Essays in Theoretical Microeconomics

Inauguraldissertation

zur Erlangung des Grades eines Doktors
der Wirtschaftswissenschaften

durch

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

vorgelegt von

Patrick Lahr

aus Frankfurt am Main

2024

Dekan:	Prof. Dr. Jürgen von Hagen
Erstreferent:	Prof. Dr. Benny Moldovanu
Zweitreferent:	Prof. Dr. Felix Bierbrauer
Tag der mündlichen Prüfung:	14. Dezember 2023

Acknowledgements

Writing this thesis was a great privilege and challenge, and I feel deeply grateful for the opportunities and support I received.

First, I thank my supervisors, Benny Moldovanu, Felix Bierbrauer, and Aleh Tsyvinski, for their guidance, encouragement, and patient support. They have introduced me to countless ideas, shared their wisdom, and encouraged me to find my path.

I am also very grateful to Tilman Börgers, Francesc Dilme, Johannes Hörner, Daniel Kraehmer, Matthias Kräkel, Stephan Lauermann, Sven Rady, Larry Samuelson, Dezsö Szalay, and Tymon Tatur, who each shared numerous comments, ideas, and pearls of wisdom with me.

I was fortunate to have Justus Winkelmann as a co-author for one chapter of this dissertation and learned much from him. Furthermore, I thank my fellow students and colleagues — especially Gregorio Curello, Alkis Georgiadis-Harris, Carl Heese, Deniz Kattwinkel, Andreas Klümper, Jan Knöpfle, Sophie Kreutzkamp, Axel Niemeyer, Lucas Pahl, Martin Pollrich, Justus Preusser, Tobias Rachidi, Paul Voss, and Alex Winter — for the fruitful exchange, the passionate debates, and not least for lots of fun.

I profited greatly from the institutions that support Economic Research in Bonn. These include the graduate school, collaborative research center, and the excellence cluster. I thank Benny Moldovanu, Thomas Dohmen, and Sven Rady for their leadership of these institutions.

Similarly, I would like to thank Silke Kinzig, Carola Braun, and Britta Altenburg and there respective teams for the management that keeps these institutions flourishing.

Last, I want to thank my family and friends, who have always supported me during this voyage. I have become significantly indebted and will aspire to repay what I have received, much of it not to those who gave to me but to others that will keep paying it forward.

Patrick Lahr
Bonn, August 2023

Contents

List of Figures

List of Tables

Introduction

This thesis comprises three self-contained essays dealing with the communication and use of information between strategic players. In Chapters 1 and 3, I follow the mechanisms design paradigm, assuming that a player (referred to as the principal or designer) can commit to how she uses the information other players (referred to as agents or voters) provide her. In contrast, Chapter 2 aligns with the cheap-talk literature.

Chapter 1 studies a model of delegation in which allocations are lotteries over a finite set of outcomes. The principal uses randomization to screen agents' von Neumann Morgenstern preferences and I permit the full domain of such preferences. I characterize all mechanisms that cannot be implemented by randomizing over other mechanisms, i.e., the extreme points of the set of mechanisms. A principal who herself maximizes utility in the von Neumann Morgenstern sense does not benefit from random mixing, as her utility is linear, so it is without loss to assume that she always choose an extreme point. It follows that every optimal mechanism using information from the agent must grant him the option to veto one outcome of his choosing. The only other potentially optimal mechanisms are constant and deterministic, dictating an outcome for the agent. A second consequence of this characterization illuminates the complexity of different classes of delegation problems. For instance, the class of problems with at most three different outcomes is simple, while all others are not.

Chapter 2, which is joint work with Justus Winkelmann, discusses a model of advice. In this model, a single decision-maker receives messages from multiple imperfectly informed experts about an unknown state of the world, with binary state and action. We assume that advisors either align with the interests of the decision maker or have state-independent preferences. These assumptions apply to various contexts, such as regulatory proceedings or peer review of research papers. The most informative equilibrium exhibits perfect communication of intermediate signals, but unaligned experts limit the transmission of the best signals. We present the concept of the "fragility of specialization" and show that more information is lost with more specialized experts. If the likelihood of diverging interests among experts is high, communication becomes binary, resembling a voting procedure.

Chapter 3 investigates utilitarian welfare maximization among strategy-proof and anonymous mechanism in a social choice model. I focus the question: Is surjectivity, i.e., the requirement that a social choice function must select any possible outcome under some preference profile, restrictive? In the case decisions between more then three outcomes and considering the full domain of preferences, the answer is affirmative, due to the Gibbard-Satterthwaite theorem. Decisions between two options typically outperform dictatorships. When preferences are single-peaked, and feasible mechanisms satisfy surjectivity, the answer remains positive. However, the most extreme outcomes need not be excluded to maximize welfare. I provide a tight welfare guarantee for surjective mechanisms as a function of the number of possible outcomes. In particular, no fraction of optimal welfare can be guaranteed when the numbers of outcomes is unbounded.

Chapter 1

An Extreme-Points Approach to Multidimensional Delegation: $1, 2, 3, \infty^{\star}$

1.1 Introduction

The allocation of decision rights is a fundamental question for organizational design. I study a simple model of such delegation in which a principal (she) has to make a potentially random decision between a finite set of outcomes, affecting both her and an agent (he). The agent is privately informed about the state of the world, which determines both players' preferences. The agent's preferences are drawn from the entire domain of von Neumann-Morgenstern (vNM) preferences over lotteries. The significance of this model is threefold.

First, it is more general than most of the literature on delegation, where both players are assumed to have concave, often quadratic loss, preferences over a continuum of states. In contrast, this model makes no restrictions on the kinds of von Neumann Morgenstern preferences players may have yet assumes a finite set of outcomes. The only general property I do exploit is the linearity of the utility functions in lotteries. As a consequence, the main

⋆ I thank my advisors, Benny Moldovanu, Felix Bierbrauer, and Aleh Tsyvinsky, for their continued guidance and support. For helpful discussions, I thank Gregorio Curello, Alkis Georgiadis-Harris, Avi Lichtig, Deniz Kattwinkel, Jan Knöpfle, Daniel Krähmer, Roger Myerson, Axel Niemeyer, Lucas Pahl, Martin Pollrich, Justus Preußer, Tobias Rachidi, Mario Schulz, Dezsö Szalay, Omer Tamuz, Paul Voß, Alex Winter and Mengxi Zhang.

qualitative behavior will be different, with players choosing extreme points of lotteries spaces rather than interior points of outcome spaces.[1]

Second, it accurately describes economic situations, such as a manager allocating projects to a worker or a municipality allocating social housing to those in need. Since those allocating need to balance supply and demand, the preferences and allocations of others act as a de-facto randomization and commitment device, with individual recipients frequently facing rationing and hence an unsure allocation.

Lastly, allowing for all von Neumann-Morgenstern preferences puts this model in a similar systematic position as the full-domain model of social choice. However, there is a strong difference: While in social choice the full domain model is known for its stark lack of appealing mechanisms, in the absence of strategic interaction between multiple agents, the full domain setup has the richest set of mechanisms. This simply follows from the fact that every mechanism on a restricted domain can be extended in potentially multiple ways to a larger domain of preferences. Therefore, whatever is true for all mechanisms on the full domain, translates to all possible domains.

My main result characterizes the set of extreme points of the incentive-compatible direct mechanisms. For any concrete problem, one of these extreme points is a, and generically the unique, second-best solution by Bauer's maximum principle. I do not study which extreme point is optimal for which problem and hence do not solve the problem of optimal delegation. In this respect, I only give a necessary condition for a mechanism to be the unique optimizer.[2] However, this characterization allows me to derive two general insights on optimal solutions.

Grant veto or dictate: An extreme point grants the agent a veto, or dictates a single outcome. Hence, the agent either has no autonomy at all or always has the choice to select a lottery s.t. a given outcome is not in its support.

Complexity of problems: For three outcomes, it is sufficient to focus on mechanisms that offer three different lotteries. This resembles the post-it price structure in the monopolistic seller problem with one good, in which two alternatives are sufficient. In stark contrast, for more than three outcomes, the set of extreme points lies dense in the set of mechanisms granting a veto. In particular, optimal mechanisms might require an arbitrary amount of different lotteries to be offered.

1. Selecting an interior outcome for sure is an extreme point of the respective lottery space.

2. If multiple extreme points are optimal, then so is every convex combination. This opposite case with a unique optimum is, however, generic.

On a conceptual level, this paper introduces new tools to the delegation literature by primarily studying mechanisms via their associated delegation sets and studying the indecomposability problem from convex geometry. The characterization relies on characterizing the extreme points of the set delegation sets, which are the extreme points of the set of convex bodies in the unit simplex. These are the indecomposable and maximal delegation sets.

I now summarize the structure of the arguments that allows me to derive the aforementioned results.

A convex set is said to be indecomposable if it has no representation as a Minkovski sum of two convex sets, both not homothetic[3] to the sum. A Minkovski sum is the set of all pairwise sums of its summands. Both indecomposable sets and Minkovski sums are widely studied objects in convex geometry. Grünbaum et al. (1967) is an excellent reference. It is well understood and easy to check algebraically whether a two-dimensional convex body or any polytope is indecomposable. Further, I call a set maximal in the simplex if it is either a vertex of the simplex or touches all facets[4].

My characterization follows from separate characterizations. Both are of independent interest. Note that the revelation and taxation principles apply in this setting. Hence any implementable decision rule can be implemented both by a direct mechanism and indirectly by letting the agent decide from a delegation set. Since the agent has linear preferences, I can restrict attention to delegation sets that are convex bodies, i.e., convex, compact, and non-empty subsets of probability distributions over k outcomes.[5] Given a direct mechanism, one can easily describe the associated delegation set and vice versa. I strengthen this equivalence of the two principles by demonstrating the existence of an isomorphism between direct mechanisms, and delegation sets that preserve convex combinations. In particular, the set of delegation sets has a convex structure, where convex combinations of delegation sets refer to the Minkovski summation as defined above. Note that I here and below refer to a convex set of convex sets. Except in specifically noted circumstances, mentions of extreme points will refer to the objects within this meta-structure. This directly yields theorem 1.3: A direct mechanism is an extreme point if and only if its associated delegation set is an extreme point.

3. Two convex set $K, L \subset \mathbb{R}^d$ are homothetic if $K = \alpha L + b$, with $\alpha \in \mathbb{R}^+$ and $b \in \mathbb{R}^d$

4. The facets of the simplex represent the subsets of all lotteries for which a given outcome has probability zero.

5. An agent can always select an option from the convex hull by randomizing over reports, yet will never have the incentive to do so.

For the second equivalence, I study the above delegation sets as geometric objects. These are convex bodies within the unit simplex. A convex body can be represented as a convex combination of other convex bodies in the simplex homothetic to the first if and only if it is non-maximal. Hence, maximality in the simplex is necessary for convex bodies to be an extreme point.

In contrast, if a convex body in the simplex is indecomposable, any representation as a convex combination must, in turn, imply that the parts are homothetic to the original. Hence maximality and indecomposability combined are sufficient for convex bodies to be an extreme point. The reverse is also true. This yields Theorem 1.7: A convex body of probability distributions over a finite set of outcomes is an extreme point of the set of all such bodies if and only if it is maximal and indecomposable. Maximality has a simple economic interpretation: A maximal delegation set either leaves no choice at all, i.e., always implements the same outcome, or it grants a veto, i.e., the agent can make sure that one outcome never realizes. Jointly with theorem 1.3, it implies my main characterization: A direct non-constant mechanism is an extreme point if it grants a veto and has an indecomposable delegation set. The above-mentioned results are then all direct consequences of the literature on indecomposable sets.

Related Literature

Holmstrom (1984) has initiated a vast field of research on delegation. Most of this literature has focused on a one-dimensional action and type space and parameterized, mostly quadratic loss utility functions, e.g., Dessein (2002), Alonso and Matouschek (2008), Amador and Bagwell (2013), and Kolotilin and Zapechelnyuk (2019).

I deviate from this main strand in two ways: The set of alternatives consists of lotteries over k outcomes, and arbitrary vNM preferences are permissible. Both the space of alternatives and preferences are, therefore, multi-dimensional and compact, and the utility functions of both players are linear.

A smaller number of publications also consider multi-dimensional types or action spaces. These include Bendor and Meirowitz (2004), Koessler and Martimort (2012), Frankel (2016) and Kleiner (2022), yet they still deviate from this present paper in the aforementioned ways.

Lastly, papers that study delegation over a finite set of outcomes include Che, Dessein, and Kartik (2013), Nocke and Whinston (2013), and Armstrong and Vickers (2010). In these models, the agent selects one of the alternatives for the principal. Upon selection, the principal receives a signal on its quality and can accept or reject the agent's recommendation. The anticipation of this signal acts as a screening device. My model shows that commitment to enacting the agent's recommendation contingent on a random event can be effective, even if this event is completely independent of the relevant state of the world.

The closest paper to the present is, however, on auctions. Manelli and Vincent (2007) characterize the extreme points of the multi-good monopolistic seller problem. Given the differences between the models, also the characterizations differ substantially. I will discuss the relationship parallel to my own results below. Kleiner, Moldovanu, and Strack (2021) characterize extreme points of the monotone functions that fulfill a majorization constraint and apply their characterization to several one-dimensional economic design problems. Both papers apply a different set of methods but share the approach to directly studying the convex structure of the set of mechanisms. In contrast, I apply an indirect approach via the convex structure of feasible delegation sets. The approaches are, of course, deeply related, yet the indirect approach allows me to apply results from convex geometry, which fits my model exactly.

This paper is also more broadly related to a wide class of models of screening without transfers. Most of these models analyze binary decisions and employ a variety of additional tools for screening. Examples include linking several independent decisions, (Jackson and Sonnenschein (2007)), using correlated information (Kattwinkel et al. (2022))

More generally, my model is deeply connected to models of multi-dimensional mechanism design and, in particular, the multi-object monopolistic seller problem, studied in, e.g., Rochet and Choné (1998), Jehiel, Meyer-Ter-Vehn, and Moldovanu (2007), Daskalakis, Deckelbaum, and Tzamos (2015), Hart and Reny (2015) and Haghpanah and Hartline (2021), since the later is a domain restriction of the present model as discussed in the introduction.

Recent work in this literature started by Hart and Nisan (2019) has focused on simple mechanisms in the sense that they have a small range of outcomes. However, there is a tight connection to the study of extreme points since extreme points are delegation set size efficient, i.e., when restricting to mechanisms with delegation sets lower than a given size, extreme points will retain their status as a sufficient candidate set. Hart and Nisan (2017) demonstrate that for two goods and arbitrary correlation structures, finite mechanisms may not secure any positive fraction of optimal revenue. Since their model is a special case of mine, this result directly translates to my model if $k \geq 8$, yet I conjecture it to be true for $k \geq 4$. Finally, Babaioff, Gonczarowski, and Nisan (2017) study how fast optimal revenue can be approximated by finite delegation set mechanisms when valuations for different goods are understood to be independent.

1.2 Model

1.2.1 Notation

$Conv(.)$ and $\overline{Conv}(.)$ denote the convex hull and closed convex hull of a set, respectively. $Ext(.)$ denotes the set of extreme points. Scalar multiplication and addition of sets of real vectors refer to the following operations:

$$\lambda M = \{\lambda m | m \in M\}$$

and

$$M + M' = \{m + m' | m \in M, m' \in M'\}.$$

The " + " here is the standard definition of the Minkovski or vector sum. $M \sim M'$ will denote that M and M' are homothetic.

1.2.2 Setting

A principal (she) selects an alternative affecting herself and an agent (he). The set of alternatives A is the set of probability distributions over some finite set of outcomes $A = \Delta\{1, \dots, k\}$. Hence an element $a \in A$ is of the form $a = (a_1, \dots, a_k)$, where $a_i \geq 0$ for $l = 1, 2, \dots, k$ and $\sum_{i=1}^{k} a_i = 1$. The agent is privately informed about his type $\theta = (\theta_1, \dots, \theta_k) \in \Theta = \{[0,1]^k : \max_i \theta_i = 1$ and $\min_i \theta_i = 0\}$ which represents his (normalized) Bernoulli utilities over outcomes. It is drawn from some prior μ on Θ. Hence his utility functions U reads.

$$U(a, \theta) = a \cdot \theta$$

I also will assume throughout that the principal is a von Neumann-Morgenstern expected utility maximizer, i.e., given a type θ, her preferences over lotteries are characterized by her Bernoulli utilities over outcomes $v(\theta) = (v_1(\theta), \dots, v_k(\theta))$. The principal's ex-ante utility function V is then given by:

$$V(a) = \mathbb{E}_\mu [a \cdot v(\theta)]$$

1.2.3 Mechanisms

The principal commits to a mechanism, and I assume the agent plays the best response.

A mechanism is given by a message space S and a choice functions $f : S \to A$ that specifies an alternative for each message sent by the agent.

Due to the revelation principle, it is without loss to restrict attention to direct mechanisms, i.e., mechanisms with $S = \Theta$ and where agents have incentives to report their true type. Formally, a mechanism satisfies incentive compatibility if

$$U(f(\theta),\theta) \;\geq\; U(f(\theta'),\theta) \text{ for each } \theta,\theta' \;\in\; \Theta. \tag{IC}$$

When the agent is indifferent between the alternative assigned to his type and another, I assume that the principal can select which best response is played by the agent.[6]

Definition 1.1. A mechanism with choice rule f satisfies principal preferred tie-breaking if the following conditions are satisfied.

i For any $\theta,\theta' \in \Theta$ s.t. $U(f(\theta),\theta) = U(f(\theta'),\theta)$ implies $V(f(\theta),\theta) \geq V(f(\theta'),\theta)$.

ii For any $\theta,\theta' \in \Theta$ s.t. $U(f(\theta),\theta) = U(f(\theta'),\theta)$ and $V(f(\theta),\theta) = V(f(\theta'),\theta)$, implies $f(\theta) \geq_{lex} f(\theta')$, where "$\geq_{lex}$" refers to the lexicographical order.

Henceforth, when considering direct mechanisms, I will restrict attention to those that satisfy incentive compatibility and principal preferred tie-breaking. I will denote the set of choice rules implemented by such mechanisms $\mathscr{F}$. As a shorthand, I will refer to a mechanism f or the set of mechanisms $\mathscr{F}$ to refer to direct mechanisms that implement the respective choice rules.

Restricting attention to $\mathscr{F}$ can not reduce the principal's utility. I can hence state her problem as follows:

$$\max_{f\in\mathscr{F}} E_\mu\left[f(\theta)\cdot v(\theta)\right]$$

1.2.4 Delegation Sets

A delegation set M is a convex, compact, non-empty subset of A. I will say that M has size n if $|Ext(M)| = n$. I will denote the set of all delegation sets with $\mathscr{M}$. Consider the following indirect mechanism: The agent has a message space $S = M$, and the alternative corresponding to his message

6. This selection follows Holmstrom (1984). Kamenica and Gentzkow (2011) use sender-preferred equilibrium as a related notion in information design.

realizes.[7] I will say that M is the delegation set of a direct mechanism $f \in \mathscr{F}$ if the allocation the agent receives under f is a best response in this indirect mechanism.

Definition 1.2. M is the delegation set of a mechanism $f \in \mathscr{F}$ if

$$f(\theta) \in \arg\max_{a \in M} U(\theta, a).$$

In principle, a mechanism could have multiple delegation sets, which is why at this point, to speak "the delegation set" is an abuse of language. Yet, I will demonstrate below that a unique delegation set is associated with every mechanism in our setting.

1.3 Extreme Points of the Mechanism Set

1.3.1 The Strong Equivalence between the Revelation and Taxation Principle

It is well known that both direct mechanisms and delegation sets are comprehensive notions to capture all implementable decision rules. In what follows, I will primarily work with delegation sets. Yet to translate my findings onto direct mechanisms, I need to show an equivalence between both approaches that extends to the convex structure of both sets. In particular, the set of delegation sets has a convex structure induced by Minkovski's addition on sets. Hence the following is a strengthened version of equivalence between the revelation and taxation principles:

Theorem 1.3. *Define* $T : \mathscr{F} \to \mathscr{M}$, *s.t.* $T(f) = \overline{Conv}(f(\Theta))$ *for all* $f \in \mathscr{F}$. *Then* T *satisfies the following properties:*

i T *is a bijection that maps* f *to the delegation set of* f.

ii $|f(\Theta)| = n < \infty$ *if and only if* $T(f)$ *is a polytope with* n *vertices.*

iii T *preserves convex combinations, i.e., for all* $f, f', f'' \in \mathscr{F}$

$$f = \lambda f' + (1-\lambda) f'' \iff T(f) = \lambda T(f') + (1-\lambda) T(f'') \tag{1.1}$$

In particular, f *is an extreme point of* $\mathscr{F}$ *if and only if* $T(f)$ *is an extreme point of* $\mathscr{M}$.

7. This indirect approach to mechanism design is referred to as the taxation principle going back to Hammond (1979).

Proof. See Appendix A □

The first two points are the equivalence of revelation and taxation principle restated in the context of this model. In contrast, the third point shows that this equivalence preserves relevant convex structure. Therefore, it is worthwhile to view both convex combinations from the agent's perspective and view them as random mixing between different mechanisms. This interpretation on the direct side is straightforward. For the equivalence to hold, the agent must achieve the same overall selection if he selects from the Minkowski sum as if he were to select from both delegations sets separately. However, by definition, the first choice is from all potential combinations, which is equivalent.

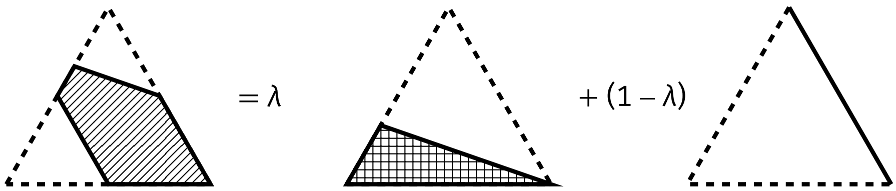


Figure 1.1. A convex combination of two delegation sets.

This result is the central connection between direct mechanisms and convex sets of probability distributions I employ in this paper.

1.3.2 Extreme Points of Delegation Sets

Extreme points of our delegation set $\mathscr{M}$ are such delegation sets that are not the Minkowski sums of appropriately scaled feasible delegation sets. Whether a given convex body has a representation as the Minkowski sum of other convex bodies is a well-studied problem that goes back to Gale (1954). Schneider (2014) is an excellent reference. The difference between both problems is the feasibility constraints for the sum and the summands present in the current setting.

To connect both problems and introduce the relevant notions, let me recall that two convex bodies $K, K' \in \mathbb{R}$ are homothetic if $K = \alpha K' + v$, for some $\alpha \in \mathbb{R}$ and $v \in \mathbb{R}^d$. Any convex body K can be written as the sum $K = (\lambda K - v) + ((1-\lambda)K + v)$ for $\lambda \in (0,1)$ and $v \in \mathbb{R}^d$. I will hence call such decompositions trivial.

Definition 1.4. A convex body $K \subset \mathbb{R}^d$ is decomposable if it has a non-trivial decomposition, i.e., if there exist convex bodies $B, C \subset \mathbb{R}^d$, B and C not homothetic to K s.t. $K = B + C$. A convex body that is not decomposable is indecomposable.

To compare the notion of $M \in \mathscr{M}$ being an extreme point of $\mathscr{M}$ or being indecomposable, suppose there exist convex bodies B, C both different from M s.t. $M = B + C$. B and C are a counterexample to M being indecomposable if they are both not homothetic to M. In contrast B and C are a counterexample to M being an extreme point of $\mathscr{M}$ if there exists a $\lambda \in (0,1)$, s.t. $\frac{1}{\lambda}B$ and $\frac{1}{1-\lambda}C$ are both feasible delegation sets, i.e. subsets of A, since then $M = \lambda \frac{1}{\lambda}B + (1-\lambda)\frac{1}{1-\lambda}C$. In particular, no notion implies the other.

To connect the two concepts, I next discuss a notion of delegation sets for which the constraint to be inside a given simplex is binding.

Definition 1.5. i A delegation set is dictating if it consists of a single vertex of the simplex.

ii A delegation set grants a veto if it has a non-empty intersection with all facets of A.

iii A delegation set M is maximal if it is dictating or granting a veto.

The next lemma establishes maximality as a necessary condition for a delegation set to be an extreme point.

Lemma 1.6. *A delegation set $M \in \mathscr{M}$ has a decomposition of the form $M = \lambda M' + (1-\lambda)M''$, with M, M', M'' all sets are non-identical and homothetic, if and only if M is non-maximal.*

In particular, for such a delegation set, there exists delegation sets $M_h, M_d \in \mathscr{M}$ and $\lambda \in (0,1)$, where $M_h \sim M$ and M and M_d consists of a single vertex of A s.t., $M = \lambda M_h + (1-\lambda)M_d$.

Proof. See Appendix A. □

A non-maximal delegation set has at least one outcome s.t. the probability of that outcome is bounded away from zero, regardless of the selection of the agent. Hence the delegation that has a convex component that disregards the agents choice. Geometrically, if $M \in \mathscr{M}$ is not maximal, it is fully enclosed in a smaller sub-simplex within A. It can be rewritten as a convex combination of a scaled-up original version and a vertex.

There is a close connection between this result and the frequent observation that "there is no distortion at the top" in several models of mechanism design, such as the multi-object monopolistic seller problem. The usual argument for this observation is that the highest type does not grant information rents to any other type. However, a distortion at the top implies a trivial decomposition of the mechanism into a scaled-up version of the same mechanism and a mechanism that never allocates a given good, contradicting optimality.

$M \in \mathscr{M}$ being maximal is a necessary condition to be an extreme point. If, in addition, it is indecomposable, it has to be an extreme point since it neither has a trivial nor non-trivial decomposition into non-identical delegation sets. In particular, it is not the sum of appropriately scaled feasible delegation sets. It is not obvious that these conditions are also jointly necessary. For example, a delegation set might be decomposable, yet the scaled parts might not be feasible for any decomposition. This is the case when feasibility depends on being a subset of, e.g., a square. However, in this setting, being maximal and indecomposable characterizes extreme points.

Theorem 1.7. *A delegation set $M \in \mathscr{M}$ is an extreme point of $\mathscr{M}$ if and only if M is maximal, i.e., dictating or granting a veto and indecomposable.*

Proof. See Appendix A □

1.3.3 Extreme Points of the set of Direct Mechanisms

Theorem 1.3 and 1.7 jointly characterize the extreme points of $\mathscr{F}$. I summarize this below.

Theorem 1.8. *A non-constant mechanism $f \in \mathscr{F}$ is an extreme point of $\mathscr{F}$ if it grants a veto and $T(f)$ is indecomposable.*

Proof. Immediate by Theorems 1.3 and 1.7.

□

At this point, everything known about indecomposable convex bodies is a statement on extreme points in our model and vice versa. A full characterization of extreme points of direct mechanisms in terms other than these would substantially extend the existing geometry literature. However, no such attempt is made here.

Indcomposability is well understood in two dimensions and for polytopes in any dimension, which correspond to finite mechanisms. In particular, simple algebraic procedures exist to check whether a polytope is indecomposable by calculating the rank of an associated matrix. SeeSmilansky (1987) for details.

In the remainder of this section, I will focus on the consequences of two results on indecomposable sets.

As with the connection of granting a veto and no distortion at the top, both results have a substantially different yet closely related corresponding phenomenon in the multi-good monopolistic seller problem.

Delegation with one or two outcomes is each trivial. In the latter, the principal has to decide between taking the decision herself or full delegation. In contrast, the case with three alternatives allows intermediate extreme points, which still all have a simple structure.

Theorem 1.9. *Suppose $k \leq 3$. Then a non-constant mechanism f is an extreme point of $\mathscr{F}$ if it grants a veto and has a menu size of at most three, i.e., $f(\Theta) \leq k$.*

Proof. See Appendix A □

Points, line segments, and triangles are the only indecomposable convex bodies in two dimensions.[8]

The case $k = 3$ corresponds to the one good case in the monopolistic seller problem. There are three possible "trades" of probabilities for two outcomes against a third. Any bundle of two trades is on itself a single trade, even though potentially for a different "price" or rate of substitution.

The situation with four or more outcomes is in sharp contrast to this classification. It is similar to how bundling opportunities and possible discounts lead to a rich set of extreme points. In both cases, an additional lever to screen can be combined in a continuum of ways.

Theorem 1.10. *Suppose $k \geq 4$. Then the non-constant extreme points of $\mathscr{F}$ with finite range are a dense subset of mechanisms granting a veto.*

Proof. See Appendix A □

There are two important consequences to this characterization. First, although extreme points can have delegation sets of infinite sizes, any such extreme point is arbitrarily close to a mechanism with a finite delegation set size. Therefore it is approximately without loss to focus attention on this better-understood class. The second consequence, however, is that even this class is so rich that it is not an easily comprehensible and sufficient candidate class for optimization, as seen for $k = 3$.

1.4 Application: Allocation of Social housing

An applicant applying at the local government for housing can be allocated option 0 no housing, at no cost or any of k housing options, with positive costs to the government. The applicant's type is his value for each option, while the local government maximizes the agent's expected value minus expected costs. Since their is no conflict of interest, when the agent prefers no housing, no housing is always part of the optimal set of choices. Indeed, whenever no housing is not offered, both the local government and agent benefit from including it.

8. The result was first mentioned in Gale (1954), yet no proof was published. It was later independently demonstrated by Meyer (1972) and Silverman (1973).

Given this observation the only dictating mechanism worth considering never offers housing. In contrast all veto mechanisms need to also offer one lottery s.t. the applicant always receives some housing option, since such a lottery is the relevant veto against no housing.

If there are only two housing options then offering no housing, an option that guarantees some housing option and potentially a third arbitrary lottery, designed to screen preferences between houses, by including a chance that no housing is offered is always enough to maximize the local governments problem. However as we have seen, with three or more options, optimality might require arbitrarily complicated choices from the applicant.

1.5 Conclusion

I have characterized the extreme points of the full domain model of delegation with a finite set of outcomes. This characterization relies on a novel connection to the indecomposability problem from convex geometry. This connection is the methodological contribution of this paper. Part of the characterization also implies that extreme points are dictating or granting a veto.

I have also applied this characterization and existing insights on the indecomposability problem to study the complexity of the model as it depends on the number of outcomes. This analysis has lead to two results with starke differences. With three outcomes, optimal mechanisms have a simple structure and offer at most three choices. With more then four outcomes the set of candidates for optimality can not be reasonably reduced further then to restrict attention to the aforementioned mechanisms that dictate or grant a veto. This robust prediction is in fact the only possible robust prediction in this setting.

The natural next step is to consider how complexity can again be reduced by limiting the model, most notably by studying how the set of extreme points is reduced, if one limits the domain of preferences.

Appendix 1.A Proofs

Proof of Theorem 1.3.

i) Fix an arbitrary $\theta \in \Theta$, then

$$f(\theta) \in \arg\max_{a \in f(\Theta)} U(\theta, a) \subset \arg\max_{a \in \overline{Conv}(f(\Theta))} U(\theta, a)$$

Hence $T(f)$ is a delegation set of f.

Next, I show that an inverse function exists. For this, define f^M s.t.

$$f^M(\theta) \in \arg\max_{a\in M} U(\theta, a)$$

and f^M satisfies principle-preferred tie-breaking. Now $T^{-1} : \mathscr{M} \to \mathscr{F}$ with $T^{-1}(M) = f^M$ is the required inverse function by construction.

ii) If $|f(\Theta)|$ is finite, all extreme points of $T(f)$ are exposed. Hence $f(\Theta) = Ext(T(f))$.

iii) I will prove sufficiency first. For this I assume fix $f, f', f'' \in \mathscr{F}$ s.t. $f = \lambda f' + (1-\lambda)f''$ for some $\lambda \in (0,1)$. I can deduce the following:

$$\begin{aligned} T(f) &= T(\lambda f' + (1-\lambda)f'') \\ &= \overline{Conv}(\lambda f'(\Theta) + (1-\lambda)f''(\Theta)) \\ &= \overline{Conv}(\lambda f'(\Theta)) + \overline{Conv}((1-\lambda)f''(\Theta)) \\ &= \lambda\overline{Conv}(f'(\Theta)) + (1-\lambda)\overline{Conv}(f''(\Theta)) \\ &= \lambda T(f') + (1-\lambda)T(f'') \end{aligned}$$

Now for the reverse direction fix $f, f', f'' \in \mathscr{F}$ s.t. $\overline{Conv}(f(\Theta)) = \lambda\overline{Conv}(f'(\Theta)) + (1-\lambda)\overline{Conv}(f''(\Theta))$ for some $\lambda \in (0,1)$. Then I can deduce:

$$\begin{aligned} f &= T^{-1}(T(f)) \\ &= T^{-1}(\overline{Conv}(f(\Theta))) \\ &= T^{-1}(\lambda\overline{Conv}(f'(\Theta)) + (1-\lambda)\overline{Conv}(f''(\Theta))) \\ &= \lambda T^{-1}\overline{Conv}(f'(\Theta)) + (1-\lambda)T^{-1}\overline{Conv}(f''(\Theta)) \\ &= (\lambda T^{-1}T(f') + (1-\lambda)T^{-1}T(f'')) \\ &= \lambda f' + (1-\lambda)f'' \end{aligned}$$

□

Proof of Lemma 1.6.

Suppose $M \in \mathscr{M}$ has an empty intersection with one facet. Without loss of generality, assume that for all $a = (a_1, \ldots, a_k) \in M$, $a_1 \neq 0$. Since M is closed, there exists an $\varepsilon > 0$, s.t. $a_1 \geq \varepsilon$ for all $a \in M$. Define

$$M_\varepsilon = \left\{ a_\varepsilon = (\frac{1}{1-\varepsilon}a_1 - \varepsilon, \frac{1}{1-\varepsilon}a_2, \ldots, \frac{1}{1-\varepsilon}a_k) | a \in M \right\}$$

M_ε is a feasible delegation set since all probabilities in all alternatives are positive and add to 1. It is then easy to check that

$$M = \varepsilon(1,0,\dots,0) + (1-\varepsilon)M_\varepsilon.$$

For the reverse, suppose $M \in \mathscr{M}$ is maximal and suppose

$$M = \lambda M' + (1-\lambda)M''$$

for some $M', M'' \in \mathscr{M}$, s.t. $M \sim M' \sim M''$. If one of the parts does not intersect a facet, so does M, but since M is maximal, so are M', M''. Yet this implies $M = M' = M''$.

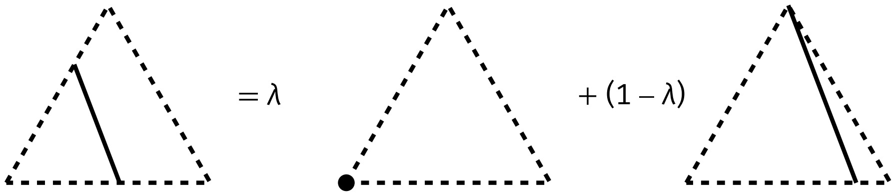


Figure 1.A.1. Illustration of the proof of Lemma 1.6

□

Proof of Theorem 1.8.

Suppose $M \in \mathscr{M}$ is indecomposable and maximal, and suppose there exists $M', M'' \in \mathscr{M}$, s.t.

$$M = \lambda M' + (1-\lambda)M''$$

since M is indecomposable $M \sim M' \sim M''$, yet since M is maximal by Lemma 1: $M = M' = M''$.

Suppose M is an extreme point. Then M is maximal by Lemma 1. Suppose there exist convex bodies $K', K'' \in \mathbb{R}^d$, s.t. $M = K' + K''$.

There are unique maximal sets $M', M'' \in \mathscr{M}$, s.t. $M' \sim K'$ and $M'' \sim K''$. Therefore

$$M = \lambda M' + (1-\lambda)M''$$

□

Proof of Theorem 1.9.

Theorem (Meyer (1972) and Silverman (1973)). In $\mathbb{R}^2$, an indecomposable compact convex set must be either a point, a line segment, or a triangle.

Given this statement, the theorem is immediate from theorem 1.9.

□

Proof of Theorem 1.10.

Theorem (Shephard (1963))**.** If all the 2-faces of a polytope P are triangles, then P is indecomposable.

This set is dense in the set of convex bodies for the Hausdorff metric. See, e.g., Schneider (2014). Given this statement, the theorem is immediate from theorem 1.9.

□

References

Alonso, Ricardo, and Niko Matouschek. 2008. "Optimal delegation." *Review of Economic Studies* 75 (1): 259–93. [6]

Amador, Manuel, and Kyle Bagwell. 2013. "The theory of optimal delegation with an application to tariff caps." *Econometrica* 81 (4): 1541–99. [6]

Armstrong, Mark, and John Vickers. 2010. "A model of delegated project choice." *Econometrica* 78 (1): 213–44. [6]

Babaioff, Moshe, Yannai A Gonczarowski, and Noam Nisan. 2017. "The menu-size complexity of revenue approximation." In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing,* 869–77. [7]

Bendor, Jonathan, and Adam Meirowitz. 2004. "Spatial models of delegation." *American Political Science Review* 98 (2): 293–310. [6]

Che, Yeon-Koo, Wouter Dessein, and Navin Kartik. 2013. "Pandering to persuade." *American Economic Review* 103 (1): 47–79. [6]

Daskalakis, Constantinos, Alan Deckelbaum, and Christos Tzamos. 2015. "Strong duality for a multiple-good monopolist." In *Proceedings of the Sixteenth ACM Conference on Economics and Computation,* 449–50. [7]

Dessein, Wouter. 2002. "Authority and communication in organizations." *Review of Economic Studies* 69 (4): 811–38. [6]

Frankel, Alex. 2016. "Delegating multiple decisions." *American economic journal: Microeconomics* 8 (4): 16–53. [6]

Gale, David. 1954. "Irreducible convex sets." *Proc. Intern. Congr. Math., Amsterdam* 2: 217–18. [11, 14]

Grünbaum, Branko, Victor Klee, Micha A Perles, and Geoffrey Colin Shephard. 1967. *Convex polytopes.* Vol. 16. Springer. [5]

Haghpanah, Nima, and Jason Hartline. 2021. "When is pure bundling optimal?" *Review of Economic Studies* 88 (3): 1127–56. [7]

Hammond, Peter J. 1979. "Straightforward individual incentive compatibility in large economies." *Review of Economic Studies* 46 (2): 263–82. [10]

Hart, Sergiu, and Noam Nisan. 2017. "Approximate revenue maximization with multiple items." *Journal of Economic Theory* 172: 313–47. [7]

Hart, Sergiu, and Noam Nisan. 2019. "Selling multiple correlated goods: Revenue maximization and menu-size complexity." *Journal of Economic Theory* 183: 991–1029. [7]

Hart, Sergiu, and Philip J Reny. 2015. "Maximal revenue with multiple goods: Nonmonotonicity and other observations." *Theoretical Economics* 10 (3): 893–922. [7]

Holmstrom, Bengt. 1984. "On the theory of delegation." In *Bayesian Models in Economic Theory,* edited by M.Boyer and R. Kihlstrom. New York: North-Holland. [6, 9]

Jackson, Matthew O, and Hugo F Sonnenschein. 2007. “Overcoming incentive constraints by linking decisions 1.” *Econometrica* 75 (1): 241–57. [7]

Jehiel, Philippe, Moritz Meyer-Ter-Vehn, and Benny Moldovanu. 2007. “Mixed bundling auctions.” *Journal of Economic Theory* 134 (1): 494–512. [7]

Kamenica, Emir, and Matthew Gentzkow. 2011. “Bayesian persuasion.” *American Economic Review* 101 (6): 2590–615. [9]

Kattwinkel, Deniz, Axel Niemeyer, Justus Preusser, and Alexander Winter. 2022. “Mechanisms without transfers for fully biased agents.” *arXiv preprint arXiv:2205.10910.* [7]

Kleiner, Andreas. 2022. “Optimal Delegation in a Multidimensional World.” *arXiv preprint arXiv:2208.11835.* [6]

Kleiner, Andreas, Benny Moldovanu, and Philipp Strack. 2021. “Extreme points and majorization: Economic applications.” *Econometrica* 89 (4): 1557–93. [7]

Koessler, Frédéric, and David Martimort. 2012. “Optimal delegation with multi-dimensional decisions.” *Journal of Economic Theory* 147 (5): 1850–81. [6]

Kolotilin, Anton, and Andriy Zapechelnyuk. 2019. “Persuasion meets delegation.” *arXiv preprint arXiv:1902.02628.* [6]

Manelli, Alejandro M, and Daniel R Vincent. 2007. “Multidimensional mechanism design: Revenue maximization and the multiple-good monopoly.” *Journal of Economic theory* 137 (1): 153–85. [7]

Meyer, Walter. 1972. “Decomposing plane convex bodies.” *Archiv der Mathematik* 23 (1): 534–36. [14, 17]

Nocke, Volker, and Michael D Whinston. 2013. “Merger policy with merger choice.” *American Economic Review* 103 (2): 1006–33. [6]

Rochet, Jean-Charles, and Philippe Choné. 1998. “Ironing, sweeping, and multidimensional screening.” *Econometrica,* 783–826. [7]

Schneider, Rolf. 2014. *Convex bodies: the Brunn–Minkowski theory.* 151. Cambridge university press. [11, 18]

Shephard, Geoffrey C. 1963. “Decomposable convex polyhedra.” *Mathematika* 10 (2): 89–95. [18]

Silverman, Ruth. 1973. “Decomposition of plane convex sets. I.” *Pacific Journal of Mathematics* 47 (2): 521–30. [14, 17]

Smilansky, Zeev. 1987. “Decomposability of polytopes and polyhedra.” *Geometriae dedicata* 24 (1): 29–49. [13]

Chapter 2

The Fragility of Specialized Advice[⋆]

Joint with Justus Winkelmann

2.1 Introduction

Consider a receiver who must decide on an action but who relies on multiple senders to obtain relevant information.[1] Given any particular question, the more specialized the senders are, the more information is concentrated among a few of them, because fewer and fewer will know anything about the subject at hand. Greater specialization can benefit decision making, because having one source that conveys perfect information is superior to having many sources whose majority advice may err. On the other hand, relying on only a few sources naturally grants each of them more influence, which increases the risk that some senders with diverging interests pose as experts. Therefore, the main questions of this article are as follows. What are optimal communication patterns, taking into account that some senders have privately known diverging interests? How do specialization and preference uncertainty affect communication, and how do they interact? Lastly, how

⋆ This manuscript has previously been circulated under the title "Fake Experts". We thank Ian Ball, Felix Bierbrauer, Jan Knöpfle, Daniel Krähmer, Stephan Lauermann, Benny Moldovanu, Martin Pollrich, Larry Samuelson, Dezsö Szalay, and Aleh Tsyvinski. Both authors wish to express their gratitude for the hospitality of the Yale Economics Department during their respective visits, amid which much of this research has been conducted. Funding by the Deutsche Forschungsgemeinschaft (DFG) through CRC TR 224 Project B1 is gratefully acknowledged.

1. Throughout the article, we refer to the receiver with the female pronoun and to the senders with the male pronoun.

can we explain that voting mechanisms are still prevalent in some situations, despite seemingly ignoring specialization?

To illustrate these questions, let us consider an example of a government agency asking for expert advice on a new regulatory decision. In these cases, experts hold information that is valuable in making the decision, yet they might also have private interests owing to financial ties to the regulated industry. Due to different specializations and experiences, it is likely that the experts in the advisory bodies are not equally well-informed. In such a situation, should the regulator give more weight to those experts who express high confidence in their positions? This allows her to account for the heterogeneity of information, but it also increases the possibility of experts with conflicts of interest exerting greater influence by falsely claiming high confidence. Or should she simply decide based on the numbers of individual votes for and against approval?

More generally, how should individuals adjust their learning when they might be lied to by interested parties? This concern extends beyond commercial interest and has recently been at the center of political debates, most notably that of “fake news.” We therefore try to understand the impact of a voter struggling to distinguish between interest-led rhetoric and reputable news. Is in-depth reporting still heard, or does public discourse reduce to deciding which side has gathered sufficient numbers?

In this article, we examine such decision problems using a multi-sender cheap talk model in which the senders receive conditionally independent information. Instead, we assume that they have either aligned interests or state-independent preferences. We believe that this assumption is reasonable in situations like the ones above, in which a clear common interest objective that might be trumped by the private considerations of the senders. As is standard in the literature, we focus throughout on the most informative equilibrium. Essential to our analysis is the interaction between the senders’ information structures and the receiver’s uncertainty over their preferences. To make things concrete, suppose the regulator from our first example has a prior of 1/2 that the approval of a new regulation is the right decision. She is advised by 5 experts. All these experts are informed conditionally iid, which can take either of the two following forms:

1. Each expert receives a binary signal that is known to match the true state of the world with a 62% chance.
2. Each expert has a 15% chance of learning the state perfectly and otherwise learns nothing.

In the absence of preference uncertainty, the regulator bases her decision under information structure 1 on a majority vote. This gives her a 72%

chance of being correct. Under Information structure 2, she follows perfectly informed experts, of whom there is at least one with a 56% chance. Otherwise, she has to toss a coin, which, in total, gives her a 78% chance of being correct.[2] However, as we show below, because gains from very precise signals are particularly vulnerable to preference uncertainty, the advantage of Information Structure 2 shrinks as preference uncertainty is gradually introduced. At around a 10% chance that a expert is partisan in either direction, both information structures allow the regulator to make the right decision with the same probability, i.e., only 68% of the time. If the chance of partisanship rises further, Information Structure 1 becomes even superior to Information Structure 2.

To conceptualize the difference between such information structures, we introduce two concepts. Any information a expert receives has both a direction, i.e., it favors either the approval or rejection of the regulation, and an intensity, i.e., how much it moves the expert's belief away from the prior. We call the mean intensity of a expert's signal his *average informativeness*. In our example, average informativeness is higher under information structure 1 because each expert's posterior is moved 0.12 away from his prior, although in the second structure this distance on average is only 0.075.

In our example, a specialized expert learns whether the question is in his field of specialization and hence updates his posterior more or less than his generalist colleague. To illustrate the concept of specialization, consider the following Information Structure 3:

3. Each expert receives a signal that matches the true state of the world with a 60% chance, but also has a 5% chance of learning the state perfectly.[3]

Information Structure 3 can be interpreted as a expert now recognizing, whether a given question is in his field of specialization in which case he is more secure in his judgment or not in which case he is less secure. For a generalist the same information would have no effect. Information Structure 3 is always weakly better than 1, but completely loses its edge when there is at least a 1 in 12 chance that a expert with either information structure is a partisan in either direction an effect we formalize in Lemma 2. Thus, preference uncertainty destroys possible gains from specialization.

As noted above, in these comparisons, we focus on the most informative equilibrium. Behavior in this equilibrium in the presence of preference

2. The two precision numbers can be derived from the following two expressions: $0.62^5 + \binom{5}{4}0.62^4 0.38 + \binom{5}{3}0.62^3 0.38^2$ and $1 - 0.85^5 + \frac{1}{2}0.85^5$ respectively.

3. As $0.05 * 0.5 + 0.95 * 1 = 0.12$ average informativeness is the same as for information structure 1.

uncertainty is characterized by senders with aligned interest stating their true beliefs, whereas partisans send messages independently of their information, mimicking senders who receive the most informative signals in their preferred direction. The receiver acts on the central trade-off to use as much information from aligned senders as possible although limiting the influence of partisans and putting caps on the influence any one message can have on her decision. However, between these caps, communication between the aligned senders and the receiver remains perfect. This is in contrast to the typical coarsening of messages in cheap talk games with known bias in which there is a coarsening affects all messages. Consequently, partisanship has two distinct effects on information transmission, which relate to the two concepts of informational content: average informativeness and specialization.

The first effect is that the information held by partisans is lost, because they send messages independent of the signal they receive. This leads to a proportional loss in average informativeness. Such a loss is due to the mere existence of partisans and is equally effective if the partisan senders preferences are known to the receiver; hence, partisan messages can be ignored. In that sense known partisanship is equivalent to having aligned but incompetent sender types.

The second effect is a loss in effective specialization caused by the indistinguishability of advisory and partisan senders. The receiver experiences uncertainty over the senders' preferences, which results in the uninformative messages of the partisans, being treated the same as the messages from senders with the most valuable signals. This leads to an effective loss in specialization, because the best signals are now diluted by signals that are, on average, uninformative, which makes the value of different messages more homogeneous overall.

If partisanship becomes sufficiently strong, all gains from specialization are wiped out. The senders reduce their messages to mere indications of the direction of their information, although the receiver bases her decision on whether the number of messages in favor of one alternative meets a fixed threshold. This binary communication between the senders and the receiver resembles a form of qualified majority voting. Consequently, even in situations in which some senders naturally have more to contribute than others, voting arises as an optimal way of communication. In these environments, the average informativeness of each sender becomes the decisive predictor of the receiver's ability to match the state with her decision, although specialization becomes worthless. More generally, in some configurations, a group of expert advisors with more heterogeneously distributed posteriors is preferable to a receiver, when partisanship is low, but performs worse when partisanship is high. We are thus worried that increasing political polarization might substantially diminish the gains society can reap from increases in the specialization of knowledge.

The article continues as follows. In the rest of this section, we review the literature. In Section 2, we introduce our model. Section 3 analyzes the special case in which all the senders have aligned interests. We use this natural benchmark to contrast our later findings. Section 4 introduces concepts to analyze both the senders' information structure and the information that is transmitted to the receiver. We apply these concepts to our general model in Section 5 and derive our main results. Section 6 concludes.

Related Literature

We place our work between the literatures on cheap talk and information aggregation in voting. The former builds on the seminal work of Crawford and Sobel (1982) and analyzes strategic communication between a better-informed sender and a receiver whose action determines the payoff of both. In their original setup, the sender has private and perfect information on a one-dimensional state of the world and a bias known to the receiver.

We depart from this classical model in three central ways, with multiple senders who are imperfectly informed and whose preferences are unknown to the receiver.

Gilligan and Krehbiel (1989) were the first to study a model with multiple senders. In their model, two privately and perfectly informed senders with publicly known biases communicate with a receiver. The focus of their analysis is the comparison of three communication protocols that comprise different forms of cheap talk. Similarly, Krishna and Morgan (2001) study a setting with two senders who sequentially send public messages to a receiver. The degree of information revelation depends on whether the senders have aligned or opposing biases.

Austen-Smith (1990) was the first to study a cheap talk problem in which the senders are imperfectly informed about a binary state of the world. Wolinsky (2002) identifies circumstances under which a cheap talk phase between senders alters the decision and solves for the most efficient communication structure.

Battaglini (2017) studies public protests as an informal means to aggregate dispersed information in democracies. As politicians generally do not commit to a reaction to protests ex-ante, such protests are, in fact, modeled by cheap talk. Information does not need to aggregate if the variation in optimal decision threshold between the politicians and the citizens is large relative to the information available to each citizen.

In Alonso, Dessein, and Matouschek (2008) and Hummel, Morgan, and Stocken (2013), uncertainty about the senders' preferences arises endogenously, because each sender is interested in the decision matching his type, although the receiver wants to match her decision to the average of the senders' types. Hence, a sender's type contains both relevant information about the state, i.e., the average of types and information about the bias, i.e., the distance of the individual sender's type from the average.

In contrast to this approach and in line with our own, Morgan and Stocken (2003) and Li and Madarász (2008) analyze a single-sender game with private bias that does not enter the receiver's payoff. In Morgan and Stocken (2003) senders are either aligned or have a non-partisan bias. The authors show that any uncertainty on the receiver's side about such a bias yields to the bunching of messages and therefore information loss. If the potential conflict of interest is sufficiently large, a case they call semi-response equilibrium of size one, the communication of their single sender, is similar to the one of a every sender individually in our model, when partisanship can only occur in one direction. In extreme cases, only information that is opposed to the potential bias is truthfully reportable. Li and Madarász (2008) find that both players can benefit from the privacy of the sender's bias. The receiver's ignorance sender's preferences allows for conflict-hiding equilibria in which a more or oppositely biased type can transmit more information, because identical messages are also transmitted by the other type, changing the expected meaning of the message. In contrast in our model not knowing the senders' types is always detrimental to the receiver, because senders with state independent preferences message independently of the signals they received. Knowing a sender's bias therefore makes it possible to ignore such messages, instead of them being garbled with informative messages from aligned agents.

Another strand of literature has assumed that one type of sender is non-strategic and always communicates thruthfully. This strand includes Sobel (1985), Benabou and Laroque (1992), Vidal (2006), and Glazer, Herrera, and Perry (2021). The last is an independent paper modelling internet recommendation systems. Their work, like ours, studies a model with multiple imperfectly informed senders, in which some senders have state-independent preferences. because we find in our model that the most informative equilibrium is one in which aligned senders have a best response to be honest, the equilibria in both models are similar. However, given the different economic motivations of the two works, the analyses building on these equilibria are distinct. For a comparison of honest sender types, versus strategic yet aligned types see also Kim and Pogach (2014).

There is also a structural similarity of our model to models of voluntary disclosure, a literature started by Dye (1985) and in particular to the of work of Jung and Kwon (1988). These similarities can be understood by focusing on the special case of our model with a single sender, that can only be a partisan towards state 0. As partisans will always send messages yielding to the lowest posterior, the resulting communication is similar to disclosure games where agents desire the receiver to have high believes about their type, yet some agents are unable to provide evidence and are hence treated equally to the the lowest types that opt not to disclose their information.

The second body of literature we relate to is on information aggregation in voting. It goes back to Condorcet (1785) and his famous jury theorem, stating that large groups of independently informed senders select the correct alternative with near certainty. He assumes that senders vote sincerely, although Feddersen and Pesendorfer (1997) establish a similar result for strategic senders. They show that when the number of voters grows large, privately held information leads to the same decision as public information.

Despite the effectiveness of voting for information aggregation in large populations, the same literature has revealed effects such as the swing voter's curse, first discussed by Feddersen and Pesendorfer (1996), which illustrates a loss of information in small populations. This loss is mainly due to the nature of the voting game, with its limited number of messages, usually two or three, and its fixed threshold. Any such voting rule can be interpreted in our model as a behavioral type of receiver to which strategic voters react optimally. We therefore believe that our model, with its strategic receiver, provides a natural benchmark for voting systems and helps distinguish which losses of information are necessary consequences of conflicts of interest and which are due to the specific features of real-world voting systems.

McMurray (2017) and Azrieli (2018) examine the limitations of elections with few available messages. McMurray (2017) studies a common interest election of ex-ante symmetric candidates by a fixed number of heterogeneously informed senders. In equilibrium, voters coordinate around specific candidates to transmit information. His model can be interpreted as a cheap talk game with a restricted number of messages. If the number of candidates becomes large, the model converges to our common interest setting. Azrieli (2018) analyzes the loss of anonymous voting rules if the senders are publicly known to be differently well-informed. The common-value analysis is also closely related to ours. However, we assume that signals are private information and focus on their interplay with private interests.

Lastly, Li, Rosen, and Suen (2001) analyze two-player decision games with known conflicts of interest. Agents can choose from a set of integer weights, and an action is taken depending on whether the sum of the weights exceeds a predetermined threshold. These rules are reminiscent of the equilibrium play of the receiver in our communication games, despite the lack of commitment in our model.

2.2 The Model

There is a set of senders $\{1, ..., n\}$ and a receiver indexed by 0. Each sender i receives a signal about the unknown state of the world $\omega = \{0, 1\}$. Signals are identically distributed and independent, conditional on the true state of the world. There is a common prior $p_0 = \mathbb{P}[\omega = 1] \in (0, 1)$ that the state of the world is 1. As signals are conditionally independent, all information is contained in the resulting distribution over posteriors, and we shift the attention completely to the latter. Each sender draws his posterior from the probability mass function μ, which is consistent with p_0. We assume that the information structure is such that it leads to a finite number of possible posteriors $\mathscr{P} = \text{supp}\ \mu$. For some results, we assume that no signal is uninformative, i.e., $\mu(p_0) = 0$. We call a distribution μ that fulfills this assumption *never-ignorant*. The receiver shares the prior but observes no signal.[4]

In addition to different signals, players are heterogeneous with respect to their preferences, as described by a parameter $\lambda \in \{0, \lambda_0, 1\}$ with $\lambda_0 \in (0, 1)$. Each sender i independently draws a preference parameter λ_i that is independent of the posteriors and distributed according to probability mass function γ. The decision-maker has the commonly known preference parameter λ_0. Each sender also draws a posterior p_i independently from both other agents and his preference type according to probability mass function μ. We call the tuple (p_i, λ_i) the type of sender i, and denote with $\mu \times \gamma$ the distribution over types.

After observing the signal, each sender i simultaneously sends a cheap talk message $t_i \in [0, 1]$ to the receiver. We denote the potentially mixed strategy by $m_i : \mathscr{P} \times \{0, \lambda_0, 1\} \to \Delta[0, 1]$, where $\Delta[0, 1]$ denotes the set of all probability measures over $[0, 1]$. Whenever the strategy of the sender specifies a finite set of messages to be sent a.s. we denote the probability that sender i with type (p_i, λ_i) sends message t_i by $m_i(p_i, \lambda_i)(t_i)$. We call a strategy *truthful for preference type* λ_i if $m_i(p_i, \lambda_i)(p_i) = 1$ for all types (p_i, λ_i). The tuple of messages of all the senders is denoted by $t = (t_1, ..., t_n)$.

We denote the belief of the receiver accounting only for sender i's message t_i by $q(t_i) = E[p_i | t_i]$ and call it the virtual posterior of sender i.[5] The posterior of the receiver incorporating the messages t of all the senders is denoted by $q(t)$. After processing all messages, the receiver takes action

4. We discuss alternatives to some of the assumptions made in this model in Appendix 2.C.

5. Different strategies m_i induce different virtual posteriors $q_i(\cdot)$. Anticipating that the senders play symmetric strategies in an optimal equilibrium, we drop the subscript i of the virtual posterior $q_i(\cdot)$ to simplify the notation.

$a \in \{0, 1\}$. Utilities for the senders and the receiver are given by

$$u(a, \omega, \lambda_i) = (1 - \lambda_i)\mathbb{1}\{a = 0\} + \lambda_i \mathbb{1}\{a = 1\} + \mathbb{1}\{a = \omega\},$$

where $\mathbb{1}$ is the indicator function, i.e., $\mathbb{1}\{A\}$ is 1 if event A is true and 0 otherwise.

A player i prefers action 1 if and only if his belief that the state of the world is 1 is larger than or equal to $1 - \lambda_i$. A higher preference parameter λ_i leads to a higher expected utility of player i given that the action is equal to 1. Senders with preference parameters 0 and 1 weakly prefer the action that matches their preference parameter irrespective of the posterior. We call senders with these preference parameters *partisans*. The remaining senders with $\lambda_i = \lambda_0$ have the same interests as the receiver. We call these senders *advisors*. Before we proceed, we summarize the timing of the game. First, nature draws a state of the world ω. Second, every sender i randomly draws a type (p_i, λ_i) according to the conditional type distribution $\mu_\omega \times \gamma$. Third, each sender i sends a message t_i to the receiver. Last, the receiver takes an action a, and payoffs are realized. We assume that the receiver does not have commitment power, i.e., she can not credibly commit to a decision rule before getting the messages of the senders.[6] Consequently, we solve for Bayesian Nash equilibria.

In the following, we split the analysis into three parts. We start by studying the common interest case in Section 2.3. Here, all the senders have aligned preferences. The special case of our setting serves as a benchmark and allows us to become familiar with how the receiver processes the signals from the senders. In Section 2.4, we focus on the information structure of the senders, introduce the concept of specialization, and illustrate its significance in the common interest case. Lastly, in Section 2.5, we apply these concepts in our analysis of the general case, in which we allow for private interests.

2.3 Common Interest

In this section, we derive a benchmark equilibrium that maximizes the utility of the receiver when all the senders have aligned preferences, i.e., $\gamma(\lambda_0) = 1$ and $\gamma(0) = \gamma(1) = 0$. In the common interest case, such an equilibrium maximizes the utility of the senders as well. The general idea of

6. In particular, this excludes equilibria of the kind discussed in Gerardi, McLean, and Postlewaite (2009).

this equilibrium is straightforward. The receiver needs to perform Bayesian updating given the senders' messages, and the senders, knowing that their information is aggregated in a statistically correct way, can state their posteriors, revealing all their information. It is as if the regulator could observe all the signals of the experts.

In the following description of this equilibrium, we focus on the statistical properties and interpretation of how the receiver updates the information and how she translates it into her decision.

Definition 2.1. A receiver follows a weighted majority rule if her strategy $a : [0,1]^n \to \{0,1\}$ is of the form

$$a(t) = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} w(t_i) > \tau \\ 0 & \text{else} \end{cases}$$

for messages $t = (t_1, ..., t_n)$ of the senders, a weighting function $w : [0,1] \to \mathbb{R}$, and a threshold τ.

Under a weighted majority rule, the receiver transforms every message t_i into a weight $w(t_i)$ and takes decision 1 if the sum of weighted messages is larger than a threshold τ. One can interpret this as the receiver giving the senders free choice over the weights in the image of w and limiting herself to the application of a simple rule. When the size of the image is equal to 2, this comes down to proposing a decision by qualified majority voting. We come back to this analogy in Section 2.5. The next proposition translates the above-described equilibrium into this language.

Proposition 2.2. *The following describes a receiver-optimal Bayesian Nash equilibrium:*

- *Advisors message truthfully, i.e.,* $m_i(p_i, \lambda_0)(p_i) = 1$.
- *The receiver follows a weighted majority rule with weighting function*

$$w(x) = \ln \frac{x}{1-x} - \ln \frac{p_0}{1-p_0}$$

and threshold $\tau = -\left(\ln \frac{\lambda_0}{1-\lambda_0} + \ln \frac{p_0}{1-p_0}\right)$.

Proof. See Appendix 2.A. □

In this optimal equilibrium, the senders play the truthful strategy to transmit their posterior to the receiver. The receiver has correct beliefs about this and can deduce from the posteriors the entirety of their information. She then translates it into the optimal decision via Bayesian updating, which we interpret as her applying a weighted majority rule, with log-likelyhood ratio weights.[7] Hence, an equilibrium with higher payoffs for the receiver cannot exist.[8]

Figure 2.1 illustrates the weighting function with prior $p_0 = \frac{3}{4}$ for the common interest case.

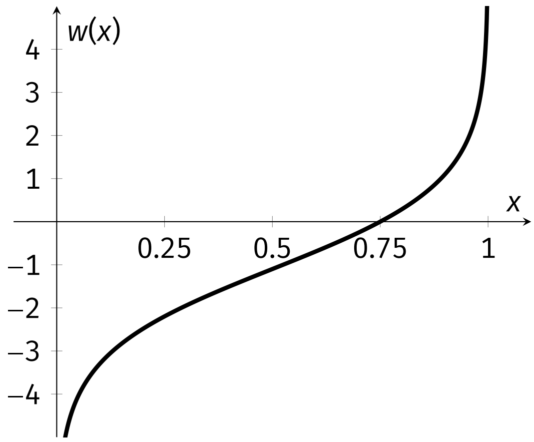


Figure 2.1. Weighting function $w(x) = \ln \frac{x}{1-x} - \ln \frac{p_0}{1-p_0}$ with prior $p_0 = \frac{3}{4}$ for the common interest case.

A posterior p_i of sender i that equals the prior p_0 gets weight 0 because it does not transmit any additional information. In contrast, a posterior $p_i \in \{0, 1\}$ means that sender i knows the state of the world perfectly. This sender's information is sufficient to make an optimal decision, and he should outweigh all other senders. Thus, as p_i goes to 1 (0), the corresponding weight tends to ∞ ($-\infty$). The unrestrictedness of the weighting function encodes the extraordinary value of perfect information.

In the next section, we refer to the receiver-optimal equilibrium when we assess different distributions of sender types. The expected utility of the receiver $u^\star(q(t))$ with the posterior $q(t)$ is given by

$$u^\star(q) = \begin{cases} \lambda_0 + q & \text{if } q > 1 - \lambda_0 \\ 2 - \lambda_0 - q & \text{else.} \end{cases}$$

We now turn to the analysis of the senders' information structure.

7. See Nitzan and Paroush (1982) and Shapley and Grofman (1984) for two classical treatments of the role of such weighting roles for optimal information aggregation in groups.

8. McLennan (1998) studies optimality of equilibria in common interest games more generally.

2.4 Specialization

The distribution over posteriors of the senders is a crucial object in our model. In this section, we develop the concepts that we use to describe them throughout. A classical concept in this regard is, Blackwell's informativeness order, introduced in, Blackwell (1950).[9]

Definition 2.3. Let μ and ν be two distributions over posteriors with cdfs F and G, respectively. We say that μ is more informative than ν, denoted by $\mu \succ \nu$, if

$$\int_0^y F(x)\mathrm{d}x \geq \int_0^y G(x)\mathrm{d}x \quad \text{for all } y \in [0,1].$$

For convenience, we have summarized some facts about the informativeness order in Appendix 2.B. An important interpretation of the above integral condition is that the more informative information structure is a mean-preserving spread of its less informative counterpart. Note that in our setting, both integrals are equal at 1 given equal priors. In this section, we build on this classical order to conceptualize specialization in knowledge:

Definition 2.4. Let μ and ν be two distributions over posteriors with cdfs F and G, respectively. We say that μ is more specialized than ν, denoted by $\mu \succ_s \nu$, if $\mu \succ \nu$ and

$$\int_0^{p_0} F(x)\mathrm{d}x = \int_0^{p_0} G(x)\mathrm{d}x. \tag{2.1}$$

We refer to the reverse order as one measure being more generalized than another.

Because of the additional equality condition, specialization is clearly a coarsening of the informativeness order. This condition is equivalent to requiring that no mass can be spread to the other side of or away from the prior. These one-sided spreads can be understood as learning about the quality of ones signal, in contrast to learning about the direction. The more specialized an agent is, the more the quality of his assessment depends on whether a given question is in his area of spezialization or not and hence the more he can learn about the quality of his judgment, based on observing the question.

9. The particular characterization used is related to the second stochastic order used in decision theory. A "riskier" distribution over posteriors contains more information.

To illustrate the difference between specialization and informativeness order, let us get back to our example from the introduction with a slight adjustment. To recall, we have assumed a symmetric prior and discussed three possible signal structures for experts advising a regulator:

1. Each expert receives a signal that matches the true state of the world with a 62% chance.
2′. Each expert has a 40% chance to learn the state perfectly.
3. Each expert receives a signal that matches the true state of the world with a 60% chance, but also has a 5% chance to learn the state perfectly.

In contrast to the introduction, we have changed the probability that a expert receives a perfect signal in 2'. Figure 2 illustrates the cdfs corresponding to these information structures.

As we can see, 3 is a mean-preserving spread of 1, wheras 2' is a mean-preserving spread of both. However, the integrals are also equal at the prior only in the case of 3 and 1. We can interpret this spread as a expert under signal structure 3 receiving an additional signal that identifies 5% of his signals as fully revealing, without ever changing the direction of any of his signals.

Another possible way to understand specialization is to pose that between two experts that have equal information once one averages above their signal in either direction the specialist is more heterogeneously informed. We fix this idea in the following definition.[10]

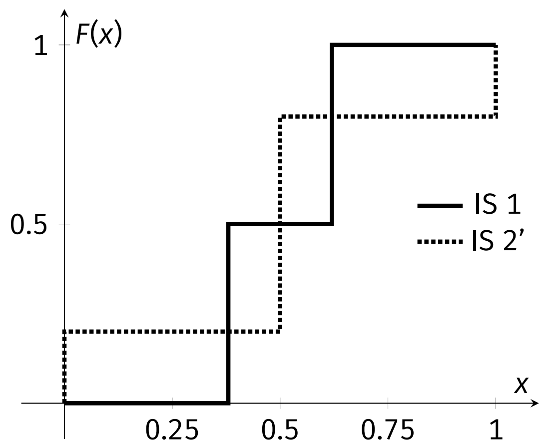
Definition 2.5. The average informativeness $\pi(\mu)$ of a sender's distribution over posteriors μ is

$$\pi(\mu) = \mathbb{E}[|p_i - p_0|].$$

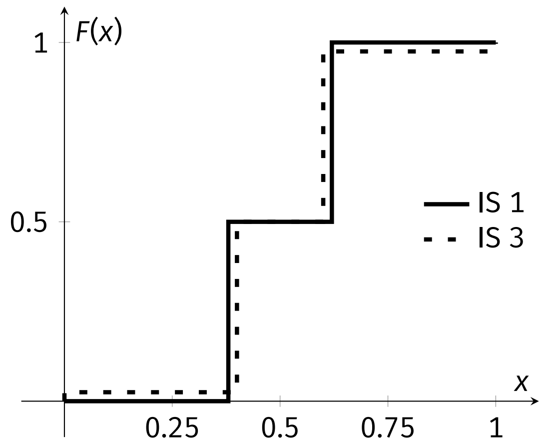
A distribution μ with $\pi(\mu) = 0$ has all mass at the prior and hence does not contain any information, whereas the maximal average informativeness is $2p_0(1-p_0)$. Note that equal or higher average informativeness of μ compared to ν is a necessary but not a sufficient condition for μ to be more informative than ν.

In our example, average informativeness in 1 and 3 equals 0.12, although it is 0.2 in 2'. The following proposition clarifies the link between average informativeness and specialization:

10. See Frankel and Kamenica (2019) for a critical discussion of the use of metrics, such as the euclidean metric used here, to quantify information. We should note however, that because we only use average informativeness in combination with the informativeness order, their criticisms generally do not apply to our setting.



(a) Information Structure 2’ is a mean-preserving spread of 1. However, 2’ is not a mean-preserving spread on both sides of the prior separately. Hence, 1 and 2’ are not comparable with the specialization order.



(b) Information Structure 3 can be constructed by applying mean-preserving spreads on both sides of the prior of Information Structure 1. Therefore, 3 is more informative and more specialized than 1.

Figure 2.2. Illustration of specialization. Distribution of virtual posteriors.

Proposition 2.6. *Let μ and ν be two distributions over posteriors. Then $\mu \succ_s \nu$ if and only if $\mu \succ \nu$ and $\pi(\mu) = \pi(\nu)$.*

Proof. See Appendix 2.A. □

By this characterization, a specialist is more informed not because his posterior is, on average, further away from the prior, but because his posteriors are more heterogeneous. The proof consists in arguing that the one-sided mean preserving spreads from the definition are exactly those that do not increase average informativeness.

In the remainder of this subsection, we connect the above insights to the receiver's utility. As more informativeness of independent individuals' distributions over posteriors produces more informativeness overall, as discussed in Blackwell and Girshick (1979) (see Proposition C in Appendix 2.B), the utility of the receiver increases with the specialization of posteriors. Thus, the next corollary to Proposition 2.2 links our discussion in this section so far with the utility of the receiver.

Corollary 2.7. *Let μ and ν be distributions over posteriors with $\mu \succ \nu$. When comparing receiver optimal equilibria under common interests, the utility of the receiver is monotone with respect to the informativeness and hence the specialization order, i.e., weakly higher if she is facing senders with distribution over posteriors μ rather than ν.*

Proof. See Appendix 2.A. □

As we learned previously that under common interests, all information held by the senders reaches the receiver and because more information is advantageous, the receiver directly benefits from more informative and therefore more specialized senders.

2.5 The Vulnerability of Specialized Advice - Private Interest Analysis

In this section, we turn to the case with private interests. In Subsection 2.5, we solve for the receiver-optimal equilibrium. In Subsection 2.5, we decompose the total loss of information into despecialization and a loss of average informativeness. Lastly, in Subsection 2.5, we present two consequences of despecialization. First, we show that voting is optimal if preferences are sufficiently heterogeneous. Second, average informativeness becomes more important and specialization less important as the number of partisans increases.

Receiver-Optimal Equilibrium

Similar to our treatment of the common-interest case we focus on a receiver optimal equilibrium. It turns out that an optimal equilibrium exists such that advisors play the truthful strategy, as in Proposition 2.2. Transmitting as much information as possible is in their best interest. Partisans interfere with this communication, by mixing over the most extreme messages in their preferred direction.

As the average posterior of a partisan sender equals the prior, their strategy shifts virtual posteriors towards the prior. They do not transmit any information to the receiver, but maximize their influence by imitating advisors with the most informative signals. Therefore, the receiver needs to discount these messages. This way, expertise bounds $\underline{b}$ and $\overline{b}$ arise. They constitute bounds on the highest (lowest) possible virtual posteriors associated with any message. The weights $w(\underline{b})$ and $(w(\overline{b}))$ are the lowest and highest weights used in the weighted majority rule of the receiver, respectively. They are endogenously determined by the sender type distribution.

Within the expertise bounds, communication between the advisors and the receiver is noise-free because partisans do not imitate advisors with imprecise signals. Thus, communication is perfect within these bounds, as in the equilibrium from Proposition 2.2. Off-equilibrium messages receive weight 0. Figure 2.3 depicts an example of a weighting function of virtual posteriors with upper and lower expertise bounds. We formalize the above discussion in our first theorem.

Theorem 2.8. *The following describes a receiver-optimal Bayesian Nash equilibrium. There exist unique expertise bounds $\underline{b}, \overline{b} \in [0,1]$, s.t.*

- *Advisors message truthfully, i.e., $m_i(p_i, \lambda_0)(p_i) = 1$.*
- *Partisans imitate and devalue expertise:*

$$m_i(p_i, 0)(t_i) = \begin{cases} \frac{\gamma(\lambda_0)\mu(t_i)(\underline{b}-t_i)}{\gamma(0)(p_0-\underline{b})} & \text{if } t_i \leq \underline{b} \\ 0 & \text{else} \end{cases}$$

$$m_i(p_i, 1)(t_i) = \begin{cases} \frac{\gamma(\lambda_0)\mu(t_i)(t_i-\overline{b})}{\gamma(1)(\overline{b}-p_0)} & \text{if } t_i \geq \overline{b} \\ 0 & \text{else} \end{cases}$$

- *The receiver uses weighted majority rule with weight function*

$$w(x) = \begin{cases} \ln \frac{\underline{b}}{1-\underline{b}} - \ln \frac{p_0}{1-p_0} & \textit{if } x < \underline{b} \\ \ln \frac{x}{1-x} - \ln \frac{p_0}{1-p_0} & \textit{if } x \in [\underline{b}, \overline{b}] \\ \ln \frac{\overline{b}}{1-\overline{b}} - \ln \frac{p_0}{1-p_0} & \textit{if } x > \overline{b} \\ 0 & \textit{else} \end{cases}$$

and threshold $\tau = -\left(\ln \frac{\lambda_0}{1-\lambda_0} + \ln \frac{p_0}{1-p_0}\right)$.

Proof. See Appendix 2.A. □

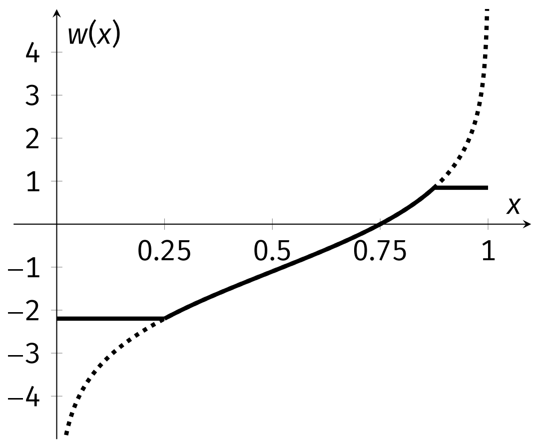


Figure 2.3. Weighting function $w(x) = \ln \frac{x}{1-x} - \ln \frac{p_0}{1-p_0}$ with prior $p_0 = \frac{3}{4}$, $\underline{b} = \frac{1}{4}$, and $\overline{b} = \frac{7}{8}$.

We briefly sketch the proof, which consists of three steps. First, we provide an explicit formula for expertise bounds and show that they exist and are unique. This is formally captured in Lemma 2.12 in the Appenddix.

Second, we verify that the senders and the receiver play a best strategy. We show that the receiver uses weights corresponding to the expected posterior for every message. It is clear that partisans' play a best response by maximizing their weight in the preferred direction. Further, aligned senders ideally want their messages to be weighted according to the untruncated weighting function, yet given the receiver's strategy, they get as close as possible by stating their true posterior.

Lastly, we show that no other equilibrium is better for the receiver. For this, we first establish that the equilibrium utilities of all agents are characterized by the strategy played by senders conditioned on being advisors. We already saw in the common interest case that the receiver's on-path strategy is fully determined by the strategies used by the senders because he acts last.

This builds on the insight that whenever an advisory type strategy leads to more informative messaging in the absence of partisan types, the same holds true, when partisans are introduced.

This argument in lemma 2.14 in the Appendix. Optimality follows in combination with Proposition 2.2 that has established, that truthful messaging by advisors is optimal, in the absence of private interests.

We end the Subsection by briefly illustrating the theorem by means of the example from the introduction. In Information Structure 3, we assume that an expert gets a perfect signal with a 5% chance and otherwise a signal that is right 60% of the time. Under a symmetric prior, the senders' posterior is 0 or 1 and 0.4 or 0.6 with probabilities 2.5% and 47.5%, respectively. Now we introduce private interests in the form of a sender being a partisan type, with an equal chance of $\frac{1}{38} \approx 2.6\%$ on either side. Theorem 2.8 predicts the following effect on the information structure:

- Senders claim with probability $\frac{1}{38} + \frac{18}{19} \cdot 2.5\% = 5\%$ a posterior of 0 or 1, respectively. The receiver treats these messages as if everyone has truthfully reported a posterior of $\frac{5}{19}$ and $\frac{14}{19}$, respectively.
- Senders claim with probability $\frac{18}{19} \cdot 47.5\% = 45\%$ a posterior of 0.4 or 0.6, respectively. The receiver treats these messages as true statements.

The example illustrates how partisans undermine the messages of the most informed sender types, although not affecting the messaging of their less informative counterparts. We decompose the associated loss of information in the next Subsection into two parts that relate to the concepts introduced in Section 2.4.

Lack of Trust versus Lack of Competence

In this Subsection, we inspect the equilibrium derived in the last Subsection with respect to the notion of specialization introduced in Section 4. To do so, we introduce an intermediate regime between private interest and common interest. In this intermediate regime, there are non-advisory types, just as before but instead of having misaligned interests, these types are aligned but incompetent, i.e., they receive no signal.[11] A receiver-optimal equilibrium is characterized by full communication, just as in Proposition 2.2, with the only difference that incompetent agents always message uninformatively.

11. The discussion below is equally valid for the case of a receiver who knows the identity of (informed) senders with partisanship. The receiver ignores any message from these sender types.

We make two principal observations. First, the possibility that a sender is incompetent has the same impact on average informativeness as the possibility that he is a partisan. Second, in contrast to this regime, partisanship also leads to an effective loss in specialization because the uninformative messages of partisans are treated the same as messages from senders with the most valuable signals. The most informative signals are thus diluted by signals that are, on average, uninformative. The distrust in messages with a reportedly high precision destroys the benefits a receiver can reap from the specialization of her senders.

Therefore, we can decompose the loss of information caused by partisanship in a loss of average informativeness and a loss in specialization, as formalized in the following theorem:

Theorem 2.9. *Let $\mu^{\gamma}_{partisan}$ and $\mu^{\gamma}_{incompetent}$ be the distribution over virtual posteriors in the equilibrium above and the most informative equilibrium with incompetent types, respectively. The loss of average informativeness in both cases is identical and equal to the probability that a sender is non-advisory:*

$$\pi(\mu^{\gamma}_{partisan}) = \pi(\mu^{\gamma}_{incompetent}) = \gamma(\lambda_0)\pi(\mu)$$

However, the virtual posteriors under partisanship are less specialized than under incompetence:

$$\mu^{\gamma}_{partisan} \prec_s \mu^{\gamma}_{incompetent}$$

Proof. See Appendix 2.A. □

Let us briefly sketch the proof. First, note that average informativeness is reduced proportionally under ignorance almost axiomatically. Second, we show that the way partisans garble advisors' best signals with their own uninformative messages is a one-sided mean-preserving contraction of the distribution of virtual posterior under ignorance. Therefore, it is less specialized and, by Proposition 2.6, implies that both distributions have the same average informativeness.

Having studied the two distinct parts of information loss, we continue by pointing out two consequences of high degrees of partisanship. We demonstrate that sufficiently heterogeneous preferences can prevent any differentiated weighting of messages. Further, we show that the value of specialization vanishes and that average informativeness becomes more important as the share of partisans rises.

A Justification for Voting

In this Subsection, we analyze the effect of a rising share of partisans by building on the results from the previous two Subsections and the concepts from Section 2.4. We demonstrate that specialization vanishes as the share of partisans rises. This is reflected by an equilibrium in which the receiver weights all messages with the same direction equally. To see why such an equilibrium might evolve, consider the following reasoning. A partisan imitates senders with the most valuable signals. The receiver devalues these messages accordingly. If the share of partisans is high enough, the weight of the most informative and second most informative signals become equal. As the share of partisans increases further although assuming that no completely uninformed senders exist, only two distinct weights remain - one for each direction.

We interpret such an equilibrium as voting because only the number of senders messaging in each direction matters. Specifically, suppose that there are weights w^0 and w^1 for the senders' messages expressing that their posteriors are left and right from the prior, respectively. According to the equilibrium, the receiver takes action 1 if $n_0 \cdot w^0 + n_1 \cdot w^1 > \tau$, where n_0 and n_1 denote the number of senders with weight w_0 and w_1, respectively. This decision rule corresponds to a qualified majority rule that we consider as a form of voting. The formal statement of the result is as follows:

Theorem 2.10. *Let μ be never-ignorant. Then there exists $c_0, c_1 \in (0,1)$ with $c_0 + c_1 < 1$, s.t. for all γ with $\gamma(0) \geq c_0$ and $\gamma(1) \geq c_1$ the receiver forms only two virtual posteriors, i.e., (qualified majority) voting is the most informative equilibrium.*

Proof. See Appendix 2.A. □

We find it instructive to illustrate this result by means of our example too. Consider Information Structure 3:

3. Each expert receives a signal that matches the true state of the world with 60%, but also has a 5% chance to learn the state perfectly.

If there are only a few partisans, all of them imitate senders with a perfect signal. But as the share of partisans rises, the weight that the receiver assigns to this message decreases and eventually equals to the weight of the initially less informative message. To be precise, equality is achieved if $\frac{1}{12}$ of the senders are partisan in either direction. Equilibrium play guarantees that the weight of both messages stays equal for any greater share of partisans. In fact, if at least $\frac{1}{12}$ of the senders are partisan on each side, Information Structure 3 collapses to Information Structure 1 in terms of the expected posteriors.

In such equilibria, the benefits of specialization are completely destroyed by partisans. The level of distrust is so high that effectively only the direction of a signal can be communicated.[12] At the same time, Theorem 2.10 underlines the robustness of voting. Thus, privately interested senders might prevent any communication of specialization and provide a justification for the prevalence of voting in many real-world information aggregation systems.

We demonstrated that specialization vanishes if there are sufficiently many partisans. Following Theorem 2.9, average informativeness decreases too. In the following we relate the information loss of both concepts to each other. Note that in an equilibrium where voting is the most efficient way to communicate, average informativeness becomes the most important statistic for the receiver because there is no more specialization. We capture this observation by the following Proposition:

Proposition 2.11. *Let μ and ν distributions over posteriors with $\pi(\mu) > \pi(\nu)$ and cdfs F and G, respectively. Then there exist $c_0, c_1 \in (0,1)$ with $c_0 + c_1 < 1$, s.t. for all γ with $\gamma(0) \geq c_0$ and $\gamma(1) \geq c_1$ and any number of senders n, we have that $\mu^\gamma \succ \nu^\gamma$ and hence the ex-ante expected utility of the receiver is weakly greater under distribution over posteriors μ than under ν.*

Proof. See Appendix 2.A. □

Proposition 2.11 implies that the comparison of two information structures in terms of the receiver's utility depends on the number of partisans. While Information Structure 2 leads to a higher utility of the receiver with only a few partisans, Information Structure 1 leads to a higher utility if the share of partisans is at least 10%. Information Structure 2 has a high degree of specialization that can only be materialized with low numbers of partisans. In contrast, Information Structure 1 has a higher average informativeness that becomes decisive if there are enough partisans. Taken together, specialization is particularly valuable with few partisans, and average informativeness is valuable with many partisans. Thus, even if partisans interfere, information aggregation guaranteeing that signals are good on average is an effective way to counteract.

12. When the receiver is not indifferent at the prior, there always exist levels of partisanship s.t. the receiver completely ignores senders, because the virtual posterior distribution becomes to uninformative. This collapse of communication corresponds to a voting threshold below 1 or above n. Our interpretation of voting is therefor most natural, when either, the receiver is close or indifferent, the number of senders is high, or the least informative signals in either direction are relatively informative. The same point applies to Proposition 2.11

2.6 Conclusion

Our goal in this research has been to understand the optimal communication of a decision maker with multiple advising experts when she faces uncertainty about experts' preferences. In particular, we have been interested in how these uncertainties affect changes among senders with different information structures.

We have found that communication that discriminates between messages, indicating different degrees of confidence, is potentially very informative for the receiver, but also highly vulnerable to strategic manipulation by partisan experts. Consequently, such communication is not optimal in a case with high levels of partisanship. In contrast, binary communication protocols such as voting prove to be very robust, explaining their prevalence as a means for information aggregation.

Our research may also lead to a new approach towards questions regarding political lobbying. Much of the literature on the subject (see, for example, Buchanan, Tollison, and Tullock (1980) and Baye, Kovenock, and De Vries (1993)) has focused on lobbying as a way in which special interest groups try to provide incentives for political actors, in order to sway them in their favored direction. It is, however, just as plausible for such groups to buy influence with advising experts to influence politicians' beliefs rather than offer direct incentives. Our work shows that this can be effective even if politicians are aware of it, as long as they remain ignorant about the exact identity of the experts who have been compromised. In particular, interest groups may seek to sometimes influence experts against their own favored decisions to create the justified belief that some experts advocating the other side are not trustworthy. When talk is cheap, trust is a valuable yet vulnerable asset.

Appendix 2.A Proofs

Proof of Proposition 2.2. In this proof, we follow closely standard arguments regarding the representation of Bayesian updating usings log-liklyhood ratios as they can be found, for example, in the proof of Theorem 1 in Nitzan and Paroush (1982), who derive the optimal non-strategic processing of signals with a symmetric prior $\lambda_0 = \frac{1}{2}$.

In the main text, we use the same notation for random variables and their realizations. For this proof, it is useful to introduce a separate notation. We use upper-case characters for random variables and lower-case characters for their realizations.

The receiver processes messages t to update her posterior. She prefers the action that yields the higher expected utility given her posterior $q(t)$. More precisely, an optimal decision rule selects action 1 if

$$\begin{aligned}
& \lambda_0 + \mathbb{P}[\omega = 1|T = t] > (1-\lambda_0) + \mathbb{P}[\omega = 0|T = t] \\
\Leftrightarrow\quad & \lambda_0\mathbb{P}[\omega = 1|T = t] > (1-\lambda_0)\mathbb{P}[\omega = 0|T = t] \\
\Leftrightarrow\quad & \lambda_0\frac{\mathbb{P}[P = p|\omega = 1]\cdot\mathbb{P}[\omega = 1]}{\mathbb{P}[P = p]} > \\
& (1-\lambda_0)\frac{\mathbb{P}[P = p|\omega = 0]\cdot\mathbb{P}[\omega = 0]}{\mathbb{P}[P = p]} \\
\Leftrightarrow\quad & \lambda_0 p_0\prod_i \mathbb{P}[P_i = p_i|\omega = 1] > \\
& (1-\lambda_0)(1-p_0)\prod_i \mathbb{P}[P_i = p_i|\omega = 0] \\
\Leftrightarrow\quad & \lambda_0 p_0\prod_i \frac{p_i}{p_0} > (1-\lambda_0)(1-p_0)\prod_i\frac{1-p_i}{1-p_0} \\
\Leftrightarrow\quad & \prod_i\left(\frac{p_i}{1-p_i}\frac{1-p_0}{p_0}\right) > \frac{1-\lambda_0}{\lambda_0}\frac{1-p_0}{p_0} \\
\Leftrightarrow\quad & \sum_i\left(\ln\frac{p_i}{1-p_i} - \ln\frac{p_0}{1-p_0}\right) > -\left(\ln\frac{\lambda_0}{1-\lambda_0} + \ln\frac{p_0}{1-p_0}\right).
\end{aligned}$$

The first equivalence is a simple algebraic consequence of the fact that the first and second factors each add to one. For the second equivalence, we apply Bayes's rule and exploit the fact that senders play the truthful strategy. In the third step, we use the conditional independence of signals. We arrive at the fifth equation by applying Bayes's rule once again. The sixth equation is a simple reformulation of the fourth. Finally, we obtain the last equation by taking the logarithm on both sides. The resulting decision rule can be interpreted as a weighted majority rule with weighting function

$$w(t_i) = \ln\frac{t_i}{1-t_i} - \ln\frac{p_0}{1-p_0}$$

and threshold $\tau = -\left(\ln\frac{\lambda_0}{1-\lambda_0} + \ln\frac{p_0}{1-p_0}\right)$.

It is optimal for the senders to play the truthful strategy because senders and the receiver have the same utility function. With the truthful strategy, the senders can transmit all available information. Any beneficial transformation of messages can be done by the receiver. □

Proof of Proposition 2.6. We rewrite $\pi(\mu)$ until we arrive at an expression from which the result is immediate:

$$\begin{aligned}
\pi(\mu) &= \mathbb{E}[|p_i - p_0|] \\
&= \int_0^1 |x - p_0|\ \mu(x)\mathrm{d}x \\
&= \int_0^{p_0} (p_0 - x)\ \mu(x)\mathrm{d}x + \int_{p_0}^1 (x - p_0)\ \mu(x)\mathrm{d}x \\
&= p_0 \cdot F(p_0) - \int_0^{p_0} x\ \mu(x)\mathrm{d}x \\
&+ \left(p_0 - \int_0^{p_0} x\ \mu(x)\mathrm{d}x\right) - p_0 \cdot \big(1 - F(p_0)\big) \\
&= 2\left(p_0 \cdot F(p_0) - \int_0^{p_0} x\ \mu(x)\mathrm{d}x\right) \\
&= 2\int_0^{p_0} F(x)\mathrm{d}x.
\end{aligned}$$

The fourth equation follows from the common prior $p_0 = \int_0^1 x\ \mu(x)\mathrm{d}x$ and the last equality from integration by parts. Therefore, two posterior distributions with the same prior have the same average informativeness if and only if they satisfy the integral condition for second-order stochastic dominance at the prior with equality. □

Proof of Corollary 2.7. In the receiver-optimal equilibrium from Proposition 1, we have seen that all the information reaches the receiver. When senders' signals are more informative and their information is independent, their joint information also becomes more informative by Proposition C in Appendix 2.B. The receiver's utility then can only increase from more information reaching her.

□

Proof of Theorem 2.8, Part A: Expertise bounds:

Lemma 2.12. *For distribution over posteriors μ with cdf F and preference distribution γ, the lower expertise bound $\underline{b}$ in the receiver-optimal equilibrium is determined by*

$$\gamma(0)\big(p_0 - \underline{b}\big) = \int_0^{\underline{b}} (\underline{b} - x)\mathrm{d}\mu = \gamma(\lambda_0) \cdot \int_0^{\underline{b}} F(x)\ \mathrm{d}x,$$

and the upper expertise bound $\bar{b}$ is determined by

$$\gamma(1)\left(\bar{b}-p_0\right) = \gamma(\lambda_0)\int_{\bar{b}}^{1}(x-\bar{b})\mathrm{d}\mu = \gamma(\lambda_0)\cdot\int_{\bar{b}}^{1} 1-F(x)\,\mathrm{d}x.$$

Both equations have a unique solution for all μ and γ.

Proof of Lemma 2.12. As all messages $t_i \leq \underline{b}$ result in the same virtual posterior $\underline{b}$, we have

$$\frac{\gamma(\lambda_0)\int_0^{\underline{b}} F(x)dx + \gamma(0)p_0}{\gamma(\lambda_0)F(\underline{b}) + \gamma(0))} = \underline{b}$$

An algebraic manipulation yields

$$\gamma(0)\left(p_0-\underline{b}\right) = \gamma(\lambda_0)\int_0^{\underline{b}}(\underline{b}-x)\mathrm{d}\mu$$

as stated in the lemma. Note that the left side of the equation is strictly decreasing in $\underline{b} \in [0, p_0]$ and is 0 only if $\underline{b} = p_0$. The right side is weakly increasing in $\underline{b}$ and is 0 for $\underline{b} = 0$. Further, both sides are continuous in $\underline{b}$. Thus, there is a unique $\underline{b}$ that fulfills the equation.

The proof for the upper expertise bound is analogous.

□

Proof of Theorem 2.8, Part B: Equilibrium. As in the proof of Proposition 2.2, we use upper-case characters for random variables and lower-case characters for their realizations.

We start to calculate the virtual posterior $q(t_i)$. The only senders that send messages within the expertise bounds are advisors. Thus, $q(t_i) = t_i$ for $t_i \in (\underline{b}, \bar{b}) \cap \mathscr{P}$. For messages $t_i \leq \underline{b}$ with $t_i \in \mathscr{P}$, the virtual posterior of the receiver is

$$\begin{aligned} q(t_i) &= \mathbb{P}[\omega = 1 | T_i = t_i] \\ &= \mathbb{P}[\omega = 1 | T_i = t_i \wedge \lambda_i = \lambda_0]\mathbb{P}[T_i = t_i | \lambda_i = \lambda_0] \\ &\quad + \mathbb{P}[\omega = 1 | T_i = t_i \wedge \lambda_i = 0]\mathbb{P}[T_i = t_i | \lambda_i = 0] \\ &= \frac{t_i\gamma(\lambda_0)\mu(t_i) + p_0\gamma(0)m(.,0)(t_i)}{\gamma(\lambda_0)\mu(t_i) + \gamma(0)m(.,0)(t_i)} \end{aligned}$$

$$
= \frac{t_i + \dfrac{\underline{b} - t_i}{p_0 - \underline{b}} p_0}{1 + \dfrac{\underline{b} - t_i}{p_0 - \underline{b} p_0}} \\
= \underline{b}.
$$

The calculation for the virtual posterior of messages $t_i \geq \underline{b}$ with $t_i \in \mathcal{P}$ is $q(t_i) = \bar{b}$ by an analogous calculation. Thus, the receiver's on-path beliefs are consistent with Bayesian updating.

The technique of Nitzan and Paroush (1982) and the proof of Proposition 2.2 teach us how to process a set of (virtual) posteriors optimally. Again, the best response of the receiver can be interpreted as a weighted majority rule with weighting function

$$
\begin{aligned}
w(x) &= \ln \frac{q(x)}{1 - q(x)} - \ln \frac{p_0}{1 - p_0} \\
&= \begin{cases}
\ln \frac{\underline{b}}{1-\underline{b}} - \ln \frac{p_0}{1-p_0} & x \in \mathcal{P} \wedge x \leq \underline{b} \\
\ln \frac{x}{1-x} - \ln \frac{p_0}{1-p_0} & x \in \mathcal{P} \wedge \underline{b} \leq x \leq \bar{b} \\
\ln \frac{\bar{b}}{1-\bar{b}} - \ln \frac{p_0}{1-p_0} & x \in \mathcal{P} \wedge \bar{b} \leq x \\
0 & \text{else}
\end{cases}
\end{aligned}
$$

and threshold $\tau = -\left(\ln \frac{\lambda_0}{1-\lambda_0} + \ln \frac{p_0}{1-p_0}\right)$.

We proceed by proving that senders play best responses. Partisans maximize the probability that the receiver takes the action that matches their preference parameter. Given the strategy of advisors and the receiver, they send a message with maximal weight in the preferred direction. In the equilibrium strategy 0- (1-) partisans mix over messages with weight $\ln \frac{\bar{b}}{1-\bar{b}} - \ln \frac{p_0}{1-p_0}$ $(\ln \frac{\underline{b}}{1-\underline{b}} - \ln \frac{p_0}{1-p_0})$ which is the highest (lowest) weight assigned by the receiver. Hence, these partisans play best responses.

The advisor and the receiver have the same utility function and prefer the same action when they have the same information. Thus, the best the advisor can do is to get as close to revealing all his information to the receiver as possible. Given any posterior the advisor holds he can either transmit his information noise-free if his posterior happens to be within the expertise bounds or their is a cap to what he can communicate, which he reaches, when sending his posterior as his message. He hence never faces the tradeoff of having to over or undershoot in what he communicates and hence truthfully messaging is optimal.

Taken together, the strategies and beliefs in Theorem 2.8 are a weak perfect Bayesian equilibrium.

□

Proof of Theorem 1, Part C: Receiver Optimality . We show that the equilibrium in Theorem 2.8 is optimal for the receiver. We proceed in two steps. First, we introduce a technique that allows us to compare equilibria in the common interest case. This represents a more complete discussion than necessary to prove Proposition 2.2. Second, we show that the comparison carries over to the case with partisans. More concretely, we show that if an equilibrium in which advisors play the truthful strategy is more informative than another one in the case without partisans, it continues to be more informative than the other one in the presence of partisans.

The receiver bases her decision on the virtual posteriors $q(t_i)$, which she infers from messages t_i of senders $i = \{1, ..., n\}$. The same set of virtual posteriors leads to the same decision. The distribution of virtual posteriors $q(t_i)$ for sender i is determined by the distribution of posteriors μ and the sender i's strategy m_i.

Definition 2.13. Let μ be a distribution of posteriors and m_i the strategy of sender i. We denote the distribution of virtual posteriors of sender i by $\mu^{\gamma}_{m_i}$ and define it by its cdf

$$F^{\gamma}_{m_i}(x) = \mathbb{P}[q(t_i) \leq x],$$

where t_i is sender i's message. We suppress superscript γ in the common interest case, i.e., we write μ_{m_i} and F_{m_i} if $\gamma(\lambda_0) = 1$.

In the following, we compare the distribution over virtual posteriors of the equilibrium in which advisors play the truthful strategy with the distributions over virtual posteriors of other equilibria. We know from Proposition 2.2 that playing the truthful strategy is part of a receiver-optimal equilibrium for the common interest case. Using the concept of distributions over virtual posteriors helps us generalize this observation to the case with partisans.

The rest of the proof consists of three steps. First, we formalize that the virtual posterior distribution of a single sender is most informative in the common interest case, when advisors play the truthful strategy. Then second we show that even though partisans do not imitate and devalue expertise in all equilibria, whenever they do not, it never changes the action the receiver takes. We can hence ignore such equilibria when checking that our optimality candidate might be dominated. Lastly, we demonstrate that given that truth telling is optimal in the absence of partisans and given that partisans imitate and devalue expertise when they are added, truth telling remains optimal. These three statements are formalized in the three following Lemmas respectively.

Lemma 2.14. *Let μ be a distribution over posteriors, $m_i^\star$ a strategy in which advisors play truthfully, and m_i' any other strategy of sender i. Then, if $\mu_{m_i^\star}$ is more informative than $\mu_{m_i'}$, it follows that $\mu^\gamma_{m_i^\star}$ is more informative than $\mu^\gamma_{m_i'}$, i.e.*

$$\mu_{m_i^\star} \succ \mu_{m_i'} \Rightarrow \mu^\gamma_{m_i^\star} \succ \mu^\gamma_{m_i'}.$$

Lemma 2.15. *Let μ be a distribution over posteriors, $m_i^\star$ the truthful strategy, and m_i' any other strategy. Then it holds that $\mu_{m_i^\star}$ is more informative than $\mu_{m_i'}$, i.e., $\mu_{m_i^\star} \succ \mu_{m_i'}$.*

Lemma 2.16. *Take any weak Bayes Nash Equilibrium s.t. some senders' strategies conditioned on partisanship do not imitate and devalue expertise. Then the receiver's action is always the same as if they were to play this strategy.*

Given these three statements, the virtual posterior of sender i is most informative if types with $\lambda_i = \lambda_0$ play the truthful strategy. Again, under Theorem 12.3.2 in Blackwell and Girshick (1979) (see Proposition B in Appendix 2.B), the sender-wise comparison carries over to the overall information structure. By Theorem 12.2.2 (4) in Blackwell and Girshick (1979) (see Proposition C in Appendix 2.B), we conclude that no better equilibrium for the receiver than that described in Theorem 2.8 exists. We now proof the three lemmas we just applied.

Proof of Lemma 2.14. To simplify the notation, we denote $\mu_{m_i^\star}$ by μ, $\mu_{m_i'}$ by ν, $\mu^\gamma_{m_i^\star}$ by μ^γ, and $\mu^\gamma_{m_i'}$ by ν^γ. Further, we denote $F_{m_i^\star}$ by F, $F_{m_i'}$ by G, $F^\gamma_{m_i^\star}$ by F^γ, and $F^\gamma_{m_i'}$ by G^γ.

To prove that μ^γ is more informative than ν^γ, we show that

$$\int_0^y G^\gamma(x)\mathrm{d}x \leq \int_0^y F^\gamma(x)\mathrm{d}x \quad \text{for all } y \in [0,1].$$

We first derive the following condition:

Let μ and ν with $\mu \succ \nu$ be distributions over posteriors with cdfs F and G, respectively. Let γ be the distribution of preference parameters. Then, the lower (upper) expertise bound $\underline{b}_\mu$ of μ is weakly smaller (greater) or equal than the lower (upper) expertise bound $\underline{b}_\nu$ of ν in the optimal equilibria with partisans.

This holds by the following argument. Suppose that $\underline{b}_\nu < \underline{b}_\mu$, and use Lemma 2.12 to see that

$$\begin{aligned}
\gamma(0)\left(p_0 - \underline{b}_\nu\right) &= \gamma(\lambda_0) \cdot \int_0^{\underline{b}_\nu} G(x)\mathrm{d}x \\
&\leq \gamma(\lambda_0) \cdot \int_0^{\underline{b}_\nu} F(x)\mathrm{d}x \\
&\leq \gamma(\lambda_0) \cdot \int_0^{\underline{b}_\mu} F(x)\mathrm{d}x = \gamma(0)\left(p_0 - \underline{b}_\mu\right).
\end{aligned}$$

Hence, $\underline{b}_\mu \leq \underline{b}_\nu$, which is a contradiction. The proof for the upper expertise bound is analogous.

Using this insight, it holds that $\underline{b}_\mu \leq \underline{b}_\nu$ and $\bar{b}_\mu \geq \bar{b}_\nu$. This allows us to check the inequality separately on the three intervals $\left[0, \underline{b}_\nu\right], \left[\underline{b}_\nu, \bar{b}_\nu\right]$, and $\left[\bar{b}_\nu, 1\right]$.

For all $y \in \left[0, \underline{b}_\nu\right]$, it holds that

$$\int_0^y G^\gamma(x)\mathrm{d}x = 0 \leq \int_0^y F^\gamma(x)\mathrm{d}x.$$

For all $y \in \left[\underline{b}_\nu, \bar{b}_\nu\right]$ it holds that

$$\begin{aligned}
\int_0^y G^\gamma(x)\mathrm{d}x &= \int_{\underline{b}_\nu}^y \gamma(0) + \gamma(\lambda_0)G(x)\mathrm{d}x \\
&= \gamma(0)(y - \underline{b}_\nu) + \gamma(\lambda_0)\int_0^y G(x)\mathrm{d}x - \gamma(\lambda_0)\int_0^{\underline{b}_\nu} G(x)\mathrm{d}x \\
&= \gamma(0)(y - p_0) + \gamma(\lambda_0)\int_0^y G(x)\mathrm{d}x \\
&\leq \gamma(0)(y - p_0) + \gamma(\lambda_0)\int_0^y F(x)\mathrm{d}x \\
&= \int_0^y F^\gamma(x)\mathrm{d}x.
\end{aligned}$$

The first equality follows from the definition of virtual posteriors and the equilibrium strategies. For the third equality, we apply Lemma 2.12. The inequality follows from the assumption that $\mu \succ \nu$.

As $G^\gamma(x) = 1$ for $x \geq \bar{b}_\nu$, it follows that for all $y \in \left[\bar{b}_\nu, 1\right]$, it holds that

$$\int_y^1 G^\gamma(x)\mathrm{d}x \geq \int_y^1 F^\gamma(x)\mathrm{d}x.$$

The expected value of both distributions is consistent with the common prior, i.e., $\int_0^1 F^\gamma(x)\mathrm{d}x = \int_0^1 G^\gamma(x)\mathrm{d}x = 1 - p_0$. Thus, we conclude that

$$\int_0^y G^\gamma(x)\mathrm{d}x \leq \int_0^y F^\gamma(x)\mathrm{d}x,$$

for all $y \in [\bar{b}_\nu, 1]$, which concludes the proof. □

Proof of Lemma 2.15. Let $\mu_{m_i^\star}$ be the distribution over virtual posteriors under the truthful strategy $m_i^\star$. Any distribution $\mu_{m_i'}$ that is induced by another strategy m_i' can be constructed from $\mu_{m_i^\star}$ by an application of garblings. We do not restrict strategies to use only a finite set of messages. Therefore, we apply a result from Blackwell (1953) that generalizes Theorem 12.2.2 in Blackwell and Girshick (1979) (see Proposition A in Appendix 2.B) to the case with continuous signals. Thereby, we conclude that $\mu_{m_i^\star}$ is more informative than $\mu_{m_i'}$. □

Proof of Lemma 2.16. We will prove this by contradiction. Let us assume that an equilibrium exists that violates the statement of the lemma. Recall that in any equilibrium, the receiver strategy is to follow a weighted majority rule. Fix a state of the world and senders' types (p, γ) s.t. if partisans were to imitate and devalue expertise, wlog action 1 would be taken, but instead, action 0 is taken. We will show that partisan types do not play a best response, thereby implying a contradiction.

In this state, if all senders with preference type 1 were to change their message to that associated with the highest virtual posterior, action 1 would be taken. If any partisan is pivotal, he is not best responding, and we directly have a contradiction.

Let us instead assume that multiple 1 partisans need to change their message, to change the action of the receiver and fix any ordering of them. Then there will be a first critical 1 partisan that can change the receiver's action provided that all partisans before him have already selected a message corresponding to the highest virtual posterior.

Now let us replace all partisan senders that were before the critical one with their respective aligned types that send the message associated with the highest virtual posterior. They need to exist in any non-babbling equilibria. This event happens with positive probability. In this situation the critical partisan is pivotal and hence his strategy was not a best response to begin with.

□

This concludes the proof of optimality. □

Proof of Theorem 2.9. The distribution over posteriors of a sender whose partisan types have been exchanged with incompetent types is given by:

$$\mu^{\gamma}_{incompetent} = \gamma(\lambda_0)\mu(x) + \big(\gamma(0) + \gamma(1)\big)\,\delta_{p_0}$$

with average informativeness:

$$\pi(\mu_{incompetent}) = \gamma(\lambda_0)\pi(\mu).$$

The distribution over virtual posteriors derived from the equilibrium in Theorem 2.8 is given by

$$\mu^{\gamma}_{partisan}(x) = \begin{cases} \gamma(\lambda_0)F(x) + \gamma(0) & \text{if } x = \underline{b} \\ \gamma(\lambda_0)\mu(x) & \text{if } x \in \big(\underline{b}, \overline{b}\big) \\ \gamma(\lambda_0)\,(1 - F(x)) + \gamma(1) & \text{if } x = \overline{b} \\ 0 & \text{else.} \end{cases}$$

We first check whether $\mu_{partisan}$ is less informative than $\mu_{incompetent}$. For this, let us denote with $F^{\gamma}_{partisan}$, $F^{\gamma}_{incompetent}$, and F the cdfs of $\mu^{\gamma}_{partisan}$, $\mu^{\gamma}_{incompetent}$, and μ, respectively. We then must show that

$$\int_0^y F_{incompetent}(x)\mathrm{d}x \geq \int_0^y F_{partisan}(x)\mathrm{d}x \;\; \forall y \in [0,1].$$

When $y \in [0, \underline{b}]$, this is true, because $F_{partisan}$ is constant and equal to 0 on this interval. The case for the interval $y \in [\overline{b}, 1]$ follows by a symmetric argument, because the integrals become equal to the prior at $y = 1$. Let us hence focus on $y \in [\underline{b}, \overline{b}]$. We then get

$$\begin{aligned} \int_0^y F_{incompetent}(x)\mathrm{d}x &= \int_0^y \gamma(\lambda_0)F(x) + (1 - \gamma(\lambda_0))\mathbb{1}\{x \geq p_0\})\mathrm{d}x \\ &= \int_0^{\underline{b}} \gamma(\lambda_0)F(x)\mathrm{d}x \end{aligned}$$

$$
\begin{aligned}
&+ \int_{\underline{b}}^{y} \gamma(\lambda_0)F(x) + (1-\gamma(\lambda_0))\mathbb{1}\{x \geq p_0\})\mathrm{d}x \\
&\geq \gamma(0)(p_0 - \underline{b}) + \int_{\underline{b}}^{y} \gamma(\lambda_0)F(x) + \gamma(0)\mathbb{1}\{x \geq p_0\})\mathrm{d}x \\
&\geq \int_{\underline{b}}^{y} \gamma(\lambda_0)F(x) + \gamma(0)\mathrm{d}x \\
&= \int_{0}^{y} F_{partisan}(x)\mathrm{d}x.
\end{aligned}
$$

The first inequality makes use of Lemma 2.12. In the last equality, we use the fact that $F_{partisan}$ is equal to 0 on $[0, \underline{b}]$. Lastly, we verify the average informativeness of $\mu_{partisan}$ to be

$$
\begin{aligned}
\pi(\mu^{\gamma}_{partisan}) &= \int_0^1 |x - p_0| \mathrm{d}\mu_{partisan} \\
&= \gamma(\lambda_0) \int_0^1 |x - p_0| \mathrm{d}\mu + \gamma(0)(p_0 - \underline{b}) \\
&\quad - \int_0^{\underline{b}} (\underline{b} - x)\mathrm{d}\mu + \gamma(1)(\bar{b} - p_0) - \int_{\bar{b}}^1 (x - \bar{b})\mathrm{d}\mu \\
&= \gamma(\lambda_0)\pi(\mu).
\end{aligned}
$$

Here the last equation is a consequence of Lemma 2.12.

□

Proof of Theorem 2.10. We prove the proposition in two steps. We start to show that by monotonicity and continuity of $\underline{b}$ and $\bar{b}$, there exists $c_0, c_1 \in (0,1)$, such that the receiver can only form two expected posteriors in the optimal equilibrium. Then, we prove that there exist c_0 and c_1 such that $c_0 + c_1 < 1$. For both parts, we use Lemma 2.12, which characterizes the lower expertise bound by the equation

$$\gamma(0)\left(p_0 - \underline{b}\right) = \gamma(\lambda_0) \cdot \int_0^{\underline{b}} F(x)\mathrm{d}x.$$

The lower expertise bound can take any value in $\underline{b} \in [0, \max\{x : F(x) = 0\}]$ if $\gamma(0) = 0$. Further, it is p_0 if $\gamma(0) = 1$. Rewriting the above equation yields

$$\frac{\gamma(0)}{\gamma(\lambda_0)} = \frac{\int_0^{\underline{b}} F(x)\mathrm{d}x}{p_0 - \underline{b}} \tag{2.A.1}$$

which exhibits that $\underline{b}$ is monotonically increasing in $\gamma(0)$, monotonically decreasing in $\gamma(\lambda_0)$, and continuous in $\gamma(0), \gamma(\lambda_0) \in (0,1)$.

As μ is never-ignorant, there exists a highest type strictly smaller than the prior, $p_L := \max\{x | x < p_0 \wedge x \in \mathscr{P}\}$. The proposition is fulfilled if the lower expertise bound equals this type $\underline{b} = p_L$. The continuity and monotonicity of $\underline{b}$ imply that the right-hand side of Equation (2.A.1) is positive and finite, and hence, $\gamma(0) < 1$ if $\underline{b} = p_L$. The proof for the upper part with type $p_H := \min\{x | x > p_0 \wedge x \in \mathscr{P}\}$ is analogous, so that constants $c_0, c_1 \in (0,1)$ are implicitly given by

$$c_0\left(p_0 - p_L\right) = \gamma(\lambda_0) \cdot \int_0^{p_L} F(x)\mathrm{d}x \tag{2.A.2}$$

$$c_1\left(p_H - p_0\right) = \gamma(\lambda_0) \cdot \int_{p_H}^{1} 1 - F(x)\mathrm{d}x. \tag{2.A.3}$$

To see that $c_0 + c_1 < 1$, divide Equations (2.A.2) by $\left(p_0 - p_L\right)$ and $\left(p_H - p_0\right)$, respectively. Adding both equations yields

$$c_0 + c_1 = \gamma(\lambda_0) \cdot \frac{\int_0^{p_L} F(x)\mathrm{d}x}{p_0 - p_L} + \gamma(\lambda_0) \cdot \frac{\int_{p_H}^{1} 1 - F(x)\mathrm{d}x}{p_H - p_0}.$$

As $\frac{\int_0^{p_L} F(x)\mathrm{d}x}{p_0 - p_L}, \frac{\int_{p_H}^{1} 1-F(x)\mathrm{d}x}{p_H - p_0} > 0$, it follows that $\gamma(\lambda_0) > 0$. This implies that $c_0 + c_1 = \gamma(0) + \gamma(1) = 1 - \gamma(\lambda_0) < 1$, which completes the proof. □

Proof of Proposition 2.11. By Lemma 2.12, we know that as long as $\gamma(0)$ and $\gamma(1)$ remain above some fixed $\epsilon > 0$, we have that as $\gamma(\lambda_0)$ converges to 0, expertise bounds converge to the prior.

Hence, there exists c_0 and c_1 for which both μ^γ and ν^γ have, at most, one mass point above and below the prior, and for large enough partisanships, the masses of these points become arbitrarily similar, as almost no aligned senders are left. However, the relative difference in the average informativeness of both virtual posterior measures stays fixed. Therefore, there exist c_0 and c_1 for which the mass points, which are not at the prior, of μ^γ are necessarily further away from the prior than those of ν^γ. However, this implies that μ^γ is more informative than ν^γ, which implies our result by Proposition 2.

□

Appendix 2.B Appendix: Basic Properties of the Informativeness Order

In this part of the appendix, we collect certain tools from the literature that we use throughout and that are related to definition 1. This allows us to compare the receiver's utility for different distributions over posteriors of senders. The methods and results in this Subsection are borrowed from Chapter 12 in Blackwell and Girshick (1979), building on majorization theory, first developed by Hardy, Littlewood, and Polya (1929). To apply their machinery to our problem, we adjust our setting, and translate our notation into theirs.

The following results rely on the assumption that the action space of the receiver is a closed bounded convex subset of $\mathbb{R}$. To fulfill this assumption, we extend the action space of the receiver from $\{0,1\}$ to $\Delta\{0,1\}$, so that her action space is the interval $[0,1]$. An action $a \in \Delta\{0,1\}$ corresponds to the probability that the receiver takes action 1. Note that we can use this extended action space throughout the whole article without changing any result. In all statements on the receiver's best response, one of the two extreme actions $\{0,1\} \subset \Delta\{0,1\}$ is optimal. We use the action space $\{0,1\}$ in the main text of the article to simplify the exposition.

To present the next results, it is also helpful to introduce some of the notation of Blackwell and Girshick (1979). For a distribution over posteriors μ, we define a $2 \times N$ matrix P, where $N = |\mathscr{P}|$ is the number of possible posteriors. The rows represent the two states of the world 0 and 1. Each column represents one possible posterior. The value P_{ij} is the probability of observing the posterior represented by column j in state i. Note that matrix P is Markov, which means that $P_{ij} > 0$ for all i and j and that $\sum_{j=1}^{N} P_{ij} = 1$ for all i. With the notation, we are equipped to remind the reader of Theorem 12.2.2 in Blackwell and Girshick (1979).

Proposition A (Blackwell and Girshick (1979))**.** Let P and Q be two $2 \times N_1$ and $2 \times N_2$ Markov matrices of distributions over posteriors μ and ν. μ is more informative than ν if and only if there is an $N_1 \times N_2$ Markov matrix M with $PM = Q$.

Matrix M is said to *garble* information by transforming matrix P to Q. This means that distribution ν can be constructed from distribution μ. This interpretation justifies the statement that μ is more informative than ν.

The next result generalizes the previous proposition by enabling the comparison of sets of distributions. Each sender sends a conditionally independent posterior. Consider two sets of senders with different distributions over posteriors. Then, Theorem 12.3.2 in Blackwell and Girshick (1979) allows us to compare the information of both groups in the following sense.

Proposition B (Blackwell and Girshick (1979)). Let $(\mu_i)_{i=1}^n$ and $(\nu_i)_{i=1}^n$ be two sets of distributions over posteriors. Suppose that μ_i is more informative than ν_i for every i. Then, the combination of distributions over posteriors $(\mu_i)_{i=1}^n$ is more informative than $(\nu_i)_{i=1}^n$.

The proposition allows us to compare the information that is transmitted to the receiver from different distributions. Theorem 12.2.2 (4) in Blackwell and Girshick (1979) allows us to use this result for a statement on the utility of the receiver.

Proposition C (Blackwell and Girshick (1979)). Let μ and ν be two distributions over posteriors such that μ is more informative than ν. Then for every continuous convex function $\phi : [0,1] \to \mathbb{R}$, we have

$$\mathbb{E}_{\mu}[\phi(x)] \geq \mathbb{E}_{\mu'}[\phi(x)].$$

Note that the utility function $u^\star(q)$ is convex in q. Thus, if there are two distributions over posteriors with $\mu \succ \nu$, the proposition implies that the expected utility for the receiver with distribution μ is at least as high, as with distribution ν.

Appendix 2.C Appendix: Discussion of Assumptions

More Actions

One can easily imagine situations in which a receiver not only has a binary choice but also might prefer to take any of a range of intermediate actions whenever his belief is sufficiently far away from either 0 or 1. In these cases, our analysis mostly generalizes as long as aligned senders still have identical utility functions as the receiver and the partisans, they still have monotone preferences with regard to the receiver's belief regardless of the true state of the world. To describe such more general models, one would likely abstract the exposition to a reduced form in which each agent's utility is simply a function of the receiver's belief. We have forgone this extra generality to present the model with a concrete and maximally simple decision problem. As a byproduct, the interpretation of our numerous illustrations is also simplified.

Also note that our focus on a binary decision problem is conservative in the role specialization and the loss of it plays for the receiver. Any additional action that does not dominate any previous actions can only make the receiver's utility more convex, hence increasing the gains from specialization.

Continuous Support

In this article, we focused on situations in which every signal only has a finite set of possible realizations. We do not see this as more or less natural than to assume a distribution over posteriors that is absolutely continuous but rather find that working with general probability measures would likely add little insight, yet complicate our exposition.

The only significant change that we anticipate, if one would pursue to rewrite this model with absolutely continuous posterior functions is a necessity to redefine *never-ignorant* as we have used it in Theorem 2.10, to mean that a neighborhood around the prior is not included in the support of posterior distribution. In essence, we require a lower bound of the information contained in every possible posterior, which in the discrete case, is given whenever the prior is not itself a possible posterior.

Heterogeneous Senders

We decided to restrict our analysis to symmetric senders for ease of exposition. However, in some of the applications, it stands to reason that the receiver can discriminate among the senders based on prior knowledge. A regulator might understand that one of his advising experts has previously worked on the approval of similar regulations and might hence believe his distribution over posteriors to be more informative than average.

As we have seen in our discussion of the symmetric senders, the receiver's learning from messages happens sender by sender, i.e., the message of one sender does not change how the message of another is processed. Consequently, all that changes in our model, when we allow for asymmetric senders, is that the weighting function that the receiver uses needs to be individualized based on each sender's posterior and preference distribution.

Learning about the informativeness and preferences of senders can, of course, only improve the receiver's situation as she can always choose to ignore that information regardless. Hence, our symmetric case can also be interpreted as a worst-case benchmark for the effect of partisanship on information transmission.

Departing from that benchmark, our previous analysis suggests that the largest gains from knowledge about the individual sender's informativeness and preferences are generated by the possibility of finding a specialist with a low probability of partisanship. As we have seen, specialization can have great benefits as long as the expert is also well trusted.

References

Alonso, Ricardo, Wouter Dessein, and Niko Matouschek. 2008. "When does coordination require centralization?" *American Economic Review* 98 (1): 145–79. [25]

Austen-Smith, David. 1990. "Information transmission in debate." *American Journal of Political Science,* 124–52. [25]

Azrieli, Yaron. 2018. "The price of 'one person, one vote'." *Social Choice and Welfare* 50 (2): 353–85. [27]

Battaglini, Marco. 2017. "Public protests and policy making." *Quarterly Journal of Economics* 132 (1): 485–549. [25]

Baye, Michael, Dan Kovenock, and Casper De Vries. 1993. "Rigging the lobbying process: an application of the all-pay auction." *American Economic Review* 83 (1): 289–94. [42]

Benabou, Roland, and Guy Laroque. 1992. "Using privileged information to manipulate markets: Insiders, gurus, and credibility." *Quarterly Journal of Economics* 107 (3): 921–58. [26]

Blackwell, David. 1950. "Comparison of experiments." Working paper. Howard University, Washington United States. [32]

Blackwell, David. 1953. "Equivalent comparisons of experiments." *Annals of Mathematical Statistics,* 265–72. [50]

Blackwell, David, and Meyer Girshick. 1979. *Theory of Games and Statistical Decisions.* Courier Corporation. [35, 48, 50, 55, 56]

Buchanan, James, Robert Tollison, and Gordon Tullock. 1980. *Toward a theory of the rent-seeking society.* 4. Texas A-&-M University Press. [42]

Condorcet, Marquis de. 1785. "Essay on the Application of Analysis to the Probability of Majority Decisions." *Imprimerie Royale Paris.* [27]

Crawford, Vincent, and Joel Sobel. 1982. "Strategic information transmission." *Econometrica,* 1431–51. [25]

Dye, Ronald A. 1985. "Disclosure of nonproprietary information." *Journal of accounting research,* 123–45. [26]

Feddersen, Timothy, and Wolfgang Pesendorfer. 1996. "The swing voter's curse." *American Economic Review,* 408–24. [27]

Feddersen, Timothy, and Wolfgang Pesendorfer. 1997. "Voting behavior and information aggregation in elections with private information." *Econometrica,* 1029–58. [27]

Frankel, Alexander, and Emir Kamenica. 2019. "Quantifying information and uncertainty." *American Economic Review* 109 (10): 3650–80. [33]

Gerardi, Dino, Richard McLean, and Andrew Postlewaite. 2009. "Aggregation of expert opinions." *Games and Economic Behavior* 65 (2): 339–71. [29]

Gilligan, Thomas, and Keith Krehbiel. 1989. "Asymmetric information and legislative rules with a heterogeneous committee." *American Journal of Political Science,* 459–90. [25]

Glazer, Jacob, Helios Herrera, and Motty Perry. 2021. "Fake reviews." *Economic Journal* 131 (636): 1772–87. [26]

Hardy, Godfrey, John. Littlewood, and George Polya. 1929. "Some simple inequalities satisfied by convex functions." *Messenger of Mathematic* 58 (58): 225–43. [55]

Hummel, Patrick, John Morgan, and Phillip Stocken. 2013. "A model of flops." *RAND Journal of Economics* 44 (4): 585–609. [25]

Jung, Woon-Oh, and Young K Kwon. 1988. "Disclosure when the market is unsure of information endowment of managers." *Journal of Accounting research,* 146–53. [26]

Kim, Kyungmin, and Jonathan Pogach. 2014. "Honesty vs. advocacy." *Journal of Economic Behavior & Organization* 105: 51–74. [26]

Krishna, Vijay, and John Morgan. 2001. "A model of expertise." *Quarterly Journal of Economics* 116 (2): 747–75. [25]

Li, Hao, Sherwin Rosen, and Wing Suen. 2001. "Conflicts and common interests in committees." *American Economic Review* 91 (5): 1478–97. [27]

Li, Ming, and Kristóf Madarász. 2008. "When mandatory disclosure hurts: Expert advice and conflicting interests." *Journal of Economic Theory* 139 (1): 47–74. [26]

McLennan, Andrew. 1998. "Consequences of the Condorcet jury theorem for beneficial information aggregation by rational agents." *American Political Science Review* 92 (2): 413–18. [31]

McMurray, Joseph. 2017. "Ideology as opinion: A spatial model of common-value elections." *American Economic Journal: Microeconomics* 9 (4): 108–40. [27]

Morgan, John, and Phillip C Stocken. 2003. "An analysis of stock recommendations." *RAND Journal of conomics,* 183–203. [26]

Nitzan, Shmuel, and Jacob Paroush. 1982. "Optimal decision rules in uncertain dichotomous choice situations." *International Economic Review,* 289–97. [31, 42, 46]

Shapley, Lloyd, and Bernard Grofman. 1984. "Optimizing group judgmental accuracy in the presence of interdependencies." *Public Choice* 43 (3): 329–43. [31]

Sobel, Joel. 1985. "A theory of credibility." *Review of Economic Studies* 52 (4): 557–73. [26]

Vidal, Jordi Blanes i. 2006. "Credibility and strategic communication: Theory and evidence from securities analysts." Working paper. [26]

Wolinsky, Asher. 2002. "Eliciting information from multiple experts." *Games and Economic Behavior* 41 (1): 141–60. [25]

Chapter 3

Bad Compromises*

3.1 Introduction

The idea of the tyranny of the majority is a significant concern within modern democracies. Limitations of majority rule are central in most political systems, from additional super-majority requirements and mandatory checks and balances to the direct unchangeability of critical aspects of constitutional frameworks. In this paper, I will discuss a related, yet much less studied, threat to welfare within democracies: The tyranny of a compromising minority.

The famous mean voter theorem implies for single-peaked-preference domains that majority rule leads to the election of the mean voter's preferred alternative. However, in principle, the preferences of the mean and average voters can be arbitrarily far apart. In some cases that I discuss below, the mean voter would prefer both alternatives to the left and the right of the median voter's first choice. I will take the mechanism design approach and ask how a robust voting system can deal with such instances. My findings are rather stark. In contrast to the gradual options of super-majority requirements that can regulate a tyranny of the majority, only the ex-ante exclusion of compromises can prevent the tyranny of the compromising minority.

I should first revisit the famous Gibbard–Satterthwaite theorem to understand this surprising result. It states that any strategy-proof and surjective mechanism is dictatorial on the complete domain of preferences over more than two alternatives. It is hence usually interpreted as a strong nega-

* I thank Felix Bierbrauer, Tilman Börgers, Johannes Hörner, Stephan Lauermann, Benny Moldovanu, Martin Pollrich and Justus Winkelmann for numerous helpful discussions.

tive result. However, suppose one is willing to preclude alternatives. In that case, it is possible to implement any qualified majority decision between two preselected alternatives, and with this limited form of preference aggregation, welfare often increases. Hence, the reduction in incentive constraints comes from precluding alternatives that potentially improve welfare.

This phenomenon extends to the domain of single-peaked preferences for which the relevant characterization of surjective mechanisms due to Moulin is usually seen as a positive result since it consists of a rich set of generalized median voting mechanisms. More specifically, I will see that in this setup, it will only ever be welfare-improving to exclude compromises, i.e., alternatives other than the left-most and right-most.

Let us illustrate this result and its relevance with the example of parliamentary elections. For most such elections, the outcomes can be classified into three categories. They can result in either a stable majority of the two blocks in the chamber or unclear majorities. Most voters favor a government by their preferred block, yet some vote for fringe parties. These fringe parties usually either represent a specific agenda incompatible with either block or are also used by voters to voice their criticism and punish the parties within the two blocks. Since those parties will not join a coalition of either side, the supporters of those parties reveal their preference for uncertain majorities. They are the compromising minority. The unstable majorities that they enforce lead to a prolonged period of finding the next government and can also lead to a government of the smallest consensus, depending on how polarized the chamber is. To illustrate the welfare effects that can arise in such a situation, I give the following numerical example. The numbers should be read as von Neumann Morgenstern utilities.

		Alternatives			
		Left	Unclear	Right	Frequencies
Types	Left Block	1	1/3	0	46- 48%
	Fringe Left	1/3	1	0	3-4%
	Fringe Right	0	1	1/3	3-4%
	Right Block	0	1/3	1	45-47%

Note that there is some aggregate uncertainty over the distribution of agents' preferences. This translates into uncertainty about whether a left-block or a right-block government is the utilitarian choice. Note, however, that there is no uncertainty that the unclear majority's outcome is the worst possible outcome, despite being the likely favored choice of the median voter.

One possible reaction to this is to be willing to deviate from majority rule and allow for biased mechanisms to prevent compromise selection. However, this leads us to a more fundamental weakness of any non-precluding mechanism. It can be biased towards either extreme outcome and hence ensures this alternatives selection with near certainty. However, no such mechanism that fulfills standard conditions and is hence a generalized median mechanism can always select the larger political block. Therefore, these non-precluding mechanisms do not allow for an optimal decision between the two block governments.

If, in contrast, the compromise outcome is precluded, then a near-optimal comparison is possible, yielding an increase in expected total welfare. Actual electoral provisions like voting thresholds, majority bonuses, and two-party systems are all designed to reduce the chance of unclear majorities. From those, only the two-party system precludes them entirely, with the others giving incentives for strategic misreporting of preferences.

Another provision in a number of modern constitutions can be interpreted as a direct application of this theory: Constructive motions of no confidence. As a lesson taken from the fragility of numerous governments during the Weimar Republic[1], the Grundgesetz of the Federal Republic of Germany first introduced the requirement that a sitting chancellor can only be removed from office by the election of his successor. It also leaves no ability for the opposition on their own to dismiss the parliament, even if they received a majority due to the deterioration of the governing coalition. This has removed the bad compromise of new elections that can generate a less stable state.[2]

I proceed with this discussion as follows. I analyze strategy-proof mechanisms and ex-ante welfare maximization in settings with arbitrary numbers of agents and alternatives where agents have single-peaked preferences over alternatives. As already discussed, utilitarian mechanisms are only sometimes surjective, i.e., they potentially never choose specific alternatives regardless of agents' reported preferences. This is particularly striking since it is often possible yet unlikely that ex-post all agents jointly rank one of these alternatives first.

I further determine the relative amount of welfare that can be guaranteed by surjective mechanisms and show that this "Guaranteed Fraction of Optimal Welfare" (GFOW) is 50% for three alternatives and converges to 0 as the number of alternatives grows large.

This analysis relies on careful discussion of previous work characterizing strategy-proof and anonymous mechanisms within this model. Moulin (1980) characterizes all surjective mechanisms that only depend on an agent's top choice from all alternatives (codomain-tops-only) as general-

1. The Weimar Republic had twenty governments in its fourteen-year existence.

2. Countries that since have adopted constructive motions of no confidence include, among others, Belgium, Spain, Hungary, Israel, Poland, and Slovenia.

ized median schemes. Such schemes select the median of agents' reported peaks and $n-1$ fixed peaks of "phantom voters". Under the same restriction, he finds that non-surjective mechanisms can be described with the help of $n+1$ phantoms. Later research by Barberà and Jackson (1994) shows that all mechanisms only depend on an agent's favorite choice from the subset of alternatives that are potential outcomes of the mechanism (image-tops-only).

The two notions of tops-only coincide for surjective mechanisms, and hence all surjective mechanisms are generalized median schemes with $n-1$ phantom voters. The same is not true for non-surjective mechanisms, i.e., some cannot be described by $n+1$ phantoms. Instead, I show that all mechanisms are generalized median voter schemes with $n-1$ phantoms on predetermined subsets of alternatives. Of those, the mechanisms that satisfy codomain-tops-only select from convex subsets.

The rest of this article is structured as follows. In the remainder of this section, I review related literature. In section 2, I introduce the basic model and characterize the set of strategy-proof mechanisms. I also contrast this result with the previous characterizations discussed above. In section 3, I show that mechanisms that map to convex proper subsets of alternatives are only sometimes uniquely optimal, define bad compromises, and give examples. In section 3.1, I discuss GFOW and its dependence on the number of original alternatives. Section 4 concludes.

Related Literature

A large body of literature has studied the implementation of desirable social choice rules. In many of these contexts, the focus has been on implementation in dominant strategies since this concept demands the least from agents' ability to predict other agents' actions. The classical *Median Voter Theorem* (Black (1948)) states that if agents are restricted to the domain of single-peaked preferences, a Condorcet winner, i.e., the alternative that would win against any other alternative in a one-to-one comparison, always exists and can be implemented. By focusing on ordinal preferences, this article and most of the subsequent literature do not consider agents' preference intensities.

In contrast, this analysis focuses on maximizing ex-ante expected welfare, an approach that goes back to Rae (1969). However, Rae's article, like most in the social welfare literature, focuses on the case of two alternatives and hence never deals with compromise alternatives.

Zeckhauser (1969) is the first who studies the implications of agents' preference intensities for a social choice problem with compromise. He observes that if a compromise alternative is the Condorcet winner, it can still lose against a proposed lottery over extreme alternatives in a majority vote. In that case, agents have non-single-peaked preferences over the set of all lotteries over alternatives, and hence the problem of cyclical majorities arises. Consequently, Zeckhauser focuses on conditions that preclude such occurrences. He finds that if agents are sufficiently averse to the risk of their least preferred choice being elected, the Condorcet winner always wins majority support against any lottery over alternatives.

Börgers and Postl (2009) study both compromises and welfare, but under the less stringent condition of Bayesian incentive compatibility. In their setup, there are three alternatives, and it is commonly known that each agent's first choice is the other agent's last choice. However, the relative value each agent assigns to the compromise is private information. They characterize the set of incentive-compatible mechanisms and determine numerically that the difference between utilitarian and first-best welfare is often small.

Lastly, the work that is closest in spirit to mine is Gershkov, Moldovanu, and Shi (2017). In this paper, the authors derive the welfare maximizing surjective mechanism in settings in which agents have both single-peaked and single-crossing preferences. My analysis also applies to their setup, but I focus on the welfare effect of not requiring surjectivity.

3.2 The Social Choice Model

I consider a social choice problem in which n agents have to choose one out of K alternatives. I denote by $\mathscr{K} = \{1, \ldots, K\}$ the set of alternatives.

Agent i's type is $\theta_i = (\theta_i^1, \ldots, \theta_i^K) \in \Theta_i$, where θ_i^k represents agent's utility if alternative k is selected. Θ_i is the same for all agents and consists of all vectors $\theta_i \in \mathbb{R}^K$ s.t. The following three requirements are fulfilled.

Firstly, preferences are single-peaked with respect to the natural order, i.e., there exists an alternative $k \in \mathscr{K}$, the peak, s.t. for all $l, m \in \mathscr{K}$, $l < m \leq k$ or $l > m \geq k$ implies $\theta_i^l < \theta_i^m$. Secondly, an agent is never indifferent, i.e., $\forall k, l \in \mathscr{K}$, $k \neq l$ I have $\theta_i^k \neq \theta_i^l$.[3] Lastly, baseline utility is normalized, i.e., $\min_{k \in \mathscr{K}} \theta_i^k = 0$.[4]

3. This simplifies the exposition since it assures that agents have single-peaked preferences over any subset of alternatives instead of single-plateaued preferences. Barberà and Jackson (1994) and Berga (1998), among others, discuss the case of single-plateaued preference and the characterization of strategy-proof tie-breaking rules.

4. This assumption is without loss of generality since I am only interested in the absolute difference the group decision makes to each agent and not each agent's baseline utility.

Agents' types $\theta = (\theta_1, \dots, \theta_n)$ are distributed according to a commonly known probability measure Ψ on $\Theta = \Theta_1 \times \dots \times \Theta_n$.

A (direct and deterministic) mechanism asks agents for their types and selects an alternative based on their reports. More formally:

Definition 3.1. A *mechanism* g is a function from Θ to $\mathcal{K}$.

To formally describe the set of mechanisms that I consider, I need a few standard definitions:

Definition 3.2.
i A mechanism g satisfies *dominant strategy incentive compatibility* (DIC) if for all i and any $\theta_i, \theta_i' \in \Theta_i$ and $\theta_{-i} \in \Theta_{-i}$ I have $\theta_i^{g(\theta_i, \theta_{-i})} \geq \theta_i^{g(\theta_i', \theta_{-i})}$.
ii A mechanism g is *anonymous* if for all permutations $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ and all reports $\theta \in \Theta$ I have $g(\theta_1, \dots, \theta_n) = g(\theta_{\sigma(1)}, \dots, \theta_{\sigma(n)})$.
iii A mechanism satisfies surjective if for all $k \in \mathcal{K}$ there exists a $\theta \in \Theta$ s.t. $|g(\Theta)| = k$.

DIC ensures that agents have a dominant strategy to report their types truthfully. A mechanism satisfies anonymity if it treats agents the same. Lastly, surjective precludes that any alternative cannot be chosen, independent of agents' reports.[5]

I denote with $t_{\mathcal{S}}(\theta_i) = \arg\max_{k \in \mathcal{S}} \theta_i^k$, agent i's top choice from a subset of alternatives $\mathcal{S} \subset \mathcal{K}$.

Definition 3.3. Let m be such that $n + m$ is odd. A mechanism g is a generalized median voter scheme on $\mathcal{S} \subset \mathcal{K}$ with m phantom voters if there exist numbers $a_1, \dots, a_m \in \mathcal{S}$ s.t. $g(\theta) = Median(a_1, \dots, a_m, t_{\mathcal{S}}(\theta_1), \dots, t_{\mathcal{S}}(\theta_n))$ for all $\theta \in \Theta$.

A generalized median voter scheme selects a subset of possible outcomes and asks agents for their most preferred alternatives from that set. Then it selects the median of these reported peaks and the previously fixed peaks of phantom voters.

I start my analysis with Moulin's characterization of anonymous, DIC, and surjective mechanisms in a strengthened form proven in Barberà and Jackson (1994).[6]

Theorem 3.4 (Moulin-Barberà-Jackson Theorem). *For any mechanism g, the following conditions are equivalent.*

5. In the context of this model, surjective is equivalent to the usually stronger notions of unanimity and Pareto-efficiency as noted for example in Ching (1997).

6. Moulin's original work also assumes that mechanisms need to be codomain-tops-only, a requirement that proved to be non-restrictive, as discussed below.

(i) *g is anonymous, DIC, and respects surjective.*

(ii) *g is a generalized median voting scheme on $\mathcal{K}$ with $n-1$ phantom voters.*

From this result, I can immediately deduce the following characterization for mechanisms that do not need to respect surjective:

Theorem 3.5. *For any mechanism g, the following conditions are equivalent.*

(i) *g is anonymous and DIC.*

(ii) *g is a generalized median voting scheme on a non-empty set $\mathcal{S} \subset \mathcal{K}$ with $n-1$ phantom voters.*

Proof. The proof can be found in the Appendix. □

One might wonder about the relationship between this characterization and another result from Moulin (1980) about $n+1$ phantom voters that seemingly analyze the same set of mechanisms. I compare the two below. The essential difference is an additional assumption in the classical paper:

Definition 3.6. A mechanism g is *image-tops-only* if for all reports $\theta', \theta'' \in \Theta$ s.t. $t_{g(\Theta)}(\theta_i') = t_{g(\Theta)}(\theta_i'')$ for all i, I have $g(\theta') = g(\theta'')$. A mechanism g is *codomain-tops-only* if for all reports $\theta', \theta'' \in \Theta$ s.t. $t_{\mathcal{K}}(\theta_i') = t_{\mathcal{K}}(\theta_i'')$ for all i, I have $g(\theta') = g(\theta'')$.

Therefore an image-tops-only mechanism only considers agents' top choices from those alternatives that are potentially chosen. In contrast, a codomain-tops-only mechanism only takes into account agents' overall most preferred choices irrespective of whether it might have been excluded ex-ante. Most of the literature uses image-tops-only, and it was proven by Barberà and Jackson (1994) that, indeed, all DIC mechanisms on the full domain of single-peaked preferences are image-tops-only. Moulin (1980), however, restrains attention to mechanisms that are codomain-tops-only. These definitions coincide if the mechanisms considered are surjective, as in the case of theorem 3.4, and hence, in this case, codomain-tops-only is also non-restrictive. If, however, surjectivity is not assumed, this restriction reduces the set of mechanisms:

Theorem 3.7 (Comparison to Proposition 2 in Moulin (1980)). *For any mechanism g, the following conditions are equivalent.*

(i) *g is anonymous, DIC and codomain-tops-only.*

(ii) *g is a generalized median voting scheme on $\mathcal{K}$ with $n+1$ phantom voters.*

(iii) g is a generalized median voting scheme on a non-empty convex set[7] $S \subset \mathcal{K}$ *with* $n-1$ *phantom voters.*

Proof. The proof can be found in the Appendix. □

Note that any mechanism that excludes an interior alternative ex-ante cannot be codomain-tops-only and DIC. Such a mechanism has to treat two preference profiles with identical first but non-identical second preferences the same, even though the first preference might have become irrelevant hence giving an agent an incentive to misreport. In contrast, if an alternative is to the left or right of all alternatives in the image of the mechanism, then an agent's most preferred alternative in the image is always the alternative closest to the original first preference. Hence the relevant information can be retrieved from the agent stating his peak on the entire domain.

To illustrate the differences between the three classes of mechanisms characterized in theorems 3.4 to 3.7, I have summarized all mechanisms for two agents and three alternatives in Table 1. In this case, all mechanisms can be represented by a triplet $[a, b, c]$[8] with $a, b, c \in \{\mathrm{x}, 0, 1, 2, 3\}$, where x represents that the respective alternative is not in $\mathcal{S}$ and $0, 1, 2$ or 3 represents the number of phantom voter peaks on a given alternative if it is in $\mathcal{S}$. The mechanisms that are characterized in theorem 3.5 but neither in theorem 3.4 nor theorem 3.7 are [1,x,0] and [0,x,1]. They are precisely those that only allow a decision between the extreme alternatives 1 and 3 and exclude the compromise choice 2 ex-ante.

7. I say that a set $\mathcal{S} \subset \mathcal{K}$ is convex if there exists an interval $I \subset \mathbb{R}$ s.t. $\mathcal{S} = \mathcal{K} \cap I$.
8. I use square bracket to distinguish mechanism from utility vectors.

Table 3.1. Table of Mechanisms for $n = 2$ and $K = 3$

surjective	Codomain-tops-only	All
[1,0,0]	[2,0,1]	[1,0,0]
[0,1,0]	[1,1,1]	[0,1,0]
[0,0,1]	[1,0,2]	[0,0,1]
	[2,1,0]	[1,0,x]
	[1,2,0]	[0,1,x]
	[0,2,1]	[x,1,0]
	[0,1,2]	[x,0,1]
	[3,0,0]	[1,x,x]
	[0,3,0]	[x,1,x]
	[0,0,3]	[x,x,1]
		[1,x,0]
		[0,x,1]

3.3 Bad compromises

Now that I have characterized the set of possible mechanisms, I can proceed with analyzing welfare maximization. I define the welfare function

$$W(g, \Psi) = \sum_{i=1}^{n} \int \theta_i^{g(\theta)} d\Psi$$

that measures the expected ex-ante welfare under type measure Ψ when mechanism g is employed.

Denote by $\mathscr{M}$ the set of all DIC and anonymous mechanisms, by $\mathscr{M}_{con}$ those that additionally satisfy co-domain tops-only or equivalently have a convex image and by $\mathscr{M}_{sur}$ the set of those that satisfy surjective. Clearly, I have $\mathscr{M}_{sur} \subset \mathscr{M}_{con} \subset \mathscr{M}$.

I can now formally state what it means for a mechanism to be welfare maximizing.

Definition 3.8. A DIC mechanism $g^* : \Theta^n \to \mathscr{K}$ is *utilitarian* for the type measure Ψ, if $g \in \arg\max_{g \in \mathscr{M}} W(g, \Psi)$.

In opposition to that, I find it useful to define a notion of the inefficiency of a mechanism that is much stronger than not always selecting Pareto-efficient outcomes. This will guide my discussion of the different subclasses of mechanisms,

Definition 3.9. A mechanism g is *inferior* to a mechanism $\hat{g}$ if for all $\theta \in \Theta$, I have that $\hat{g}(\theta)$ is a weak Pareto-improvement over $g(\theta)$ and there exists a $\theta \in \Theta$ s.t. $\hat{g}(\theta)$ is a strict Pareto-improvement over $g(\theta)$.

An inferior mechanism is unappealing in a strong sense. All non-surjective codomain-tops-only mechanisms are of this type.

Theorem 3.10 (No preclusion of extremes). *For every mechanism $g \in \mathcal{M}$ s.t. $1 \notin g(\Theta)$ or $K \notin g(\Theta)$ there exists a mechanism $\hat{g} \in \mathcal{M}$ s.t. g is inferior to $\hat{g}$*

Proof. The proof can be found in the Appendix. □

Corollary 3.11. *There exists a utilitarian mechanism s.t. 1 and K are in its image, and further, there is always a utilitarian mechanism that is not in $\mathcal{M}_{con} \setminus \mathcal{M}_{sur}$.*

Proof. The proof can be found in the Appendix. □

If one were to maximize welfare on $\mathcal{M}_{con}$, it would be without loss to restrict oneself further to $\mathcal{M}_{sur}$, which gives us an additional reason why the previous literature has focused on the latter. From the example in the introduction, I know that welfare can be improved when one extends the set of feasible mechanisms to $\mathcal{M}$, allowing for the ex-ante restriction to non-convex subsets. To describe these cases, I define:

Definition 3.12. An alternative $k \in \mathcal{K}$ is a bad compromise if it is not in the image of any utilitarian mechanism.

Next, I will give a quantitative analysis of the welfare impact of bad compromises.

3.3.1 Welfare Guarantees for Surjective Mechanisms

Given surjectivity's natural appeal, one might be tempted to disregard the welfare loss to allow for a mechanism that gives more choices to voters. To inform this discussion, I study the total welfare fraction at stake.

To make this analysis, I define it as

$$WELF(\Psi) = \max_{g \in \mathcal{M}} W(g, \Psi)$$

utilitarian welfare and as

$$SWELF(\Psi) = \max_{g \in \mathcal{M}_{sur}} W(g, \Psi)$$

welfare under the best surjective mechanism. I capture worst-case considerations with the following definition.

Definition 3.13. The *Guaranteed Fraction of Optimal Welfare for K alternatives (GFOW(K))* describes the minimal relative welfare achieved by the best surjective mechanisms relative to utilitarian welfare, i.e.:

$$GFOW(K) = \inf_{\Psi} \frac{SWELF(\Psi)}{WELF(\Psi)}$$

.

In the following theorem, I give a tight welfare bound:

Theorem 3.14 (Welfare Bound). *The maximal welfare guarantee is* $GFOW(K) = \frac{1}{\lfloor \frac{K+1}{2} \rfloor}$*. In particular, no positive fraction can be guaranteed independent of the number of alternatives.*

Proof. The proof can be found in the Appendix. □

The ability to exclude an alternative ex-ante ensures that compromises do not prevent an optimal decision between alternatives that are not direct neighbors. In the extreme, welfare only depends on selecting the best odd alternative, while all even alternatives are bad compromises. Excluding compromises allows one always to select the welfare optimal odd alternative, while surjective mechanisms are constant mechanisms that only allow the right alternative to be chosen in one case. $\lfloor \frac{K+1}{2} \rfloor$ is the count of odd numbered alternatives. Hence the worst case is if all odd-numbered alternatives contribute equally to total welfare.

3.4 Concluding Remarks

In this research, I studied utilitarian welfare within the framework of a traditional social choice problem with single-peaked preferences. My exploration primarily revolved around assessing the implications of enforcing surjective mechanisms that do not exclude any outcome in advance. Two key findings emerged from this study. Firstly, it is never advantageous to preclude the selection of an extreme option. However, welfare might improve upon the exclusion of a compromise. Secondly, the extent of welfare achievable via surjective mechanisms diminishes with an increasing number of alternatives and is not bound away from zero.

Appendix 3.A Proofs

Proof of Theorem 3.5. The idea of the proof is to use codomain restrictions to "surjectify" anonymous and DIC mechanisms and then use the Moulin-Barberà-Jackson characterization on them.

For this fix, any anonymous and DIC mechanism g. I can now define $\hat{g}$ s.t. $g(\theta) = \hat{g}(\theta)\ \forall \theta \in \Theta$, with $\hat{g}$ corestricted to g's image, i.e. $\hat{g} : \Theta \to g(\Theta)$. Then $\hat{g}$ is surjective by construction. Further, the utility vectors in Θ induce the full domain of single-peaked preference profiles over $g(\Theta)$ due to our non-indifference condition. Hence any mechanism $\hat{g}$ can be described as the corestricted version of a generalized median voter scheme on $g(\Theta)$ with $n-1$ phantom voters. If I now reverse the corestriction, I find that g itself is a generalized median voter scheme on $g(\Theta)$ with $n-1$ phantom voters. If I reverse the steps in the previous argument, I can conclude the second direction.

□

Proof of Theorem 3.7. The equivalence between (i) and (ii) is shown in Moulin (1980). Hence I only prove the equivalence of (ii) and (iii). For this, I argue in three steps: First, I show that any generalized median voting scheme on $\mathscr{K}$ with $n+1$ phantoms has a convex image. Then I demonstrate that it is also a generalized median voting scheme on that image with $n+1$ phantoms. Lastly, I show that two of these phantoms are superfluous and that hence the mechanism is indeed a generalized median voting scheme on its convex image with $n-1$ phantoms. Simply reversing the above steps proves the other direction.

Now suppose that g is a generalized median voting scheme on $\mathscr{K}$ with $n+1$ phantoms and let $a = (a_1, \dots, a_{n+1}) \in \mathscr{K}^{n+1}$ be the vector of these phantoms' peaks. I denote with $l(a) = \min_{i \in \{1,\dots,n+1\}} a_i$ and $r(a) = \max_{i \in \{1,\dots,n+1\}} a_i$ the location of the leftmost and rightmost phantom voter peak respectively. Now I show that $g(\Theta) = [l(a), r(a)] \cap \mathscr{K}$ and that hence the image of g is indeed convex. Note that any alternative to the left of $l(a)$ or to the right of $r(a)$ is never selected by the mechanism. This is true since a majority of all votes are on the right or left of this alternative, respectively, regardless of the reports of the agents. On the other hand, g selects any alternative within $[l(a), r(a)] \cap \mathscr{K}$ if all agents report their peak on this alternative. This is the case since, for each of these alternatives, there is either a phantom at that position or at least one at alternatives, both to the left and to the right. In either case, the alternative with the unanimous support of voters is the generalized median.

Next, I prove that g is also a generalized median voter scheme on $[l(a), r(a)] \cap \mathscr{K}$ with $n+1$ phantoms. For this, observe that reporting a type

θ_i with a peak to the left of $l(a)$ or right of $r(a)$ always implies the same outcome as reporting on $l(a)$ and $r(a)$ respectively. Further single-peaked preferences imply that for any such type, I have $t_{[l(a),r(a)]\cap\mathscr{K}}(\theta_i) = l(a)$ or $t_{[l(a),r(a)]\cap\mathscr{K}}(\theta_i) = r(a)$ respectively.

Lastly, g is also a generalized median voter scheme on $[l(a), r(a)] \cap \mathscr{K}$ with $n-1$ phantoms. To see this, imagine deleting one phantom from each $l(a)$ and $r(a)$. Whichever alternative has been the generalized median on $[l(a), r(a)] \cap \mathscr{K}$ before remains the generalized median. □

Proof of Theorem 3.10. The proof is constructive: Fix any mechanism g and suppose that without loss of generality $1 \notin g(\Theta)$. By theorem 3.5, g is a generalized median voter scheme on its image $g(\Theta)$ characterized by a vector $(a_j)_{j\in\{1,\dots,n-1\}}$ of phantom voter peaks. Define now another mechanism $\hat{g}$ as a generalized median voter scheme on $g(\Theta) \cup 1$ with the same phantom voters peaks $(a_j)_{j\in\{1,\dots,n-1\}}$.

Moving from g to $\hat{g}$ counts reports as peaks at 1 instead of the left-most alternative in g's image. However, this change in peaks can change the position of the median peak of the $2n-1$ real and phantom voters in one case. It is the situation where all n actual voters get reattributed to 1, and hence in this case, $\hat{g}$ selects the unanimously preferred alternative.

□

Proof of Corollary 3.11. First, note that any mechanism in $\mathscr{M}_{con} \setminus \mathscr{M}_{sur}$ does not contain 1 and K in its image. Suppose the corollary is false. Then there would exist a mechanism whose image does not contain 1 or K and that generates higher welfare than any mechanism that does. However, that would be a direct contradiction to theorem 3.7.

□

Proof of Theorem 3.14. First I show $GFOW(K) \leq \dfrac{1}{\left\lfloor \frac{K+1}{2} \right\rfloor}$.

Denote with $K_{odd} = \{k \in \mathscr{K} : k\ odd\}$ the set of odd alternatives. Note that $|K_{odd}| = \left\lfloor \frac{K+1}{2} \right\rfloor$ is the number of odd numbered alternatives.

For each element of $o \in K_{odd}$ we now consider a state of the world $\theta_o^* \in \Theta$ with the following properties:

i In state θ_o^* aggregate welfare for alternative o is 1, while it is less then $\varepsilon > 0$ for all different alternatives. This can be done with just a single agent listing o as his top choice.

ii In state o $r_o \leq n$ agents have their peak on alternative $o+1$ and $r_{(o)}$ is a strictly decreasing sequence. Similarly, in state o $n - r_o - 1$, agents have their peak on alternative $o-1$. Naturally $r_1 = n-1$ and $r_K = 0$, for K odd and $r_{K-1} = 0$ for K even. We assume that all but agents lists o at least as their second choice.

Consider Ψ_{odd} a uniform measure over such states. Every mechanism with image K_{odd} achieves the expected value of 1 since, in each state, the alternative with aggregate welfare of 1 has unanimous support on that image.

In contrast, consider mechanisms with the full image. For alternative o to be selected in state θ_o^*, mechanisms need to have an exact difference of phantom voters strictly to the right versus strictly to the left of alternative o. Furthermore, this difference is strictly decreasing. Hence each surjective mechanism can satisfy this criterion at most for a unique o. Since we have chosen the bound ε arbitrarily, this yields the desired result.

For the reverse direction, we need to focus on the set of disconnected subsets. For a set $T \subset \mathcal{K}$ a subset $S \subset T$ is disconnected if $\min_{s\in S'} s - 1, \max_{s\in S'} s - 1 \notin T$. For each mechanism $g \in \mathcal{M}$ and a disconnected subset $S \subset g(\Theta)$ there exists a surjective mechanism $g_S \in \mathcal{M}_{sur}$ s.t. $g_S(\theta) = g(\theta)$ for all $\theta \in \Theta$ s.t. $g(\theta) \in S$. To construct this mechanism, take the vector of phantom voters of g and place all phantom voters to the left of $min_{s\in S'} s$ on this alternative as well and the same on the right for $\max_{s\in S'} s$. This surjective mechanism receives at least as much welfare from selecting alternatives from that subset as g. Taking the most welfare-contributing subset S^* and its corresponding mechanism g_{S^*}, g_{S^*} secures at least an equal fraction of total welfare. Noting that $|K_{odd}|$ is the maximal possible number of disconnected subsets gives the desired result.

□

References

Barberà, Salvador, and Matthew Jackson. 1994. "A characterization of strategy-proof social choice functions for economies with pure public goods." *Social Choice and Welfare* 11 (3): 241–52. [64–67]

Berga, Dolors. 1998. "Strategy-proofness and single-plateaued preferences." *Mathematical Social Sciences* 35 (2): 105–20. [65]

Black, Duncan. 1948. "On the rationale of group decision-making." *Journal of Political Economy,* 23–34. [64]

Börgers, Tilman, and Peter Postl. 2009. "Efficient compromising." *Journal of Economic Theory* 144 (5): 2057–76. [65]

Ching, Stephen. 1997. "Strategy-proofness and "median voters"." *International Journal of Game Theory* 26 (4): 473–90. [66]

Gershkov, Alex, Benny Moldovanu, and Xianwen Shi. 2017. "Optimal voting rules." *Review of Economic Studies,* 688–717. [65]

Moulin, Hervé. 1980. "On strategy-proofness and single peakedness." *Public Choice* 35 (4): 437–55. [63, 67, 72]

Rae, Douglas W. 1969. "Decision-rules and individual values in constitutional choice." *American Political Science Review* 63 (01): 40–56. [64]

Zeckhauser, Richard. 1969. "Majority rule with lotteries on alternatives." *Quarterly Journal of Economics,* 696–703. [64]