VOTING WITH RESTRICTED COMMUNICATION

ANDREAS HAUPT

The theory of voting thus far assumes abundant capacity for the voters to communicate their type (revelation principle). In many situations, limitations of intellectual capacity or practical considerations allow the agents to only send a finite number of different messages. This project aims to study voting with such a restriction.
Related Literature. A body of literature studied classical economic models (matching, auction design) with limited communication of the agents to the mechanism, [McA02, HMO11, BNS07, BSXY11], see appendix B.

1. Model

We consider a one-dimensional model similar to [GMS17a]. The primitives of the model are a natural number k, the number of different messages an agent can send, and a real-valued, finite mean type distribution F. *Voters* $i = 1, 2, \ldots, n$ have independent, identically distributed types $\theta_1, \theta_2, \ldots, \theta_n \sim F$. Given type θ a voter has quadratic utility $u_\theta \colon \mathbb{R} \to \mathbb{R}, u_\theta(x) = -(\theta - x)^2$.[1] The principal announces an anonymous mechanism $\phi \colon \{1, 2, \ldots, k\}^n \to \mathbb{R}$. Subsequently, the voters simultaneously send one of k possible *messages* $m(\theta_i) \in \{1, 2, \ldots, k\}$, $i = 1, 2, \ldots, n$, m being the measurable *strategy* for each agent. Finally, $\phi(m(\theta_1), m(\theta_2), \ldots, m(\theta_n)) = x$, $x \in \mathbb{R}$ is assigned as the outcome. The principal wishes to maximise total expected welfare $\mathbb{E}\left[\sum_{i=1}^n u(\phi(m(\theta_1), m(\theta_2), \ldots, m(\theta_n))\right]$. We call mechanisms together with a dominant-strategy equilibrium $(m, m, \ldots, m)$ *implementable*. Call welfare-maximising mechanisms *constrained first-best* and implementable welfare-maximising mechanisms *constrained second-best*.

We stress that cards do not have a meaning a priori. The players need to select an equilibrium. This coordination problem is solved if the principal communicates at the outset the equilibrium that shall be played.

2. Questions and Loose Ends

2.1. **Characterisation.** What are constrained best mechanisms?

JEL *Classification.* D83,D72.
Key words and phrases. mechanism design, quantisation, voting.

[1] The assumption of quadratic preferences is restrictive, but necessary for a result to hold. We comment on this issue in Remark 1 in appendix A.

2.1.1. *Constrained First-Best Mechanisms.* As a new result, we characterise the first-best mechanism.

Proposition 1. *Let* $a_1 \leq a_2 \leq \cdots \leq a_k \in \mathbb{R}$ *be the centers of a Lloyd-Max optimal partition* [Llo82, Max60], *i.e. for* $\beta_i := \frac{a_i + a_{i+1}}{2}$, $i = 1, 2, \ldots, k-1$, $\beta_0 = -\infty, \beta_k = \infty$ *(*β *for* boundary*), we have* $a_i = \mathbb{E}[\theta_i | \theta_i \in (\beta_i, \beta_{i+1})]$.

Then the mechanism in which voters types in partition interval (β_i, β_{i+1}) *send message* i *and the conditional means* $\mathbb{E}[\theta | \theta \in (\beta_i, \beta_{i+1})]$ *of the blocks are averaged is the constrained first-best mechanism.*

This extends [BSXY11]'s result and uses additional proof techniques. The proof is in appendix A.

2.1.2. *Constrained Second-Best-Mechanisms.* We claim that [Sap09]'s result for maximal domains of single-crossing preferences has a counterpart here.

Conjecture 1. *The second-best mechanisms lie in the following class of voting rules: There are* k *possible outcomes* $a_1 \leq a_2 \leq \cdots \leq a_k \in \mathbb{R}$. *Let* $\beta_i, i = 0, 1, \ldots, k$ *be as in Proposition 1. Voter types in* (β_i, β_{i+1}) *send message* i *and the mechanism is a median rule with phantom voters* $x_1, x_2 \ldots, x_{n-1} \in \{a_1, a_2, \ldots, a_n\}$, $\text{median}(a_{m(\theta_1)}, a_{m(\theta_2)}, \ldots, a_{m(\theta_n)}, x_1, x_2, \ldots, x_{n-1})$.

The following observations/claims lead to Conjecture 1. We mark claims we are not certain to hold with an asterisk.

(1) $m^{-1}(\{i\})$ are intervals.
(2) The mechanisms from Conjecture 1 are implementable.
(3) Implementable mechanisms ϕ only attain k values, $\langle \text{range}(k) \rangle \leq k$.*
(4) There is a correspondence between implementable mechanisms and strategyp-proof mechanisms in a game where types $\rho_1, \rho_2, \ldots \rho_n \in \{1, 2, \ldots, k\}$ are sampled where $\rho_i = j$, w.p. $F(m^{-1}(\{j\}))$ the mass of the types in the original game that vote for j, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, k$.*
(5) In the equivalent setup, equivalent mechanism design problem has a maximal set of single-crossing preferences.*

2.2. **Comparative Statics.** In contrast to the related paper [GMS17b], in our model, k is part of the information structure. Therefore, comparative statics in k are meaningful.

2.2.1. *Incentives vs. Information.* In models of restricted communication, decreasing k (a) gives less information for the social planner, but also (b) less incentive constraints—a tradeoff. Can it be quantified?[2]

[2]One should, however, keep in mind that the welfare will always increase, as a mechanism can "emulate" any mechanism with a smaller allowed number of messages by treating to possible messages equal.

2.2.2. *Optimal Limits of Communication.* Given a distribution of types (e.g. wealth in western countries): How does welfare vary in k? Is there a remarkable increase in welfare that makes a certain number of messages preferable?

Appendix A. The First-Best Mechanism

Notation 1. 1 *is an all-one column vector,* $[k] := \{1, 2, \dots, k\}$*,* $\{a_i\}_{i\in A}$ *denotes a set with elements* a_i*,* $i \in A$*.*

Proposition 1. *Let* $a_1 \leq a_2 \leq \cdots \leq a_k \in \mathbb{R}$ *be the centers of a Lloyd-Max optimal partition* [Llo82, Max60]*, i.e. for* $\beta_i := \frac{a_i+a_{i+1}}{2}$*,* $i = 1, 2, \dots, k-1$*,* $\beta_0 = -\infty, \beta_k = \infty$ *(*β *for* boundary*), we have* $a_i = \mathbb{E}[\theta_i|\theta_i \in (\beta_i, \beta_{i+1})]$*.*

Then the mechanism in which voters types in partition interval (β_i, β_{i+1}) *send message* i *and the conditional means* $\mathbb{E}[\theta|\theta \in (\beta_i, \beta_{i+1})]$ *of the blocks are averaged is the constrained first-best mechanism.*

A similar problem to finding the first-best in our model is studied in the electrical engineering literature. This literature calls the problem “average consensus”. Paper [XBK07] studies it for square-loss, which is our quadratic utility subset. The result are dissimilar, however, as the problem there is to make a decentralised decision.

Proof. For convenience, we repeat the welfare maximisation problem: Let F be a probability distribution and $\theta_1, \theta_2 \dots, \theta_n \sim F$ iid random variables (r.v.s). We would like to solve

$$x \max_{\substack{m\colon \mathbb{R}\to[k] \\ \phi\colon [k]^n\to\mathbb{R}}} \mathbb{E}\left[\sum_{i=1}^{n} -(\phi(m(\theta_1), m(\theta_2), \dots, m(\theta_n)) - \theta_i)^2\right]. \tag{1}$$

Problem (1) equals

$$\mathbb{E}\left[\sum_{i=1}^{n} -(\overline{\theta} - \theta_i)^2\right] + \max_{\substack{m\colon \mathbb{R}\to[k] \\ \phi\colon [k]^n\to\mathbb{R}}} \mathbb{E}\left[\sum_{i=1}^{n} -(\phi(m(\theta_1), m(\theta_2), \dots, m(\theta_n)) - \overline{\theta})^2\right],$$

where $\overline{\theta} = \frac{1}{n}\sum_{i=1}^{n}\theta_i$ (caution: it is not the highest possible type). The equality follows from the vector orthogonality relation $x - \overline{x}1 \perp c1$ for any $x \in \mathbb{R}^n$ and $c \in \mathbb{R}$ upon choosing $x = (\theta_1, \theta_2, \dots, \theta_n)$ and $c = \phi(m(\theta_1), m(\theta_2), \dots, m(\theta_n))$. For maximisation we can neglect the first, constant summand. Pulling in a minus, this can be equivalently written as a minimisation,

$$\min_{\substack{m\colon \mathbb{R}\to[k] \\ \phi\colon [k]^n\to\mathbb{R}}} \mathbb{E}\left[\sum_{i=1}^{n} (\phi(m(\theta_1), m(\theta_2), \dots, m(\theta_n)) - \overline{\theta})^2\right].$$

Let us denote $A_i = m^{-1}(\{i\})$, $i \in [k]$. There is a one-to-one correspondence between ordered partitions $\mathcal{A} := \{A_1, A_2, \dots, A_k\}$ of measurable sets in the real line and measurable functions $m\colon \mathbb{R} \to [k]$.

We now give an alternative characterisation of the r.v.

$$Z = \phi(m(\theta_1), m(\theta_2), \dots, m(\theta_n)).$$

The factorisation into functions m, ϕ is equivalent to constancy of Z on events $\{\theta_i \in A_j\}$, $i \in [n]$, $j \in [k]$ which is by probability theory equivalent to $\sigma(\{\{\theta_i \in A_j\} | i \in [n], j \in [k]\})$-$\mathbb{R}$-measurability of Z (where $\sigma(\bullet)$ denoted the σ-algebra spanned by a set of measurable sets). Hence, we can rewrite the optimisation problem as

$$\min_{\substack{\mathcal{A}=\{A_1,A_2,\dots,A_k\}\\ \text{partition}}} \left(\min_{Z \sigma(\{\{\theta_i \in A_j\}\}_{\substack{i\in[n]\\ j\in[k]}})-\mathbb{R}\,\text{mb.}} \mathbb{E}[(Z-\overline{\theta})^2]\right).$$

Fix a partition $\mathcal{A}$ and consider the inner minimisation problem. From probability theory we know that the conditional expectation is the L^2-projection of a random variable onto the subspace of measurable functions w.r.t. the σ-algebra that is conditioned on, [Bil95, Theorem 34.16]:[3]

$$Z = \mathbb{E}[\overline{\theta} | \sigma(\{\{\theta_i \in A_j\}\}_{\substack{i\in[n]\\ j\in[k]}})]$$

Then

$$Z = \frac{1}{n}\sum_{i'=1}^{n} \mathbb{E}[\theta_{i'} | \sigma(\{\{\theta_i \in A_j\}\}_{\substack{i\in[n]\\ j\in[k]}})] = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[\theta_i | \sigma(\{\{\theta_i \in A_j\}\}_{j\in[k]})].$$

The first equality is the linearity of conditional expectation [Bil95, Theorem 34.2 (ii)], the second follows by independence of the θ_i, $i \in [n]$, . Substituting into (1), we obtain

$$\begin{aligned}&\min_{\substack{\mathcal{A}=\{A_1,A_2,\dots,A_k\}\\ \text{partition}}} \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}(X_n - \mathbb{E}[X_n|\sigma(\{\{X_i \in A_j\}\}_{j\in[k]})])\right)^2\right]\\ &\qquad = \min_{\substack{\mathcal{A}=\{A_1,A_2,\dots,A_k\}\\ \text{partition}}} \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[(X_n - \mathbb{E}[X_n|\sigma(\{\{X_i \in A_j\}\}_{j\in[k]})])^2].\end{aligned}$$

The reformulation is possible, as the X_i are independent and, hence, uncorrelated. Finally, as the X_i are identically distributed, there will be a maximiser for each summand with the same partition A_j, $j \in [k]$. To find such a partition, one has to solve

$$\min_{\substack{\mathcal{A}=\{A_1,A_2,\dots,A_k\}\\ \text{partition}}} \mathbb{E}[(\theta_1 - \mathbb{E}[\theta_1|\sigma(\{\{\theta_1 \in A_j\}\}_{j\in[k]})])^2].$$

But this is maximised for the Lloyd-Max optimal partition as stated in the theorem (see [Llo82, BSXY11]). □

[3] L^2-integrability of Z follows from the observation that Z only attains $k^n < \infty$ values.

Remark 1. *Crucial for the proof is that (a) conditional expectation is a best-approximation in a measurable subspace and (b) $X \mapsto \mathbb{E}[X|\mathcal{F}]$ for a σ-algebra $\mathcal{F}$ is linear. One cannot replace the L^2 distance in (a) easily, as for other distance measures, e.g. L^p, $p \neq 2$, property (b) need not hold.*

Appendix B. Related Literature

For matching, [McA02] studied the utility that can be obtained if agents only report a limited number of choices, yielding *coarse* schemes. This has been extended to a model with private information in [HMO11]. Following [BNS07], the literature studied revenue maximisation with private information, reducing the problem largely to [Mye81]. [BSXY11] is particularly close to our setup. They studied a linear-quadratic model of welfare maximising allocation of a good to one agent with a limit on the number of menus the seller can offer. Using results from quantisation [Llo82], the authors characterise first-best mechanisms.

University of Bonn, Bonn 53115, Germany
Email address: `s6anhaup@uni-bonn.de`