

# Strategy-proof and Efficient Mediation: An Ordinal Market Design Approach\*

Onur Kesten<sup>†</sup> Selçuk Özyurt<sup>‡</sup>

October 2021

## Abstract

Mediation, also known as assisted negotiation, is the preferred alternative dispute resolution approach that has given rise to a multi-billion-dollar industry worldwide. Online dispute resolution providers, in particular, rely heavily on mechanized e-Negotiation systems. We develop a novel ordinal framework where negotiators with conflicting preferences seek resolution over multiple issues. The mediation process is represented by a mechanism with voluntary participation. We characterize the full class of efficient, individually rational, and strategy-proof mediation protocols. A necessary and sufficient condition for the existence of such protocols is the so-called quid pro quo property that allows negotiators to compromise between issues.

**JEL Codes:** C78; D47; D74; D78; D82

**Keywords:** Market design, online dispute resolution, mediation, negotiation, e-Negotiation systems, alternative dispute resolution, logrolling, quid pro quo, outside options, strategy-proofness, efficiency.

## 1 INTRODUCTION

The classic bestseller book *Getting to Yes*, by Roger Fisher and William Ury, is one of the most famous references on the topic of negotiation. The authors identified conflict as a growth industry, and the last few decades have proved them right. Judicial systems of developing and emerging economies are often challenged with a large backlog of cases, and efficiency concerns fueled implementation of reforms that have focused on increased usage of alternative dispute resolution (ADR) processes.<sup>1</sup> Mediation is often the preferred form

---

\*We thank Johannes Hörner, Mehmet Ekmekçi, Kalyan Chatterjee, Hülya Eraslan, Paola Manzini, Utku Ünver, Tayfun Sönmez, Rakesh Vohra, George Mailath, Vijay Krishna, Hervé Moulin, Larry Ausubel, Michael Ostrovsky, Ed Green, Ron Siegel, Luca Rigotti, Sevgi Yüksel, Alexey Kushnir, Alex Teytelboym, William Thomson, Peter Troyan, Charlie Holt, Ruben Juarez, Francis Bloch, Leeat Yariv, Laura Doval, Piotr Dworzak, Nicholas Yannellis, and all seminar participants at Stanford (NBER Market Design Workshop), NYU, Rice, Maryland, Boston College, University of Pittsburgh, Carnegie Mellon, Penn State (PETCO), Virginia, Durham, Sydney, University of Technology Sydney, Bilkent, ITU, Paris Dauphine, York, Dalhousie, Sussex, and SAET for many useful discussions and suggestions. Selçuk Özyurt especially thanks M. Remzi Sanver, David Pearce, Eric Maskin, Harvard University Department of Economics, Tepper School of Business, and Sabancı University Economics Group for their mentoring, support and hospitality during this project, and the European Commission for the funding from European Union Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 659780.

<sup>†</sup>(Corresponding author) University of Sydney, School of Economics (onur.kesten@sydney.edu.au)

<sup>‡</sup>York University, Department of Economics (ozyurt@yorku.ca)

<sup>1</sup>See, for example, Ali (2018) for an extensive account of the recent mediation reforms.

of ADR due to its cost-effectiveness (timewise and financially),<sup>2</sup> flexibility,<sup>3</sup> and confidentiality. Since 1990s a significant number of countries have implemented both mandatory and voluntary mediation programs to improve the efficacy of their legal systems.<sup>4</sup>

Mediation is a consensual negotiation process in which a neutral third party (i.e., mediator) assists disagreeing parties to identify the underlying interests, issues, and solutions, and help them reach an agreement short of litigation. Notwithstanding the practical conveniences it affords, the mediation process is often considered less formal and less transparent than binding adjudication processes such as litigation and arbitration. Traditional legal theorists argue<sup>5</sup> that the low visibility and lack of formal rules and structure in traditional mediation reduce the rights of less powerful participants. In a seminal work, LaFree and Rack (1996) provide empirical evidence from the small claims court mediation program in Bernalillo County in Albuquerque, New Mexico, and conclude that ethnicity and gender could be more important determinants in informal mediation than they are in adjudication. In particular, they report that white males receive significantly more favorable outcomes in mediation than minority females.<sup>6</sup>

A structured and rigorous view of mediation is pioneered in online dispute resolution (ODR) that often rely on automation. ODR systems resolve disputes that arise both online and off-line. In a standard ODR system, parties interact through an online platform and the mediator is usually a patented software, also known as e-Negotiation system (ENS), that follows predetermined sets of rules embodied in a mechanized algorithm. During the Internet “bubble” of 1999-2000, many ODR start-ups appeared and then disappeared, but since then interest in ODR has grown and its focus has expanded (Wahab et al., 2012). Over 134 ODR platforms currently operate worldwide,<sup>7</sup> while SquareTrade, Cybersettle and SmartSettle are the oldest and probably the most famous ones. It is estimated that e-commerce platforms like eBay, Paypal, Uber and Amazon resolved more than a billion disputes in 2017 through their ODR systems (Habuka and Rule, 2017). Since its founding, Cybersettle handled over 200,000 claims combined value in excess of \$1.6B, and the City of New York uses the system since 2004 to speed their settlement process for a backlog of

---

<sup>2</sup>According to Hadfield (2000), it costs a minimum of \$100,000 to litigate a straightforward business claim in the US, whereas a mediation session varies from few hours to a day and even the most reputable mediators charge around \$10,000 - \$15,000 for a day. Also, disputants do not pay any fees for experts, witnesses, document preparation, investigation, or paralegal services, which easily make the costs pile up.

<sup>3</sup>It is impossible to discuss a legally “irrelevant” issue in litigation/arbitration. In mediation, however, parties can discuss and negotiate issues that are not directly linked to the case.

<sup>4</sup>Several provinces of Canada, most notably Ontario, refer civil actions, which are subject to case management, to mandatory mediation. The mediation is conducted by a private-sector mediator and the disputants are responsible with the corresponding fees: <https://www.attorneygeneral.jus.gov.on.ca/english/courts/manmed/notice.php>. In the US, 63 federal district courts authorized the required use of mediation, out of which 12 courts mandated the use of mediation for some or all civil cases. In UK, the Small Claims Mediation Scheme is funded by HMCS (Her Majesty’s Courts Service) and provides a free service for small claims cases operating in all court centers. If the parties’ claim does not exceed £10,000 and agree to mediate, then a phone-based or face-to-face mediation session is arranged. In Singapore, Australia, Italy, and India court-annexed mediation takes place in the courts after parties have commenced legal proceedings, and serves as the primary method of civil dispute resolution. See Ali (2018) for an extended discussion on mandatory mediation practices in the US, UK, and aforementioned other countries.

<sup>5</sup>See, for example, Damaska (1975).

<sup>6</sup>In a similar vein, many others emphasize the factors that can cause disputant dissatisfaction that are under the direct control of mediators. As a remedy, Tyler and Huo (2002) advocate the use of fair procedures that are described as those in which decisions are viewed as *neutral, objective, and consistent*.

<sup>7</sup>See <http://odr.info/provider-list/> (last visit August 12, 2021).

40,000 personal injury claims.<sup>8</sup> Government use of ODR promises to be a very large market as well (Wahab et al., 2012). Government agencies, such as the National Mediation Board<sup>9</sup> and the Office of Government Information Services<sup>10</sup> in the United States, are adopting and promoting ODR as an effective method of resolving problems with citizens. In the US and Canada, 27 courts either partially or fully integrated ODR into their systems.<sup>11</sup> Rapid technological developments and worldwide changes brought by Covid-19 pandemic have shown that ODR is the inevitable future of dispute resolution in the new millennium.<sup>12</sup>

In principle, ODR systems are ideal platforms to deliver impartial, consistent, and fair outcomes since the human factor (i.e., the mediator) is taken out of the equation and replaced with a set of reliable and objective rules and procedures. The task of automating a negotiation process is not a simple one, and this is evidenced by a myriad of systems (mostly still research efforts) around the world (Thiessen et al. 2012). However, these systems are vulnerable to strategic manipulation, and negotiators face a daunting task of finding optimal strategies. This weakness is acknowledged by the experts in the field:

*“A concern with the use of ENS is the possible effects of gaming and cheating. By supplying false information concerning the range of issues over which they are willing to negotiate, the results will be distorted.”* Thiessen et al. (1998).

These distortions may cause severe inefficiencies: Negotiators may fail to achieve the best possible solution, although they successfully reach a resolution. Thiessen et al. (1998), the founders of the popular SmartSettle algorithms, defend the current systems in this account as follows:

*“It is not clear from various experiments carried out that these distortions will always be to the benefit of the cheater. It may turn out, however, that if everyone cheats, the alternatives ENS generates may be acceptable and therefore useful in the negotiation process even though they may not be truly equivalent or efficient.”*

ODR systems would help parties resolve disputes that are otherwise doomed to fail, and this may be a significant efficiency gain. Nevertheless, we find the view in the quote above rather optimistic. Systems that are prone to gaming may produce systematically unfair outcomes. Infrequent users (e.g., customers in e-commerce disputes, individual plaintiffs against companies and government agencies) may be hurt under current systems when they face experienced users who have accumulated enough expertise about how to game the system. Although the design of *fair* and *efficient* e-Negotiation systems is an active research area within the management information systems literature, existing frameworks do not take incentive considerations into account. We are inspired by the structured mediation programs that are offered by the ODR systems, and follow a market design approach to

<sup>8</sup>Online Dispute Resolution Advisory Group 2015 report: <https://www.judiciary.uk/wp-content/uploads/2015/02/Online-Dispute-Resolution-Final-Web-Version1.pdf> (last visit August 12, 2021).

<sup>9</sup>See [https://nmb.gov/NMB\\_Application/](https://nmb.gov/NMB_Application/) (Last visit August 14, 2021).

<sup>10</sup>See <https://www.archives.gov/ogis> (Last visit August 14, 2021).

<sup>11</sup>See <http://odr.info/courts-using-odr/> (last visit August 12, 2021).

<sup>12</sup>One of the pioneers and the first director of the ODR systems at eBay and PayPal, Colin Rule, famously wrote (see Rule (2014)) “*Now that society has embraced technology so thoroughly, the key question for dispute resolution professionals is, how can we leverage technology to best assist parties in resolving their disputes? Online Dispute Resolution is no longer a novelty—it is now arguably the future of Alternative Dispute Resolution.*”

develop a tractable framework in search for efficient, incentive compatible, and impartial mediation (recommendation) mechanisms.

Mechanism design has been successful in many applications, most notably in market design for auctions and matching. Unlike the traditional mechanism design approach to bargaining (i.e., Myerson and Satterthwaite (1983)) we adopt an *ordinal* approach in the context of mediation for three reasons. First, rather than restricting players' preferences to a specific transferable utility setting, we maintain a basic common implication of any monotonic preferences in a conflict situation. In doing so, we characterize all classes of preferences that would support a possibility result and thereby allow for both transferable and nontransferable utility.<sup>13</sup> Second, it is genuinely simple to implement ordinal mechanisms, which is particularly important when agents are boundedly rational.<sup>14</sup> Third, the ordinal approach together with dominant strategy implementation makes it possible to avoid the famous Wilson critique by providing “detail-freeness” and “robust incentives” to participants.<sup>15</sup>

The backbone of our formal setting consists of two negotiators that are in a dispute over an issue  $X$ . Each negotiator has a commonly known ranking over the discrete set of solutions (alternatives) that are available for issue  $X$ , and the negotiators' rankings are diametrically opposed. In keeping with the practice of ODR, we assume that negotiators have private bargaining ranges for issue  $X$ . In other words, negotiators come to the mediation table with a privately known “least acceptable outcome.” We capture such circumstances by assuming that each negotiator has an *outside option*, what is referred as the BATNA (Best Alternative to a Negotiated Agreement) in the field.<sup>16</sup> How a negotiator ranks her outside option is her private information. Depending on the type realizations, the private bargaining ranges may not overlap and a mutually acceptable alternative need not always exist. For this reason, we say that issue  $X$  has *uncertain gains from mediation*. A mediation rule is a systematic way of choosing an outcome for any reported pair of types of the two negotiators. By using a standard revelation principle, we denote the mediation mechanism by the mediation rule whose outcome may be vetoed by either negotiator, in which case the mediation fails and both negotiators receive the outside option.

In line with ODR practices, we constrain our attention to mechanisms that never suggest an alternative falling outside of the declared bargaining ranges. This requirement corresponds to individual rationality of the mediation mechanism. We easily conclude that it is impossible to find a “good” mediation mechanism in a single-issue dispute, which is *Pareto efficient*, *individually rational*, and *strategy-proof* (i.e., no negotiator ever gains from a misreport). This conclusion should not be surprising. For mediation practitioners, single

---

<sup>13</sup>Although money is an important issue in disputes, it is rarely the only issue (Malhotra and Bazerman, 2008). This underscores the necessity of a setting that can also admit nontransferable utility specifications.

<sup>14</sup>There is a large body of experimental evidence that finds that the representation of preferences by VNM utility functions may be inadequate; see, for example, Kagel and Roth (2016). This literature argues that the formulation of rational preferences over lotteries is a complex process that most agents prefer not to engage in if they can avoid it.

<sup>15</sup>While stressing the powerful insights that mechanism design offers in bargaining problems, Ausubel et al. (2002) voice a similar concern: “... *Despite these virtues, mechanism design has two weaknesses. First, the mechanisms depend in complex ways on the traders' beliefs and utility functions, which are assumed to be common knowledge. Second, it allows too much commitment. In practice, bargainers use simple trading rules—such as a sequence of offers and counteroffers—that do not depend on beliefs or utility functions.*”

<sup>16</sup>See, for example, Fisher and Ury (1981).

issue disputes are often bound to fail because they are extremely competitive; parties see them as win-lose contests. From a mechanism design perspective, this negative result is also not surprising in light of the well-known impossibility of Myerson and Satterthwaite (1983) that there is no incentive compatible and efficient bilateral trading mechanism.<sup>17</sup>

According to negotiation experts, success in mediation lies in parties' ability of expanding the pie and finding integrative (win-win) outcomes, which necessitates the idea of multi-issue negotiation and logrolling (Malhotra and Bazerman, 2008). From a technical perspective, relaxing the tension between efficiency, individual rationality and strategy-proofness by introducing an additional issue makes perfect sense as negotiators could then be asked to consider concessions in one issue for a favorable treatment (i.e., compensation) in the other so they will be disincentivized from gaming the system. However, we show that the impossibility continues if each additional issue also exhibits uncertain gains from mediation. This is simply because it is impossible to guarantee that the new issue always contains mutually acceptable alternatives that can be used for compensation. This case and the corresponding impossibility result is deferred to Section 6.

Naturally, we then study a framework with two issues,  $X$  and  $Y$ , where issue  $Y$  exhibits *certain gains from mediation*: Namely, it is common knowledge that a set of mutually acceptable alternatives exist for issue  $Y$ . As is the case for issue  $X$ , negotiators' preferences over the alternatives for issue  $Y$ , except the outside option, are diametrically opposed. Therefore, types of each negotiator differ only by their rankings of the outside option in issue  $X$ . The mediation mechanism maps the negotiators' private information to a recommendation bundle (i.e., a solution for each issue) that never includes an unacceptable alternative. Efficiency, individual rationality and strategy-proofness of the mediation mechanism necessitate a certain discipline on negotiators' preferences over bundles, and a key assumption for this is *monotonicity*; bundles with better alternatives are always better.

Our first main result is a complete characterization of the class of strategy-proof, efficient, and individually rational mediation mechanisms (Theorem 1). These mechanisms operate through an exogenously specified *precedence order* (i.e., sequential hierarchy) over a special set of bundles, which we call *logrolling bundles*. As the precedence order varies, the characterized class of mechanisms span what we refer to as the *family of logrolling mechanisms*. A visual characterization of this family demonstrates that a mediation mechanism belongs to the family if and only if its matrix representation can be partitioned into rectangular regions (Theorem 3). The visual characterization simplifies the mechanics of the logrolling mechanisms and transform them into easy-to-read diagrams that can be sequentially implemented as a menu of offers. This should make mediation more accessible and comprehensible.

On the contrary, algorithms of current ODR platforms are unknown to users (see Section 1.1 for a brief overview). Some platforms (e.g., Cybersettle) choose not to disclose their algorithms because of patent infringement concerns. Others (e.g., SmartSettle) also adopt this "black box" approach because their algorithms involve sophisticated and unintuitive integer optimization techniques. Designers of SmartSettle admit that their multi-issue e-

---

<sup>17</sup>For this reason, we briefly discuss and illustrate the impossibility in the single-issue case with the help of a simple example in Section 2.

Negotiation system cannot be used to its full potential by novices without the assistance of a facilitator (Lodder and Thiessen, 2003). This may pose a serious concern from an economic design perspective because it is widely acknowledged in the literature that the simplicity and the transparency of the underlying mechanics of a mechanism is crucial for its efficacy.<sup>18</sup>

Our second main result (Theorem 2) is a complete characterization of the class of preference domains that admit strategy-proof, efficient, and individually rational mediation mechanisms. The necessary and sufficient condition is the so-called *quid pro quo* property. This property imposes a form of substitutability between issues  $X$  and  $Y$ .<sup>19</sup> It entails that issue  $Y$  is rich enough so that a negotiator is able to make concessions in issue  $X$  for a more preferred alternative in issue  $Y$ ; e.g., for any pair of (acceptable) alternatives  $x$  and  $x'$  of  $X$ , there exists a corresponding pair of alternatives  $y$  and  $y'$  of  $Y$  such that when bundled together,  $(x', y')$  is preferred over  $(x, y)$ , although  $x$  is preferred over  $x'$ . Such reversals in the preference domain should induce a partial order and a semilattice structure on issue  $X$ . An important takeaway from Theorem 2 is that not all multi-issue disputes admit good mechanisms, and in this sense, not all cases are *solvable*, despite the best efforts of the designers. It all boils down to the negotiators' underlying interests and substitutability of the issues.<sup>20</sup> Quid pro quo constitutes the limits of solvable disputes. If we map our discrete model into a classic exchange economy, where alternatives represent quantities of goods, then quid pro quo is compatible with well known utility functions, such as CES and quasi-linear. In such an environment, quid pro quo hinges on the availability of a set of (logrolling) bundles and certain restrictions on the rates of substitution at these bundles.

Finally, we introduce an ordinal fairness criterion that is useful in judging impartiality of the mediation processes. The family of logrolling mechanisms nests interesting special members. When the precedence order is in line with the preference of a given negotiator over the logrolling bundles, we obtain the corresponding *negotiator-optimal mechanism*. A negotiator-optimal mechanism represents situations when a mediator may be categorically biased toward one party in the dispute. It turns out there is a central member of the family of strategy-proof, efficient, and individually rational mediation mechanisms that satisfies our fairness criterion (Theorem 4). This is the so-called *constrained shortlisting* mechanism, which recommends the *median logrolling bundle* when it is mutually acceptable, and when it is not, favors the least-accepting negotiator.

## 1.1 ONLINE DISPUTE RESOLUTION AND E-NEGOTIATION SYSTEMS

Before presenting our model and theoretical results, we provide a brief overview of the fundamental aspects of e-Negotiation protocols. The essential features of these protocols

<sup>18</sup>To this end, Li (2017) introduces the notion of obvious strategy-proofness which allows to distinguish even among those mechanisms that are strategy-proof. See also Pycia and Troyan (2019) and Pycia and Ünver (2020).

<sup>19</sup>It can be viewed as the nontransferable utility analogue of the *possibility of compensation* assumption in a transferable utility model, see, e.g., Thomson (2016).

<sup>20</sup>Consider, for example, a scenario where alternatives of issue  $Y$  have little appeal for the negotiators compared to those in issue  $X$  (e.g., preferences are lexicographic over the two issues). Then there is little reason to suspect that the impossibility in the single-issue case will be overturned in the two-issue world. In fact, quid pro quo property fails to hold in such preference domains.

will also motivate some of our modeling choices.<sup>21</sup> ODR protocols usually take issues and possible alternatives in every single issue as given. This information is often solicited when parties describe the dispute when they first request the service of the ODR platform. The negotiation problem is then created by populating this information on the system. Aside from this preliminary step, protocols generally involve three common steps: (1) *elicitation*, (2) *proposal*, and (3) *ratification*. The process ends either when parties unanimously accept a proposal, or if no mutual agreement is reached after several iterations of an “*elicitation-proposal-ratification*” cycle.

The goal of the elicitation step is to obtain parties’ private information. First, parties are asked about their bargaining ranges for each issue, with the understanding that the mediator’s proposal will never include an alternative outside this range. Negotiators’ bargaining ranges are elicited by asking each negotiator to choose an alternative that is least acceptable for her. Given the negotiators’ positions (i.e., how negotiators rank the alternatives) the system infers the set of acceptable alternatives once they declare their bargaining ranges. ODR platforms commonly make the implicit assumption that negotiators’ ordinal rankings over alternatives are monotonic in the sense that an alternative that delivers more is always better. This automatically implies that negotiators’ rankings of the available alternatives in a given issue are diametrically opposed.<sup>22</sup> Depending on the ODR platform, parties may be further asked to report their preferences over issues to indicate how they trade off one issue against another. SmartSettle, for example, elicits cardinal preferences in the form of utility (satisfaction) points over issues and alternatives. Namely, each negotiator is asked to “bid” a point value (between 0 and 100) for each alternative in each issue.

Protocols usually differ in how they process all this information to make recommendations, and majority of the existing e-Negotiation protocols use a combination of cooperative game-theoretic and optimization techniques. In all existing protocols, a common theme is that any aspect of a user’s input (e.g., bargaining ranges) may be modified at any time during the negotiation before a proposal is accepted by both negotiators. To illustrate we discuss three examples.

The protocols used by the popular SmartSettle system are based on optimization algorithms that use mixed-integer programming techniques.<sup>23</sup> The system categorizes solution packages (bundles of alternatives) unacceptable if they include alternatives that are outside of the declared bargaining ranges or utility points fall short of the minimum scores privately declared by the disputants. The system never recommends these bundles.

A second well-known example is the Family Winner by Bellucci and Zeleznikow (2005), which is based on cooperative game-theory. The protocol was first developed to model Australian family law based on the repository of cases in the Australian Institute of Family Studies. It was subsequently applied to international disputes, enterprise bargaining, and company mergers. The algorithm is a point allocation procedure that aims to distribute items to the negotiators on the premise of who values the issue the most. At the outset the

---

<sup>21</sup>This overview is particularly inspired by SmartSettle (Thiessen and Loucks, 1992), which is successfully commercialized in multi-issue dispute resolution and offers a great deal of publicly available information regarding its protocols.

<sup>22</sup>See Section 2 for an illustration of this implication, which also forms the basis of our model.

<sup>23</sup>Although the complete details of the algorithm is not publicly available, a subset of the linear equations are provided in Thiessen and Loucks (1992).

negotiators are required to distribute 100 points across the range of issues. The algorithm first identifies the issue that the disputants are furthest apart and allocates the item in this issue to the party who values it the most. Then it finds the next issue where the disputants are furthest apart and allocates the item in that issue to the party who values it the most, and so on.

A third example is SquareTrade, a platform that has been eBay’s contractor on dispute resolution and the leading ODR provider for consumer mediation since 1999.<sup>24</sup> The main difference of the SquareTrade from the previous two examples is that it does not involve any preference elicitation. Specifically, the SquareTrade dispute resolution process consists of two stages. In first stage it presents the claimant a list of possible solutions (alternatives or bundles) and asks her to select the ones that she finds acceptable. Upon agreeing to participate in the process, the other party is then asked to do the same. If at least one solution is mutually acceptable, then the process ends. Incidentally, this process is a simpler and less refined version of the central mechanism (i.e., the logrolling mechanism) we propose and characterize in this paper.<sup>25</sup> If parties cannot reach an agreement in the first stage, then the protocol allows parties to exchange visible optimistic proposals, defining the bargaining range in the second stage. The system then generates suggestions that fall into the bargaining range. Parties may continue to exchange visible proposals or contribute their own suggestions to the mixture of standing proposals. The process terminates when all parties accept at least one standing proposal.

A strand of market design literature that can offer valuable insights in the context of ODR includes the recent works on *multi-item assignment*, such as course allocation at business schools. A common allocation method in practice is a *course-bidding mechanism* where students are asked to allocate an artificial currency endowment across different courses, and courses are assigned to highest bidders. Both theoretical and experimental research have shown that such auction mechanisms can perform rather poorly due to the perverse incentives they generate.<sup>26</sup> A major insight from that context immediately carries over to dispute resolution: When a bidding mechanism is used, a disputant can find it strategically advantageous not to “waste” points on less contested issues despite having a truly high valuation of these issues. This in turn translates into strategic reports that are not representative of true preferences. Such incentive shortcomings of point-based course allocation systems played role in the recent replacement of the course-bidding mechanism at the Wharton School (University of Pennsylvania) with the Approximate Competitive Equilibrium from Equal Incomes (A-CEEI) mechanism of Budish (2011), which has better incentive properties.

## 2 AN EXAMPLE

This section is intended for readers who may not be proficient with the methodology of market/mechanism design, or those who would like to grasp the basic insights behind our results. Proficient readers may prefer to skip to the next section.

---

<sup>24</sup>SquareTrade has resolved millions of disputes across 120 countries in 5 different languages. See <https://www.worldarbitration.center/on-line-disputes/> (Last visit September 13, 2021.)

<sup>25</sup>See Section 4.4 for a detailed discussion.

<sup>26</sup>See, for example, Sönmez and Ünver (2010) and Krishna and Ünver (2008).

In the simplest possible form, consider a single issue dispute where there are only two available solutions,  $x_1$  and  $x_2$ . It is possible to think of this case as a dispute over the division of some jointly owned assets in the dissolution of a partnership, where alternative  $x$  denotes the percentage of the assets negotiator 1 gets, and so 2 gets the remaining  $100 - x$  percent. Negotiator 1 prefers alternative  $x_1$  to  $x_2$  and negotiator 2 prefers  $x_2$  to  $x_1$ .<sup>27</sup> The ranking of the *outside option*, denoted by  $o$ , is each negotiator's private information and each negotiator ranks any "unacceptable" alternative below the outside option. Hence, this private outside option determines a negotiator's *bargaining range*.<sup>28</sup> Let set  $\bar{X} = \{x_1, x_2, o\}$  denote the set of all possible outcomes of the mediation process. Each negotiator has two types, where  $L_i$  and  $M_i$  refer to the "least-accepting" and "most-accepting" types of negotiator  $i$ , respectively.<sup>29</sup>

$L_1$	$M_1$	$L_2$	$M_2$
$x_1$	$x_1$	$x_2$	$x_2$
$o$	$x_2$	$o$	$x_1$
$x_2$	$o$	$x_1$	$o$

The mediation mechanism  $f$  maps the negotiators' private information to an outcome in  $\bar{X}$ , and we use the following matrix to represent it:

	$M_2$	$L_2$
$L_1$	$f_{1,1}$	$f_{1,2}$
$M_1$	$f_{2,1}$	$f_{2,2}$

where  $f_{\ell,j} \in \bar{X}$  for all  $\ell, j \in \{1, 2\}$ .

Since a mutually acceptable alternative does not always exist (e.g., when negotiator types are  $L_1$  and  $L_2$ ), we say that issue  $X$  has *uncertain gains from mediation*. A mediation process that respects negotiators' bargaining ranges (i.e., always suggests an *individually rational* outcome) should, therefore, suggest the outside option at this type profile (i.e.,  $f_{1,2} = o$ ). If the mediation process is *efficient*, then we should have  $f_{1,1} \neq o$ . Furthermore, if the process is expected to be individually rational, then we have  $f_{1,1} = x_1$ . Likewise, an efficient and individually rational mediation process must satisfy  $f_{2,2} = x_2$  and  $f_{2,1} \in \{x_1, x_2\}$ . Therefore, there are only two (deterministic) mechanisms satisfying individual rationality and efficiency in this simple framework.

However, neither of these mechanisms is immune to strategic manipulation. Namely, it is not a dominant strategy for the negotiators to report their types truthfully. To see this point, suppose that  $f_{2,1} = x_1$ . In this case, the most-accepting type of negotiator 2 (i.e.,  $M_2$ ) would misreport her type when negotiator 1 is of type  $M_1$  because she guarantees her favorite outcome  $x_2$  by reporting  $L_2$  instead. Symmetrically, if  $f_{2,1} = x_2$ , then the most-accepting type of negotiator 1 (i.e.,  $M_1$ ) would lie about her type. We conclude, therefore, that it is impossible to find an efficient, individually rational, and strategy-proof mediation mechanism in a single-issue dispute. It is straightforward to extend this impossibility to the case with more than two alternatives.<sup>30</sup>

<sup>27</sup>Here we implicitly assume that each negotiator only cares about her own share.

<sup>28</sup>In line with the interpretation we adopted for this specific example, the outside option may be a vector indicating each negotiator's BATNA.

<sup>29</sup>We assume, without loss of generality, that there is at least one acceptable alternative for each negotiator.

<sup>30</sup>This impossibility also prevails when we allow stochastic mechanisms. In that case, the only difference in the argument would be that  $f_{2,1}$  is a lottery over  $x_1$  and  $x_2$ . However, the above deviations would still remain profitable.

Next, consider the framework with two issues,  $X$  and  $Y$ , with the set of outcomes  $\bar{X} = \{x_1, x_2, o_x\}$  and  $\bar{Y} = \{y_1, y_2, o_y\}$ , respectively. Suppose issue  $Y$  represents the time of dissolution of the partnership, where outcomes  $y_1, y_2$  and  $o_y$  denote “sooner”, “later”, and “never”, respectively. It is common knowledge that negotiator 1 prefers  $x_1$  to  $x_2$  and  $y_1$  to  $y_2$ , negotiator 2’s preferences are diametrically opposed in each issue, and both negotiators find  $y_1$  and  $y_2$  acceptable. Hence, issue  $Y$  has *certain gains from mediation*. Thus, publicly known preferences are as follows:

<i>Negotiator 1</i>		<i>Negotiator 2</i>	
$X$	$Y$	$X$	$Y$
$x_1$	$y_1$	$x_2$	$y_2$
$x_2$	$y_2$	$x_1$	$y_1$
	$o_y$		$o_y$

As before, types of each negotiator differ only by their rankings of the outside option in issue  $X$ , so each negotiator has two types. Now the mediation mechanism maps the negotiators’ private information to a bundle  $(x, y)$  in  $\bar{X} \times \bar{Y}$ . We maintain individual rationality (i.e., mechanism  $f$  should never offer a bundle that includes an unacceptable alternative). This implies that  $f_{1,2} = (o_x, y)$ , where  $y$  is either  $y_1$  or  $y_2$ . We now also assume that negotiators’ preferences over bundles are *monotonic* (i.e., bundles with (weakly) better alternatives are always more preferred).

Under these assumptions, by the efficiency, individual rationality and strategy-proofness of  $f$ , we must have  $f_{1,1} = (x_1, y_2)$  and  $f_{2,2} = (x_2, y_1)$ . To see this, consider  $f_{1,1}$ . An efficient mechanism that respects the declared bargaining ranges should suggest a bundle with  $x_1$  at this entry because only  $x_1$  is declared mutually acceptable in issue  $X$ . If  $f_{1,1}$  is  $(x_1, y_1)$ , then monotonicity and strategy-proofness will necessitate that both  $f_{2,1}$  and  $f_{2,2}$  are  $(x_1, y_1)$ . However,  $f_{2,2}$  corresponds to the profile where  $x_1$  is declared unacceptable by negotiator 2. Thus, the mechanism must bundle  $x_1$  either with  $y_2$  or  $o_y$ . Monotonicity and efficiency eliminate the latter possibility. Hence,  $f_{1,1}$  must be  $(x_1, y_2)$ , and symmetrically  $f_{2,2} = (x_2, y_1)$ . A similar logic implies that  $f_{2,1}$  must be either  $(x_1, y_2)$  or  $(x_2, y_1)$ . We call these two bundles *logrolling bundles* since they pair a higher ranked alternative in one issue with a (relatively) lower ranked alternative from the other.

A good mechanism  $f$  will always suggest a logrolling bundle whenever a mutually acceptable alternative exists. Thus, strategy-proofness requires that if  $f_{2,1} = (x_1, y_2)$ , then  $M_2$  should prefer  $(x_1, y_2)$  over  $(x_2, y_1)$  so that she prefers to reveal her type truthfully, and if  $f_{2,1} = (x_2, y_1)$ , then  $M_1$  should prefer  $(x_2, y_1)$  over  $(x_1, y_2)$  so that she prefers to reveal her type truthfully. This discipline on preferences is the essence of what we call as the *quid pro quo* property, which turns out to be equivalent to a certain semilattice structure when there are more than two alternatives.

### 3 THE MAIN SETUP: MULTI-ISSUE MEDIATION

The first part of this section presents the terminology and notation adopted throughout this paper. We defer our interpretation and discussion of key assumptions to the second part.

### 3.1 PRELIMINARIES

There are two negotiators,  $N = \{1, 2\}$ . We refer to negotiators as “she” and to the mediator as “he”. Let  $\bar{X} = \{x_1, \dots, x_m, o_x\}$  and  $\bar{Y} = \{y_1, \dots, y_n, o_y\}$  denote the finite set of potential outcomes in the **main issue** and **second issue**, respectively, where  $n \geq m \geq 2$ . The sets  $X = \bar{X} \setminus \{o_x\}$  and  $Y = \bar{Y} \setminus \{o_y\}$  are the sets of available **alternatives**. We denote the **outside options** in issue  $X$  and  $Y$  by  $o_x$  and  $o_y$ , respectively. For any issue  $Z \in \{\bar{X}, \bar{Y}\}$ , let  $\Theta_i^Z$  denote the set of all linear orders (i.e., preference relations) of negotiator  $i$  over issue  $Z$ , and  $\theta_i^Z$  denote an ordinary element of the set  $\Theta_i^Z$ .<sup>31</sup>

It is publicly known that for all  $k = 1, \dots, |Z|$  and all  $Z \in \{X, Y\}$ ,  $z_k \theta_1^Z z_{k+1}$  for negotiator 1,  $z_{k+1} \theta_2^Z z_k$  for negotiator 2, and  $y_k \theta_i^Y o_y$  for all  $i \in N$ . Namely, the negotiators’ preferences over the alternatives are diametrically opposed in each issue, and any alternative in issue  $Y$  is acceptable for both negotiators (in which sense  $Y$  exhibits *certain gains from mediation*). The ranking of the outside option in issue  $X$ ,  $o_x$ , is the negotiators’ private information, and so it exhibits *uncertain gains from mediation*. Therefore,  $\Theta_i = \Theta_i^X$  denotes the set of all **types** of negotiator  $i$  and  $\Theta = \Theta_1 \times \Theta_2$  is the set of all type profiles. Without loss of generality, we ignore those types that declare all alternatives of  $X$  unacceptable.<sup>32</sup> We use  $\theta_i$  for the rest of the paper instead of  $\theta_i^X$  to indicate negotiator  $i$ ’s preferences over the outcomes in issue  $X$ . Whenever we need to distinguish  $i$ ’s preferences over issue  $X$  and  $Y$ , we then use  $\theta_i^X$  and  $\theta_i^Y$ , respectively.

For any negotiator  $i$  and type  $\theta_i \in \Theta_i$ , let  $A(\theta_i) = \{x \in X \mid x \theta_i o_x\}$  denote  $i$ ’s **bargaining range** (i.e., set of **acceptable** alternatives) in issue  $X$ . For any type profile  $(\theta_1, \theta_2) \in \Theta$ , let the set  $A(\theta_1, \theta_2) = \{x \in X \mid x \theta_i o_x \text{ for all } i \in N\}$  denote the set of all **mutually acceptable** alternatives in issue  $X$ . In case we need to specify the **bargaining range of type**  $\theta_i$  of player  $i$ , where alternative  $x \in X$  is her least acceptable alternative, we use  $\theta_i^x \in \Theta_i$ . Namely, for any  $x' \in X$ ,  $x \theta_i^x x'$  implies  $o_x \theta_i^x x'$ .

A mediation mechanism  $f : \Theta \rightarrow \bar{X} \times \bar{Y}$  asks each negotiator to report her type and proposes a **bundle**  $(x, y)$  that specifies an outcome for each issue. For convenience, a mediation mechanism  $f$  can be equivalently represented by an  $m \times m$  matrix  $f = [f_{\ell,j}]_{(\ell,j) \in M^2}$ , where  $f_{\ell,j} = f(\theta_1^{x_\ell}, \theta_2^{x_j})$  and  $M = \{1, \dots, m\}$ . The rows of this matrix correspond to types of negotiator 1 and the columns to types of negotiator 2. The row (respectively, column)  $\ell$  indicates the type of negotiator 1 (respectively, 2) that finds all alternatives  $\{x_k \in X \mid k \leq \ell\}$  (respectively,  $\{x_k \in X \mid k \geq \ell\}$ ) acceptable: See Figure 1 for an illustration. For any reported pair of types  $(\theta_1^{x_\ell}, \theta_2^{x_j})$ , mechanism  $f$  chooses an outcome  $f_{\ell,j} \in \bar{X} \times \bar{Y}$ . We use  $f_\theta^Z$  or  $f_{\ell,j}^Z$  to denote the alternative that the mediation mechanism  $f$  offers in issue  $Z \in \{\bar{X}, \bar{Y}\}$  when type profile is  $\theta = (\theta_1^{x_\ell}, \theta_2^{x_j})$ . Therefore,  $f(\theta) = f_{\ell,j} = (f_{\ell,j}^X, f_{\ell,j}^Y) = (f_\theta^X, f_\theta^Y)$ .

The set of all bundles is denoted by  $\bar{X} \times \bar{Y}$  and  $\mathfrak{R}$  denotes the set of all linear orders over  $\bar{X} \times \bar{Y}$ . Relation  $R$  is a standard element of the set  $\mathfrak{R}$ , and for any two bundles  $b, b' \in \bar{X} \times \bar{Y}$ ,  $b R b'$  means “ $b$  is at least as good as  $b'$ .” Let  $P$  denote the strict counterpart of  $R$ .<sup>33</sup> A preference (extension) map is a correspondence  $\Lambda$  that assigns to every negotiator  $i$  and

<sup>31</sup>A binary relation  $\theta$  on set  $Z$  is called a **linear order** on  $Z$  if  $\theta$  is complete, transitive, reflexive, and antisymmetric.

<sup>32</sup>Therefore, there are  $m$  possible orderings in  $\Theta_i^X$  and a unique preference ordering in  $\Theta_i^Y$ .

<sup>33</sup>That is,  $b P b'$  if and only if  $b R b'$  but not  $b' R b$ .

type  $\theta_i \in \Theta_i$  a nonempty set  $\Lambda(\theta_i) \subseteq \mathfrak{R}$  of admissible orderings over bundles.  $\bigcup_{\theta \in \Theta} \Lambda(\theta) \subseteq \mathfrak{R}^2$  is the domain of admissible preference profiles that is restricted by the preference map  $\Lambda$ .<sup>34</sup> Let  $B(\theta_i) = \{(x, y) \in X \times Y \mid x \theta_i o_x\}$  denote the set of all **acceptable bundles** for type  $\theta_i$ . A bundle  $b$  **Pareto dominates** a bundle  $b'$  if for all  $(\theta_i, \theta_{-i}) \in \Theta$ , where  $b, b' \in B(\theta_i)$  for  $i = 1, 2$ ,  $b R_i b'$  for all  $i \in N$  and  $R_i \in \Lambda(\theta_i)$ , and  $b P_i b'$  for some  $i \in N$  and  $R_i \in \Lambda(\theta_i)$ . Let  $\Lambda(\theta_i)|_B$  denote the restriction of all admissible preference relations over  $X \times Y$  to a subset  $B \subseteq X \times Y$ .<sup>35</sup>

**Definition 1.** A preference map  $\Lambda$  is **regular** if the following hold for all  $i \in N$  and  $\theta_i \in \Theta_i$ :

i. **Monotonicity (M):** For any  $x, x' \in \bar{X}$  and  $y, y' \in \bar{Y}$  with  $(x, y) \neq (x', y')$ ,

$$(x, y) P_i (x', y') \text{ for all } R_i \in \Lambda(\theta_i) \text{ whenever } x \theta_i^X x' \text{ and } y \theta_i^Y y'.$$

ii. **Consistency (C):** For any  $\theta'_i \in \Theta_i$  with  $B(\theta_i) \subseteq B(\theta'_i)$ ,

$$\Lambda(\theta'_i)|_{B(\theta_i)} = \Lambda(\theta_i)|_{B(\theta_i)}.$$

iii. **Bargaining ranges (BR):** For any  $y \in Y$ ,  $y' \in \bar{Y}$ , and  $x, x' \in X$  with  $x \theta_i o_x \theta_i x'$ ,

$$(x, y) R_i (o_x, y') R_i (x', y) \text{ for all } R_i \in \Lambda(\theta_i).$$

The mediation mechanism  $f$  is **strategy-proof** if for all  $i \in N$  and all  $\theta_i \in \Theta_i$ ,  $f(\theta_i, \theta_{-i}) R_i f(\theta'_i, \theta_{-i})$  for all  $R_i \in \Lambda(\theta_i)$ ,  $\theta'_i \in \Theta_i$  and all  $\theta_{-i} \in \Theta_{-i}$ . It is **individually rational** if for all  $i \in N$  and all  $(\theta_i, \theta_{-i}) \in \Theta$ ,  $f(\theta_i, \theta_{-i}) R_i (o_x, o_y)$  for all  $R_i \in \Lambda(\theta_i)$ . Finally, the mediation mechanism  $f$  is **efficient** if there exists no  $(\theta_i, \theta_{-i}) \in \Theta$  and  $(x', y') \in X \times Y$  such that  $(x', y') R_i f(\theta_i, \theta_{-i})$  for all  $R_i \in \Lambda(\theta_i)$  and all  $i \in N$ , and for at least one  $i$  and  $R_i \in \Lambda(\theta_i)$ ,  $(x', y') P_i f(\theta_i, \theta_{-i})$ .

A function  $t : X \rightarrow Y$  is called **order-reversing** if for all  $x, x' \in X$  and all  $i \in N$ ,  $x \theta_i^X x' \iff t(x') \theta_i^Y t(x)$ . For any  $\ell, j$  with  $1 \leq j \leq \ell \leq m$  the nonempty subset  $X_{j\ell} = \{x_k \in X \mid j \leq k \leq \ell\}$  of  $X$  is called a **connected** subset of  $X$ . Put differently,  $X_{j\ell}$  is the set of mutually acceptable alternatives in issue  $X$  at the type profile  $(\theta_1^{x_\ell}, \theta_2^{x_j})$ . For any nonempty subset  $S$  of  $X$  and a partial order  $\succeq$  on  $X$ , let  $\mathbf{max}_S \succeq$  denote the maximal element in  $S$  with respect to  $\succeq$ .<sup>36</sup> Namely, if  $x_S^* = \mathbf{max}_S \succeq$ , then  $x_S^* \succeq x$  for all  $x \in S$ . Note that such a maximal element is not guaranteed to exist under an arbitrary (e.g., incomplete) partial order.

Finally, the tuple  $(X, \succeq)$  is called a **poset** (short for partially ordered set) if  $\succeq$  is a partial order on  $X$ . For a poset  $(X, \succeq)$ , we say that an element  $x \in X$  is an **upper bound of a subset**  $S \subseteq X$  when  $x \succeq x'$  for all  $x' \in S$ . The **least upper bound** of  $S$  is the upper bound of  $S$  that is less than or equal to every upper bound of  $S$ . Namely,  $x$  is a least upper bound of  $S$  if  $x' \succeq x$  for all upper bounds  $x'$  of  $S$ . Given a doubleton  $\{x, x'\} \subseteq X$ , let the join of  $x$  and  $x'$ , denoted by  $x \vee x'$ , be the least upper bound of the doubleton. A

<sup>34</sup>For all  $\theta = (\theta_1, \theta_2) \in \Theta$  we have  $\Lambda(\theta) = \Lambda(\theta_1) \times \Lambda(\theta_2)$ .

<sup>35</sup>The restriction of a binary relation over a set  $X \times Y$  to a subset  $B \subseteq X \times Y$  is the set of all pairs of bundles  $(b, b')$  in the relation for which  $b$  and  $b'$  are in  $B$ .

<sup>36</sup>A binary relation  $\succeq$  on set  $X$  is called a *partial order* on  $X$  if  $\succeq$  is transitive, reflexive, and antisymmetric.

poset  $(X, \supseteq)$  is called a **join semilattice** if every doubleton  $\{x, x'\} \subseteq X$  has a least upper bound in  $X$ .

### 3.2 DISCUSSION

We study disputes with two issues and finite sets of alternatives, but our results can easily be extended to cases with more than two issues (see Section 6.3) or continuum of alternatives (see the Supplementary Appendix). In keeping with the practice of mediation, we assume that each negotiator has a bargaining range for each issue, and a negotiator’s bargaining range for an issue is determined by how she ranks her outside option for that issue. The outside option in issue  $Y$  need not be the worst outcome for our results to go through. All that is needed is the availability of a sufficiently large set of alternatives in  $Y$  that are efficient and individually rational.<sup>37</sup> The assumption that preferences over alternatives in each issue are diametrically opposed is without loss of generality. Under efficiency, any dispute where preferences over alternatives are not diametrically opposed can be equivalently represented by a “reduced dispute” where “reduced preferences” are diametrically opposed (see Section 6.1).

Although we notationally distinguish between the two outside options,  $o_X$  and  $o_Y$ , our model allows for interdependence of the resolutions in these two issues. We say that two issues are **joint** if whenever the outside option in one issue is selected, the outside option in the second issue must also be selected.<sup>38</sup> Otherwise, we say that the two issues are **separate**.<sup>39</sup> Since whether issues are joint or separate makes little difference in our analysis, we assume for expositional simplicity that the issues are separate in the remainder of the paper. Nevertheless, we delegate the discussion of the resulting differences to footnotes.

Mediation would potentially be a very complicated, multistage game between the negotiators and the mediator. The mediation protocol, whatever the details may be, produces proposals for agreement that are always subject to unanimous approval by the negotiators. That is, before finalizing the protocol, each negotiator has the right to veto the proposal and exercise her outside option. A version of the revelation principle, which we prove in the Supplementary Appendix, guarantees that we can stipulate the following type of a direct mechanism without loss of generality when representing mediation.

The direct mechanism consists of two stages, an *announcement* stage and a *ratification* stage, and it is characterized by a mediation rule  $f : \Theta \rightarrow \bar{X} \times \bar{Y}$ . After being informed of her type, each negotiator  $i$  privately reports her type  $\hat{\theta}_i$  to the mediator who then proposes a bundle  $f(\hat{\theta}_1, \hat{\theta}_2) \in \bar{X} \times \bar{Y}$ . In the ratification stage, each party simultaneously and independently decides whether to accept or veto the proposed bundle. If both negotiators accept the proposed bundle, then it becomes the final outcome. If either or both negotiators veto the proposal, each party gets the outside option for both issues (i.e.,  $(o_X, o_Y)$ ). Such two-stage mechanisms will be called *direct mechanisms with veto rights*. We seek direct

<sup>37</sup>More formally, the number of efficient and mutually acceptable alternatives in  $Y$  should be no less than the cardinality of  $X$ .

<sup>38</sup>For example, in a multi-unit trade negotiation between a buyer and a seller over price and quantity, if the two sides cannot agree on price, then the quantity issue becomes obsolete, and so may be set to zero.

<sup>39</sup>For example, in a family dispute where the issues are division of assets and terms of child custody, despite failing to reach an agreement in the former issue, the two sides may have a mutual interest in a quick resolution in the latter.

mechanisms with veto rights in which truthful reporting of types at the announcement stage is a *dominant strategy equilibrium* and the mediator’s proposals are never vetoed in equilibrium. It immediately follows from the definitions that such an equilibrium exists if and only if the mediation rule  $f$  is strategy-proof and individually rational.<sup>40</sup> For the rest of the paper we denote a mediation mechanism by the mediation rule  $f$  and, with slight abuse of language, refer to both as a mediation mechanism.

In our formulation, the mediation mechanism asks each negotiator to report her bargaining range as her type rather than her full-fledged preferences over all bundles. Negotiators’ underlying preferences over bundles are then assumed to be compatible with the reported types and to satisfy certain regularity conditions. While an appealing alternative and worthy of future investigation, learning parties’ preferences over bundles together with bargaining ranges requires a two-layer information elicitation. From a both practical and theoretical standpoint, this is a much difficult task since it entails further complications and restrictions on preference reporting language.<sup>41</sup> Our formulation is in line with some of the existing ODR practices<sup>42</sup> and our objective of detail-freeness.<sup>43</sup> Our simpler approach provides an added advantage of studying the trade-offs between incentives and efficiency in dispute resolution in isolation while searching for family of preference domains that admit positive results. This is a particularly useful exercise to be able to design an effective preference reporting language.

To obtain the set of possible preferences compatible with the reported types, we invoke a preference map that satisfies three regularity conditions.<sup>44</sup> Our interpretation of the preference map is that although negotiator  $i$  knows her exact preferences over the bundles, the opponent and the mediator (or the designer) believe that  $i$ ’s preferences belong to the set  $\Lambda(\theta_i)$  conditional on  $i$ ’s type being  $\theta_i$ . In a cardinal setup, where the modeler assumes a specific utility function for each negotiator,  $\Lambda$  would generate a unique ordering for each type  $\theta_i$ . However, our setup allows for more general domain specifications. For example, if the modeler only knows that the negotiators’ preferences are consistent with expected utility theory or additively separable, then  $\Lambda$  is a multi-valued function. Our results hold for *all* “regular” preference maps.

Monotonicity is a standard requirement and simply demands that a bundle with (weakly) better outcomes in both issues is always more preferred. Consistency requires that all types of the same negotiator rank the acceptable bundles in the same way. Bargaining ranges is a requirement we adopt in keeping with the practice of mediation to respect the bargaining

<sup>40</sup>This result immediately follows from the proof of the revelation principle we provide in the Supplementary Appendix.

<sup>41</sup>Such concerns are actively debated and in need of further research (Milgrom 2011). It also is an open question in the related problem of combinatorial assignment problems (Budish et al. 2017).

<sup>42</sup>Many ODR platforms including the previously discussed SquareTrade and e-commerce platforms that are used by Amazon, eBay, Netflix, and Walmart generate a menu of recommendations without direct preference elicitation.

<sup>43</sup>Asking negotiators to report full preferences over  $m \times n$  bundles even in the simplest case with two issues is arguably impractical and cognitively complex. Such pursuit of full preference elicitation conflicts with the ease and convenience expected from the mediation process, which is also admitted by the designers of some ODR platforms (e.g., Lodder and Thiessen, 2003).

<sup>44</sup>Using preference maps to deduce complete preferences is a common tool in social choice theory pioneered by Barberà (1977) and Kelly (1977) as a way to explore the strategy-proofness of social choice correspondences. For such an analysis to be carried out, individual preferences over sets are required. A typical approach is to infer this information from individual preferences over alternatives through certain extension axioms which assign to every ordering over alternatives a list of acceptable orderings over sets.

ranges reported by the negotiators. It allows us to incorporate the negotiators' bargaining ranges into their preferences and requires that each negotiator prefers bundles that are in agreement with her bargaining range to those that violate them. Specifically, a bundle with an acceptable alternative in the main issue is always preferred over a bundle with the outside option, which in turn is always preferred over a bundle with an unacceptable alternative, regardless of the alternatives chosen for the second issue. In other words, a bundle is unacceptable to a negotiator if it includes an alternative outside her bargaining range in issue  $X$ . This requirement together with individual rationality ensures that a proper mediation mechanism never proposes bundles that are not compatible with the bargaining range declared by either negotiator.

Our definitions for strategy-proofness, efficiency, and individually rationality are standard. Strategy-proofness requires truthful revelation of one's type to be her dominant strategy regardless of her underlying preferences and the type of the opposite negotiator. Individual rationality guarantees an outcome at least as good as what each negotiator would receive if she were to walk away from mediation. Efficiency says that it should not be possible to find an alternative proposal that would make both parties better off at all possible preferences and one party strictly better off at some preference profile.

## 4 MAIN RESULTS

### 4.1 STRATEGY-PROOF MEDIATION

We start with the characterization of the necessary conditions on strategy-proof, efficient, and individually rational mediation mechanisms.

**Theorem 1.** *Suppose that the preference map  $\Lambda$  is regular and  $f$  is a strategy-proof, efficient, and individually rational mediation mechanism. Then there exists an injective and order-reversing function  $t : X \rightarrow Y$ , a partial order  $\succeq$  on  $X$ , and an alternative  $y \in Y$  such that*

$$f_{\ell,j} = f(\theta_1^{x_\ell}, \theta_2^{x_j}) = \begin{cases} (x_{X_{j\ell}}^*, t(x_{X_{j\ell}}^*)) & \text{if } j \leq \ell \\ (o_x, y) & \text{otherwise;} \end{cases}$$

where  $x_{X_{j\ell}}^* = \mathbf{max}_{X_{j\ell}} \succeq$  is well-defined.

Theorem 1 states that a desired mediation mechanism must always make selections from a special set of bundles when the set of mutually acceptable alternatives in issue  $X$  is nonempty. At these bundles, for each alternative in  $X$ , there is a corresponding distinct alternative in  $Y$  with which it must be paired, and the order-reversing property implies that a more preferred alternative from issue  $X$  must be paired with a less preferred alternative from issue  $Y$ . We interpret these bundles as representing possible ‘‘compromises’’ between the two issues. As such, we henceforth call a bundle  $(x^*, t(x^*)) \in X \times Y$  a **logrolling bundle**. For a given order-reversing function  $t$ , let  $B^t$  be the set of all the logrolling bundles. When  $n = |Y| = |X| = m$ , this set is unique and  $t(x_k) = y_{m-k+1}$ . Otherwise (i.e., when  $n > m$ ), there can be multiple such  $t$ 's, hence multiple classes of mechanisms.

The set  $B^t$  of logrolling bundles constitutes the ‘‘backbone’’ of every strategy-proof, efficient, and individually rational mechanism in the sense that the diagonal of any such

mechanism (i.e., when  $\ell = j$  or there is a unique mutually acceptable alternative) must always be comprised of these bundles. The mediator has discretion over the choice of the **precedence order**  $\triangleright$  on  $X$ . Which logrolling bundle is selected below the diagonal (i.e., when  $\ell > j$  or there are multiple mutually acceptable alternatives) depends on the chosen precedence order  $\triangleright$ . Specifically, the mediator selects the highest precedence alternative among the set of mutually acceptable alternatives in the main issue and pairs it with its corresponding alternative in the second issue according to  $t$ . Intuitively, the logrolling bundles on the diagonal “propagate” in the southwestern direction following the precedence order  $\triangleright$ . Theorem 3 provides a complementary visual characterization of these mechanisms based on this insight.

When there is no mutually acceptable alternative in issue  $X$  (i.e., when  $\ell < j$ ), the mediation mechanism always chooses a designated disagreement bundle at which the outside option in  $X$  is coupled with some efficient alternative in  $Y$ . In this case, the mediation mechanism provides only a partial resolution to the dispute because of the severity of the disagreement on issue  $X$ .

For the rest of the paper, we refer to  $f$  as a **logrolling mechanism** if it satisfies the properties described in Theorem 1, and denote it by  $f^\triangleright$ . The choice of the set of logrolling bundles together with the precedence order characterizes each mediation mechanism. Before giving a sketch of the proof of Theorem 1, we provide an example of these mechanisms.

**Example 1 (A logrolling mechanism):** Suppose the main issue  $X$  consists of five alternatives, i.e.,  $m = 5$ , and the second issue  $Y$  has at least five alternatives. Take a possible set of logrolling bundles

$$B^t = \{(x_1, t(x_1)), (x_2, t(x_2)), (x_3, t(x_3)), (x_4, t(x_4)), (x_5, t(x_5))\}$$

for some injective and order-reversing function  $t : X \rightarrow Y$ . Let us construct the logrolling mechanism  $f^\triangleright$  associated with the precedence order  $\triangleright$  where

$$\triangleright: x_5 \triangleright x_1 \triangleright x_4 \triangleright x_2 \triangleright x_3.$$

The main diagonal is filled with the members of the set of logrolling bundles,  $B^t$ , e.g., we have  $f_{1,1}^\triangleright = (x_1, t(x_1))$  in the first diagonal entry,  $f_{2,2}^\triangleright = (x_2, t(x_2))$  in the second diagonal entry, and so on. Suppose we would like to determine  $f_{3,1}^\triangleright$ . The set of mutually acceptable alternatives are  $X_{13} = \{x_1, x_2, x_3\}$ . The highest precedence alternative in this set is  $x_1$ . Thus,  $f_{3,1}^\triangleright = (x_1, t(x_1))$ . Similarly, to determine  $f_{4,2}^\triangleright$  we maximize  $\triangleright$  on  $X_{24} = \{x_2, x_3, x_4\}$ , which yields  $x_4$ . Hence,  $f_{4,2}^\triangleright = (x_4, t(x_4))$ .

Alternatively, we can start from the diagonal and let the logrolling bundles spread in the southwestern direction following  $\triangleright$ . Since alternative  $x_5$  has higher precedence than all other alternatives in  $X$ , the corresponding logrolling bundle claims all the entries to its southwest, which amounts to the set of all entries on the bottom row to the left of  $f_{5,5}^\triangleright$ . The second-highest precedence belongs to  $x_1$ ; and the corresponding logrolling bundle similarly claims all the unfilled entries to its southwest. Thus, starting from the entry  $f_{1,1}^\triangleright$  on the main diagonal, all the remaining empty entries on the first column fill up with  $(x_1, t(x_1))$ .

Finally, whenever the negotiators have no mutually acceptable alternative in  $X$ , suppose the mechanism picks the bundle  $(o_X, y)$  for some  $y \in Y$ . The following matrix shows this logrolling mechanism.

	$\theta_2^{x_1}$	$\theta_2^{x_2}$	$\theta_2^{x_3}$	$\theta_2^{x_4}$	$\theta_2^{x_5}$
$\theta_1^{x_1}$	$(x_1, t(x_1))$	$(o_X, y)$	$(o_X, y)$	$(o_X, y)$	$(o_X, y)$
$\theta_1^{x_2}$	$(x_1, t(x_1))$	$(x_2, t(x_2))$	$(o_X, y)$	$(o_X, y)$	$(o_X, y)$
$\theta_1^{x_3}$	$(x_1, t(x_1))$	$(x_2, t(x_2))$	$(x_3, t(x_3))$	$(o_X, y)$	$(o_X, y)$
$\theta_1^{x_4}$	$(x_1, t(x_1))$	$(x_4, t(x_4))$	$(x_4, t(x_4))$	$(x_4, t(x_4))$	$(o_X, y)$
$\theta_1^{x_5}$	$(x_5, t(x_5))$				

Figure 1: A standard member of the logrolling mechanisms family

**Sketch of the proof of Theorem 1:** The proof follows four main steps: (1) establishing an injective and order-reversing map  $t$  from  $X$  to  $Y$ , and thus the set  $B^t$ ; (2) proving that each entry of the lower half of the matrix  $f$  comes from the set  $B^t$ ; (3) establishing the binary relation  $\succeq$  over  $X$  that is transitive and antisymmetric; and (4) proving that each entry of the lower half of the matrix  $f$  is in fact the maximal element of a particular subset of  $X$  with respect to the partial order  $\succeq$ .

These four steps prove particular claims by utilizing the following core idea, which we call the **weak axiom of revealed precedence (WARP)**: If two distinct alternatives  $x, x'$  in issue  $X$  are mutually acceptable at some type profile and  $f$  suggests  $x$  (as part of a bundle) at that profile, then it cannot be the case that  $f$  suggests  $x'$  at another type profile where both  $x$  and  $x'$  are mutually acceptable. Therefore, whenever the set of mutually acceptable alternatives in issue  $X$  is nonempty, a strategy-proof, efficient and individually rational mediation mechanism behaves as if it is a single valued “choice mechanism” that satisfies the weak axiom of revealed preference (see Rubinstein 2012).

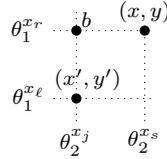


Figure 2

The intuition behind WARP is simple. Suppose it does not hold, e.g., Figure 2 indicates some entries at the lower half of the matrix  $f$ , where distinct alternatives  $x$  and  $x'$  of issue  $X$  are mutually acceptable by all the type profiles represented in this figure. Note that all three bundles, i.e.,  $b$ ,  $(x, y)$ , and  $(x', y')$ , are acceptable by both types of negotiator 1 because  $f$  is individually rational and type  $\theta_1^{x_l}$  is more accepting than type  $\theta_1^{x_r}$ . Strategy-proofness implies that  $b R_1 (x', y')$  for all  $R_1 \in \Lambda(\theta_1^{x_r})$ , which is also true for all  $R_1 \in \Lambda(\theta_1^{x_l})$  since preferences are consistent. The converse is also true, i.e.,  $(x', y') R_1 b$  for all  $R_1 \in \Lambda(\theta_1^{x_l})$ ; since  $f$  is strategy-proof and  $f$  suggests the bundle  $(x', y')$  when player 1 announces her type

as  $\theta_1^{x_\ell}$ . Thus, we must have  $b = (x', y')$  since preferences over bundles are antisymmetric. By repeating the symmetric arguments for negotiator 2 and recalling that  $b$  and  $(x', y')$  are the same, we conclude that all these three bundles must be the same, contradicting our presumption that  $x$  and  $x'$  are distinct.

Individual rationality and efficiency of  $f$  (together with **BR** and **M**) imply that every alternative  $x_k \in X$  must appear on the main diagonal of  $f$  once. We construct the injective and order-reversing function  $t : X \rightarrow Y$  by setting  $t(f_{k,k}^X) = f_{k,k}^Y$  for  $k = 1, \dots, m$ . WARP implies that any entry on the second diagonal of  $f$  is equal to the main diagonal entry that is located either to its right or above. If, for example,  $f_{2,1}$  and  $f_{1,1}$  are the same, then negotiator 2 would profitably deviate if the function  $t$  is not order-reversing. This is true because negotiator 2 would get better alternatives in both issues by deviating to type  $\theta_2^{x_2}$  rather than declaring her true type  $\theta_2^{x_1}$ . We denote the set of all bundles on the main diagonal by  $B^t$  (Step 1). Given that, WARP implies that each entry of the lower half of the matrix  $f$  is equal to an entry on the main diagonal of  $f$  (Step 2).

Much like the case in rationalizable choice mechanisms, WARP implies that  $f$  behaves as if it follows a binary relation (which we call a precedence order) over the set of alternatives in issue  $X$  such that it always picks the alternative in issue  $X$  that is revealed to be “better” than any other mutually acceptable alternative. Therefore, we construct the partial order as follows. Take any type profile  $\theta = (\theta_1^{x_\ell}, \theta_2^{x_j})$  that corresponds to an entry in the lower half of the matrix  $f$  and consider the set of all acceptable alternatives in issue  $X$  at that profile (i.e.,  $X_{j\ell}$ ). We say  $f_{\ell,j}^X \succeq x$  whenever  $x \in X_{j\ell}$ . It follows from construction that the binary relation  $\succeq$  is antisymmetric and transitive and each entry of the lower half of the matrix is indeed the maximal element of a particular subset of  $X$  with respect to the partial order.

By individual rationality,  $f$  must choose  $o_x$  above the diagonal when there is no mutually acceptable alternative in  $X$ . Then the fact that there must be a unique designated disagreement bundle is shown by iterating strategy-proofness along the rows and columns above the diagonal.

## 4.2 FULL CHARACTERIZATION AND QUID PRO QUO

Theorem 1 characterizes the necessary conditions that a strategy-proof, efficient, and individually rational mediation mechanism must satisfy. However, a logrolling mechanism is not guaranteed to be strategy-proof in general. Therefore, we now search for a condition on preferences that guarantees strategy-proofness. Since the class of logrolling mechanisms contains the only candidates that can achieve the properties in Theorem 1, ensuring that a logrolling mechanism is strategy-proof automatically entails imposing a discipline on preference profiles regarding how negotiators rank the logrolling bundles. To this end, we define a key notion.

**Definition 2.** *The preference map  $\Lambda$  satisfies **quid pro quo** if there exists an injective and order-reversing function  $t : X \rightarrow Y$  and a partial order  $\succeq_t$  over  $X$  such that:*

- i. For any distinct  $x, x' \in X$ ,  $x \succeq_t x'$  if there exists  $i \in N$  such that  $x' \theta_i x$  and*
  1.  $(x, t(x)) R_i (x', t(x'))$  for all  $R_i \in \Lambda(\theta_i)$  and  $\theta_i \in \Theta_i$  satisfying  $x, x' \in A(\theta_i)$ ,

2. there is no  $y \in Y$  with  $t(x) \theta_i^Y y \theta_i^Y t(x')$  such that  $(x', y)$  Pareto dominates  $(x, t(x))$ .

ii. The poset  $(S, \succeq_t)$  is a join semilattice for all connected  $S \subseteq X$ .

Given a preference map  $\Lambda$  satisfying quid pro quo, let  $\Pi_\Lambda$  denote the set of all partial orders induced by  $\Lambda$ . Namely,  $\Pi_\Lambda$  is the set of all partial orders  $\succeq_t$  over  $X$  such that the order-reversing function  $t$  and  $\succeq_t$  satisfy Definition 2.

Quid pro quo is a property for the domain of preference profiles (not just for individual preferences). It says that the preference domain should permit the possibility that the negotiators are willing to make compromises in issue  $X$  for a more favorable treatment in issue  $Y$ . Put differently, it should be possible to find some alternatives in issue  $Y$  that are sufficiently attractive for at least one of the negotiators to reverse her ranking of some alternatives in issue  $X$  when they are bundled together. Specifically, condition (i.1) says that for some pairs of acceptable alternatives  $x, x'$  in  $X$ , there must be a negotiator such that although she ranks  $x'$  above  $x$ , there is a pair of alternatives  $y, y'$  in  $Y$  with the property that she (weakly) prefers  $(x, y)$  to  $(x', y')$  at all admissible preferences. Such possibility of a preference reversal induces the partial order  $x \succeq_t x'$ . Condition (i.2) ensures that all logrolling bundles, generated by  $t$ , are efficient.<sup>45</sup> These preference reversals define a partial order on  $X$  and condition (ii) requires that this partial order together with any connected subset of  $X$  form a join semilattice. We are now ready to provide a full characterization result.

**Theorem 2.** *Given a regular preference map  $\Lambda$ , there exists a mediation mechanism  $f$  satisfying strategy-proofness, efficiency, and individual rationality if and only if  $\Lambda$  satisfies quid pro quo and there is a partial order  $\succeq_t \in \Pi_\Lambda$  such that  $f = f^{\succeq_t}$ .*

Theorem 2 states that quid pro quo is both necessary and sufficient for the existence of strategy-proof, efficient, and individual rational mediation mechanisms, and any such mechanism must be a logrolling mechanism associated with a precedence order  $\succeq_t$  which is induced by the preference map  $\Lambda$ .

### *A Practical Depiction of Quid Pro Quo*

Definition 2 expresses quid pro quo property based on an order-theoretic semilattice structure. An equivalent and arguably more intuitive description uses a recursive and algorithmic process on the set of logrolling bundles, which we present through a simple example. This alternative structure may be practically useful since it provides hints to approximate the core principles of quid pro quo property and Theorem 2.

The essence of quid pro quo is that the preference domain (i.e., the preference map  $\Lambda$ ) allows an “*elimination tournament*” of the form discussed below among the set of logrolling bundles, where there is always a winner of each matchup at each round. Furthermore, each round of this tournament represents the corresponding diagonal of the strategy-proof, efficient, and individual rational mediation mechanism to be constructed.

<sup>45</sup>Condition (i.2) is redundant when  $|Y| = |X|$ , and satisfied otherwise when working with standard continuous utility functions such as those in Example 2.

As an example, consider the mediation problem with three alternatives in each issue. The tournament always starts with all logrolling bundles ordered from  $b_1$  to  $b_3$  (see the left side of Figure 3), where  $b_k = (x_k, y_{4-k})$ .<sup>46</sup> In the first round of the tournament, each logrolling bundle matches up with its neighbors (i.e., both  $b_1$  and  $b_3$  match only with  $b_2$ ). In the matchup between  $b_k$  and  $b_{k+1}$ , the “winner” is  $b_{k+1}$  if (and only if) negotiator 1 unambiguously ranks  $b_{k+1}$  over  $b_k$ . Here we say “*negotiator  $i$  unambiguously ranks bundle  $b$  over  $b'$* ” if  $b R_i b'$  for all  $R_i \in \Lambda(\theta_i)$  and all  $\theta_i$  who deems both bundles acceptable. On the other hand, the winner of this matchup is  $b_k$  if (and only if) negotiator 2 unambiguously ranks  $b_k$  over  $b_{k+1}$ .<sup>47</sup> When both negotiators are able to compare these two bundles as required (i.e., negotiator 1 unambiguously ranks  $b_{k+1}$  over  $b_k$  and negotiator 2 unambiguously ranks  $b_k$  over  $b_{k+1}$ ), then the winner can be any of these two bundles. In this case, the mediator (i.e., the partial order  $\succeq_t$  that we create along the way) has the freedom to choose either one of these two bundles to proceed to the next round.

In our example, we suppose that the negotiators’ preferences are such that negotiator 2 unambiguously ranks  $b_1$  over  $b_2$  and negotiator 1 unambiguously ranks  $b_3$  over  $b_2$ . Therefore,  $b_1$  and  $b_3$  win over  $b_2$  and move to the next round. In the second round, the winners of the first round (i.e.,  $b_1$  and  $b_3$ ) match up (see the second row on the left side of Figure 3). Once again, the winner will be  $b_3$  (respectively  $b_1$ ) if negotiator 1 (respectively 2) can unambiguously rank  $b_3$  over  $b_1$  (respectively  $b_1$  over  $b_3$ ).

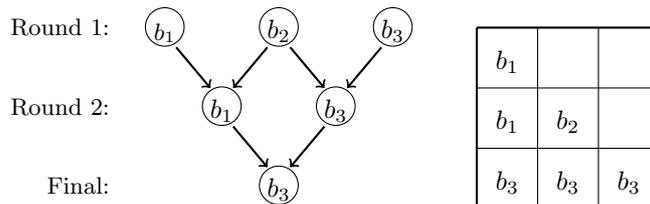


Figure 3: An example for the elimination tournament and the matrix representation for the corresponding mediation rule

If no negotiators can compare  $b_1$  and  $b_3$  as required, then the process fails. In this case, we go back to the previous round(s). If the mediator had the freedom to choose the winner of any matchups in the earlier rounds, then we replace the winner(s) of these matchups and reiterate the entire process. However, if the mediator had no freedom to choose the winner in earlier rounds, or if none of these reiterations yield a matchup in the second round with a winner, then the process fails. This means that the domain of preferences does not satisfy *quid pro quo*. Suppose, for the sake of the argument, bundles  $b_1$  and  $b_3$  are unambiguously ranked by both negotiators in the desired way. Then either bundle can be the winner of the second round. In the illustration above, we have chosen  $b_3$  as the winner of the tournament.

The matchup configurations in this entire tournament is in fact the join semilattice structure imposed by (ii) of Definition 2. Theorem 2 says that we can use this tournament structure in creating logrolling mechanisms that are efficient, individually rational and strategy-proof. The winners of each round fill up the corresponding diagonals. For the

<sup>46</sup>When  $|X| = |Y|$ , then order-reversing function  $t$  is unique. For cases where  $|X| < |Y|$ , the process may start with any such  $t$ . If the logrolling bundles in  $B^t$  fail to satisfy (i.2), then  $t$  should be replaced and the entire process should be repeated with the new logrolling bundles.

<sup>47</sup>These are the conditions implied by (i.1) of Definition 2.

tournament described above, the order of the logrolling bundles in the first round gives the placement order of these bundles (from the top corner to the bottom corner) along the first diagonal, the order in the second round gives the placement order along the second diagonal, and the last winner,  $b_3$ , fills up the bottom left entry of this matrix (which is the last diagonal). The constructed mechanism corresponds to the logrolling mechanism  $f_t^{\succeq}$  with  $\succeq_t: b_3 \ b_1 \ b_2$  where  $t(x_k) = y_{4-k}$  for  $k = 1, 2, 3$ . Recall that both  $b_1$  and  $b_3$  were winners in the second round in our example, so the preference domain also admits a second logrolling mechanism  $f_t^{\succeq'}$  with  $\succeq'_t: b_1 \ b_3 \ b_2$ .

Although our ordinal approach does not make any explicit assumptions about (or seek to elicit) negotiators' cardinal preferences, this certainly does not preclude the possibility that the negotiators are inherently endowed with such preferences. In fact, many standard utility functions are compatible with the quid pro quo condition. This is illustrated in Example 2 for the Cobb-Douglas and quasi-linear preferences when alternatives represent quantities.

**Example 2 (Quid Pro Quo Under Standard Preferences):** Suppose the negotiators are in dispute over dividing 10 units of good  $X$  and 6 units of good  $Y$ . Also suppose there are five possible ways of dividing good  $X$  where an alternative  $x \in X = \{1, 3, 5, 7, 9\}$  denotes negotiator 1's share of good  $X$ .<sup>48</sup> Each negotiator's private bargaining range is determined by her least acceptable amount of good  $X$ . Symmetrically, an alternative  $y \in Y = [0, 6]$  denotes negotiator 1's share of good  $Y$ . Therefore, when bundle  $(x, y)$  is chosen, negotiator 1 gets  $(x, y)$  and 2 gets  $(10-x, 6-y)$ . It is commonly known that negotiator  $i$ 's consumption utility from a bundle with an acceptable amount of good  $X$  and  $Y$  is given by some function  $U_i$  (and, is otherwise zero when offered an unacceptable amount of good  $x$  outside her bargaining range) which may also depend on some privately known parameter(s). The private parameter in the utility function represents the uncertainty others are facing with regards to the negotiators' full fledged preferences over bundles as captured by the preference map  $\Lambda$ . Consider a linear order-reversing function  $t(x) = \frac{11-x}{2}$  with which the mediator chooses alternatives from issue  $Y$ . This leads to one possible set of logrolling bundles where  $B^t = \{(1, 5), (3, 4), (5, 3), (7, 2), (9, 1)\}$ .<sup>49</sup>

**Cobb-Douglas:** Suppose  $U_i(x, y) = x^{a_i} y^{b_i}$  where  $\frac{a_i}{b_i} \in [\frac{7}{12}, 1]$  for  $i = 1, 2$ ; but the exact values of  $a_i$  and  $b_i$  are only privately known. Although negotiator 1 ceteris paribus prefers higher values of  $x$  (i.e.,  $x' \theta_1 x$  when  $x' > x$ ), it can be verified that  $(5, 3) R_1 (7, 2) R_1 (9, 1)$ . This induces the partial order  $\succeq_t$  such that  $5 \succeq_t 7 \succeq_t 9$ . Similarly, although negotiator 2 ceteris paribus prefers lower values of  $x$  (getting more of good  $X$ ), we have  $(5, 3) R_2 (3, 4) R_2 (1, 5)$ . This induces the partial order  $\succeq_t$  such that  $5 \succeq_t 3 \succeq_t 1$ . A logrolling mechanism associated with  $B^t$  and partial order  $\succeq_t$  is strategy-proof. Figure 6 in Section 5 illustrates one such mechanism.

<sup>48</sup>While we have chosen this particular division possibility for a simple illustration with symmetric division options, it is also possible to choose  $X = [0, 10]$  or some other discrete set. See the continuous analogue of our model in the Supplementary Appendix.

<sup>49</sup>When either issue is a continuum, there are infinitely many (possibly non-linear) order-reversing (i.e., strictly decreasing) functions and infinitely many possible sets of logrolling bundles with which the mediator can choose to form the basis of a logrolling mechanism.

**Quasi-linear:** Suppose  $U_i(x, y) = x^{\alpha_i} + y$  where  $\alpha_i \in (0, \frac{2}{3}]$  for  $i = 1, 2$ ; but the exact value of  $\alpha_i$  is only privately known. Although negotiator 1 ceteris paribus prefers higher values of  $x$ , she prefers  $(x, t(x))$  over  $(x', t(x'))$  for all  $x, x' \in X$ . Based solely on 1's preferences, this induces a complete order on  $X$  where  $x \succeq_t^1 x'$  when  $x' > x$ . Similarly, negotiator 2's preferences over bundles satisfies an analogous reversal. Based solely on 2's preferences, this induces a complete order on  $X$  where  $x \succeq_t^2 x'$  when  $x > x'$ . Consequently, any linear order on  $X$  satisfies Condition (i1) in Definition 2. Indeed, any logrolling mechanism associated with  $B^t$  and any well-defined partial order on  $X$  is strategy-proof under these preferences.

**Remark:** It is straightforward to generalize the above example to any well-behaved (e.g., continuous and monotonic) utility functions and arbitrary division possibilities. All that matters for the quid pro quo condition is how the curvature of the utility function (i.e., rate of substitution between the two issues) compares with the slope of the order-reversing (decreasing)  $t$  function that governs the set of logrolling bundles.<sup>50</sup> Consequently, for any given pair of differentiable and increasing utility functions, it is possible to choose an arbitrary division possibility (i.e., order-reversing function) so long as these slope comparisons have the appropriate direction, and vice versa. More generally, when division possibilities are governed by a linear  $t$  function as in Example 2, quid pro quo holds when preferences are convex. Also see the continuous analogue of our model in the Supplementary Appendix allowing for non-linear division possibilities.

### *Sketch of the Proof of Theorem 2*

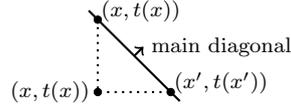
Consider the “if” part. The mediation mechanism  $f^{\succeq_t}$  satisfies individual rationality because it never suggests an unacceptable alternative. For efficiency, we consider a bundle  $b = (x_b, t(x_b))$  that  $f^{\succeq_t}$  suggests at some type profile, corresponding to the lower half of the matrix. Suppose for a contradiction that another mutually acceptable bundle  $a = (x_a, t(x_a)) \in B^t$  Pareto dominates  $b$ . Because  $f^{\succeq_t}$  suggests  $b$ , we must have  $x_b \succeq_t x_a$ . Moreover, since  $\Lambda$  satisfies quid pro quo, there must exist a negotiator  $i$  such that  $x_a \theta_i x_b$ ; but  $b R_i a$  for all consistent  $R_i$ , contradicting that  $a$  Pareto dominates  $b$ . If, however,  $a \notin B^t$ , then by condition (i.2) of Definition 2, there is no alternative  $y \in Y$  that can be matched with  $x_a$  so that  $a$  Pareto dominates  $b$ . Hence,  $f$  must be efficient.

Regarding strategy-proofness of  $f$ , a profitable deviation is never possible, by the **BR** property, from or to a type profile in which  $f^{\succeq_t}$  suggests  $(o_x, y)$ . So, consider a type profile where  $f^{\succeq_t}$  suggests  $b$ . Any deviation of, say, negotiator 1 to a less-accepting type to get  $a$ , which is located on the same column with  $b$  but on a lower row, is never profitable. This is true because (1) we have  $x_b \succeq_t x_a$  since  $f^{\succeq_t}$  suggests  $b$  when both these bundles are mutually acceptable; (2) bundle  $a$  must appear above bundle  $b$  on a lower row on the main diagonal (due to the transitivity of  $\succeq_t$ ), namely  $x_a \theta_1 x_b$ ; and thus (3)  $b R_1 a$  by quid pro quo and (1) above. A similar reasoning proves that negotiator 1 has no incentive to deviate to a more-accepting type. Hence,  $f^{\succeq_t}$  is strategy-proof.

Consider now the “only if” part. By Theorem 1, strategy-proofness, efficiency, and individual rationality of  $f$  imply an injective and order reversing function  $t$  and a partial

<sup>50</sup>For example, if the privately known parameters are outside the given interval in Example 2, then quid pro quo condition can be satisfied by choosing a flatter or a steeper  $t$  function.

order  $\succeq$  such that  $f = f^\succeq$ . To prove  $\succeq \in \Pi_\Lambda$ , and so  $\Lambda$  satisfies quid pro quo, we show that  $\succeq$  and  $t$  satisfy Definition 2. Condition (i.2) is simply implied by the efficiency of  $f$ . For condition (i.1) take any  $x, x'$  with  $x \succeq x'$ . By the construction of  $\succeq$  in the proof of Theorem 1, we know that  $x \succeq x'$  implies that  $f$  must be suggesting a bundle with  $x$  at some type profile where both  $x$  and  $x'$  are mutually acceptable. Assuming, w.l.o.g., that  $x \theta_1 x'$ , WARP requires that  $f(\theta_1^x, \theta_2^x) = (x, t(x))$  (see the figure below). Then, it is easy to verify that strategy-proofness of  $f$  implies  $(x, t(x)) R_2 (x', t(x'))$ , as required by condition (i.1).



Finally, the collection of sets  $X_{j\ell}$  where  $1 \leq j \leq \ell \leq m$  constitutes the set of all connected subsets of  $X$ , and every doubleton  $\{x, x'\} \subset X_{j\ell}$  has a least upper bound in  $X_{j\ell}$ , which is  $x_{X_{j\ell}}^*$ . Thus,  $(S, \succeq)$  is a semilattice for all connected subsets of  $X$ , as required by condition (ii).

### 4.3 A VISUAL CHARACTERIZATION OF THE CLASS OF LOGROLLING MECHANISMS

To provide further insight into the logrolling mechanisms that are characterized by Theorems 1 and 2, we offer a geometric analysis of these mechanisms. We first take a mechanism  $f = [f_{\ell,j}]_{(\ell,j) \in M^2}$  and introduce a couple of definitions to represent different rectangular and triangular regions of the matrix. In the following two definitions we slightly abuse notation and terminology in order to keep track of the entries contained in a rectangular/triangular region. Namely, we use  $f_{\ell,j}$  to refer to entry  $(\ell, j)$  of the matrix rather than the specific bundle that mechanism  $f$  assigns to that entry.

**Definition 3.** Consider the entry  $f_{k,k}$  for some  $k \in M$  and an entry that lies (weakly) to its southwest,  $f_{\ell,j}$  with  $1 \leq j \leq k \leq \ell \leq m$ . The **rectangle** induced by  $f_{k,k}$  and  $f_{\ell,j}$ , denoted by  $\square_{\ell,j}^k$ , is the set of all entries in the rectangular region of the matrix (inclusively) enveloped between rows  $k$  and  $\ell$  and columns  $k$  and  $j$ . Namely,  $\square_{\ell,j}^k = \bigcup_{\substack{j \leq s \leq k \\ k \leq t \leq \ell}} \{f_{t,s}\}$ .

**Definition 4.** The **triangle** induced by an entry  $f_{\ell,j}$  with  $1 \leq j \leq \ell \leq m$ , denoted by  $\triangle_{\ell,j}$ , is the set of all entries in the triangular region of the matrix that is (inclusively) enveloped by the entry  $f_{\ell,j}$ , row  $\ell$ , column  $j$ , and the main diagonal. Namely,  $\triangle_{\ell,j} = \bigcup_{j \leq k \leq \ell} \{f_{k,j}, f_{k,j+1}, \dots, f_{k,k}\}$ .

A rectangle/triangle is merely a collection of entries of the matrix induced by mechanism  $f$  (i.e., sets of pairs of indexes). Note that an entry on the main diagonal is a special triangle (and also a special rectangle) that consists of a singleton entry. Furthermore, the entire main diagonal of the matrix and all the entries to its southwest constitute the largest possible triangle  $\triangle_{m,1}$ . Given a triangle  $\triangle_{\ell,j}$ , its entries that lie on the main diagonal are said to be on the *hypotenuse* of  $\triangle_{\ell,j}$ . A partition of the lower half of the matrix is called a *rectangular*

(triangular) partition if and only if it is the union of disjoint rectangles (triangles).<sup>51</sup>

**Theorem 3 (Visual Characterization).** *The following statements are equivalent for the lower half of the mediation mechanism  $f$ , corresponding to type profiles in which a mutually acceptable alternative from the main issue exists:*

- (i)  $f$  is a logrolling mechanism.
- (ii) The triangle  $\Delta_{m,1}$  has a rectangular partition such that  $f$  assigns a unique bundle from the set of logrolling bundles  $B^t$  to each rectangle in this partition.<sup>52</sup>

Part (ii) of Theorem 3 states that a logrolling mechanism  $f$  can be represented as the union of  $m$  disjoint rectangular regions. Each rectangle has a distinct corner entry on the main diagonal that contains the logrolling bundle that fills up the entire rectangle. Procedurally, these rectangles are obtained as follows. Given the precedence order on  $X$ , start with the logrolling bundle with the highest-precedence alternative (i.e., highest-precedence bundle). Starting from the entry of this bundle on the hypotenuse of the largest triangle,  $\Delta_{m,1}$ , let it fill up all the entries located to its southwest. This creates the first and largest rectangle  $\square$ , and leads to a triangular partition of  $\Delta_{m,1} \setminus \square$ . Next, pick any triangle from this partition and let the highest-precedence bundle on the hypotenuse of this triangle fill up all the entries that are located to its southwest. This leads to a second rectangle  $\square'$  as well as a unique triangular partition of  $\Delta_{m,1} \setminus \{\square, \square'\}$ . The process can be iterated in this fashion until the entire triangle  $\Delta_{m,1}$  is partitioned into  $m$  disjoint rectangles in  $m$  steps. Figure 4a provides an illustration of one such partitioning, where  $b_k = (x_k, t(x_k))$  for  $k = 1, \dots, 9$ . This process effectively traces the semilattice  $(X, \supseteq)$  in Figure 4b. Conversely, any such geometric set, namely any rectangular partition of  $\Delta_{m,1}$ , can be used to construct a precedence order and a corresponding logrolling mechanism.

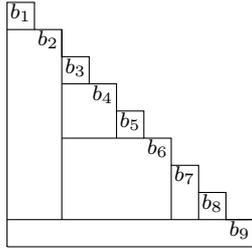


Figure 4a: A rectangular partitioning of  $f^\supseteq$  with  $m=9$

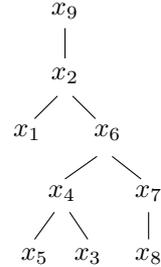


Figure 4b: A semilattice  $(X, \supseteq)$

#### 4.4 A PRACTICAL FORMULATION OF LOGROLLING MECHANISMS

As briefly discussed in Section 1.1, dispute resolution protocols of some ODR platforms (e.g., SquareTrade) is based on generating a menu of recommendations from which negotiators are asked to make selections sequentially. In light of Theorem 3, the working principles

<sup>51</sup>Note that a rectangular partition consists of  $m$  disjoint rectangles. For example,  $\{\square_{k,1}^k\}_{k=1}^m$  and  $\{\square_{m,k}^k\}_{k=1}^m$  are two obvious rectangular partitions of  $\Delta_{m,1}$ . These two partitions correspond respectively to what we will later refer to as the negotiator 1- and negotiator 2-optimal mechanisms.

<sup>52</sup>More formally, for any  $\square$  in the partition of  $\Delta_{m,1}$  and any bundles  $b, b' \in \square$ ,  $b = b'$ ; but for any distinct pair  $\square, \square'$  in the partition of  $\Delta_{m,1}$ ,  $(x, y) \in \square$  and  $(x', y') \in \square'$  implies  $x \neq x'$  and  $y \neq y'$ .

of the logrolling mechanisms generate a similar interpretation that is also reminiscent of the divide-and-choose mechanisms in fair division literature.

In particular, a logrolling mechanism can be thought to operate as a “*shortlisting mechanism*” in a decentralized fashion: One negotiator offers a shortlist of bundles as acceptable solutions for the dispute, the mediator communicates these options to the other negotiator who then chooses her favorite bundle from this list. To see this, observe that when negotiator 1 reports her type as  $\theta^{x_\ell}$ , it can be viewed as negotiator 1 forming a shortlist consisting of all the bundles on row  $\ell$ . When faced with the list of bundles negotiator 1 offers, negotiator 2 indeed picks the bundle  $f_{\ell,j}$  since it is her favorite acceptable bundle on row  $\ell$  by strategy-proofness. If the roles of the negotiators in this procedure were reversed, then the outcome would still be the same by symmetric arguments.<sup>53</sup>

For a more specific example, consider the logrolling mechanism depicted in Figure 4a. Suppose that negotiator 1 is of type  $\theta_1^{x_3}$ . Then we can think of her as proposing the shortlist  $\{b_2, b_3, (o_x, y)\}$  to the other negotiator when she truthfully declares her type. The corresponding shortlists for other announcements as  $\theta_1^{x_5}$  and  $\theta_1^{x_7}$  are  $\{b_2, b_4, b_5, (o_x, y)\}$  and  $\{b_2, b_6, b_7, (o_x, y)\}$ , respectively.

Under this interpretation, a logrolling mechanism specifies a set of shortlisted bundles that a negotiator can offer to the other party for each possible type she reports. By reporting a more-accepting type, the proposer may add new bundles or remove some from her shortlist. Theorem 3 implies that as negotiators declare more-accepting types, suggested shortlists must satisfy some kind of regularity in the sense that a previously removed bundle can never be added back to the shortlist. For the logrolling mechanism depicted in Figure 4a, for instance, if negotiator 1 switches from  $\theta_1^{x_3}$  to  $\theta_1^{x_5}$ , she adds bundles  $b_4$  and  $b_5$  to the shortlist and removes  $b_3$ . If she switches from  $\theta_1^{x_5}$  to  $\theta_1^{x_7}$ , then she adds  $b_6$  and  $b_7$  and removes  $b_4$  and  $b_5$  from the shortlist. Note for this logrolling mechanism that once bundles  $b_3$ ,  $b_4$  or  $b_5$  are removed, they are never added back in.

## 5 SPECIAL MEMBERS OF THE LOGROLLING FAMILY

We next visit interesting members of the logrolling family. At the outset we assume that preference domain is such that all members of the family are strategy-proof, e.g., negotiators have quasi-linear preferences.<sup>54</sup> Three notable members of the family are worth pointing out. A **negotiator-optimal mechanism** represents a situation of extreme partiality to one side of the dispute and is constructed by using the precedence order implied by a negotiator’s preferences over the logrolling bundles. Specifically, the negotiator 1-optimal mechanism takes

$$\succeq^1: x_m \succeq^1 x_{m-1} \succeq^1 \dots \succeq^1 x_1,$$

<sup>53</sup>One drawback of the divide-and-choose rule in the context of fair division is that its outcome depends on the order of agents. Divide-and-choose also violates strategy-proofness unlike the logrolling mechanism.

<sup>54</sup>This assumption renders a more meaningful comparison of the members possible. Specifically, we assume that quid pro quo is satisfied in the following strong sense. The preference map  $\Lambda$  admits an injective and order-reversing function  $t : X \rightarrow Y$  such that for all  $i$ ,  $\theta_i \in \Theta_i$ ,  $R_i \in \Lambda(\theta_i)$  and all  $x, x' \in A(\theta_i)$  with  $x' \theta_i x$ , we have  $(x, t(x)) R_i (x', t(x'))$ . In this case, the negotiators’ favorite logrolling bundles are at the opposite corners of the diagonal.

whereas the negotiator 2-optimal mechanism takes

$$\triangleright^2: x_1 \triangleright^2 x_2 \triangleright^2 \dots \triangleright^2 x_m.$$

In case of severe disagreement, i.e., when there is no mutually acceptable alternative in issue  $X$ , the corresponding designated bundle includes the favored negotiator's most-preferred alternative in issue  $Y$ . The two dual mechanisms are shown below for the case of  $m = n = 5$ .

	$\theta_2^{x_1}$	$\theta_2^{x_2}$	$\theta_2^{x_3}$	$\theta_2^{x_4}$	$\theta_2^{x_5}$
$\theta_1^{x_1}$	$(x_1, y_5)$	$(o_X, y_1)$	$(o_X, y_1)$	$(o_X, y_1)$	$(o_X, y_1)$
$\theta_1^{x_2}$	$(x_2, y_4)$	$(x_2, y_4)$	$(o_X, y_1)$	$(o_X, y_1)$	$(o_X, y_1)$
$\theta_1^{x_3}$	$(x_3, y_3)$	$(x_3, y_3)$	$(x_3, y_3)$	$(o_X, y_1)$	$(o_X, y_1)$
$\theta_1^{x_4}$	$(x_4, y_2)$	$(x_4, y_2)$	$(x_4, y_2)$	$(x_4, y_2)$	$(o_X, y_1)$
$\theta_1^{x_5}$	$(x_5, y_1)$				

Figure 5a: *Negotiator 1-optimal mechanism*

	$\theta_2^{x_1}$	$\theta_2^{x_2}$	$\theta_2^{x_3}$	$\theta_2^{x_4}$	$\theta_2^{x_5}$
$\theta_1^{x_1}$	$(x_1, y_5)$	$(o_X, y_5)$	$(o_X, y_5)$	$(o_X, y_5)$	$(o_X, y_5)$
$\theta_1^{x_2}$	$(x_1, y_5)$	$(x_2, y_4)$	$(o_X, y_5)$	$(o_X, y_5)$	$(o_X, y_5)$
$\theta_1^{x_3}$	$(x_1, y_5)$	$(x_2, y_4)$	$(x_3, y_3)$	$(o_X, y_5)$	$(o_X, y_5)$
$\theta_1^{x_4}$	$(x_1, y_5)$	$(x_2, y_4)$	$(x_3, y_3)$	$(x_4, y_2)$	$(o_X, y_5)$
$\theta_1^{x_5}$	$(x_1, y_5)$	$(x_2, y_4)$	$(x_3, y_3)$	$(x_4, y_2)$	$(x_5, y_1)$

Figure 5b: *Negotiator 2-optimal mechanism*

A negotiator-optimal mechanism always chooses the corresponding negotiator's most-preferred bundle among the mutually acceptable logrolling bundles. The analogous shortlisting mechanism is rather simple: the favored negotiator's shortlist includes only two bundles, which are her favorite logrolling bundle and the designated disagreement outcome.<sup>55</sup> Clearly, these two polar members of the family of logrolling mechanisms are highly unattractive in practice.<sup>56</sup> Fortunately, there is a remarkable member of this family that treats negotiators symmetrically.

Impartiality entails focusing on a central element of the set of logrolling bundles as a compromise. It is then intuitive for the mediator to recommend a *median* logrolling bundle when it is mutually acceptable, or seek a bundle as close to it as possible when it is not. Within the family of logrolling mechanisms, this is achieved simply by assigning the highest precedence to a median logrolling bundle, and the next precedence to those bundles that are closest to the chosen median, and so on, and lowest precedence to the extremal logrolling bundles. Based on similar logic, when there is no mutually acceptable alternative in  $X$ , the designated bundle chosen by an impartial mediator should naturally include a median alternative in  $Y$ . This motivates the following type of mechanism, which we call a **constrained shortlisting (CS) mechanism**.

**Definition 8.** Let  $k \in \{\bar{k}, \underline{k}\}$  be the index of a median alternative, where  $\bar{k} = \lceil \frac{m+1}{2} \rceil$  and  $\underline{k} = \lfloor \frac{m+1}{2} \rfloor$ . A mechanism is a constrained shortlisting mechanism, denoted  $f^{CS} = [f_{\ell,j}]_{(\ell,j) \in M^2}$ , if it is a logrolling mechanism that is associated with a precedence order  $\triangleright$ , where  $x_k \triangleright x_{k-1} \triangleright \dots \triangleright x_1$  and  $x_k \triangleright x_{k+1} \triangleright \dots \triangleright x_m$ , and  $f_{\ell,j}^{CS} = (o_X, y_k)$  whenever  $\ell < j$ .

<sup>55</sup>Alternatively, the non-favored negotiator's shortlist includes all of her acceptable logrolling bundles and the designated disagreement outcome.

<sup>56</sup>Note that despite their polarity, these mechanisms are not dictatorial. Unlike a dictatorship, they remain individually rational and never get vetoed in equilibrium. Nevertheless, they hint at the possibility of the mediator having the power to tilt the balance in a dispute despite using a mechanism that meets our desiderata (i.e., efficiency, individual rationality, and strategy-proofness).

When the number of alternatives is odd, there is a unique constrained shortlisting mechanism. When the number of alternatives is even, however, a constrained shortlisting mechanism prescribes one of four possible types of outcomes.<sup>57</sup> Figure 6 illustrates the constrained shortlisting mechanism for the case of  $m = n = 5$ .

When the number of alternatives is odd, the CS mechanism is a symmetric member of the logrolling mechanisms family.<sup>58</sup> In the lower half of the matrix, it acts as a negotiator-optimal mechanism whenever the median alternative in issue  $X$  is not mutually acceptable and recommends the median logrolling bundle whenever the set of mutually acceptable alternatives includes the median alternative. In other words, when both negotiators find at least half of the alternatives in  $X$  acceptable, the mechanism chooses the median logrolling bundle; and, when one negotiator finds at least half of the alternatives acceptable while the other finds less than half of the alternatives acceptable, the mechanism chooses the less-accepting negotiator's favorite logrolling bundle.

	$\theta_2^{x_1}$	$\theta_2^{x_2}$	$\theta_2^{x_3}$	$\theta_2^{x_4}$	$\theta_2^{x_5}$
$\theta_1^{x_1}$	$(x_1, y_5)$	$(o_X, y_3)$	$(o_X, y_3)$	$(o_X, y_3)$	$(o_X, y_3)$
$\theta_1^{x_2}$	$(x_2, y_4)$	$(x_2, y_4)$	$(o_X, y_3)$	$(o_X, y_3)$	$(o_X, y_3)$
$\theta_1^{x_3}$	$(x_3, y_3)$	$(x_3, y_3)$	$(x_3, y_3)$	$(o_X, y_3)$	$(o_X, y_3)$
$\theta_1^{x_4}$	$(x_3, y_3)$	$(x_3, y_3)$	$(x_3, y_3)$	$(x_4, y_2)$	$(o_X, y_3)$
$\theta_1^{x_5}$	$(x_3, y_3)$	$(x_3, y_3)$	$(x_3, y_3)$	$(x_4, y_2)$	$(x_5, y_1)$

Figure 6: *Constrained shortlisting mechanism*

In discrete resource allocation problems where agents are endowed with ordinal preference rankings, fairness properties (together with efficiency) have often proved difficult to attain in the absence of monetary transfers or a randomization device. It is nevertheless worthwhile to investigate whether it is possible for a member of the logrolling mechanisms family to achieve alternative fairness requirements beyond symmetry. We next formulate one such ordinal fairness notion as a normative requirement for our context.

Given the negotiators' preferences over alternatives (not including the outside option), let  $r_i(z) \in M$  denote negotiator  $i$ 's ranking of an acceptable alternative  $z \in Z \in \{X, Y\}$ . For a normalization, we re-assign ranks 1 through  $m$  to the chosen alternatives in  $Y$  and set the ranking of the outside option to be zero.<sup>59</sup> Given the logrolling mechanism  $f = [f_{\ell,j}]_{(\ell,j) \in M^2}$ , the *rank variance of the bundle*  $f_{\ell,j}$  is defined as<sup>60</sup>

$$\text{var}(f_{\ell,j}) \equiv \sum_{i \in N} (r_i(f_{\ell,j}^X))^2 + (r_i(f_{\ell,j}^Y))^2.$$

<sup>57</sup>In this case, the mechanism depends on whether  $x_{\bar{k}}$  or  $x_{\underline{k}}$  has the highest precedence and whether  $y_{\bar{k}}$  or  $y_{\underline{k}}$  is included in the designated disagreement bundle.

<sup>58</sup>When the number of alternatives is even, no logrolling mechanism is fully symmetric.

<sup>59</sup>This normalization is clearly not without loss, but simplifies the notation significantly as it treats issue  $Y$  as though it also has  $m$  alternatives. Nevertheless, the rank minimizing logrolling mechanism in the absence of this normalization is merely a "shifted" version of a CS mechanism where the magnitude of the shift depends on the order-reversing function  $t$ .

<sup>60</sup>This formulation assigns equal weights to both issues. One may also consider assigning different weights to different issues. Theorem 4 remains unchanged in that case due to the symmetric structure of the logrolling bundles under the normalization above.

Then, the **rank variance** of a mechanism  $f$  is the total sum of the rank variance of all possible outcomes of  $f$ , and defined as

$$\text{Var}(f) \equiv \sum_{\ell=1}^m \sum_{j=1}^m \text{var}(f_{\ell,j}).$$

Intuitively, the larger the differences between the two negotiators' rankings of the alternatives in a given bundle, the higher the rank variance of that bundle. For example, while never recommended by a logrolling mechanism, the bundles  $(x_1, y_1)$  and  $(x_m, y_m)$  have the highest rank variance. Despite making one negotiator as well off as possible, they make the opposite negotiator as worse off as possible. In this sense, the larger the rank variance of a mediation mechanism, the more skewed it is toward extremal bundles.

**Theorem 4.** *A mediation mechanism minimizes rank variance within the class of logrolling mechanisms if and only if it is a constrained shortlisting mechanism.*

## 6 DISCUSSION AND EXTENSIONS

In this section we provide a general discussion of our main model in light of the results obtained so far. To this end, first, we elaborate on some of our essential modeling assumptions, discuss the role they play in driving the positive results of our paper, and offer directions in which they can be extended to cases not covered in the main exposition. Second, drawing on our findings, we consider how one can go about formulating the mediation problem in a standard Bayesian setting such as that of Myerson and Satterthwaite (1983) [henceforth MS] and offer a reconciliation of the possibility results in our setup with the impossibility result in the MS setting.

### 6.1 MODELING CONFLICTING PREFERENCES

We argue that diametrically opposed preferences in each issue is without loss of generality. When describing a dispute, using diametrically opposed preferences over alternatives is intuitive. However, it is conceivable that many other situations, where preferences are not necessarily diametrically opposed, could also depict a dispute. Consider, for example, a case where the set of available alternatives is  $X = \{x_1, x_2, x_3, x_4, x_5\}$  and the negotiators' preferences are as follows:

$$\begin{aligned} \theta_1 &: x_1 \ x_2 \ x_3 \ x_4 \ x_5 \\ \theta_2 &: x_3 \ x_5 \ x_4 \ x_2 \ x_1 \end{aligned}$$

These preferences are not diametrically opposed, but they are certainly conflicting to some extent as the agents cannot agree on their best alternative. Notice, however, that alternatives  $x_4$  and  $x_5$  are (Pareto) dominated by  $x_3$ . So, if selecting an efficient outcome by the mediation protocol is desired, then the presence of these two alternatives is irrelevant for the problem and can be eliminated from the preferences. Thus, this particular dispute problem can be transformed into a reduced problem where the only available alternatives are  $x_1, x_2$ , and  $x_3$  and the negotiators' preferences over these three are diametrically opposed.

Proposition S.1. in the Supplementary Appendix shows that this observation generalizes to any (discrete) set of alternatives and any preference profile. A similar result, which we omit for brevity, also holds for two-person, multi-issue disputes whenever preferences over bundles satisfy monotonicity.

## 6.2 SYMMETRIC TREATMENT OF THE OUTSIDE OPTIONS

In this section we consider the case where the outside option in issue  $Y$  is also treated as each negotiator’s private information. We let  $\Theta_i = \Theta_i^X \times \Theta_i^Y$  denote the set of all **types** for negotiator  $i$ , and  $\Theta = \Theta_1 \times \Theta_2$  be the set of all type profiles. Therefore, the mediation mechanism  $f$  maps  $\Theta$  into  $\bar{X} \times \bar{Y}$ . We need to adjust the regularity assumption concerning the negotiators’ preferences over bundles. Specifically, we need to modify the bargaining ranges condition since both issues can now potentially have unacceptable alternatives. The complete formalization and the proof of the next result is deferred to the Supplementary Appendix.

**Proposition 1.** *Under the symmetric treatment of the outside options, there is no mediation mechanism  $f$  that is efficient, individually rational, and strategy-proof.*

## 6.3 MORE THAN TWO ISSUES OR NEGOTIATORS

Our two-issue model is without loss of generality. If there are more than two issues in the dispute, then we can regroup these issues under two types of categories depending on whether an issue has certain or uncertain gains from mediation. In particular, let category- $X$  be the collection of issues that exhibit uncertain gains from mediation (i.e., a negotiator’s least acceptable alternative is her private information), and category- $Y$  be the collection of issues that exhibit certain gains (i.e., it is common knowledge that efficient and mutually acceptable alternatives exist). Under this regrouping, each negotiator now faces a vector of alternatives for each category. The negotiators’ preferences over these vectors (of alternatives) need not be diametrically opposed in general. However, as long as the negotiators’ preferences are monotonic, by applying the transformation discussed in Section 6.1, we can eliminate all inefficient vectors. This brings us back to an environment analogous to our main model, in which preferences over vectors are diametrically opposed. When there are multiple parties involved in a dispute, as it would be the case for community/public disputes, we can similarly regroup them to be represented by either negotiator, effectively treating them as clones of the two negotiators. Nevertheless, there might be cases where negotiators’ preferences are extremely disperse, and so grouping them into two “representative” agents may not be feasible. Such mediation environments are both practically and theoretically more complex than the one we study here and in need of further research.

## 6.4 ISSUE-WISE VOTING IN THE DIRECT MECHANISM WITH VETO RIGHTS

In our main model, we assume that negotiators simultaneously and independently decide in the ratification stage whether to accept or veto the proposal. It does not matter whether voting is simultaneous or sequential. However, it is critical that the parties vote on the proposed bundle as a whole rather than voting on each issue separately. An alternative

consideration would be to allow the negotiators vote separately for each individual issue. In this case, revealing one's type truthfully in the announcement stage may no longer be an optimal strategy even if the mediation mechanism is a logrolling mechanism. Consider, for example, a strategy-proof negotiator-1 optimal mechanism when  $m = n = 2$ :

	$M_2$	$L_2$
$L_1$	$(x_1, y_2)$	$(o_x, y_1)$
$M_1$	$(x_2, y_1)$	$(x_2, y_1)$

Suppose negotiator 1 reports her type as  $L_1$  and negotiator 2's true type is  $L_2$ . When negotiator 2 reports truthfully, the mechanism picks the disagreement bundle  $(o_x, y_1)$ , and both  $o_x$  and  $y_1$  would prevail in the ratification stage when voted individually. Suppose negotiator 2 instead reports  $M_2$ , in which case the mechanism would pick the logrolling bundle  $(x_1, y_2)$ . In the ratification stage, negotiator 2 would veto the unacceptable alternative  $x_1$ , making the final outcome  $(o_x, y_2)$ . Namely, negotiator 2 would gain by misreporting in the announcement stage. With similar logic, all logrolling mechanisms can be shown to be manipulable under issue-wise voting. Hence, there is no dominant strategy incentive compatible and efficient direct mechanism with veto rights under issue-wise voting.

The general impossibility of truthfully eliciting negotiators' private information under issue-wise voting underlines the importance of jointly resolving the two issues. In particular, bundling alternatives from different issues allows the negotiators to trade favors, which our analysis reveals to be manifested by the logrolling bundles. Consequently, to achieve dominant strategy incentives together with efficiency, it is paramount that the ratification stage only allows for voting on proposed bundles as a whole.

## 6.5 RECONCILIATION WITH THE NEGATIVE RESULTS IN BAYESIAN SETTINGS

The influential work of MS is an important milestone in showing the difficulty of efficient trade in bargaining problems with asymmetric information. It is useful to discuss the underlying factors that are absent in the MS model, which may account for the possibility results in our model. Briefly, the mechanism design problem in MS concerns a bilateral trade between a buyer and a seller, who have private information about their valuations of a good. The mechanism has two components: the probability of trade,  $p$ , and the transfer,  $x$ , both of which are functions of the traders' reports. If no trade occurs, then  $x = p = 0$  (the outside option), and so both traders receive zero utility. The utility functions are  $U_b = v_b p - x$  for the buyer and  $U_s = x - v_s p$  for the seller, where the valuations  $v_b, v_s$  are the traders' private information.

"Budget balancedness" is automatically satisfied in our setup, and so, "budget imbalance" is not the driving force for our possibility result. The buyer (seller) prefers lower (higher) transfers in MS and it is a priori uncertain whether a transfer leading to a mutually beneficial trade exists. Moreover, the quasi-linear utility functions in MS also satisfy the monotonicity and the quid pro quo assumptions. Despite all these similarities, the impossibility of MS is in agreement with our results (particularly the impossibilities we refer in Section 2 and Section 6.2) because the MS model translates as a single-issue mediation problem in our setup, where the transfer is the issue with uncertain gains. Efficiency in

MS implies that the probability of efficient trade is generically either 0 or 1, depending on whether or not the buyer's valuation is higher than the seller's valuation. This means that probability of trade cannot be considered as a second issue since we require the second issue to have at least as many alternatives as the main issue.

What is needed for a possibility is a new issue with a large set of efficient alternatives as in the case of issue  $Y$  in our model. To provide an illustration of the above points, in the following example we offer a simple adaptation of the MS setup in our model and demonstrate how one can overcome the impossibility by adding an extra issue:

**Example 3 (Possibility in the augmented MS framework):**

Suppose that the seller and the buyer now negotiate not only over the terms of trade but also over the division of a unit surplus. We refer to the latter as issue  $Y$ . The valuations of the good to the buyer and the seller are  $v_b$  and  $v_s$ , respectively. We assume that each negotiator knows her valuation and believes that the opponent's valuation is distributed over  $[0, 1]$  with some probability distribution. The mediator privately solicits the traders' valuations and recommends a quadruple  $(p, x, y_s, y_b)$ , where  $p$  denotes the probability of trade,  $x$  is the transfer, and  $y_s$  and  $y_b$  are respectively the seller's and the buyer's share of the unit surplus. The preferences of the two traders are as follows:  $U_b = pv_b - x + u_b(y_b)$  and  $U_s = x - pv_s + u_s(y_s)$ . For simplicity, suppose that  $u_b(y) = u_s(y) = y$  and each trader has only two types,  $v_b, v_s \in \{0.2, 0.6\}$ .

Efficiency implies that  $p = 1$  if  $v_b \geq v_s$ ,  $p = 0$  if  $v_s < v_b$ , and  $y_b + y_s = 1$ . Individual rationality implies that the traders' utilities are nonnegative. Therefore, the following mechanism is strategy-proof, efficient, and individually rational:<sup>61</sup>

	$v_b = 0.6$	$v_b = 0.2$
$v_s = 0.6$	$p = 1$ $x = 0.6$ $y_s = 0.3$ $y_b = 0.7$	<i>No trade</i> $y_s = 0.5$ $y_b = 0.5$
$v_s = 0.2$	$p = 1$ $x = 0.4$ $y_s = 0.5$ $y_b = 0.5$	$p = 1$ $x = 0.2$ $y_s = 0.7$ $y_b = 0.3$

## 7 RELATED LITERATURE

The law and economics literature on settlement negotiations under asymmetric information is extensive.<sup>62</sup> Our approach is fundamentally different from this literature both conceptually and methodologically. In a settlement negotiation, communications between parties revolve around evidence, rule of law, and witnesses, and when negotiations fail, trial is generally the next step. By contrast, we study what is referred as *facilitative mediation* in which the goal is not to determine who is right or who has a stronger case, but rather to explore mutually acceptable resolutions. In this type of mediations, mediators never invite

<sup>61</sup>The seller of type  $v_s = 0.2$  has no incentive to mimic type  $v_s = 0.6$ . This is true because the seller's payoff under truth-telling (which is 0.7 regardless of the buyer's type) is higher than or equal to her deviation payoffs 0.7 (if the buyer is of type  $v_b = 0.6$ ) and 0.5 (if the buyer is of type  $v_b = 0.2$ ). Similarly, the seller of type  $v_s = 0.6$  has no incentive to mimic type  $v_s = 0.2$ . Her payoff under truth-telling is either 0.3 (if the buyer is type  $v_b = 0.6$ ) or 0.5 (if the buyer is type  $v_b = 0.2$ ). However, her deviation payoffs are 0.3 regardless of the buyer's type. Symmetric arguments apply for the buyer.

<sup>62</sup>See, for example, Daughety and Reinganum (2017) and Wickelgren (2013) for two comprehensive accounts of this literature.

parties to present their evidences or cases or to allow parties to discuss their interpretation of the law. This solution-oriented approach is what makes facilitative mediation as the de facto dispute resolution method in e-commerce. To the best of our knowledge, we are not aware of any paper, in the economics, law, management, or information systems literature, that investigates the role of incentives in facilitative mediation.

From a modeling perspective, standard settlement negotiation models involve at least one party having private information about some aspects of the case. Parties' strengths determine the outcome of the trial and the value of the outside option for each side. As a direct implication of this modeling choice, settlement negotiation processes may reveal some information about parties' strengths, which would mean updated beliefs and expectations about outside options. A model where parties can influence other parties' beliefs, and so preferences, creates a highly adversarial environment. However, the whole point behind facilitative mediation is to prevent the formation of such environments. In our model, parties' preferences (i.e., acceptable alternatives) do not change with the opponent's private information. Finally, settlement negotiations are generally modeled through the lens of non-cooperative bargaining theory, whereas we model facilitative mediation as a mechanism design problem.

Our model is more comparable to multi-issue bargaining problems with exogenous outside options. Our modeling of separate and joint outside options is in line with how issues are addressed in political bargaining; see, e.g., Chen and Eraslan (2014, 2017) for similar interpretations to us. Mediation has been studied in the traditional bargaining literature with incomplete information, which is primarily based on a cardinal approach. A central question is whether private information prevents the bargainers from reaping all possible gains from trade. The mechanism design approach to this problem was pioneered by Myerson and Satterthwaite (1983), which shows that for a model with transferable utility there is no ex post efficient, individually rational, Bayesian incentive compatible, and budget balanced mechanism when there is uncertainty about whether gains are possible.<sup>63</sup> On the topic of mediation, specifically, there are very few papers, most of which model mediation as a settlement negotiation. For a model featuring a continuum of types, Bester and Warneryd (2006) show that asymmetric information about relative strengths as an outside option in a conflict may render agreement impossible even if there is no uncertainty about the agreement being efficient. In their model, conflict shrinks the pie and agreement on a peaceful settlement is always ex post efficient. Following Bester and Warneryd (2006), Hörner et al. (2015) compare the optimal mechanisms with two types of negotiators under arbitration, mediation, and unmediated communication. They show that there is no ex post efficient and Bayesian incentive compatible mechanism: the optimal mechanism is necessarily inefficient. Compte and Jehiel (2009) consider bargaining problem where outside options are private but correlated, and parties have a veto right. They show that inefficiencies are inevitable whatever the exact form of correlation, which resonates with the negative result in our benchmark model of single-issue mediation.

Obtaining a possibility result in our model hinges crucially on the availability of (at

---

<sup>63</sup>The MS impossibility crucially depends on types being independent. Subsequently, it was shown that efficient trade may be possible when types are correlated (e.g., Gresik (1991) and McAfee and Reny (1992)).

least) a second issue. Linking multiple decisions/issues to overcome welfare and incentive constraints has been a useful tool in many economic applications such as bundling of goods by a monopolist (e.g., McAfee et al. 1989), agency problems (e.g., Maskin and Tirole 1990), and logrolling in voting (e.g., Wilson 1969). A common insight in these approaches is based on applying a law of large numbers theorem to ensure that truth telling incentives are restored in a sufficiently large market. In this vein, Jackson and Sonnenschein (2007) show that by linking different issues in many situations, including the bilateral bargaining setting of MS, it is possible to achieve outcomes that are approximately efficient in an approximately incentive compatible way as the number of issues goes to infinity.<sup>64</sup> In contrast with these approaches, we establish efficiency in dominant strategies with only two issues in an application where the number of potential issues is inherently limited.

With some caveats, a dispute resolution problem can also be interpreted as a type of fair division problem involving indivisible items. Logrolling mechanisms allow one negotiator to effectively reduce the set of possible outcomes to a shortlist, from which the other negotiator makes her favorite selection. In that sense, logrolling mechanisms are reminiscent of the well-known biblical rule of divide-and-choose, which has been extensively studied in fair cake-cutting problems. Two advantages of a logrolling mechanism relative to divide-and-choose is that it is strategy-proof (whenever preferences satisfy *quid pro quo*) and its outcome is independent of the ordering of the negotiators. More generally, the fair division literature almost exclusively focuses on fairness and efficiency issues due to inherent incompatibilities with strategy-proofness similar to those in the multi-unit assignment context; see, e.g., Brams and Taylor (1996).

Assignment problems have proved useful in achieving strategy-proofness and efficiency via non-dictatorial mechanisms in a number of applications. In this context, ordinal mechanisms are well known to achieve better incentive properties than their cardinal contenders in these problems.<sup>65</sup> In early work, Zhou (1990) showed that no cardinal mechanism is strategy-proof, efficient, and symmetric. By contrast, ordinal mechanisms such as the random priority, are well known to attain the three properties. In two-sided one-to-one and many-to-one matching problems, however, a stable mechanism can be strategy-proof only for one side of the market (see e.g., Roth and Sotomayor 1990).<sup>66</sup>

In assignment/matching problems, an agent's outside option is private consumption whereas in our model it creates an externality on the other negotiator, e.g., whenever either negotiator chooses to exercise her outside option by vetoing the proposal, the other negotiator is automatically compelled to also exercise her outside option. When outside options do not exist and the issues are discrete, our setting roughly resembles a type of multi-unit assignment problem (e.g., course allocation) in which only certain assignments

<sup>64</sup>Jackson et. al. (2020) derives a similar conclusion in a noncooperative bilateral bargaining game setting.

<sup>65</sup>For example, the most prominent cardinal mechanism in the context of unit-assignment problems (possibly allowing for stochastic assignments), the competitive equilibrium from equal incomes solution (Hylland and Zeckhauser 1979), is not strategy-proof. This difficulty of achieving strategy-proofness is generally attributed to the tension with efficiency since cardinal mechanisms achieve stronger welfare properties (e.g., maximization of utilitarian welfare) than ordinal mechanisms.

<sup>66</sup>Similar to the literature on linking decisions discussed below, a common method of circumventing these impossibilities is to resort to large market arguments by allowing for the number of participants and resources to grow. Such methods are obviously inapplicable in the context of mediation.

are feasible.<sup>67</sup> As discussed in Section 1.1, some of the existing point allocation-based ODR mechanisms are akin to course-bidding mechanisms that have been identified as deficient in this literature. Nonetheless, the multi-unit assignment setting provides little reason to remain optimistic for positive results. The literature contains a series of papers that show impossibility results. The main result of this literature is that the only strategy-proof and efficient mechanisms are serial dictatorships; e.g., see Pápai (2001), Klaus and Miyagawa (2002), and Ehlers and Klaus (2003).<sup>68</sup> Clearly, dictatorship mechanisms have little appeal in a dispute resolution situation.<sup>69</sup>

Absent outside options, our model also resembles a voting setting where a voting scheme aggregates individual preferences (Black 1948). This type of voting domains allows to overcome the Gibbard-Satterthwaite impossibility, and the famous median voter theorem states that the majority-rule voting system that selects the Condorcet winner (i.e., the outcome most preferred by the median voter) is strategy-proof; see Moulin (1980) for a classic generalization of this result. The constrained shortlisting mechanism can be viewed as similar to a Condorcet winner in the sense that it recommends the median logrolling bundle when the median is mutually acceptable for both negotiators and the closest logrolling bundle to it when it is not. Nevertheless, this connection is superficial as our model differs in several ways from a voting framework. In these voting models, there are several voters whose bliss point (peak value) is their private information, whereas in our model there are two agents (the negotiators) whose peaks in each issue are publicly known. What is private information here consists of two negotiators' outside options, which have no analogue in a voting model. Consequently, there is no clear way to adopt such voting schemes in our setup, as they would violate individual rationality. Moreover, the above analogy between the two types of models applies only when each issue is considered separately, since negotiators' underlying joint preferences over bundles in our two-issue model are not necessarily single-peaked.<sup>70</sup>

A novelty of our approach that distinguishes it from the literature on fair division, market design, and voting is that we do not ask agents to report their preferences. Instead, we maintain the view that a dispute is solvable so long as the underlying preference structure allows for it. This not only frees us from further complications related to the choice of a

---

<sup>67</sup>Suppose there are two agents, each of whom needs to be assigned two objects, one from each of two sets  $A$  and  $B$ , where an alternative in issue  $X$  (respectively  $Y$ ) represents a specific pair of objects from set  $A$  (respectively  $B$ ) that must be assigned simultaneously. Suppose, for example, set  $A$  contains three objects in the order of decreasing desirability,  $a$ ,  $b$ , and  $c$ , where  $a$  and  $c$  are in unit supply and  $b$  has two copies. Then issue  $X$  can be viewed as consisting of the following object pairings  $X = \{(a, c), (b, b), (c, a)\}$ . That is, if one agent gets  $a$ , the other must get  $c$ , and  $b$  cannot be assigned together with any other object.

<sup>68</sup>Results continue to be negative even with stochastic mechanisms (Kojima, 2009). In the course allocation context, two notable contributions that identify non-dictatorship mechanisms are Sönmez and Ünver (2010) and Budish (2011). The former paper argues for eliciting bids from students together with ordinal preferences over courses and then using a Gale-Shapley mechanism where bids are interpreted as course priorities. The mechanism is strategy-proof only if the bids are treated as exogenously given. The latter paper proposes an approximately efficient mechanism that is strategy-proof in a large market.

<sup>69</sup>Worse still, dictatorships violate individual rationality in our mode, i.e., such recommendations will be vetoed in equilibrium. A constrained dictatorship where one negotiator maximizes her welfare among the set of mutually acceptable outcomes would satisfy individual rationality, but such a mechanism is easily manipulable.

<sup>70</sup>An alternative view could be based on a multi-issue voting setting. However, in multidimensional voting models where people vote on several issues, a main conclusion is that strategy-proofness effectively requires each dimension to be treated separately in the sense that each dimension should admit its own generalized median voter schemes. Our strategy-proofness result, by contrast, depends critically on having more than one dimension and relies heavily on leveraging the exchangeability between the two issues.

suitable preference reporting language, but is also consistent with some of the ODR practices as well as the recommendation systems used in ecommerce.

Finally, with the hope of arriving at possibility results, there is a tradition of searching for strategy-proof mechanisms in restricted economic environments that make it possible to escape Arrow-Gibbard-Satterthwaite impossibilities. Well-known examples include VCG mechanisms (Vickrey 1961, Groves 1973, and Clarke 1971) for public goods and private assignment with transfers; the uniform rule (Sprumont 1991) for the distribution of a divisible private good under single-peaked preferences; generalized median-voters (Moulin 1980); proportional-budget exchange rules (Barberà and Jackson 1995) that allow for trading from a finite number of prespecified proportions (budget sets); deferred acceptance (Gale and Shapley 1962) and top trading cycles (Shapley and Scarf 1974, Abdülkadiroglu and Sönmez 2003) and hierarchical exchange and brokerage (Pápai 2001, Pycia and Ünver 2017). We also add to this literature by introducing and characterizing an entirely new class of strategy-proof and efficient mechanisms.

## 8 CONCLUSION

Mediation is a preferred alternative dispute resolution method thanks to the cost-effectiveness, speed, and convenience it affords to all parties involved. The need for structured and rigorous mediation protocols in practice has often been stressed by researchers and practitioners alike. Online dispute resolution platforms are often based on automation and rely on mechanized negotiation protocols. However, existing dispute resolution protocols fail to account for the incentives faced by disputants. Taking a foundational market design approach to this problem, we sought systematic mechanisms for delivering consistent, transparent, and objective recommendations while giving proper incentives to the disputants to be truthful. Without putting any restrictions on preferences, we considered mechanisms that have a simple preference reporting language; negotiators only report their bargaining ranges (i.e., least acceptable alternatives) in the main issue. It turns out that complementing the main issue with a second one—a piece of advice often voiced by pioneers in the field—is key to achieving strategy-proof, efficient, and individually rational mechanisms. Any such mechanism belongs to the family of logrolling mechanisms, which require that the mediator’s recommendation must always be a logrolling bundle (a bundle that complements a more preferred alternative in one issue with a less preferred alternative from the other) when a mutual agreement is feasible. A sufficient and necessary condition for strategy-proofness is the quid pro quo property of preferences that necessitates the alternatives in the second issue to be interesting enough relative to those in the main issue. The constrained shortlisting mechanism is the central member within the characterized class and makes recommendations as close to the median logrolling bundle as possible.

Our approach can also be viewed as a novel attempt to marry the two distinct literatures of bargaining and assignment. Although the design of facilitative mediation protocols has not been previously considered in the former, this literature emphasizes the tensions due to private information and outside options in mechanism design with transferable utility. The latter literature offers blueprints for designing robust protocols in assignment problems that often arise in practice. The multiple-assignment nature of the problem at hand in our

study, however, is less than encouraging in light of the abundance of negative results in that literature. Our analysis confirms these challenges in that possibility results in our framework are also elusive unless the outside options in the two issues are treated asymmetrically. We argued that ordinal mechanisms coupled with strategy-proofness can help obtain detail-free and genuinely simple protocols for mediating disputes. Notwithstanding our emphasis on ordinality, the framework developed in this paper can accommodate both transferable and nontransferable utility settings since we do not directly elicit preferences.

While it would be premature to conclude that logrolling mechanisms are ready-to-use protocols for immediate practical applications, our theoretical analysis may help shed light on the fundamental forces at work when efficiency is sought together with robust incentives. An interesting open question is how to incorporate full preference elicitation from negotiators into the mechanism design problem. Further research is needed on this front since allowing negotiator types to also include preferences would readily give rise to a more sophisticated preference reporting language than ours as well as new incentive and welfare considerations. Although we leave this direction for future investigation, we contend that the class of mechanisms characterized here would constitute an ideal starting point for developing more sophisticated dispute resolution protocols.

## 9 APPENDIX

**Proof of Theorem 1:** Suppose that  $\Lambda$  is regular and the mediation mechanism  $f$  is strategy-proof, efficient, and individually rational.

**The case where  $\ell < j$ :** Individual rationality and regularity imply  $f_{\ell,j}^X = o_X$ . Then regularity and efficiency require  $f_{\ell,j} = (o_X, y)$  for some  $y \in Y$ . By strategy-proofness and monotonicity, we must have  $f_{\ell',j} = (o_X, y)$  for all  $\ell' < j$ . Similarly,  $f_{\ell,j'} = (o_X, y)$  for all  $\ell < j'$ . Fixing  $j$  (and  $\ell$ ) and applying the same argument for all remaining rows and columns yields  $f_{\ell,j} = (o_X, y)$  whenever  $\ell < j$ .

**The case where  $\ell \geq j$ :**

**Lemma 1 (WARP).** *If  $x, x' \in A(\theta) \cap A(\theta') \neq \emptyset$ ,  $x \neq x'$  and  $x = f_{\theta}^X$ , then  $f_{\theta'}^X \neq x'$ .*

*Proof.* Let  $\theta = (\theta_1, \theta_2)$  and  $\theta' = (\theta'_1, \theta'_2)$  be two type profiles that correspond to the lower half of the matrix form of  $f$  (i.e.,  $A(\theta) \cap A(\theta') \neq \emptyset$ ). Suppose for a contradiction that there are two distinct alternatives  $x, x' \in A(\theta) \cap A(\theta')$  such that  $x = f_{\theta}^X$  and  $f_{\theta'}^X = x'$ . Suppose, without loss of generality, that  $\theta$  and  $\theta'$  correspond to different rows and columns (if they are on the same row or column, then we can just skip this step with bundle  $b$ ). Start from the type profile that corresponds the higher row; namely the profile where negotiator 1's type is more accepting. Suppose, without loss of generality, that  $\theta'$  is the higher row type profile. Let  $b = f(\theta_1, \theta'_2)$  be the bundle that is on the same column with  $\theta'$  and on the same row with  $\theta$ . All three bundles,  $b$ ,  $f(\theta)$ , and  $f(\theta')$ , are acceptable by both types of negotiator 1 because  $f$  is individually rational and  $\theta'_1$  is more accepting than type  $\theta_1$ . Strategy-proofness implies that  $b R_1 f(\theta')$  for all  $R_1 \in \Lambda(\theta_1)$ . This comparison is also true for all  $R_1 \in \Lambda(\theta'_1)$  since preferences are consistent. Similarly, we must have  $f(\theta') R_1 b$  for all  $R_1 \in \Lambda(\theta'_1)$  since  $f$  is strategy-proof. The last two comparisons imply that  $b = f(\theta')$  because preferences over bundles are antisymmetric. By repeating the symmetric arguments for negotiator 2 and recalling that  $b = f(\theta')$  and both  $x, x'$  are acceptable by types  $\theta_2$  and  $\theta'_2$ , we conclude that these two bundles (i.e.,  $f(\theta)$  and  $f(\theta')$ ) must be the same, contradicting our presumption that  $x$  and  $x'$  are distinct alternatives.  $\square$

**Lemma 2 (Existence of  $t$ ).** *There exists an injective order-reversing function  $t : X \rightarrow Y$  such that  $f_{k,k} = (x_k, t(x_k))$  for every  $k = 1, \dots, m$ .*

*Proof.* Row and column  $k$  correspond to the preference profile  $(\theta_1^{x_k}, \theta_2^{x_k})$  where the only mutually acceptable alternative in issue  $X$  is  $x_k$ . Therefore, for any  $1 \leq k \leq m$ , efficiency and individual rationality of  $f$  and regularity of preferences imply  $f_{k,k}^X = x_k$  and  $f_{k+1,k}^X \in \{x_k, x_{k+1}\}$  whenever  $k \neq m$ . Therefore, strategy-proofness of  $f$ , monotonicity of preferences, and WARP imply that  $f_{k+1,k} \in \{f_{k,k}, f_{k+1,k+1}\}$ .

Next, we claim that  $f_{k+1,k+1}^Y \theta_1^Y f_{k,k}^Y$  for each  $k = 1, \dots, m-1$ : If it is true, then we are done with the proof of Lemma 2 by setting  $t(x_k) = f_{k,k}^Y$  for all  $k$ , including  $t$  being injective because  $\theta_1^Y$  is transitive and irreflexive. To prove the last claim, take any  $k$  satisfying  $1 \leq k \leq m-1$ . If  $f_{k+1,k} = f_{k+1,k+1}$ , then strategy-proofness and monotonicity of preferences of negotiator 1 require that  $f_{k+1,k+1}^Y \theta_1^Y f_{k,k}^Y$  because otherwise negotiator 1 would profitably deviate by declaring his type as  $\theta_1^{x_k}$  rather than  $\theta_1^{x_{k+1}}$  as she would be getting better alternative in issue  $X$  and better or the same alternative in issue  $Y$ . On the other hand, if  $f_{k+1,k} = f_{k,k}$ , then strategy-proofness and monotonicity of preferences of negotiator 2 require that  $f_{k,k}^Y \theta_2^Y f_{k+1,k+1}^Y$ , which implies  $f_{k+1,k+1}^Y \theta_1^Y f_{k,k}^Y$  as the negotiators' preferences over the alternatives in issue  $X$  are diametrically opposed. This completes the proof.  $\square$

Therefore, a strategy-proof, efficient, and individually rational  $f$  implies an injective order-reversing function  $t : X \rightarrow Y$  and a nonempty set of bundles  $B^t = \{(x, t(x)) | x \in X\}$ , which constitutes the set of bundles on the main (first) diagonal. For any  $1 \leq j \leq \ell \leq m$  let  $B_{j\ell}^t = \{(x_k, t(x_k)) \in B^t \mid j \leq k \leq \ell\}$  denote the bundles on the main diagonal between row  $j$  and  $\ell$ .

**Construction of  $\succeq$ :** Take any type profile  $\theta = (\theta_1^{x_\ell}, \theta_2^{x_j})$  where  $1 \leq j \leq \ell \leq m$ . We say  $f_{\ell,j}^X \succeq x$  whenever  $x \in X_{j\ell}$ . WARP implies that  $\succeq$  is antisymmetric and reflexive, but not necessarily complete.

**Lemma 3.** *The binary relation  $\succeq$  is transitive. That is, for any triple  $x, x', x'' \in X$  where  $x \succeq x'$  and  $x' \succeq x''$ , we have  $x \succeq x''$ . Furthermore, for all  $1 \leq j < \ell \leq m$ ,  $f_{\ell,j} = (x_{X_{j\ell}}^*, t(x_{X_{j\ell}}^*)) \in B_{j\ell}^t$  where  $x_{X_{j\ell}}^* = \max_{X_{j\ell}} \succeq$ .*

*Proof.* Suppose for a contradiction that  $f_{\ell,j} \notin B_{j\ell}^t$ . By efficiency and individual rationality  $f_{\ell,j}^X \in \{x_j, x_{j+1}, \dots, x_\ell\} = X_{j\ell}$ . Let  $f_{\ell,j}^X = x_k$  for some  $j \leq k \leq \ell$ . Lemma 2 shows that  $f_{k,k}^X = x_k$ , and thus by WARP we must have  $f_{s,r}^X = x_k$  for all  $k \leq s \leq \ell$  and  $k \leq r \leq j$ . Consider  $f_{k,j}$ : By strategy-proofness of  $f$  and monotonicity of the preferences, we must have  $f_{k,j} = f_{k,k}$ . Similar arguments imply  $f_{\ell,j} = f_{k,j}$ . Hence,  $f_{\ell,j} = f_{k,k} \in B_{j\ell}^t$  and  $f_{\ell,j}^Y = t(f_{\ell,j}^X)$ .

Now, suppose for a contradiction that there exists three distinct  $x, x', x'' \in X$  such that  $x \succeq x'$ ,  $x' \succeq x''$ , and  $x'' \not\succeq x$ . Let  $a, b$ , and  $c$  in  $B^t$  denote  $(x, t(x))$ ,  $(x', t(x'))$ , and  $(x'', t(x''))$  respectively. Also suppose, without loss of generality, that  $a$  appears above bundle  $b$  and  $b$  appears above bundle  $c$  on the main diagonal. Similar to the arguments above, strategy-proofness of  $f$ , monotonicity of preferences and WARP imply that for some  $\ell, j$ , and  $k$ ,  $f_{\ell,j} = a$ ,  $f_{k,j} = c$ , and  $f_{k,\ell} = b$  (see Figure 7). Therefore,  $f$  selects  $b$  at entry  $(k, \ell)$  while both  $x'$  and  $x''$  are mutually acceptable, but selects  $c$  at entry  $(k, j)$  while these alternatives are still mutually acceptable, contradicting with WARP. Finally, by construction of  $\succeq$  we know that  $f_{\ell,j}^X = x_{X_{j\ell}}^* \succeq x$  for all  $x \in X_{j\ell}$ , which completes the proof.  $\square$

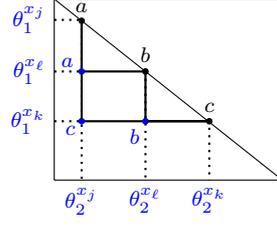


Figure 7

**Proof of Theorem 2:**

*Proof of ‘if’:* Suppose that a regular preference map  $\Lambda$  satisfies quid pro quo. By Definition 2 there exists an injective order-reversing function  $t : X \rightarrow Y$  and a partial order  $\succeq_t$  over  $X$  such that  $\succeq_t \in \Pi_\Lambda$ . Define the mediation mechanism  $f^{\succeq_t}$  by using the partial order  $\succeq_t$  as follows:

$$f_{\ell,j}^{\succeq_t} = \begin{cases} (x_{X_{j\ell}}^*, t(x_{X_{j\ell}}^*)), & \text{if } j \leq \ell \\ (o_X, y), & \text{otherwise} \end{cases}$$

where  $x_{X_{j\ell}}^* = \mathbf{max}_{X_{j\ell}} \succeq_t$  and  $y \in Y$ .

First note that  $X_{j\ell}$  is a connected subset of  $X$  for all  $1 \leq j \leq \ell \leq m$ , and  $\succeq_t$  is a semilattice for all connected subsets of  $X$ . Thus,  $\mathbf{max}_{X_{j\ell}} \succeq_t$  uniquely exists. Next, we prove that  $f^{\succeq_t}$  is individually rational, efficient, and strategy-proof. The mechanism  $f^{\succeq_t}$  never suggests an unacceptable alternative, and thus, it is individually rational by the regularity of preferences.

To show efficiency, first consider the type profile where both negotiators deem all alternatives acceptable in issue  $X$ , i.e.,  $(\theta_1^{x_m}, \theta_2^{x_1})$ . Let  $f^{\succeq_t}$  propose a bundle  $b = (x_b, t(x_b))$  from the set  $B^t = \{(x, t(x)) \mid x \in X\}$  at that profile. If instead the negotiators receive another bundle from  $B^t \setminus \{b\}$  at that profile, then one of the negotiators would certainly get worse off. Suppose for a contradiction that there is another bundle  $(x_a, t(x_a)) = a \in B^t \setminus \{b\}$  that Pareto dominates  $b$ . Because  $f^{\succeq_t}$  suggests  $b$ , we must have  $x_b = \mathbf{max}_X \succeq_t$ , and so  $x_b \succeq_t x_a$ . Moreover, because  $\Lambda$  satisfies quid pro quo there must exist a negotiator  $i$  where  $x_a \theta_i x_b$  and  $b R_i a$  for all admissible  $R_i$ , and the comparison is strict for some  $R_i$  as admissible preferences are consistent and bundles  $a$  and  $b$  are distinct. The last statement contradicts the presumption that  $a$  Pareto dominates  $b$ .

Again at that type profile, i.e.,  $(\theta_1^{x_m}, \theta_2^{x_1})$ , if the negotiators had a bundle with the outside option in issue  $X$ , rather than  $b$ , then both negotiators would be worse off because of the **BR** property. Finally, if the negotiators had any other bundle, say  $c = (x_c, y)$  where  $y \neq t(x_c)$ , which is neither from the set  $B^t$  nor a bundle with the outside option in issue  $X$ , then one of the negotiators would certainly get worse off. To prove this point, suppose for a contradiction that the bundle  $c$  Pareto dominates  $b$ . Because  $x_b = \mathbf{max}_X \succeq_t$ , it must be the case that  $x_b \succeq_t x_c$ . Suppose, without loss of generality, that  $x_c \theta_1 x_b$ . Since  $\Lambda$  satisfies quid pro quo and  $x_c \theta_1 x_b$ , we must have  $b R_1 (x_c, t(x_c))$  for all consistent  $R_1$ 's. There are three exhaustive cases regarding the value of  $y$  and we consider all in turn:

First consider the case where  $t(x_c) \theta_1^Y y$ . In this case, monotonicity implies that  $(x_c, t(x_c)) R_1 (x_c, y) = c$  for all admissible  $R_1$ 's, and thus by transitivity we have  $b R_1 (x_c, t(x_c)) R_1 c$  for all admissible  $R_1$ , where the comparison is strict for some  $R_1$  as admissible preferences are consistent and bundles  $b$  and  $c$  are distinct. The last statement contradicts the presumption that bundle  $c$  Pareto dominates  $b$ . On the other hand, if  $y \theta_1^Y t(x_b)$ , then by monotonicity  $b = (x_b, t(x_b)) P_2 (x_c, t(x_b)) P_2 (x_c, y) = c$  for all admissible  $R_2$ , contradicting the presumption that bundle  $c$  Pareto dominates  $b$ . Finally, if  $t(x_b) \theta_1^Y y \theta_1^Y t(x_c)$ , then the bundle  $(x_c, y)$  cannot Pareto dominate the bundle  $b$  because  $\Lambda$  satisfies quid pro quo, which proves our initial claim that  $b$  is not

Pareto dominated.

Thus, no other bundle makes one negotiator better off without hurting the other when both of the negotiators deem all alternatives acceptable. We can directly apply the same logic to all type profiles that the negotiators deem less alternatives acceptable. Finally, for those type profiles where there is no mutually acceptable alternative in issue  $X$ , in which case the mechanism suggests  $(o_X, y)$  for some  $y \in Y$ , any other bundle will include an alternative that is unacceptable in issue  $X$  by at least one of the negotiators who will veto the proposal. Thus, by regularity, at least one negotiator would be worse off if  $f^{\succeq_t}$  had been proposing something other than  $(o_X, y)$ . Hence, the mechanism  $f^{\succeq_t}$  is efficient.

Next, we prove that the mechanism  $f^{\succeq_t}$  is strategy-proof, but first establish some facts about the structure of this mechanism. If  $a = f_{\ell,j}^{\succeq_t}$  and  $b = f_{r,s}^{\succeq_t}$  are two bundles (i.e., bundle  $a$  appears on row  $\ell$  and column  $j$  whereas bundle  $b$  appears on row  $r$  and column  $s$ ), then we say bundle  $a$  appears above (below) bundle  $b$  whenever  $\ell < r$  ( $\ell > r$ ). Likewise, we say bundle  $a$  appears the right (left) of bundle  $b$  if  $j > s$  ( $j < s$ ).

Given the mediation mechanism  $f^{\succeq_t}$  and a bundle  $a$  that appears on the main diagonal (i.e.,  $a = f_{k,k}^{\succeq_t}$  for some  $k \in M$ ) define  $V(a)$  to be the **value region of bundle  $a$** , which is the submatrix of  $[f_{\ell,j}^{\succeq_t}]_{(\ell,j) \in M^2}$  excluding all the rows lower than row  $k$  and all the columns higher than column  $k$ . Namely,  $V(a) = [f_{\ell,j}^{\succeq_t}]_{(\ell,j) \in (M^k, M_k)}$  where  $M^k = \{k, \dots, m\}$  and  $M_k = \{1, \dots, k\}$ . Furthermore, if bundle  $b = f_{r,r}^{\succeq_t}$  appears on the main diagonal with  $r \in M$  and  $r > k$ , then  $V(a) \cap V(b) = [f_{\ell,j}^{\succeq_t}]_{(\ell,j) \in (M^r, M_r)}$  where  $M^r = \{r, \dots, m\}$ . In Figure 8, the value region of bundle  $a$  is region  $I$  and  $III$ , the value region of bundle  $b$ ,  $V(b)$ , is region  $II$  and  $III$ , and  $V(a) \cap V(b)$  is region  $III$ .

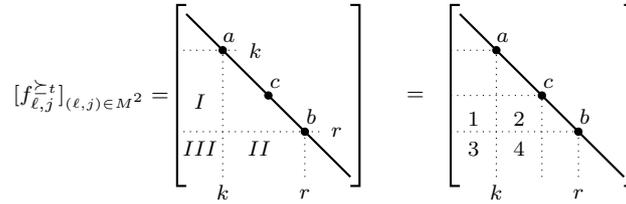


Figure 8

**Lemma 4.** For the mediation mechanism  $f^{\succeq_t}$  and for any two bundles  $a, b \in B^t$ ,

- (i)  $a$  never appears outside of its value region  $V(a)$ ,
- (ii)  $a$  and  $b$  both never appear in  $V(a) \cap V(b)$ , and
- (iii) if both  $a$  and  $b$  appear on the same column (or row), where  $a$  is above  $b$  (or  $a$  is on the left of  $b$ ), then on the main diagonal, bundle  $a$  appears above bundle  $b$ .

*Proof.* Let  $(x_a, t(x_a)) = a = f_{k,k}^{\succeq_t}$  and  $(x_b, t(x_b)) = b = f_{r,r}^{\succeq_t}$  be two distinct bundles for some  $k, r \in M$ . The first claim follows directly from the construction of  $f^{\succeq_t}$ , the fact that  $t$  is one-to-one and that  $x_a \notin X_{j\ell}$  for any  $j \leq \ell < k$  and  $k < j \leq \ell$ . Transitivity of  $\succeq_t$  implies the second claim but deserves a proof. Suppose first that  $a$  and  $b$  appear on the same column in region  $III$ , say column  $s$ , and  $a$  is located above bundle  $b$  on this column, namely  $a$  is on row  $r_a$  and  $b$  is on row  $r_b$  where  $r \leq r_a < r_b \leq m$ . Starting from column and row  $r$  (i.e., from bundle  $b$ ) as we move from column  $r$  to column  $s$  along the row  $r$ , transitivity of  $\succeq_t$  and the fact that the set  $X_{jr}$  is getting larger as  $j$  increases from  $r$  to  $s$  imply that the first components of the bundles on the row  $r$ , which includes the bundle  $f_{r,s}^{\succeq_t}$ , are either ranked higher than  $x_b$  with respect to  $\succeq_t$  or equal to  $x_b$ . Now starting from column  $s$  and row  $r$  (i.e., the bundle  $f_{r,s}^{\succeq_t}$ ) and move toward row  $r_a$  along column  $s$ . Transitivity of  $\succeq_t$  and the fact that the set  $X_{s\ell}$  is getting larger as  $\ell$  increases from  $r$  to  $r_a$  imply

that the first component of the bundle on row  $r_a$  and column  $s$  (i.e., the bundle  $a$ ) is ranked higher than  $x_b$  with respect to  $\succeq_t$ . Namely,  $x_a \succeq_t x_b$  must hold.

Continue iterating from where we left off with the same logic. Starting from column  $s$  and row  $r_a$  (i.e., the bundle  $a$ ) as we move from row  $r_a$  to  $r_b$  along the column  $s$ , first components of the bundles, including the bundle at row  $r_b$  (i.e.,  $b$ ) are either ranked above  $x_a$  or equal to  $x_a$ . Thus, we must have  $x_b \succeq_t x_a$ , contradicting the fact that  $\succeq_t$  is antisymmetric and bundles  $a$  and  $b$  are distinct. If bundle  $b$  is above bundle  $a$  on column  $s$ , then we start the iteration from  $f_{k,k}^{\succeq_t} = a$ . Therefore,  $a$  and  $b$  cannot appear on the same column in region *III*. Symmetric arguments suffice to show that they cannot appear on the same row in region *III* either.

Therefore, suppose that  $a$  and  $b$  appear on different rows and columns. With similar arguments as above, if we start iteration from  $f_{r,r}^{\succeq_t} = b$  and go left on the same row and then go down to bundle  $a$  in region *III*, we conclude that  $x_a \succeq_t x_b$ . However, when we start iteration from  $f_{k,k}^{\succeq_t} = a$  and move down the same column and then go left to bundle  $b$  in region *III*, we conclude that  $x_b \succeq_t x_a$ , which yields the desired contradiction. Hence, either bundle  $a$  or  $b$ , whosever first component is ranked first with respect to  $\succeq_t$ , may appear in region *III*, but not both.

The proof of condition (iii) uses (ii). Suppose for a contradiction that  $a$  and  $b$  appear on the same column  $s$ , where  $b$  is above  $a$  (i.e.,  $r_b < r_a$ ) and  $a$  appears above  $b$  on the main diagonal. In Figure 8,  $a$  and  $b$  can appear on the same column with  $r_b < r_a$  only in region *III*, which contradicts what we just proved above. We can make symmetric arguments for rows as well.  $\square$

We are now ready to show that the mechanism  $f^{\succeq_t} = [f_{\ell,j}^{\succeq_t}]_{(\ell,j) \in M^2}$  is strategy-proof. Consider, without loss of generality, deviations of negotiator 1 only. If  $\ell < j$ , then  $A(\theta_1^{\ell}, \theta_2^{x_j}) = \emptyset$ . Negotiator 1 may receive a different bundle by deviating to a type that is represented by a higher (numbered) row, say  $\theta_1^{x_k}$  where  $k > \ell$ .  $A(\theta_2^{x_j})$  is fixed because negotiator 2's type is fixed. Because the negotiators' preferences over issue  $X$  are diametrically opposed and  $f^{\succeq_t}$  is individually rational, the alternative in issue  $X$  at type profile  $(\theta_1^{x_k}, \theta_2^{x_j})$  will be unacceptable for negotiator 1's true type,  $\theta_1^{\ell}$ . Thus, by the **BR** property, negotiator 1 has no profitable deviation from a type profile  $(\theta_1^{\ell}, \theta_2^{x_j})$  with  $\ell < j$ .

On the other hand, if  $\ell = j$ , then negotiator 1 can deviate to (i) a lower row and receive  $(o_x, y)$ , which is worse than  $f_{\ell,\ell}^{\succeq_t} = (x_\ell, t(x_\ell))$  by **BR**, or (ii) a higher row and receive a bundle that suggests an unacceptable alternative in issue  $X$ . Thus, the **BR** property implies that negotiator 1 has no profitable deviation in that case either.

Finally, suppose that  $j < \ell$ . Let  $c = (x_c, t(x_c)) \in B_{j,\ell}^t$  denote the bundle negotiators get under truthful reporting. If negotiator 1 deviates to a row where  $f^{\succeq_t}$  takes the value  $(o_x, y)$ , then he clearly is worse off, by **BR** property. If he deviates to a lower numbered row and receives, say, bundle  $a = (x_a, t(x_a)) \in B^t \setminus \{c\}$ , then  $a$  appears above bundle  $c$  on the first diagonal, by the third condition of Lemma 4. In this case, it is true that  $x_c \succeq_t x_a$  because  $f^{\succeq_t}$  suggests  $c$  at some type profile where both  $a$  and  $c$  are acceptable and  $f^{\succeq_t}$  always selects the bundle with the best first component. Moreover, by the construction of  $f^{\succeq_t}$  and the fact that  $a$  appears above bundle  $c$  on the first diagonal we have  $x_a \theta_1 x_c$ . Since  $\Lambda$  satisfies quid pro quo, we must have  $c R_1 a$  for all admissible  $R_1$ . Thus, there is no profitable deviation for negotiator 1 by declaring a lower numbered row and getting  $a$  instead of  $c$ .

However, if negotiator 1 declares a higher numbered row and gets a different bundle  $b \in B^t \setminus \{c\}$  (see in Figure 8), then by the third condition of Lemma 4 bundle  $c$  must be located on the first diagonal above bundle  $b$ . As it is clearly visible in Figure 8, Lemma 4 implies that negotiator 1's true preferences must give him the bundle  $c$  in region 1 or 2 and the deviation bundle  $b$  must be in region 3 or 4 because they cannot coexist in region 3 or 4. However, bundle  $b$  includes alternative  $x_r$  from issue  $X$ , which is an unacceptable alternative for all types that lie above row  $r$ , including

negotiator 1's true type. Thus, by the **BR** property, negotiator 1 has no profitable deviation in that case either. Hence,  $f^{\succeq t}$  is strategy-proof.

**Proof of 'only if':** Now suppose that  $\Lambda$  is regular and there exists a mediation mechanism  $f$  that is strategy-proof, efficient, and individually rational. By Theorem 1 we know that there exists an injective order-reversing function  $t : X \rightarrow Y$ , a partial order  $\succeq$  on  $X$  and  $y \in Y$  such that

$$f = f_{\ell,j}^{\succeq} = \begin{cases} (x_{X_{j\ell}}^*, t(x_{X_{j\ell}}^*)), & \text{if } j \leq \ell \\ (o_x, y), & \text{otherwise} \end{cases}$$

where  $x_{X_{j\ell}}^* = \mathop{\max}_{X_{j\ell}} \succeq$ . Then we need to prove that  $\succeq \in \Pi_\Lambda$ , and thus  $\Lambda$  satisfies quid pro quo.

To prove that  $\succeq$  and  $t$  satisfy the first part of Definition 2, let  $x_\ell, x_j \in X$  be two distinct alternatives for some  $j, \ell \in M$  and  $x_\ell \succeq x_j$ . Suppose, without loss of generality, that  $x_j \theta_1 x_\ell$ , namely  $j < \ell$ . Recall the construction of  $\succeq$  in the proof of Theorem 1: we define  $x_\ell \succeq x_j$  if  $f_{\ell,j} = (x_\ell, t(x_\ell))$ . Strategy-proofness of  $f$  and consistency of preferences require that  $f_{\ell,j} R_1 f_{j,j}$ , or equivalently  $(x_\ell, t(x_\ell)) R_1 (x_j, t(x_j))$  for all admissible  $R_1 \in \Lambda(\theta_1)$  and  $\theta_1 \in \Theta_1$  satisfying  $x_j, x_\ell \in A(\theta_1)$ , as required by part (i). To prove part (ii) suppose for a contradiction that there is some  $y \in Y$  with  $t(x_\ell) \theta_1^Y y \theta_1^Y t(x_j)$  such that  $(x_j, y)$  Pareto dominates  $(x_\ell, t(x_\ell)) = f_{\ell,j}$ . Because both of these bundles are acceptable at the profile  $(\theta_1^{x_\ell}, \theta_2^{x_j})$ , the existence of such bundle, i.e.,  $(x_j, y)$ , contradicts with the presumption that  $f$  is efficient.

We now prove that  $\succeq$  and  $t$  satisfy the second part of Definition 2. First recall that all sets of the form  $X_{j\ell}$  with  $1 \leq j \leq \ell \leq m$  designate all the connected subsets of  $X$ . By Theorem 1 we already know that every doubleton  $\{x, x'\} \subseteq X_{j\ell}$  has a least upper bound in  $X_{j\ell}$ , which is  $x_{X_{j\ell}}^*$ , and thus the poset  $(S, \succeq)$  is a semilattice for all connected subset  $S$  of  $X$ . Hence,  $\succeq \in \Pi_\Lambda$ , and thus  $\Lambda$  satisfies quid pro quo. ■

## References

- [1] Ali, S. F.. (2018). Court mediation reform: Efficiency, confidence and perceptions of Justice. Edward Elgar Publishing.
- [2] Abdülkadiroglu, A., and Sönmez, T. (2003). School choice: A mechanism design approach. *American Economic Review*, 93(3), 729-747.
- [3] Ausubel, L. M., Cramton, P., and Deneckere, R. J. (2002). Bargaining with incomplete information. *Handbook of game theory with economic applications*, 3, 1897-1945.
- [4] Barberà, S. (1977). The manipulation of social choice mechanisms that do not leave too much to chance. *Econometrica* 45:1573-1588.
- [5] Barberà, S., and Jackson, M. O. (1995). Strategy-proof exchange. *Econometrica*, 63(1): 51-87.
- [6] Bellucci, E., and Zeleznikow, J. (2005). Developing Negotiation Decision Support Systems that support mediators: a case study of the Family Winner system. *Artificial Intelligence and Law*, 13(2), 233-271.
- [7] Bester, H., and Warneryd, K. (2006). Conflict and the social contract. *Scandinavian Journal of Economics*, 108(2), 231-249.
- [8] Black, D. (1948). On the rationale of group decision-making. *Journal of Political Economy*, 56(1), 23-34.
- [9] Brams, S. J., and Taylor, A. D. (1996). *Fair Division: From cake-cutting to dispute resolution*. Cambridge University Press.
- [10] Budish, E. (2011). The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. *Journal of Political Economy*, 119(6), 1061-1103.
- [11] Budish, E., Cachon, G. P., Kessler, J. B., and Othman, A. (2017). Course match: A large-scale implementation of approximate competitive equilibrium from equal incomes for combinatorial allocation. *Operations Research*, 65(2), 314-336.
- [12] Chen, Y. and Eraslan, H. (2014). Rhetoric in legislative bargaining with asymmetric information. *Theoretical Economics*, 9(2), pp.483-513.

- [13] Chen, Y. and Eraslan, H. (2017). Dynamic agenda setting. *American Economic Journal: Microeconomics*, 9(2), pp.1-32.
- [14] Clarke, E. H. (1971). Multipart pricing of public goods. *Public choice*, 11(1), 17-33.
- [15] Compte, O., and Jehiel, P. (2009). Veto constraint in mechanism design: inefficiency with correlated types. *American Economic Journal: Microeconomics*, 1(1), 182-206.
- [16] Damaska, M. (1975). Presentation of Evidence and Fact finding Precision, *University of Pennsylvania Law Review*, 1083-1106.
- [17] Daughety, A. F., and Reinganum, J. F. (2017). Settlement and trial. *The Oxford Handbook of Law and Economics*, 3, 229-246.
- [18] Ehlers, L., and Klaus, B. (2003). Coalitional strategy-proof and resource-monotonic solutions for multiple assignment problems. *Social Choice and Welfare*, 21(2), 265-280.
- [19] Fisher, R. U., and Ury, W. (1981). *Getting to Yes: Negotiating Agreement Without Giving In*. A Perigee Book/Penguin Group.
- [20] Gale, D., and Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1), 9-15.
- [21] Gresik, T. A. (1991). Ex ante incentive efficient trading mechanisms without the private valuation restriction. *Journal of Economic Theory*, 55(1), 41-63.
- [22] Groves, T. (1973). Incentives in teams. *Econometrica*, 617-631.
- [23] Habuka, H., and Rule, C. (2017). The promise and potential of online dispute resolution in Japan. *International Journal of Online Dispute Resolution*, 4, 74.
- [24] Hadfield, G., K. (2000). The Price of Law: How the Market for Lawyers Distorts the Justice System. *Michigan Law Review*, 98(4), 953 - 1006.
- [25] Hylland, A., and Zeckhauser, R. (1979). The efficient allocation of individuals to positions. *Journal of Political Economy*, 87(2), 293-314.
- [26] Hörner, J., Morelli, M., and Squintani, F. (2015). Mediation and peace. *The Review of Economic Studies*, 82(4), 1483-1501.
- [27] Jackson, M.O. and Sonnenschein H.F. (2007): Overcoming Incentive Constraints by Linking Decisions. *Econometrica*, 75 (1), 241 - 258.
- [28] Jackson, M. O., Sonnenschein, H., Xing, Y., Tombazos, C., and Al-Ubaydli, O. (2020). The efficiency of negotiations with uncertainty and multi-dimensional deals. Working Paper.
- [29] Kagel, J. H., and Roth, A. E. (2016). *The handbook of experimental economics* (Vol. 2). Princeton University Press.
- [30] Kelly, J.S. (1977). Strategy-proofness and social choice functions without single-valuedness. *Econometrica* 45: 439-446.
- [31] Klaus, B., and Miyagawa, E. (2002). Strategy-proofness, solidarity, and consistency for multiple assignment problems. *International Journal of Game Theory*, 30(3), 421-435.
- [32] Kojima, F., (2009). Random assignment of multiple indivisible objects. *Mathematical Social Sciences*, 57(1), pp.134-142.
- [33] Krishna, A., and Ünver, M. U. (2008). Research note—improving the efficiency of course bidding at business schools: Field and laboratory studies. *Marketing Science*, 27(2), 262-282.
- [34] LaFree, G., and Rack, C. (1996). The Effects of Participants' Ethnicity and Gender on Monetary Outcomes in Mediated and Adjudicated Civil Cases, *Law and Society Review*, 30 (4): 767-798.
- [35] Li, S. (2017). Obviously strategy-proof mechanisms. *American Economic Review*, 107(11), 3257-87.
- [36] Lodder, A., and Thiessen, E. (2003). The role of artificial intelligence in online dispute resolution. In *Workshop on online dispute resolution at the international conference on artificial intelligence and law*, Edinburgh, UK.
- [37] Malhotra, D., and Bazerman, M. H. (2008). *Negotiation genius: How to overcome obstacles and achieve brilliant results at the bargaining table and beyond*. Bantam.
- [38] Maskin, E., and Tirole, J. (1990). The principal-agent relationship with an informed principal: The case of private values. *Econometrica*, 379-409.
- [39] McAfee, R. P., and Reny, P. J. (1992). Correlated information and mechanism design. *Econometrica*, 395-421.
- [40] McAfee, R. P., McMillan, J., and Whinston, M. D. (1989). Multiproduct monopoly, commodity bundling, and correlation of values. *The Quarterly Journal of Economics*, 104(2), 371-383.
- [41] Milgrom, P. (2011). Critical issues in the practice of market design. *Economic Inquiry*, 49(2), 311-320.

- [42] Moulin, H. (1980). On strategy-proofness and single peakedness. *Public Choice*, 35(4), 437-455.
- [43] Myerson, R. B., and Satterthwaite, M. A. (1983). Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 29(2), 265-281.
- [44] Pápai, S. (2001). Strategyproof and nonbossy multiple assignments. *Journal of Public Economic Theory*, 3(3), 257-271.
- [45] Pycia, M., and Troyan, P. (2021). A theory of simplicity in games and mechanism design. University of Zurich, Department of Economics, Working Paper, (393).
- [46] Pycia, M., and Ünver, M. U. (2017). Incentive compatible allocation and exchange of discrete resources. *Theoretical Economics*, 12(1), 287-329.
- [47] Pycia, M., and Ünver, M. U. (2020). Arrowian efficiency and auditability in the allocation of discrete resources. CEPR Discussion Paper Series, (DP15377).
- [48] Roth, A. and Sotomayor, M. (1990). *Two-sided Matching: A Study in Game-theoretic Modeling and Analysis*. Econometric Society Monographs, Cambridge University Press.
- [49] Rubinstein, A. (2012). *Lecture notes in microeconomic theory: the economic agent*. Princeton University Press.
- [50] Rule, C. (2014). Technology and the future of dispute resolution. *Dispute Resolution Magazine*, 21, 4.
- [51] Shapley, L., and Scarf, H. (1974). On cores and indivisibility. *Journal of Mathematical Economics*, 1(1), 23-37.
- [52] Sönmez, T. and Ünver, M.U., (2010). Course bidding at business schools. *International Economic Review*, 51(1), pp.99-123.
- [53] Sprumont, Y. (1991). The division problem with single-peaked preferences: a characterization of the uniform allocation rule. *Econometrica*, 59, 509-519.
- [54] Thiessen, E. M., and Loucks, D. P. (1992). Computer Assisted Negotiation of Multi-objective Water Resources Conflicts 1. *JAWRA Journal of the American Water Resources Association*, 28(1), 163-177.
- [55] Thiessen, E. M., Loucks, D. P., and Stedinger, J. R. (1998). Computer-assisted negotiations of water resources conflicts. *Group Decision and Negotiation*, 7(2), 109-129.
- [56] Thiessen, E. M., Miniato, P., and Hiebert, B. (2012). *ODR and eNegotiation (Chapter 16): Online dispute resolution: theory and practice: a treatise on technology and dispute resolution*. Eleven International Pub..
- [57] Thomson, W. (2016). Fair allocation, in *The Oxford Handbook of Well-Being and Public Policy*, ed. by M. D. Adler and M. Fleurbaey, Oxford University Press.
- [58] Tyler, T. R., and Huo, Y. J. (2002). *Russell Sage Foundation series on trust. Trust in the law: Encouraging public cooperation with the police and courts*. New York, NY, US: Russell Sage Foundation.
- [59] Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *The Journal of Finance*, 16(1), 8-37.
- [60] Zhou, L. (1990). On a conjecture by Gale about one-sided matching problems. *Journal of Economic Theory*, 52(1), 123-135.
- [61] Wahab, M. S. A., Katsh, M. E., Rainey, D. (2012). *Online dispute resolution: theory and practice: a treatise on technology and dispute resolution*. Eleven International Pub..
- [62] Wickelgren, A. L. (2013). *Law and economics of settlement*. In *Research Handbook on the Economics of Torts*. Edward Elgar Publishing.
- [63] Wilson, R. (1969). An axiomatic model of logrolling. *The American Economic Review*, 59(3), 331-341.