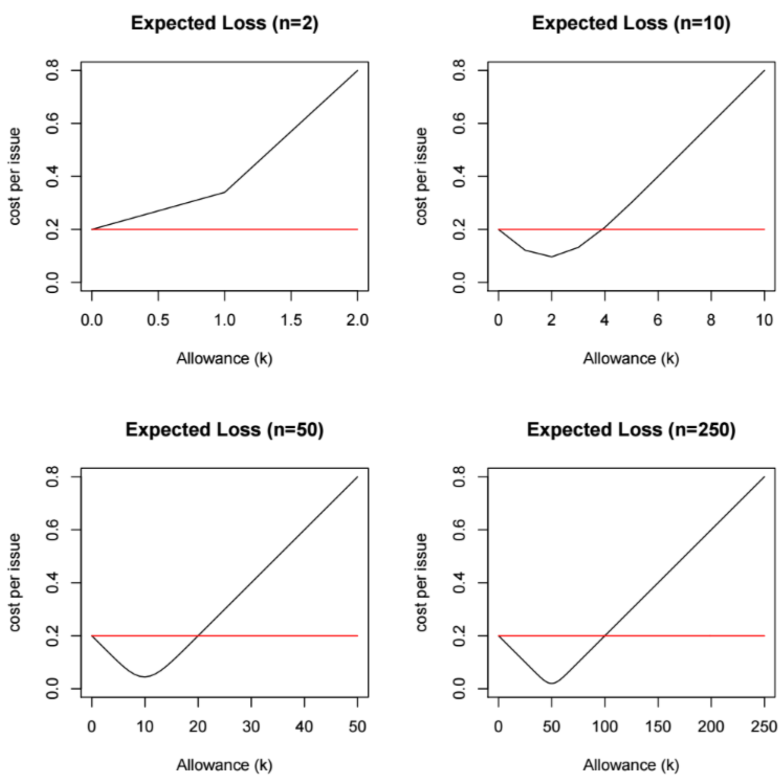


Figure 2: These four plots demonstrate the behavior of the expectation of the loss function as $n = |S|$ grows. All plots were generated assuming the state was a vector of i.i.d. Bernoulli random variables where $\Pr(\theta_s = 1) = 1/5$. The horizontal axis features the number of propositions the receiver will accept and expected costs are calculated assuming that no propositions of good quality are not accepted when a proposition of bad quality is.

In general, the receiver optimal outcome described in Theorem 3, even when a sequential equilibrium for certain distributions, seems to defy a common sense notion of credibility when the receiver is not restricted in capacity. The equilibrium is enforced by the receiver having particularly pessimistic beliefs about the type of a sender who might send a message off the equilibrium path. Intuitively, all types of senders should be equally likely to deviate from the equilibrium path given identical preferences and equilibrium payoffs. Though withstanding common refinements like the intuitive criterion and even neologism-proofness arguments.[5] Consider the following example, where we examine the optimal size $|S|$ for the receiver if we select the sender-optimal equilibrium.

**Example 3.** Assume symmetric sender preferences. Let $\theta$ be a random vector drawn according to a binomial distribution so that $\Pr(\theta_s = 1) = 1/4$ for any $s \in S$ and suppose $\zeta(x, y) = x + y$. By Theorem 3, the receiver optimal sequential equilibrium includes accepting $k^* = \lfloor |S|/4 \rfloor$ or $\lceil |S|/4 \rceil$ propositions, using established results for binomial distributions, Kaas and Buhrman (1980). By Theorem 4, the expected loss per issue goes to zero in this receiver-optimal equilibrium. However, within the class of sender-optimal equilibria, the receiver is best off when $|S| = 6$ and the expected loss per issue is 0.154. To demonstrate this, suppose by way of contradiction, there exists another $|S|$ where the receiver may obtain a lower expected loss in the sender-optimal equilibrium. By Theorem 2, this requires there to be a $k$ such that $\mathbb{E}L(d, \theta) = f(k) \leq 0.154$ and $f(k+1) \geq f(0) = 0.25$ so that $f(k+1) - f(k) \geq 0.096$. For a particular $|S| = n$, the increase in expected loss per issue is

$$\begin{aligned} f(k+1) - f(k) &= \frac{1}{n}\left(\Pr(\sum_S \theta_s < k+1) - \Pr(\sum_S \theta_s \geq k+1)\right) \\ &= \frac{1}{n}\left(1 - 2\Pr(\sum_S \theta_s \geq k+1)\right). \end{aligned}$$

The increase in the expected loss is bounded by $\frac{1}{n}$ so we only need to check $n \leq 11$ to look for cases where $f(k+1) - f(k) \geq 0.096$. It is verified in the appendix that $|S| = 6$ is worst-case optimal for the receiver.

[5]Neologism proofness is preserved because, in transitioning from an equilibrium with $k$ acceptances to $k+1$, everyone must send a different kind of message, not just that some set of types would deviate to a single new message.

Increasing the number of propositions reduces the randomness faced by a receiver, but this benefit is seen to be quickly dominated by the sender's increasing ability to push the receiver to the edge of indifference as the number of propositions becomes large. Indeed, as the number of propositions becomes arbitrarily large, the receiver realizes no benefit from influential communication.

**Theorem 5.** *Assume $\theta$ is an exchangeable sequence, the sender has symmetric preferences, and $\zeta$ is supermodular and concave. In the limit, the receiver incurs the same expected loss in the sender-optimal influential equilibrium as in the pooling equilibrium.*

# 4 Discussion

Theorem 2 shows that as the number of issues becomes arbitrarily large, so could the number of sequential equilibria. Theorem 5 underscores the potential value of a receiver's ability to set a restrictive capacity in order to favorably reduce the set of equilibria. The ability to commit to a capacity is not infeasible. Indeed, colleges commit to a capacity by building dorms and a legislator who invites a biased expert for advice might do well to give the expert a short amount of to present their case–effectively limiting the number of propositions which may be recommended.

The college admissions process is often seen as an engine of social mobility and, remaining unpredictable, has spurred a cottage industry in high-priced consulting, with services routinely priced in the five figures. This model, in its application to admissions, of course ignores important parts of any application and a more detailed model would be necessary to draw firm conclusions. However, this model might still provide a useful framework in thinking about how colleges and high schools interact, as a high school counselor coordinates the application process. The lack of reputation effects in this model is in fact realistic, as no colleges compare a students projected success at the admissions stage (through the influence of the cheap talk components of an application) with their later realized success to my knowledge. We could think of prep schools and more average public schools as playing this game under different conditions. Namely, a prep school counselor works with a larger set of propositions, as such a school is likely to have a better pool of college-ready candidates (those who meet threshold requirements on grades, etc). This makes influential communication more likely, and if we are to take Theorem 5 even slightly seriously, prep school students should be overrepresented in colleges and perhaps of lower quality. Strong inferences would be unwise, but historical data suggests that the average prep school student's college GPA is lower than that of the average public school product Zweigenhaft (1993). Colleges, not being unsophisticated, have used docket systems, whereby they essentially allocate so many spots for each docket, where a docket defines (usually geographic) boundaries of comparison. As discussed by Karen (1990), colleges consider dealing with prep school dockets as more difficult, as these schools are more likely to challenge decisions.

# References


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

_ and _ , "Persuasion by cheap talk," *The American Economic Review*, 2010, pp. 2361–2382.

Che, Yeon-Koo, Wouter Dessein, and Navin Kartik, "Pandering to persuade," *The American Economic Review*, 2013, *103* (1), 47–79.

Frankel, Alex, "Selecting Applicants," 2016.

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

_ and _ , "A model of persuasion with boundedly rational agents," *Journal of Political Economy*, 2012, *120* (6), 1057–1082.

_ and _ , "Complex Questionnaires," *Econometrica*, 2014, *82* (4), 1529–1541.

Kaas, Rob and Jan M Buhrman, "Mean, median and mode in binomial distributions," *Statistica Neerlandica*, 1980, *34* (1), 13–18.

Kallenberg, O, "Foundations of Probability," 1997.

Karen, David, "Toward a political-organizational model of gatekeeping: The case of elite colleges," *Sociology of Education*, 1990, pp. 227–240.

Ostrovsky, Michael and Michael Schwarz, "Information disclosure and unraveling in matching markets," *American Economic Journal: Microeconomics*, 2010, *2* (2), 34–63.

Rivera, Lauren A, "Hiring as cultural matching the case of elite professional service firms," *American Sociological Review*, 2012, *77* (6), 999–1022.

_ , *Pedigree: How elite students get elite jobs*, Princeton University Press, 2015.

Zweigenhaft, Richard L, "Prep school and public school graduates of Harvard: A longitudinal study of the accumulation of social and cultural capital," *Journal of Higher Education*, 1993, pp. 211–225.

# Appendix A Proofs

*Proof of Lemma 1.* We introduce a more general lower envelope on the expected loss function. $\hat{L} : \mathbb{Z}_+ \to \mathbb{R}_+$ gives the lowest cost possible given a decision $d$ that accepts exactly $k$ propositions. That is,

$$\hat{L}(k) := \min_{d \in D : |d|_1 = k} \mathbb{E}L(d, \theta).$$

We will show that this is convex in $k$ when $\theta$ is an exchangeable random sequence and costs $\zeta$ are convex and submodular.

First, we show that the net benefit from accepting a particular proposition $s$ is decreasing in $d_{-s}$.

Observe that the marginal benefit of accepting an additional proposition, eliminating potential type 1 errors, is a probability weighted sum of terms $\zeta(k, j) - \zeta(k-1, j)$ where $1 \leq k \leq n$ and $0 \leq j \leq n-1$.

The marginal cost, introducing possible type 2 errors, will be a probability weighted sum of terms $\zeta(l, m) - \zeta(l, m-1)$ with $0 \leq l \leq n$ and $1 \leq m \leq n$.

If we consider switching from rejecting all to accepting just one proposition, then $j = 0$ in the above terms.

By submodularity, the marginal benefit $\zeta(k, j) - \zeta(k-1, j)$ is decreasing in $j$ given a constant $k$. Given a constant $j$, this difference increases with $k$, by convexity. As $d_{-s}$ increases, $k$ must weakly decrease and $j$ must weakly increase. Therefore, the marginal benefit of accepting a particular proposition decreases in $d_{-s}$.

As more propositions are accepted, this weakly decreases the number of type I errors. Thus, for a fixed $m$, the difference $\zeta(l,m) - \zeta(l,m-1)$ grows by submodularity as we consider decisions that accept more propositions. As more propositions are accepted, this weakly increases the type II errors. Thus, for a fixed $l$, the difference $\zeta(l,m) - \zeta(l,m-1)$ grows by convexity. Altogether, the marginal cost must increase in $d_{-s}$.

Therefore, the net benefit of accepting proposition $s$ decreases with $d_{-s}$.

Now, consider the set of sequences $\{d_k\}_{k=1}^n$ so that $\mathbb{E}L(d_k,\theta) = \hat{L}(k)$. Exchangeability guarantees that the receiver is indifferent between all decisions that accept the same number of propositions, so we must have an increasing sequence in this set. By the above, the net benefit of accepting a proposition $s$ decreases in $d_{-s}$, so $\hat{L}(k)$ must therefore be convex. Furthermore, because the lower envelope $\hat{L}$ can be attained by an increasing sequence, a local minimum must also be a global minimum.

□

*Proof of Lemma 2.* Observe that the marginal benefit of accepting an additional proposition, eliminating potential type 1 errors, is a probability weighted sum of terms $\zeta(k,j)-\zeta(k-1,j)$ where $1 \le k \le n$ and $0 \le j \le n-1$.

The marginal cost, introducing possible type 2 errors, will be a probability weighted sum of terms $\zeta(l,m) - \zeta(l,m-1)$ with $0 \le l \le n$ and $1 \le m \le n$.

If we consider switching from rejecting all to accepting just one proposition, then $j = 0$ in the above terms. By submodularity, this maximizes $\zeta(k,j) - \zeta(k-1,j)$ given a constant $k$. Given a constant $j$, this difference is maximized by taking $k$ as large as possible, by convexity. Thus, the marginal benefit of accepting a particular proposition is maximized when no other propositions have been accepted, as this will minimize type 2 while maximizing type 1 errors.

As more propositions are accepted, this weakly decreases the number of type I errors. Thus, for a fixed $m$, the difference $\zeta(l,m) - \zeta(l,m-1)$ grows by submodularity as we consider decisions that accept more propositions.

As more propositions are accepted, this weakly increases the type II errors. Thus, for a fixed $l$, the difference $\zeta(l, m) - \zeta(l, m-1)$ grows by convexity. Altogether, the marginal cost is minimized when no other propositions have been accepted.

Therefore, if for some proposition $s$, if $\mathbb{E}L(\mathbf{0}) \leq \mathbb{E}L(e_s)$, it must be that $\mathbb{E}L(\mathbf{0}) \leq \mathbb{E}L(d + e_s)$ for all $d \in D$ such that $d_s = 0$. This proves the statement.

□

*Proof of Lemma 3.* We define $\Delta f(k) = f(k+1) - f(k)$. For convexity, it is enough to show that $\Delta f(k)$ is increasing. Observe

$$\Delta f(k) = \sum_{i=0}^{k} \Pr(\sum \theta_s = i)[\zeta(0, k+1-i) - \zeta(0, k-i)] + \sum_{i=k+1}^{n} \Pr(\sum \theta_s = i)[\zeta(i-k-1, 0) - \zeta(i-k, 0)]$$

As we increase $k$, then the first part of the sum will increase as the differences will grow. Similarly, the magnitude of the difference in the second part of the will will decrease, increasing the sum itself. □

*Proof of Theorem 1.* Part I: We show that $f(0) < f(1)$ implies there will be no influential equilibrium. Suppose there is an influential equilibrium. Then, the receiver must find it a best response to accept at least one proposition in response to some message. Furthermore, all sender types must earn the same payoff in equilibrium, so all must send a message that is accepted with a positive probability. So, the receiver must incur an expected loss of at least $f(1)$. By hypothesis, $f(0) < f(1)$, so the receiver has a profitable deviation. They can reject all propositions for a reduction in costs. This contradicts the existence of an influential equilibrium. Therefore, there can be no influential equilibrium.

Part II: If $f(0) \geq f(1)$, then an influential equilibrium exists.

Without loss of generality, let the propositions be ordered $s_1, s_2, \ldots, s_n$ in order of decreasing utility according to a sender's preferences so that $u(e_1) \geq \cdots \geq u(e_n)$ where $e_k$ is the decision vector where $d_s = 1$ for $s = s_k$ and $d_s = 0$ otherwise. In place of a message, $m = e_k$ is the message which recommends only issue $s_k$.

We define an upper contour set for every issue $s_k$; let $U_k = \{s_1, \ldots, s_{k-1}\}$. We partition the type space $\Theta$ into subsets. For $1 \leq k \leq n$,

$$\Theta_k = \{\theta \in \Theta : \theta_s = 0 \text{ for all } s \in U_k \text{ and } \theta_k = 1\}$$

and define $\Theta_{n+1} = \{\mathbf{0}\}$.

We create a messaging strategy whereby all types recommend a single proposition and the receiver is indifferent between accepting just the recommended proposition or rejecting all propositions.

Begin by putting $\sigma_\theta(e_k) = 1$ for all $\theta \in \Theta_k, k = 1, \ldots, n$. Then for $\theta = \mathbf{0}$, we choose $\sigma_{\mathbf{0}}(e_1) = \min\{1, \hat{\sigma}_{\mathbf{0}}(e_1)\}$, where $\hat{\sigma}_{\mathbf{0}}(e_1)$ solves

$$\sum_{i=1}^{n} \Pr(\sum_{\theta \in \Theta_1} \theta_s = i)[\zeta(i, 0) - \zeta(i-1, 0)] = \hat{\sigma}_{\mathbf{0}}(e_1) \Pr(\theta = \mathbf{0})\zeta(0, 1).$$

The lefthand side of the equation is the marginal benefit from switching from $d = \mathbf{0}$ to $d = e_1$ in response to the message $e_1$. The righthand side is the marginal cost.

Proceeding to the $k^{\text{th}}$ most preferred issue, $s_k$. Set

$$\sigma_{\mathbf{0}}(e_k) = \min\left\{1 - \sum_{s \in U_k} \sigma_\theta(e_s), \hat{\sigma}_{\mathbf{0}}(e_k)\right\}$$

where $\hat{\sigma}_{\mathbf{0}}(e_k)$ solves

$$\sum_{i=1}^{n} \Pr(\sum_{\theta \in \Theta_k} \theta_s = i)[\zeta(i, 0) - \zeta(i-1, 0)] = \hat{\sigma}_{\mathbf{0}}(e_k) \Pr(\theta = \mathbf{0})\zeta(0, 1).$$

Put $T = \min\left\{k : \sigma_{\mathbf{0}}(e_{r_{k+1}}) = 0\right\}$, where we define $\sigma_{\mathbf{0}}(e_{n+1}) = 0$. In words, $T$ identifies the least preferred proposition that the lowest type sender will attempt to recommend to the receiver. If $T = n$, then we are basically done. Set $\sigma_R^{e_k}(e_k) = \frac{u(e_n)}{u(e_k)} = 1 - \sigma_R^{e_k}(\mathbf{0})$ and set off-path beliefs $\mu^m(\theta = \mathbf{0}) = 1$ for all unsent messages $m$.

If $T < n$, we proceed as follows.

First, if $\sigma_{\mathbf{0}}(e_{r_T}) < \hat{\sigma}_{\mathbf{0}}(e_T)$, then there is some action $d \geqq e_T$ that is strictly preferred to $\mathbf{0}$. So we revise the strategies of other types $\theta \in \Theta_k$ for some $T < k \leq n$ to send message $e_T$ also.

Now there are three cases to consider regarding the receiver's best response(s) to $m = e_T$.

Case 1: There is some best response $d$ where $0 < u(d) \leq u(e_{T-1})$ or best responses $d, d'$ with $u(d) < u(e_{T-1})$ and $u(d') > u(e_{T-1})$.

In this case, the receiver should respond with a (possible) mixture of best response that gives an expected utility greater than zero but weakly less than $u(e_{T-1})$. Call this expected utility value $x$. Set $\sigma_R^{e_k}(e_k) = \frac{x}{u(e_k)} = 1 - \sigma_R^{e_k}(\mathbf{0})$ for all $k - 1, \ldots, T-1$ and set off-path beliefs $\mu^m(\theta = \mathbf{0}) = 1$ for all unsent messages $m$.

Case 2: There is a singleton best response set of $d = \mathbf{0}$, yielding $u(d) = 0$.

It must be that, given the message, $\mathbb{E}L(\mathbf{0}) - \mathbb{E}L(e_k) < 0$ for $k = 1, \ldots, n$. We revise messaging strategies for types $\theta \in \Theta_k$ for $T + 1 \leq k \leq n$. Begin with $\sigma_\theta(e_T) = 0$ for all $\theta \in \Theta_k$ for $T + 1 \leq k \leq n$. Now, it must be that $\mathbb{E}L(\mathbf{0}) - \mathbb{E}L(e_k) \geq 0$ for $k = T$.

Note $\mathbb{E}L(\mathbf{0}) - \mathbb{E}L(e_T)$ must be continuous in $\sigma_\theta$ for each $\theta \in \Theta$. We proceed sequentially, beginning with $\theta \in \Theta_{T+1}$, increasing $\sigma_\theta(e_T)$ until it equals 1 or $\mathbb{E}L(\mathbf{0}) - \mathbb{E}L(e_k) \leq 0$ for all $k$ and with equality for some $k = T, \ldots, n$. When $\sigma_\theta(e_T) = 1$ for all $\theta \in \Theta_r$, $r = T + 1, \ldots, n$, and not $\mathbb{E}L(\mathbf{0}) - \mathbb{E}L(e_k) \leq 0$ for all $k$ and with equality for some $k = T, \ldots, n$, we increase $\sigma_\theta(e_T)$ for $\theta \in \Theta_{r+1}$ until $\sigma_\theta(e_T) = 1$ or $\mathbb{E}L(\mathbf{0}) - \mathbb{E}L(e_k) \leq 0$ for all $k = 1, 2, \ldots, n$. and $\mathbb{E}L(\mathbf{0}) - \mathbb{E}L(e_k) = 0$ for some $k = T, \ldots, n$. We are guaranteed, by continuity, that this must happen for some $r = T, \ldots, n$ since this case began with $\mathbb{E}L(\mathbf{0}) - \mathbb{E}L(e_k) < 0$ for $k = 1, \ldots, n$ and no types $\theta \in \Theta_k$ such that $k < T$ will ever send $m = e_T$.

Define $\Theta_R$ to be the subset of $\Theta$ such that $\sigma_\theta(e_T) = 0$ for all $\theta \in \Theta_k$ such that $k > R$.

Then we create a new message $m = e_R$. Let $\sigma_\theta(e_T) = 1 - \sigma_\theta(e_R)$ for $\theta \in \Theta_R$ and $\sigma_\theta(e_R) = 1$ for all $\theta \in \Theta_k$ such that $k > R$. We now consider the receiver's best response to this message and proceed similarly according to the three cases presently established. This procedure is guaranteed to end in case 1 due to the availability of types in $\Theta_n$.

Case 3: For all best responses $d$ to message $m$, $u(d) > u(e_{T-1})$.

Here we revise the strategies of all those not sending the message, $m$, in question. For such types $\theta$, define a new message strategy $\hat{\sigma}_\theta$ where $\hat{\sigma}_\theta(m) = \alpha\sigma_\theta(m)$. Note this does not change the conditional probabilities on types given a particular previously defined message, and so the receiver's best responses are not affected. We know, by hypothesis, that if we choose $\alpha = 1$, the best response will be $d = \mathbf{0}$. Now, using the continuity of $\mathbb{E}L(\mathbf{0} \mid m \text{ received}) - \mathbb{E}L(e_k \mid m \text{ received})$ in $\sigma_\theta$, we know there must exist some $\alpha$ such that $\mathbb{E}L(\mathbf{0} \mid m \text{ received}) \leq \mathbb{E}L(e_k \mid m \text{ received})$ for all $k = 1, \ldots, n$ and with equality for at least one value of $k$. Then, by Lemma 2, we are in Case 1 and finished with the equilibrium construction.

□

*Proof of Theorem 2.* Part I. A sequential equilibrium is constructed as follows. Choose a $k \in \mathbb{Z}_+$. Let $\binom{n}{k}$ unique messages be sent, each identifying $k$ of $n = |S|$ propositions for acceptance. Let types $\theta'$ such that $\sum_S \theta'_s = k$ fully separate, reporting their type truthfully with $m = \theta'$, thus creating the $\binom{n}{k}$ distinct messages. Then consider the $\binom{n}{l}$ unique types $\theta$ such that $\sum_S \theta_s = l > k$ where the probability of any one of these types is therefore $\Pr(\sum_S \theta_s = l)/\binom{n}{l}$. For a $\theta'$ such that $\sum_S \theta'_s = k$, there will be $\binom{n-k}{l-k}$ types $\theta$ such that $\theta \geq \theta'$. One such $\theta$ will randomize uniformly across $\binom{l}{k}$ different messages, choosing $k$ of the $l$ dimensions in which $\theta_s = 1$. Next, we consider types $\theta$ such that $\sum_S \theta_s = j < k$. For a particular, $j$, there are $\binom{n}{j}$ unique and equally probable such types so that the probability of any one of these types is $\Pr(\sum_S \theta_s = j)/\binom{n}{j}$. For a $\theta'$ such that $\sum_S \theta'_s = k$, there will be $\binom{k}{j}$ types $\theta$ such that $\theta' \geq \theta$. A particular $\theta$ will then randomize uniformly across $\binom{n-j}{k-j}$ messages, one for each type $\theta'$ where $\sum_S \theta'_s = k$ and $\theta' \geq \theta$. Together,

$$\sigma_\theta(\theta') = \begin{cases} \frac{1}{\binom{l}{k}} & \text{if } \sum_S \theta_s = l \geq \sum_S \theta'_s = k \text{ and } \theta \geqq \theta' \\ \frac{1}{\binom{n-j}{k-j}} & \text{if } \sum_S \theta_s = j < \sum_S \theta'_s = k \text{ and } \theta \leq \theta' \\ 0 & \text{otherwise.} \end{cases}$$

Now, $\Pr(\theta_s = 1|m_s = a) = \Pr(\theta_{s'} = 1|m_{s'} = a)$ for any issues $s, s' \in S$ where $a \in \{0, 1\}$. On path, this dictates the receiver's beliefs. Off the equilibrium path, let $\mu^m(\theta = \mathbf{0})$ for all unsent messages $m$.

Now we consider when it is an equilibrium

First, Suppose $k$ minimizes $f$. Upon receiving a message, the receiver sets $d_s = 1$ for all $s$ such that $m_s = 1$. The senders do not wish to deviate because all on-path messages give the same payoff and any off-path message gives a zero payoff. By minimizing $f$, we have also minimized $\mathbb{E}L$, so the receiver is best responding and this must be an equilibrium.

Next, suppose simply that $f(k) < f(0)$ and that costs are concave, exhibiting decreasing first forward differences, and are supermodular.

The marginal benefit from rejecting an issue is a probability weighted sum of terms $\zeta(j, l + 1) - \zeta(j, l)$ and the marginal cost is a probability weighted sum of terms $\zeta(m + 1, p) - \zeta(m, p)$. Suppose it is profitable to reject one of the issues. Then, as we continue to reject more, this introduces more type 1 errors. By supermodularity, this increases the marginal benefit. It introduces fewer type 2 errors. By concavity, this also increases the marginal benefit. The marginal cost decreases by concavity and by supermodularity. So if $d - e_s$ is preferred to $d$, then $\mathbf{0}$ is preferred to $d$. This can only be the case when $f(k) > f(0)$.

Next, we see if the receiver would like to accept a proposition $s$ for which $m_s = 0$. The marginal benefit would be a probability weighted sum of terms $\zeta(m + 1, p) - \zeta(m, p)$. As we continue to accept more issues, this will decrease $m$ and increase $p$. So this difference grows in by supermodularity and concavity. The marginal cost is a probability weighted sum of terms $\zeta(j, l + 1) - \zeta(j, l)$. As we accept more propositions, $j$ decreases and $l$ increases. By concavity and supermodularity, these costs decrease as we accept more propositions. By construction, all unrecommended propositions are equally likely to be of good quality, so when the receiver prefers to accept some proposition $s$, the receiver must prefer to accept all unrecommended propositions, and therefore all $n$ propositions. By assumption, the distribution is such that this cannot be the case. Therefore, the receiver will prefer to reject all propositions not recommended by the sender.

Therefore, we have an equilibrium where the receiver follows the sender's recommendation whenever $f(k) \leq f(0)$.

If $f(0) < f(k)$, then the global minimum is attained by simply rejecting all issues, and so there can be no influential equilibrium.

$\square$

*Proof of Theorem 3.* We need only show that $k^*$ satisfies the conditions established in Proposition 2 and then that the resulting strategy minimizes the expected loss.

$$
\begin{aligned}
f(k) &= \frac{c}{|S|}\sum_{i=0}^{k}(k-i)\Pr(\sum_S \theta_s = i) + \frac{1}{|S|}\sum_{i=k+1}^{|S|}(i-k)\Pr(\sum_S \theta_s = i) \\
f(k)-f(k-1) &= \frac{c}{|S|}\Pr(\sum_S \theta_s \le k-1) - \frac{1}{|S|}\Pr(\sum_S \theta_s > k-1).
\end{aligned}
$$

The last expression is negative when $\Pr(\sum_S \theta_s < k) \le \frac{1}{1+c}$. Note $f$ is continuous with a first derivative in $k$ equal to $\frac{c}{|S|}\Pr(\sum_S \theta_s \le k) - \frac{1}{|S|}\Pr(\sum_s \theta_s > k)$ for $k$ not an integer. This is increasing in $k$, hence $f$ is convex in $k$. So, $k^* = \max\{k \in \mathbb{Z}_+ : \Pr(\sum_S \theta_s < k) \le \frac{1}{1+c}\}$ minimizes $f$. Note, $f(0) = \frac{1}{|S|}\mathbb{E}\sum_S \theta_s$, so this $k^*$ must satisfy the conditions from Proposition 2 and therefore there exists a corresponding sequential equilibrium.

$\square$

*Proof of Theorem 4.* Define $X_s = \theta_s - \mathbb{E}\theta_s$.

Then, for any sequence of independent random variables, $\lim_{|S|\to\infty}\sum_S n^{-2}\mathbb{E}X_s^2 < \infty$. Therefore, $\frac{1}{n}\sum_S X \to 0$ almost surely (Kallenberg (1997)). Accordingly, $\lim_{|S|\to\infty}\Pr(\sum_S \theta_s \ne \mathbb{E}\sum_S \theta_S) \to 0$.

Now, set $k^\star = \text{round}(\mathbb{E}\sum_S \theta_s)$.

Case 1: iid

For an exchangeable sequence, there is an equilibrium where the receiver never makes mistakes of both type 1 and type 2. We have

$$
\mathbb{E}L(d^\star,\theta) \le \frac{1}{n}\Pr(\sum_S \theta_s \ne \mathbb{E}\sum_S \theta_s)\max\{\zeta(n-k^\star,0),\zeta(0,k^\star)\} + \frac{1}{n}\Pr(\sum_S \theta_s = \mathbb{E}\sum_S \theta_s)\alpha(n),
$$

where $\alpha(n)$ is a bounded rounding error.

The first term goes to zero when $\zeta(0,n), \zeta(n,0) \in \mathcal{O}(n)$. and the second term always goes to zero. Because the expected loss goes to zero, we know $L(d^*, \theta) = 0$ with probability one.

Case 2: Independent

We have

$$\mathbb{E}L(d^\star, \theta) \leq \frac{1}{n} \Pr(\sum_S \theta_s \neq \mathbb{E} \sum_S \theta_s) \zeta(n,n) + \frac{1}{n} \Pr(\sum_S \theta_s = \mathbb{E} \sum_S \theta_s) \alpha(n),$$

where $\alpha(n)$ is a bounded rounding error.

The first term goes to zero when $\zeta(0,n), \zeta(n,0) \in \mathcal{O}(n)$. and the second term always goes to zero. Because the expected loss goes to zero, we know $L(d^*, \theta) = 0$ with probability one.

□

*Proof of Corollary 1.* There will be $K$ separate problems concerning proposition sets $S_1, S_2, \ldots, S_K$. We can construct a partitioned message space $M_1 \times M_2 \times \cdots \times M_K$ so that is as if the receiver and sender play the game $K$ times. By theorem 4, the receiver will make zero mistakes per proposition in the limit for all subsets that become arbitrarily large. For those subsets $S_i$ that do not, the receiver will incur make potentially $|S_i|$ type 1 and type 2 errors for each subset. If there are no more than $l$ issues belong to all of these subsets in the limit, then the expected cost will be

$$\mathbb{E}L \leq \frac{1}{n} \Pr(\sum_S \theta_s \neq \mathbb{E} \sum_S \theta_s) \zeta(n,n) + \frac{1}{n} \Pr(\sum_{S_j} \theta_s = \mathbb{E} \sum_{S_j} \theta_s \text{ for all } S_j \text{ unbounded}) \zeta(l,l),$$

and for $\theta$ exchangeable,

$$\mathbb{E}L \leq \frac{1}{n} \Pr(\sum_S \theta_s \neq \mathbb{E} \sum_S \theta_s) \max\{\zeta(0,n), \zeta(n,0)\} +$$

$$\frac{1}{n} \Pr(\sum_{S_j} \theta_s = \mathbb{E} \sum_{S_j} \theta_s \text{ for all } S_j \text{ unbounded}) \zeta(l,l),$$

$\zeta(l,l)$ is a finite number and so the second term disappears $n \to 0$. By the law of large numbers, the first term in the sum will go to zero when $\zeta$ follows the appropriate linear growth condition. □

*Proof of Theorem 5.* Parameterize the function $f$ by the number of propositions, $n$. Then we examine the behavior of $f(k;n)$. We know for the receiver optimal equilibrium, $\mathbb{E}L(d^\star,\theta) = f(k^\star) \leq f(0)$ $f(k^\star;n) \to 0$ as $n \to \infty$. We assume $f(n;n) \geq f(0;n)$. Using continuity of $f(x;n)$ in $x$ and the intermediate value theorem, there exists an $x_n \in \mathbb{R}_+, x_n \geq k^*$ such that $f(x_n;n) = f(0;n)$ when $f$ is extended to a real valued domain and $f(x_n;n)$ the appropriate linear combination of $f(\lfloor x_n \rfloor;n)$ and $f(\lceil x_n \rceil;n)$. Then take $k_n = \lfloor x_n \rfloor$. So,

$$
\begin{aligned}
&f(x_n,;n) - f(k_n;n) = \\
&\frac{1}{n}\sum_{i\leq k_n} \Pr(\sum_S \theta_s = i)[\zeta(0,x_n-i)-\zeta(0,k_n-i)]+\frac{1}{n}\sum_{i>k_n} \Pr(\sum_S \theta_s = i)[\zeta(i-x_n,0)-\zeta(i-k_n,0)] \\
&\leq \frac{1}{n}\Pr(\sum_S \theta_s \leq k_n)[\zeta(0,1)-\zeta(0,0)]+\frac{1}{n}\Pr(\sum_S \theta_s > k_n)[\zeta(i-1,0)-\zeta(0,0)],
\end{aligned}
$$

where the last inequality is due to decreasing first forward differences in each component. This term is bounded annd so the difference will converge to zero. This shows $f(k_n;n) \nearrow f(x_n;n)$ as $n \to \infty$. By Theorem 2, $k_n$ must correspond to a sequential equilibrium for any $n$.

□

# Appendix B Table

Table 1: This table follows Example 3 by showing the value of $f(k)$, the expected loss from the receiver setting an allowance $k$ on the number of propositions to be accepted. All numbers are based on $\theta$ being drawn from a Binomial$(n, p)$ distribution with $p = 0.25$.

$f(k; n)$

| $n \backslash k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | **0.25** | 0.3125 | 0.75 | | | | | | | |
| 3 | 0.25 | **0.1979167** | 0.4270833 | 0.75 | | | | | | |
| 4 | 0.25 | **0.1582031** | 0.2773438 | 0.5019531 | 0.75 | | | | | |
| 5 | 0.25 | 0.1449219 | **0.1980469** | 0.3566406 | 0.5503906 | 0.75 | | | | |
| 6 | 0.25 | 0.1426595 | **0.1539714** | 0.2641602 | 0.4182943 | 0.5834147 | 0.75 | | | |
| 7 | 0.25 | 0.1452811 | 0.1295515 | **0.2028111** | 0.3255092 | 0.4646868 | 0.6071603 | 0.75 | | |
| 8 | 0.25 | 0.1500282 | 0.1167984 | **0.1614342** | 0.2579803 | 0.3761559 | 0.5000992 | 0.6250038 | 0.75 | |
| 9 | 0.25 | 0.1555744 | 0.1112052 | 0.1335780 | **0.2078612** | 0.3080995 | 0.4169896 | 0.5278024 | 0.6388897 | 0.75 |
| 10 | 0.25 | 0.1612627 | 0.1100677 | 0.1151863 | **0.1703613** | 0.2547359 | 0.3507904 | 0.4500893 | 0.5500061 | 0.6500002 |
| 11 | 0.25 | 0.1667700 | 0.1116968 | 0.1035515 | 0.1423342 | **0.2124021** | 0.2970698 | 0.3866041 | 0.4772972 | 0.5681833 |