AN EXPERIMENT ON TWO-PERSON SOCIAL CHOICE

by
SONNUR BAŞ

Submitted to the Graduate School of Social Sciences
in partial fulfilment of
the requirements for the degree of Master of Arts

Sabancı University
July 2023

AN EXPERIMENT ON TWO-PERSON SOCIAL CHOICE

Approved by:

Prof. Özgür Kıbrıs .....
(Thesis Supervisor)

Assoc. Prof. Mehmet Barlo .....

Asst. Prof. Aylin Ece Çiçek .....

Date of Approval: July 18, 2023

ABSTRACT

AN EXPERIMENT ON TWO-PERSON SOCIAL CHOICE

SONNUR BAŞ

ECONOMICS M.A. THESIS, JULY 2023

Thesis Supervisor: Prof. Özgür Kıbrıs

Keywords: social choice, experimental economics, decision making, arbitration, compromise

This study empirically compares four different mechanisms recommended in the arbitrator assignment process: Alternate Shortlisting (ASL), Compromise Rule of k Names (CRk), Gradual Veto (GV), and Shortlisting Mechanism (SL). These four mechanisms differ according to the number of steps they contain, whether they include a first-mover, and whether they contain vetoes. Data were collected in a computer laboratory environment with four different z-Tree treatments in which 290 undergraduate students from Sabancı University participated. A comparison was made on notions such as Truthfulness, Pareto Optimality, Inequality, First Mover Advantage, and Equal Loss Principle. All performance benchmarks are tested with t-tests and OLS regressions with clustered standard errors by session. We found that the Gradual Veto mechanism slightly outperforms the sequential mechanisms regarding playing truthful behavior, but the same success is not valid in achieving an efficient outcome. SL and ASL are more successful for players staying in their upper-half alternatives. The Gradual Veto mechanism creates a more equal environment for matched players. Among sequential mechanisms, ASL and CRk have lower First Mover Advantage than SL. Also, we found some significant relationship between participant characteristics and round outcomes.

ÖZET

İKI KIŞILIK SOSYAL SEÇIM ÜZERINE BIR DENEY

SONNUR BAŞ

EKONOMİ YÜKSEK LİSANS TEZİ, TEMMUZ 2023

Tez Danışmanı: Prof. Dr. Özgür Kıbrıs

Anahtar Kelimeler: sosyal seçim, deneysel ekonomi, karar verme, arabuluculuk, uzlaşma

Bu çalışma, arabulucu atama sürecinde önerilen dört farklı mekanizmayı ampirik olarak karşılaştırmaktadır: Alternatif Kısa Listeleme (ASL), k Adın Uzlaşma Kuralı (CRk), Kademeli Veto (GV) ve Kısa Listeleme (SL) Mekanizması. Bu dört mekanizma, içerdikleri adım sayısına, sıralı olup olmamasına ve veto içerip içermemesine göre farklılık göstermektedir. Sabancı Üniversitesi'nden 290 lisans öğrencisinin katıldığı dört farklı z-Tree deneyi ile bilgisayar laboratuvarı ortamında veriler toplanmıştır. Doğruluk, Pareto Optimalliği, Eşitsizlik, İlk Oyuncu Avantajı ve Eşit Kayıp İlkesi gibi kavramlar üzerinde bir karşılaştırma yapılmıştır. Tüm performans kıyaslamaları, t-testi ve oturum bazında kümelenmiş standart hatalı OLS regresyonu ile test edilmiştir. Doğrucu davranma konusunda vetolu simültane mekanizmanın sıralı mekanizmadalardan biraz daha iyi performans gösterdiğini ancak aynı başarının Pareto Optimal bir sonuca ulaşmada geçerli olmadığını gördük. SL ve ASL, oyuncuların üst yarı alternatiflerinde kalabilmesi adına daha başarılıdır. GV mekanizması, eşleşen oyuncular için daha eşit bir ortam yaratır. Sıralı mekanizmalar arasında ASL ve CRk, SL'den daha düşük ilk oyuncu avantajına sahiptir. Ayrıca, katılımcı özellikleri ile tur sonuçları arasında bazı anlamlı ilişkiler bulduk.

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my advisor, Özgür Kıbrıs, for his invaluable guidance, support, and motivation throughout this journey. What I learned from him is not limited to my thesis subject. His experiences and sincere advice will also guide the rest of my life.

I am also profoundly grateful to the jury members, Mehmet Barlo and Aylin Ece Çiçek, whose valuable contributions elevated the quality of my thesis.

I sincerely thank the lab manager, Aydın Yörük, for his support and for ensuring a smooth lab experience. I would also like to acknowledge the Sabancı University Economics faculty for their constant support and contribution to my education. Also, I am fortunate to have had the rare privilege of friendship with my colleague and classmate, Bakican. His companionship throughout this journey has been invaluable.

I am grateful to my family, mother Durdane, father Mustafa, and sister İlknur, for their endless love, encouragement, and support. Without them, none of this would have been possible.

To my family
Who endlessly supports me

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

A group of decision-makers trying to reach a compromise on a set of alternatives is a common problem. Applications range from project evaluations and committee decisions to the selection of arbitrators in conflict resolution. In arbitration, the disputing agents select a neutral third party, an arbitrator, to make a binding decision. In dispute resolution, arbitrators are assigned to cases in a decentralized way that takes into account the parties' preferences (Afacan, Anbarci, and Kıbrıs 2022). According to the United States Arbitration Act, if in the agreement provision be made for a method of naming or appointing an arbitrator or arbitrators or an umpire, such method shall be followed. For this purpose, many mechanisms have been implemented or proposed as arbitrator selection methods. The literature on arbitrator selection has proposed various mechanisms which may vary according to the number of steps they include, whether they are simultaneous or sequential, and whether they contain a veto. To add to this literature, in this thesis, We created an experimental test environment inspired by this problem and the proposed mechanisms.

The reason for suggesting new mechanisms and comparisons is to create the "ideal" mechanism in the joint decision and compromising process. While making such comparisons, the previous literature employed some notions like Reaching Pareto Optimal outcome, Inequality, First Mover Advantage (for Sequential games), Playing Truthfully, Reaching Equilibrium outcome, Median Satisfaction Test (De Clippel, Eliaz, and Knight 2014), Unanimity Compromise Rule (Hurwicz and Sertel 1999), Equal Loss Principle (Chun 1988), Fallback Bargaining Rule (Brams and Kilgour 2001), and Fairness and Reciprocity. These notions, which measure whether the social choice is a better and fair outcome, are heavily used in many different contexts and appear in the literature of arbitrator selection mechanisms as expected. Considering these notions, I will empirically compare four different mechanisms according to their success in delivering a desirable outcome.

This paper aims to compare and contrast four different mechanisms, where the

setting consists of two agents and a set of five alternatives. The first mechanism is the Gradual Veto Mechanism (GV), where agents simultaneously veto one alternative until 0 or 1 alternative remains. The second mechanism is Shortlisting Mechanism (SL), where the first-mover agent shortlists three alternatives, and the second-mover agent makes the final decision among shortlisted candidates. The third and fourth mechanisms come from the *Rule of k Names* family proposed by Barberà and Coelho (2010). The third mechanism is the Compromise Rule of k Names (CRk), where the first-mover agent picks a number $k \in \{3, 4, 5\}$. Following this, the second-mover agent chooses to be a "Proposer" or a "Chooser". Depending on this decision, the "Proposer" agent shortlists k alternatives, and the "Chooser" agent makes the final decision among them. Alternate Shortlisting (ASL) is the final mechanism. Here, the first-mover shortlists at least three alternatives. Following this, the second-mover agent chooses to "Accept" or "Reject". If she accepts, she makes the final decision among that proposed set. If she rejects it, she proposes a new set with cardinality one greater than the old one. The first-mover makes the final decision among this new set.

By comparing these four mechanisms, I hope to shed light on the strengths and weaknesses of each mechanism and provide insights into how the collective decision process can be improved. This paper contributes to the literature by testing these four mechanisms that have yet to be empirically compared. Although they have been studied theoretically and favored in some regards, no empirical comparison has been made regarding the aforementioned notions.

We found that the Gradual Veto mechanism outperforms the sequential mechanisms regarding playing truthful behavior, but the same success is not valid in achieving an efficient outcome. SL and ASL are more successful for players staying in their upper-half alternatives. The Gradual Veto mechanism creates a more equal environment for matched players. Among sequential mechanisms, ASL and CRk have lower First Mover Advantage than SL. Also, we found some significant relationship between participant characteristics and round results. We tried to make a performance comparison according to which criteria the mechanism would like to be selected and applied.

This paper comprises four main sections. The first section provides a comprehensive literature review on the two-person social choice and arbitrator selection. The second section briefly defines the four mechanisms subject to the experiment. The third section describes the experimental design, summary statistics of participants, and methodology used in the study. The fourth section begins with the comparison results on the Truthfulness, Pareto Optimality, Median Satisfaction Test, Fallback

Bargaining, Inequality, First Mover Advantage, Equal Loss Principle, Payoffs, Fairness & Reciprocity. Finally, this section also includes a correlational analysis to test whether there is a significant relationship between participant characteristics and strategic behavior. The final section concludes the paper with a summary of the essential findings and implications of the study. Experiment instructions, the end of the study questionnaire, and Sabancı University Ethics Committee (SUREC) approval can be found in the Appendix.

2. LITERATURE REVIEW

The arbitrator behavior and arbitrator selection processes have long been the subject of the economics literature. As one of the earliest empirical studies, Bloom and Cavanagh (1986) looks at the preferences of unions and employers for different arbitrators under New Jersey's Fire and Police Arbitration Law data. Since their estimation results and veto/rank mechanism outputs are close, they conclude that most information about the parties' preferences comes from the vetoed arbitrators and not from the rank order of the non-vetoed arbitrators. The following literature dealt with arbitrator selection in several categories: proposing new mechanisms, creating properties and models to compare mechanisms, and empirically testing proposed mechanisms.

First, if we look at the proposed mechanisms, we come across various mechanisms with and without a veto. Van der Linden (2017), one of the studies on the veto power of a mechanism, presents impossibility results showing that even limited veto power makes many mechanisms of interest manipulable. However, there are many veto mechanisms in the literature, and Laslier, Nunez, and Sanver (2021) propose strike mechanisms as a solution to the classical problem of Hurwicz and Schmeidler (1978)and Maskin (1999), according to which, in two-person societies, no Pareto efficient rule is Nash-implementable. In the strike mechanism, each player simultaneously casts these vetoes, and the mechanism randomly selects one alternative among the non-vetoed ones. Another veto mechanism is the Voting by Alternating Offers and Vetoes (VAOV) mechanism proposed by Anbarci (2006). In the VAOV scheme, the two players take turns to make offers until an alternative is accepted; any offer rejected by a player is taken out of consideration, and if no offer is accepted, the last remaining alternative is the outcome. VAOV and similar mechanisms are essential in this respect: Anbarci (2006) shows that equilibrium outcome sets converge to the Equal Area solution's outcome if the alternatives are distributed uniformly over the comprehensive utility possibility set and as the number of alternatives tends to infinity. (The outcome of the Equal Area solution is the intersection of the Pareto

frontier and the straight line that goes through the disagreement point and cuts S into two equal areas.)

Looking at the sequential mechanisms in the literature, we can start with the Rule of k names proposed by Barberà and Coelho (2010). In the Rule of k names, one of the parties proposes a shortlist, and the other chooses from it (Barberà and Coelho 2022). Then a single individual from outside the committee selects one of the listed names for the office. Later, Barberà and Coelho (2022) proposed three mechanisms based on the Rule of k names. The characteristics of these mechanisms are that they contain few steps, weakly implement the Unanimity Compromise Set, and are robust to the strategic inclusion of candidates. They argue that rules of k names and shortlisting are good methods to achieve compromise but can be improved. They say the presence of the first mover in Shortlisting Mechanism makes unfair treatment too obvious. A comparison in terms of Inequality and First Mover Advantage is required to test this argument empirically. The fact that we will compare two mechanisms from the Rule of k names, namely Compromise Rule of k Names and Alternate Shortlisting, and the Shortlisting Mechanism, helps us to test this argument.

As another branch of social choice theory and arbitrator selection literature, some papers propose and model several properties and notions, which are also employed in this paper. Hurwicz and Sertel (1999) proposed the Majoritarian Compromise related to Unanimity Compromise and Fallback Bargaining notions. The compromise rule can equivalently be interpreted as maximizing the welfare of the worst-off agent when each agent's payoff from an alternative (Kıbrıs and Sertel 2007). This paper examines whether outcomes satisfy the Fallback Bargaining property, in which bargainers begin by indicating their preference rankings over alternatives. They then fall back, in lockstep, to less and less preferred alternatives – starting with first choices, then adding second choices, and so on – until an alternative is found on which all bargainers agree. (Brams and Kilgour 2001). Apart from optimality and efficiency-related notions, literature also considers the inequality among players. Chun (1988) proposes the Equal Loss Principle to check whether agents lose from the best-ranked alternative equally. Furthermore, the Minimal Rawlsian Principle (sometimes called Minimal Satisfaction Test (De Clippel and Eliaz 2012)) ensures that the chosen alternative is in the upper half of both agents' preferences.

Finally, we can start with De Clippel, Eliaz, and Knight (2014) to review studies that empirically test some arbitrator selection mechanisms in the laboratory environment. De Clippel, Eliaz, and Knight (2014) compare Shortlisting Mechanism and Veto-Rank mechanism regarding efficiency and minimal satisfaction test. They conclude that the Veto-Rank mechanism includes a significant pattern of strategic

behavior and that the SL mechanism outperforms VR in terms of efficiency. Also, they look at fairness and reciprocity and try to find some punishment behaviors. They argue that fairness concerns seem to affect the behavior of some participants, whose decisions may be reconciled with a theory of intentions-based reciprocity. With a similar experimental design, Bol, Laslier, and Núñez (2022) evaluate the performance of three bargaining mechanisms: Shortlisting Mechanism, The simultaneous mechanism, where each player simultaneously vetoes two options out of the five available ones, and the outcome is a uniform lottery over non-vetoed options and Gradual Veto mechanism, where at each step and simultaneously, each subject vetoes one option. While comparing the mechanism, they look at the sum of subjects' monetary payoffs (called efficiency in this context), inequality, first mover advantage, and rates of reaching Pareto Optimal & Fallback Bargaining alternatives. They conclude that while the simultaneous mechanism performs poorly both in terms of efficiency and inequality, the gradual vetoes mechanism achieves high levels of efficiency and even reduces inequality compared to the non-symmetric shortlisting mechanism.

This paper synthesizes the work done so far in the arbitrator selection literature. We took some of the mechanisms proposed in the literature and may be superior to others by their solutions. We included the properties usually used in mechanism comparisons. We compare the mechanisms that have yet to be empirically compared in an experimental setting and properties/notions that have not been included in such an empirical analysis. A mechanism from the Rule of k names family has not been compared with the Shortlisting mechanism before. Therefore, we included the Alternate Shortlisting and Compromise Rule of k Names mechanisms proposed by Barberà and Coelho (2022) in the analysis. Although Shortlisting mechanism and Gradual Veto have been compared before, this comparison is incomplete in some respects; for example, it does not include an Equal Loss Principle analysis. We would like to contribute to the mechanism comparison literature by comprehensively analyzing different mechanisms by different properties.

3. EXPERIMENTAL DESIGN

The experimental sessions were held at Sabancı University, Faculty of Arts and Social Sciences. Sabancı University uses the SONA system where registered students can earn Research Credits (RC) as bonus points in some courses in exchange for participating in psychology, marketing, and economics experiments. The study announcement is published on the SONA website, and students can sign up voluntarily. 170 undergraduate students from the SONA pool participated. Additionally, students of some ECON mass courses were invited in exchange for the opportunity to earn bonus points in their courses. 120 students enrolled in these courses also participated in the experiment. Participants were invited to the Faculty of Arts and Social Sciences Computer Lab. The session began with the instructions being read aloud when all participants arrived at the lab, followed by the treatment played on computer screens with the mouse. When the treatment was completed, the session was completed with the participants filling out a personal questionnaire. The all-inclusive session took approximately 50-55 minutes to complete.

This experiment is registered in The American Economic Association's registry for randomized controlled trials with the RCT ID AEARCTR-0010662

Each mechanism was designed with z-Tree (Fischbacher 2007). The language of the treatments was Turkish, and only students whose native language is Turkish could participate. An even number of students attended the sessions. Which mechanism to play was chosen randomly before the session. Each participant was introduced to a single mechanism, implying that it is a between-subject design. Participants played the chosen mechanism for 40 rounds by randomly matching another participant. It is not a setting where the participants are constantly matched with the same participant because the arbitrator selection and conflict resolution literature is based on a decision-making process without communication and post-history. Each round ended with selecting one of the alternatives $\{a, b, c, d, e\}$. The participants won Experimental Money (EM) corresponding to that alternative in their preference ordering, which varies between 0 EM and 400 EM. One of the 40 earnings was

randomly chosen when the study was completed. This amount was added to the Experimental Money earned from participation, which is 300, fixed. Finally, this summation was converted into Research Credits (RC) for students from the SONA pool and into bonus points for ECON mass course students. Table 3.1 shows the conversion between Experimental Money and Research Credits and bonus points.[1]

Table 3.1 Experimental money, research credits and bonus points conversion

EM From Random Round	Total EM Earned	Total RC Earned	Total Bonus Earned
400	700	3.5	2
300	600	3	1.7
200	500	2.5	1.4
100	400	2	1.1
0	300	1.5	0.8

Upon completing the experimental study, participants filled out a short questionnaire on their demographic information, like age, sex, gender, and birthplace. In addition to these questions, several questions were asked on religion, trust, and political views. They were asked an open-ended and optional question for the religion, *"What are our religious beliefs?"*. For the trust, we employed the question statement frequently used by General Social Survey (GSS) (Davis and Smith 1991) or the World Values Survey (WVS), which is *"Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?"* (Naef 2009). For political view, *"Some people think of political attitudes as being on the "left" or "right". Here is a scale stretching from left to right, with "1" as being the most to the left and "10" being the most to the right. When you think of your political opinions, where would you put yourself on this scale?"* (Lambert 1983). Table 3.2 summarizes the descriptive statistics of the participants. The average age of the participants is 22, and slightly less than half are female. Although this is an economics experiment, there is a diverse profile of participants from different majors. Only 9% of the participants are enrolled in the Economics program. The remaining students are enrolled in departments such as Computer Science, Psychology, Management, and Industrial Engineering. Life satisfaction is about 7 out of 10, trust toward others is 5 out of 10, and participants have located themselves on the 3.8 of the political spectrum on average. These values are quite similar across treatments.

Forty rounds consist of 4 sections. The rules of the game do not change between

[1]When students collect 10 Research Credits, they can use these points to get a bonus. The recommended research credit to be awarded after a 1-hour study is 2.5.

Table 3.2 Summary statistics for the participants to the experiment

	ASL	CRk	GV	SL	All Sample
Participants	78	74	70	68	290
Age	22.5	22.1	21.8	21.9	22.1
Female	38	33	30	38	139
ECON Major	6	6	7	7	26
CS Major	30	27	25	24	106
Life Satisfaction	7	7.2	7.1	6.6	6.7
Trust Toward Others	5.3	5	5	4.9	5
Political View	4.1	3.8	3.4	3.8	3.8

the sections; the same mechanism is still played. The only thing that changes is the preference orderings. Therefore, the participant's earnings assigned to each alternative and the matched participant's earnings assigned to each alternative may change. At the beginning of each round, participants are divided into two groups. One preference ranking was assigned to one half, and another preference ranking was assigned to the other half. During these ten rounds, players were randomly matched with a participant from the other group. Hence, they had the same preference ranking for ten rounds, also the player they matched with had the same preference ranking whole the time. The payoff table describing the participants' preference ordering and corresponding EMs was always shown on the screen. It is a setting with complete information because, in real life, it can be easily assumed that the two agents making such a decision have an idea of both their own preferences and the preferences of the other agent. The preference order, used by De Clippel, Eliaz, and Knight (2014) for the first time is shown in 3.3

Table 3.3 Preference profiles

Pf 1		**Pf 2**		**Pf 3**		**Pf 4**		**EM**
Pl. 1	Pl. 2	Pl. 1	Pl. 2	Pl. 1	Pl. 2	Pl. 1	Pl. 2	
a	e	a	b	a	c	a	e	400
b	d	b	a	b	b	b	c	300
c	c	c	c	c	a	c	a	200
d	b	d	d	d	d	d	b	100
e	a	e	e	e	e	e	d	0

As De Clippel, Eliaz, and Knight (2014) suggested, each preference profile has different characteristics regarding the level of conflict and equilibrium strategies. The first preference profile has a total conflict; the only alternative that does not create inequality between players is *c*. The second preference profile has a partial conflict with the first two-ranked alternatives. Especially in the mechanisms that involve veto, any miscoordination on first-two ranked alternatives might result

in choosing alternative c or worse. The third preference profile also has a partial conflict in the upper half, but this time, there is a compromise point b. The fourth preference profile is added to the analysis for the empirical consistency with De Clippel, Eliaz, and Knight (2014). We have created various orders of playing in which these preference profiles are played. Playing a random order in each session makes controlling for a learning spillover effect across preference profiles easier. The four orders of playing are $Pf1-Pf2-Pf3-Pf4$, $Pf1-Pf3-Pf2-Pf4$, $Pf4-Pf2-Pf3-Pf1$, $Pf4-Pf3-Pf2-Pf1$. In short, we changed the order of $Pf1-Pf4$ and $Pf2-Pf3$.

Participants had a time limit for making each decision, usually no more than 30 seconds. If no decision was made within the time limit, the round was automatically skipped, and its data was not included in the analysis. About 8% of total observations were dropped due to timeout. The instructions provided to participants and survey questions are in the Appendix.

4. MECHANISMS

These mechanisms consist of two agents with some strict preference rankings over a set of five alternatives, namely a, b, c, d, e

4.1 Shortlisting Mechanism

At the start of each round, one of the matched participants will be chosen as Player 1 and the other as Player 2. The game will be played according to this. It is equally likely that being selected as Player 1 or Player 2 by the computer for each round.

- **Step 1:** Player 1 starts the game and offers 3 of 5 alternatives to Player 2.
- **Step 2:** Player 2 chooses one of these three alternatives, and the round is completed by choosing this alternative.

In Shortlisting Mechanism, the truthful strategy is to shortlist the top three ranked alternatives for Player 1 and choose the top-ranked alternative among shortlisted alternatives for Player 2.

4.2 Compromise Rule of k Names (CRk)

At the start of each round, one of the matched participants will be chosen as Player 1 and the other as Player 2, and the game will be played according to this. It is equally likely that being selected as Player 1 or Player 2 by the computer for each round.

- **Step 1:** Player 1 moves first and chooses value $k \in \{3, 4, 5\}$.

- **Step 2:** Once this choice is made public, Player 2 decides whether to act as the Proposer or the Chooser. Hence, the game will follow one of the 2 cases described below depending upon Player 2's decision.
  - If Player 2 decides to act as the Proposer:
    - **Step 3:** Player 2 selects k candidates out of 5 available alternatives. Remember that k is the value that was chosen by Player 1 at Step 1.
    - **Step 4:** Player 1 is informed of Player 2's shortlist. Player 1 chooses the final alternative out of this shortlist.
  - If Player 2 decides to act as the Chooser:
    - **Step 3:** Player 1 selects k candidates out of 5 available alternatives. Remember that k is the value that was chosen by Player 1 in Step 1
    - **Step 4:** Player 2 is informed of Player 1's shortlist. Player 2 chooses the final altenatives out of this shortlist.

4.3 Alternate Shortlisting

At the start of each round, one of the matched participants will be chosen as Player 1 and the other as Player 2. The game will be played according to this. It is equally likely that being selected as Player 1 or Player 2 by the computer for each round.

- **Step 1:** Player 1 proposes a set of minimum three alternatives from the set $\{a, b, c, d, e\}$.
- **Step 2:** Player 2 is informed of Player 1's proposal. Player 2 has two alternatives. Either she Accepts Player 1's proposal or she Rejects it. Hence, the game will follow one of the 2 cases described below depending upon Player 2's decision.
  - If Player 2 accepts:
    - **Step 3:** Player 2 chooses the final alternative out of this proposed set.
  - If Player 2 rejects:
    - **Step 3:** Player 2 proposes an alternative set in which the cardinality must be one more than Player 1's proposal.

    - **Step 4:** Player 1 chooses the final alternative from this new proposed set.

4.4 Gradual Veto

- **Step 1:** Both Players simultaneously veto an alternative from 5 alternatives.
- **Step 2:** Once vetoed alternative/s are eliminated from available set, both Players simultaneously veto an alternative from the remaining alternatives again. This veto process might take several steps until 1 or 0 alternative is available. There are two ways of determining the final outcome depending on the number of remaining alternatives.
    - If exactly one alternative remains:
        - This one alternative is the outcome.
    - If all alternatives have been vetoed, no alternatives left:
        - This means that exactly two alternatives were non-vetoed at the previous stage, and both were vetoed at the very last step. The outcome is a random draw over these two alternatives.

In the Gradual Veto Mechanism, vetoing the worst-ranked alternative among available ones are the truthful strategy.

5. RESULTS

This section starts with the empirical comparison of the mechanisms in terms of Truthfulness, Reaching Equilibrium Outcome, Efficiency, Minimal Satisfaction Test, Fallback Bargaining, Inequality, First Mover Advantage, Equal Loss Principle, and Fairness & Reciprocity. Results are given by categorizing Preference Profiles as being early or late. Early means the first five periods of a Preference Profile, in which participants may experience getting used to the payoff table. Late means the last five periods of a Preference Profile, in which participants might have a meaningful convergence in their strategies. Following the descriptive numbers, I made performance comparisons with different t-tests by separating and not separating according to Preference Profiles. Finally, the correlational analysis of the relationship between participant characteristics and strategic behaviors and outcomes is shown.

5.1 Truthfulness

When comparing mechanisms, one of the most important aspects is truthtelling and strategizing. In this section, we define the decisions for each mechanism step that players must make if they completely stick to their preference orderings. Mechanisms where truthful behavior is not common and players hide their true preferences by acting strategically are unfavorable. Now, here is the list of how we define truthful behavior:

- **ASL**
  - First mover proposes a set that only includes her best 3, 4, or 5 alternatives.
  - Second mover accepts a set if this set only includes her best 3, 4, or 5 alternatives.

- If the second mover rejects and decides to propose a new set, she only proposes the set that includes her best 4 or 5 alternatives.
- If the second mover accepts and chooses an alternative from the proposed set, she only chooses her best-ranked alternative among available ones.
- If the second mover rejects and proposes a new set, the first mover only chooses her best-ranked alternative among available ones.

- **CRk**
  - If the second mover decides to propose, she only proposes the set that includes her best k alternatives.
  - If the second mover decides to choose, the first mover only proposes the set that includes her best k alternatives.
  - If the second mover decides to propose, the first mover only chooses her best-ranked alternative among the available ones.
  - If the second mover decides to choose, she only chooses her best-ranked alternative among the available ones.

- **GV**
  - Players veto their worst-ranked alternatives among available ones.

- **SL**
  - The first mover shortlists her best three alternatives.
  - The second mover chooses her best-ranked alternative among the shortlisted alternatives.

Here, we made this analysis on a decision basis rather than a round basis. It means that we evaluate every decision separately as long as the *if* conditions above are satisfied. We were not interested in whether the outcome was a product of complete truth. And all decisions apart from the above are classified as somehow strategizing. We did not make a separate analysis of the levels of reasoning and strategies. However, there is a significant difference between early decisions and late decisions ($p < 0.01$.). The early decision means the very first decision of the players, usually including a proposal. The late decision means the decisions that end the round, usually choosing the final alternative. The details of the early/late decisions are explained in the Appendix. So, we represent another comparison just interested in the early decisions.

Table 5.1 Two sample t-test results for the truthfulness

		Mean 1	Mean 2	p-value
Pf 1	SL vs GV	0.78	0.77	0.5
	GV vs ASL	0.77	0.61	0***
	ASL vs CRk	0.61	0.55	0.02**
Pf 2	GV vs ASL	0.52	0.52	0.9
	ASL vs SL	0.52	0.48	0.12
	SL vs CRk	0.48	0.44	0.2
Pf 3	GV vs CRk	0.5	0.43	0***
	CRk vs SL	0.43	0.43	0.9
	SL vs ASL	0.43	0.41	0.4
Pf 4	GV vs ASL	0.61	0.56	0.04**
	ASL vs SL	0.56	0.5	0.03**
	SL vs CRk	0.5	0.37	0***
All	GV vs SL	0.6	0.55	0***
	SL vs ASL	0.55	0.52	0.07**
	ASL vs CRk	0.52	0.45	0***

***p <0.01, **p <0.05, *p<0.1

Table 5.1 shows the means of the truthful behavior. Most truthful behavior can be observed in Pf 1. SL has significantly the highest number. Shortlisting the best three alternatives fully coincides with shortlisting other player's worst three alternatives. And the-second mover is pushed to choose the middle alternative, which is also a truthful behavior. In Pf 2, there is no significant difference among mechanisms. In Pf 3, GV has a significant advantage, and other sequential mechanisms are similar. Vetoing the worst-ranked alternative makes meeting at compromise point b easier. In Pf 4, all means are significantly different, and GV has the highest number. Overall, ASL and CRk are falling behind the SL and GV. This analysis includes all decisions. However, there is a significant change in the behavior through a round. As the end of the round approaches, players are already creating the environment to get their higher-ranked alternatives and pushing others to play truthfully to get that alternative. In sequential mechanisms, shortlisting the best-ranked alternative and other's worst alternatives is a strategic behavior. Later, the second mover has to choose the first mover's best-ranked alternative. Now, it is time to decompose the decision type and look at the beginning of the game, where certain strategies might be set.

Table 5.2 shows the comparison of early decision truthfulness. The ordering of mechanisms is generally unchanged, but there is a significant decrease in the means. Shortlisting the best-three alternatives is not common in Pf 2, Pf 3, and Pf 4 for SL. In Pf 3, only 10% of the proposals are a, b, c. In the mechanism where the size

Table 5.2 Two sample t-test results for the early decisions truthfulness

		Mean 1	Mean 2	p-value
Pf 1	SL vs GV	0.73	0.67	0.03**
	GV vs ASL	0.67	0.63	0.3
	ASL vs CRk	0.63	0.59	0.2
Pf 2	GV vs ASL	0.35	0.32	0.3
	ASL vs CRk	0.32	0.24	0.02**
	CRk vs SL	0.24	0.15	0.003***
Pf 3	CRk vs GV	0.31	0.27	0.12
	GV vs ASL	0.27	0.13	0***
	ASL vs SL	0.13	0.1	0.16
Pf 4	GV vs ASL	0.59	0.4	0***
	ASL vs CRk	0.4	0.27	0***
	CRk vs SL	0.27	0.17	0***
All	GV vs ASL	0.46	0.36	0***
	ASL vs CRk	0.36	0.35	0.7
	CRk vs SL	0.35	0.27	0***

***p <0.01, **p <0.05, *p<0.1

of the proposed set can be more than three, the truthful behavior at the beginning is more common. In GV, truthful behavior increases at the end of the game. It means that the very first behavior includes some strategic thinking. Once they create this strategic environment, they veto the worst alternatives later. Overall, SL early decisions have significantly more strategic thinking. This pattern decreases in ASL and CRk, but no significant difference exists. Regarding GV, almost half of the early decisions are truthful, which is significantly more common than other mechanisms.

5.2 Reaching Equilibrium Outcome

From the equilibrium characterizations of Barberà and Coelho (2022) and Bol, Laslier, and Núñez (2022), we identified the equilibrium outcomes for Pf 1, Pf 2, and Pf 3. We did not analyze equilibrium strategies separately since thick best responses and multiple steps make it harder to comment on whether the participants follow equilibrium strategies. Also, as empirically consistent with Bol, Laslier, and Núñez (2022), we did not include analysis on Pf 4. Pf 4 took place in De Clippel and Eliaz (2012) due to Nash-Equilibrium analysis of the Veto-Rank mechanism. Since we do not have that mechanism, we did not include Pf 4 in our equilibrium outcome analysis.

Table 5.3 Two sample t-test results for reaching equilibrium outcome

		Mean 1	Mean 2	p-value
Pf 1	SL vs GV	0.76	0.73	0.31
	GV vs CRk	0.73	0.63	0***
	CRk vs ASL	0.63	0.51	0***
Pf 2	GV vs ASL	0.59	0.46	0***
	ASL vs CRk	0.46	0.4	0.03**
	CRk vs SL	0.4	0.39	0.8
Pf 3	GV vs ASL	0.7	0.7	0.7
	ASL vs CRk	0.7	0.5	0***
	CRk vs SL	0.5	0.2	0***
All	GV vs ASL	0.68	0.55	0***
	ASL vs CRk	0.55	0.51	0***
	CRk vs SL	0.51	0.45	0***

***p <0.01, **p <0.05, *p<0.1

Table 5.3 shows the performances of the mechanisms in reaching the equilibrium outcome. ASL is one of the successful mechanisms in reaching equilibrium outcomes, except Pf 1. In Pf 1, SL has the highest percentage, where %76 of the outcomes are Equilibrium outcomes. Overall, all comparisons are significant, and GV is the most successful. Even though in Pf 1 of SL players can meet at c, which constitutes an equilibrium outcome, SL has the lowest percentage in reaching an equilibrium outcome overall.

5.3 Efficiency

By simply adopting the definition of efficiency, an alternative is considered efficient (Pareto optimal) if there is no feasible alternative that can make some agent better off without making someone else worse off. Table 5.4 shows efficient alternatives for each Preference Profile.

Table 5.4 Efficient alternatives

Pf 1		**Pf 2**		**Pf 3**		**Pf 4**		**EM**
Pl. 1	Pl. 2	Pl. 1	Pl. 2	Pl. 1	Pl. 2	Pl. 1	Pl. 2	
a	**e**	**a**	**b**	**a**	**c**	**a**	**e**	400
b	**d**	**b**	**a**	**b**	**b**	b	**c**	300
c	**c**	c	c	**c**	**a**	**c**	**a**	200
d	**b**	d	d	d	d	d	b	100
e	**a**	e	e	e	e	**e**	d	0

In Preference Profile 1, all alternatives are efficient since there is a total conflict. In Preference Profile 2, only a, b are efficient. Any miscoordination or desire to punish the other player might result in failing efficiency. As discussed in behavioral concerns, attempting to eliminate the alternative worth 400 EM for the other participant to get 400 EM can suddenly drop the round payoff to 200 EM or less. In Preference Profile 3, also c is an efficient alternative. This partial conflict with a compromise point might decrease the risk of reaching an inefficient alternative. In Preference Profile 4, only b, d are inefficient alternatives.

Table 5.5 Percentage (%) of reaching efficient alternative

		Pf 1	Pf 2	Pf 3	Pf 4
ASL	Early 5	100	84	95	75
	Late 5	100	80	96	85
CRk	Early 5	100	75	83	79
	Late 5	100	68	76	75
GV	Early 5	100	49	86	62
	Late 5	100	69	88	51
SL	Early 5	100	79	82	85
	Late 5	100	79	84	81

Table 5.5 shows the percentages of reaching an efficient alternative. At first glance, sequential mechanisms give better results than the mechanism with veto. The most obvious difference is seen in Pf 2 and Pf 4. A strategy to veto the best alternatives of the matched player can lead to inefficient outcomes. The GV mechanism facilitates access to efficient outcomes in Pf 3, where the focal compromise alternative exists. One explanation can be the player who has the power to shortlist moves away from the compromise point.

Table 5.6 compares the percentages of reaching the PO alternative to see whether a mechanism performs significantly better. In Pf 2, SL and ASL outperform CRk, but there is no significant difference between SL and ASL. Also, CRk is giving significantly more successful results compared to GV. In Pf 3, ASL is relatively successful again. This time, GV is significantly better than SL, and SL is significantly better than CRk. The compromise point in Pf 3 has a large effect in the sudden rise of the GV after Pf 2. In Pf 4, there is no significant difference between percentages. In a nutshell, ASL is one of the most efficient mechanisms for all Preference Profiles, followed by SL. On the other hand, CRk is usually falling behind these two. GV is successful only in Pf 3, with the compromise point in the upper half, but by looking at all observations, CRk is significantly better than GV.

Table 5.6 Two sample t-test results for the efficiency

		Mean 1	Mean 2	p-value
Pf 2	ASL vs SL	0.82	0.79	0.12
	SL vs CRk	0.79	0.72	0***
	CRk vs GV	0.72	0.6	0***
Pf 3	ASL vs GV	0.95	0.87	0***
	GV vs SL	0.87	0.83	0.03**
	SL vs CRk	0.83	0.79	0.08*
Pf 4	SL vs ASL	0.83	0.8	0.24
	ASL vs CRk	0.8	0.77	0.19
	CRk vs GV	0.77	0.56	0***
All	ASL vs SL	0.9	0.86	0***
	SL vs CRk	0.86	0.82	0***
	CRk vs GV	0.82	0.76	0***

***p <0.01, **p <0.05, *p<0.1

5.4 Minimal Satisfaction Test

An alternative in the upper half of both players' rankings satisfies the Minimal Satisfaction Test (MST).

Table 5.7 Minimal satisfaction test alternatives

Pf 1		**Pf 2**		**Pf 3**		**Pf 4**		**EM**
Pl. 1	Pl. 2	Pl. 1	Pl. 2	Pl. 1	Pl. 2	Pl. 1	Pl. 2	
a	e	**a**	**b**	**a**	**c**	**a**	e	400
b	d	**b**	**a**	**b**	**b**	b	**c**	300
c	**c**	**c**	**c**	**c**	**a**	**c**	**a**	200
d	b	d	d	d	d	d	b	100
e	a	e	e	e	e	e	d	0

Table 5.7 shows alternatives satisfying MST for each Preference Profile. In Preference Profile 1, MST alternatives are a strict subset of efficient alternatives since only c satisfies MST. In Pf 1, all alternatives are efficient and have the same total payoff. But we need MST analysis to understand whether players are being pushed to their worst-ranked alternatives. In Pf 2, we are expanding the initial efficient set. Miscoordination may result in c being chosen and inefficiency, but it is an alternative that satisfies MST since c is in the upper half for both players. In Pf 3, efficient alternatives and MST alternatives are the same. In Pf 4, only a, c are satisfying MST.

Table 5.8 Percentage (%) of reaching MST alternatives

		Pf 1	Pf 2	Pf 3	Pf 4
ASL	Early 5	43	98	95	69
	Late 5	57	97	96	78
CRk	Early 5	57	91	83	74
	Late 5	68	82	76	71
GV	Early 5	70	85	86	61
	Late 5	77	89	88	49
SL	Early 5	66	91	82	80
	Late 5	85	95	84	81

Table 5.8 shows the percentages of reaching an alternative satisfying MST. In Preference Profile 1, knowing that only c satisfies MST, GV and SL are relatively successful in reaching it. Later in the SL mechanism, the strategy of equalizing and meeting in the middle prevailed. On the other hand, antisymmetric outcomes are reached with higher percentages in CRk and ASL mechanisms. In Pf 2, we can see a sudden increase compared to PO percentages. This increase comes from including c in the desired set. The mechanism by which c is selected with the highest percentage is the Gradual Veto mechanism. This confirms our hypothesis. In this simultaneous veto mechanism, players can act to eliminate the other player's highest-rank alternative. In other mechanisms that include a shortlist, seeing alternatives with higher payoffs motivates them to choose a or b more easily. In Pf 4, the percentage reaching MST alternative is less than the PO alternative since e does not satisfy MST. The least decrease is seen in the Gradual Veto mechanism. This reinforces the hypothesis that veto-containing mechanisms are important in eliminating the other player's highest-ranked alternative, implying that alternative e is already less likely to be selected.

Table 5.9 compares the percentages of reaching the alternatives satisfying MST to see whether a mechanism performs significantly better. To start with Pf 1, SL and GV are doing better. This is the only Preference Profile that ASL is giving the worst success. Only half of the outcomes satisfy MST, meaning that c is selected. ASL is significantly better than all mechanisms in all preference profiles except Pf 1. If we look at the overall results, ASL and SL are the first two again, and GV and CRk are following them. Due to antisymmetric outcomes of ASL in Pf 1, SL is significantly better than ASL in terms of MST.

Table 5.9 Two sample t-test results for the MST

		Mean 1	Mean 2	p-value
Pf 1	SL vs GV	0.76	0.73	0.3
	GV vs CRk	0.73	0.62	0***
	CRk vs ASL	0.62	0.5	0***
Pf 2	ASL vs SL	0.98	0.93	0***
	SL vs CRk	0.93	0.87	0***
	CRk vs GV	0.87	0.87	0.9
Pf 3	ASL vs GV	0.95	0.87	0***
	GV vs SL	0.87	0.83	0.03**
	SL vs CRk	0.83	0.79	0.08*
Pf 4	SL vs ASL	0.81	0.74	0.01**
	ASL vs CRk	0.74	0.72	0.5
	CRk vs GV	0.72	0.55	0***
All	SL vs ASL	0.83	0.8	0.003***
	ASL vs GV	0.8	0.76	0***
	GV vs CRk	0.76	0.75	0.5

***p <0.01, **p <0.05, *p<0.1

5.5 Fallback Bargaining Test

The fallback bargaining Test (FB) gives the highest possible level of utility to the worse-off subject. Bol, Laslier, and Núñez (2022). In this context, the alternatives with the highest rank for the worst-off players are the Fallback Bargaining alternatives. Table 5.10 shows alternatives satisfying FB for each Preference Profile.

Table 5.10 Fallback bargaining alternatives

Pf 1		**Pf 2**		**Pf 3**		**Pf 4**		**EM**
Pl. 1	Pl. 2	Pl. 1	Pl. 2	Pl. 1	Pl. 2	Pl. 1	Pl. 2	
a	e	**a**	**b**	a	c	**a**	e	400
b	d	**b**	**a**	**b**	**b**	b	**c**	300
c	**c**	c	c	c	a	**c**	**a**	200
d	b	d	d	d	d	d	b	100
e	a	e	e	e	e	e	d	0

In Preference Profile 1, like MST, only c is the Fallback Bargaining alternative since the highest rank of the worst-off player can be 3. In Preference Profile 2, the highest rank of the worst-off player is 2, and a, b are FB alternatives. In Pf 3, we can see a difference compared to PO and MST since Fallback Bargaining only admits b. In Pf 3, Fallback Bargaining success can be used to measure the power of compromise.

Table 5.11 shows the percentages of reaching an alternative satisfying Fallback Bar-

gaining.

Table 5.11 Percentage (%) of reaching fallback bargaining alternatives

		Pf 1	Pf 2	Pf 3	Pf 4
ASL	Early 5	43	84	66	69
	Late 5	57	80	73	78
CRk	Early 5	57	75	49	73
	Late 5	68	68	51	71
GV	Early 5	70	49	67	61
	Late 5	77	69	74	49
SL	Early 5	66	79	44	80
	Late 5	85	79	41	81

Fallback Bargaining analysis is a synthesis of PO and MST analyses. Pf 1 is the same as MST, and Pf 2 is the same as PO. In Pf 3, only b satisfies the Fallback Bargaining, the focal compromise outcome that gives the highest payoff. In Pf 3, more than %80 of the observations resulted in a, b, c selected in all mechanisms. Now, it is time to distinguish whether players compromise. In SL, rounds that b is selected are less than half. Having an opportunity to increase the size of the shortlisted set, ASL and CRk give relatively better results than SL in terms of compromise. In Pf 3, the most successful mechanism is the GV.

5.6 Inequality

As measures of Inequality among matched participants, we employed three notions: Average Payoff Difference by Round, First Mover Advantage for sequential mechanisms, and Equal Loss Principle. Table 5.12 shows the absolute value of the payoff difference.

It can be seen that the Gradual Veto mechanism is usually doing a better job in terms of equality. Especially in Pf 3, it facilitates access to alternative b, the only Fallback Bargaining alternative, creating more equal results. Also, there is a common trend for all mechanisms, such that inequality between matched participants decreases in later periods, except Shortlisting Mechanism. In Shortlisting mechanism, only three alternatives are shortlisted, and sometimes shortlisting other participant's worst three alternatives results in punishment. This behavior pattern will be explained in the next sections.

Table 5.12 Average payoff difference by round

		Pf 1	Pf 2	Pf 3	Pf 4
ASL	Early 5	161	84	58	174
	Late 5	115	80	46	166
CRk	Early 5	116	75	67	167
	Late 5	80	68	49	158
GV	Early 5	73	49	39	162
	Late 5	55	69	28	159
SL	Early 5	88	79	77	170
	Late 5	44	79	85	163

Table 5.13 Two sample t-test results for the inequality

		Mean 1	Mean 2	p-value
Pf 1	ASL vs CRk	135	97	0***
	CRk vs SL	97	65	0***
	SL vs GV	65	64	0.87
Pf 2	ASL vs SL	82	79	0.11
	SL vs CRk	79	72	0.01**
	CRk vs GV	72	59	0***
Pf 3	SL vs CRk	81	58	0***
	CRk vs ASL	58	52	0.2
	ASL vs GV	52	34	0***
Pf 4	ASL vs SL	170	167	0.5
	SL vs CRk	167	163	0.3
	CRk vs GV	163	160	0.5
All	ASL vs SL	104	97	0.02**
	SL vs CRk	97	96	0.7
	CRk vs GV	96	77	0***

***p <0.01, **p <0.05, *p<0.1

Table 5.13 compares the absolute value of payoff differences per round to see whether a mechanism creates a significantly equal environment. First, in all Preference Profiles, GV payoffs are significantly more equal. Only SL payoffs in Pf 1 and CRk payoffs in Pf 3 are near GV. In Pf 1, ASL is significantly the most unequal mechanism. Since each sequential alternative has a payoff difference of 100, the difference between the order of the alternatives selected in the ASL mechanism are more than one on average. In Pf 1, CRk is more unequal than SL and GV. In Pf 2, ASL is the most unequal one but this time SL is as unequal as ASL. In Pf 3, compromise point b effect shows itself in the GV inequality, which is significantly the lowest. In Pf 4, there is no significant relationship. Overall, ASL is the most unequal one, SL and CRk are close to each other and GV is significantly more equal.

5.7 First Mover Advantage

Another inequality parameter is the First Mover Advantage. While recommending Rule of k Names family, Barberà and Coelho (2018) suggested that SL mechanism migh have a heavier first mover advantage, but CRk and ASL can decrease it. Table 5.14 shows the payoff difference between the first and second mover. Even though we observe negative first mover advantage sometimes, SL and ASL generally have some FMA. Only in Pf 4, the difference is quite volatile and sometimes favors the second mover.

Table 5.14 Average first mover advantage by round

		Pf 1	Pf 2	Pf 3	Pf 4
ASL	Early 5	16	-2	-9	-25
	Late 5	7	20	0	27
CRk	Early 5	-11	-4	-8	-27
	Late 5	-7	-13	-9	-45
SL	Early 5	-2	11	9	32
	Late 5	-5	11	34	53

Table 5.15 Two sample t-test results for the first mover advantage

		Mean 1	Mean 2	p-value
Pf 1	ASL vs SL	11	-4	0.12
	SL vs CRk	-4	-9	0.5
Pf 2	SL vs ASL	11	9	0.6
	ASL vs CRk	9	-8	0***
Pf 3	SL vs ASL	21	-5	0***
	ASL vs CRk	-5	-9	0.4
Pf 4	SL vs ASL	43	3	0***
	ASL vs CRk	3	-36	0***
All	SL vs ASL	18	5	0.001***
	ASL vs CRk	8	-15	0***

***p <0.01, **p <0.05, *p<0.1

Table 5.15 shows the p-values to see whether a mechanism favors the first mover significantly. In Pf 1 and Pf 2, there is no significant relationship on the top. However, in Pf 3 and Pf 4, SL has a huge first-mover advantage. Overall, in SL, the first-mover completes the rounds with more payoffs than the second-mover compared to the ASL and CRk. This difference is significant. What is interesting here is that, on average, CRk has some negative first-mover advantage, which is significantly

lower than ASL. ASL has an FMA value of 5, which can be interpreted as the most neutral mechanism among sequential ones.

5.8 Equal Loss Principle

An alternative satisfies the Equal Loss Principle if it equalizes the loss of both agents from the top point. In that regard, Table 5.16 shows the alternatives that satisfy ELP.

Table 5.16 Equal loss principle alternatives

Pf 1		**Pf 2**		**Pf 3**		**Pf 4**		**EM**
Pl. 1	Pl. 2	Pl. 1	Pl. 2	Pl. 1	Pl. 2	Pl. 1	Pl. 2	
a	e	a	b	a	c	a	e	400
b	d	b	a	**b**	**b**	b	c	300
c	**c**	**c**	**c**	c	a	c	a	200
d	b	**d**	**d**	**d**	**d**	d	b	100
e	a	**e**	**e**	**e**	**e**	e	d	0

In Pf 1, only c satisfies ELP. In Pf 2, the lower half is giving the same payoff. In Pf 3, focal compromise point b and worst-two ranked alternatives d, e satisfy ELP. In Pf 4, there is no chance that players to equalize their payoffs.

Table 5.17 Percentage (%) of reaching equal loss principle alternatives

		Pf 1	Pf 2	Pf 3	Pf 4
ASL	Early 5	43	16	71	0
	Late 5	57	20	77	0
CRk	Early 5	57	25	66	0
	Late 5	68	32	76	0
GV	Early 5	70	51	81	0
	Late 5	77	31	86	0
SL	Early 5	66	21	61	0
	Late 5	85	21	57	0

Table 5.17 shows the percentages of reaching an alternative satisfying ELP for each mechanism. Usually, reaching ELP outcomes increases in later periods. It can be due to a desire to compromise or to punish first-movers and locate themselves in an equal place. In Pf 2, ELP outcomes are relatively low since pairs can stay in the best-two alternatives a, b and has different payoffs.

Table 5.18 Two sample t-test results for the equal loss principle

		Mean 1	Mean 2	p-value
Pf 1	SL vs GV	0.76	0.73	0.3
	GV vs CRk	0.73	0.63	0***
	CRk vs ASL	0.63	0.5	0***
Pf 2	GV vs CRk	0.41	0.28	0***
	CRk vs SL	0.28	0.21	0.001***
	SL vs ASL	0.21	0.18	0.11
Pf 3	GV vs ASL	0.83	0.74	0***
	ASL vs CRk	0.74	0.71	0.2
	CRk vs SL	0.71	0.59	0***
All	GV vs CRk	0.51	0.41	0***
	CRk vs SL	0.41	0.39	0.18
	SL vs ASL	0.39	0.39	0.7

***p <0.01, **p <0.05, *p<0.1

Table 5.18 shows the p-values of the comparison of ELP outcomes. In Pf 2, Pf 3, and Pf 4, GV is the most successful mechanism that can equalize playoffs. In Pf 1, SL has a higher percentage than GV, but this difference is insignificant. Usually, ASL falls behind the other mechanisms. Looking at all observations, in the GV mechanism, more than half of the rounds are completed with equal payoffs. Other mechanisms are more or less giving the same result.

5.9 Payoffs

Instead of looking at Pareto Optimal alternatives, Bol, Laslier, and Núñez (2022) defines efficiency as the summation of the payoffs. Similarly, looking at the average payoff might be a good idea to capture the total welfare. Table 5.19 shows the average payoffs by rounds.

Table 5.19 Average payoff by round

		Pf 1	Pf 2	Pf 3	Pf 4
ASL	Early 5	200	324	288	244
	Late 5	200	316	291	249
CRk	Early 5	200	299	263	243
	Late 5	200	278	243	236
GV	Early 5	200	252	269	223
	Late 5	200	288	272	206
SL	Early 5	200	308	261	252
	Late 5	200	313	265	251

There is no significant pattern of the change in average payoff between early and late periods. Table 5.20 shows the means and p-values of the differences.

Table 5.20 Two sample t-test results for the average payoff

		Mean 1	Mean 2	p-value
Pf 2	ASL vs SL	320	310	0.03**
	SL vs CRk	310	289	0***
	CRk vs GV	289	270	0.002***
Pf 3	ASL vs GV	289	270	0***
	GV vs SL	270	263	0.16
	SL vs CRk	263	253	0.08*
Pf 4	SL vs ASL	252	247	0.4
	ASL vs CRk	247	240	0.3
	CRk vs GV	240	214	0***
All	ASL vs SL	266	257	0.001***
	SL vs CRk	257	246	0***
	CRk vs GV	246	239	002**

***p <0.01, **p <0.05, *p<0.1

In Pf 2 and Pf 3, ASL significantly produces the highest payoff on average. In Pf 2, all differences are significant, and GV falls behind all other mechanisms. This lower average payoff is related to the vetoing a, b and regressing to c. Overall, all average payoffs are between 266 and 239, significantly different from each other. This result is completely the same as the Efficiency result in Section 5.3. ASL produces the most efficient outcomes and gives the highest payoffs to the participants. GV is the most inefficient one and has the lowest average payoff.

5.10 Fairness & Reciprocity

When the experiment was completed, and the participants were asked their opinions in an off-the-record way, we realized that sometimes they acted with feelings of punishment. For punishment behavior, we tested whether the second player desires to punish when the other player's worst alternatives are offered. This analysis was performed only for ASL, CRk, and SL since the punishment behavior scenarios are similar across sequential mechanisms. We looked at the case where the second player was offered the set containing the worst alternatives for each Pf. Then, we represent the percentages of the second player choosing the alternative that did not bring the highest payoff, which also brings a lower payoff to the first mover. Punishment decision cases are summarized in the Appendix. Table 5.21 shows the punishment

Table 5.21 Punishment behavior

		Mean 1	Mean 2	p-value
Pf 2	ASL vs SL	0.21	0.19	0.7
	SL vs CRk	0.19	0	0***
Pf 3	SL vs ASL	0.3	0.1	0***
	ASL vs CRk	0.1	0	0***
Pf 4	SL vs ASL	0.17	0.1	0.09**
	ASL vs CRk	0.1	0	0***
All	SL vs ASL	0.21	0.15	0.04**
	ASL vs CRk	0.15	0	0***

***p <0.01, **p <0.05, *p<0.1

behavior percentages.

Only in Pf 2 there is no significant difference between SL and ASL. In other preference profiles, there is a significant punishment behavior in SL compared to others. Especially in Pf 3, Pl 1 shortlisting a, d, e or Pl 2 shortlisting c, d, e results in d, e being chosen with 30%. Overall, we do not see any punishment behavior in CRk, and in ASL, punishment behavior is significantly lower than CRk. One explanation can be the following: the mechanisms that allow expanding the set or rejecting and re-proposing might create a less sensitive environment for punishment. However, in SL, the shortlisting only consists of 3 elements, and there is no way to change it. This certainty might increase the desire to punish directly.

5.11 Session Effect: OLS Regression Analysis

So far, we have done all the analysis based on the assumption that sessions do not have an effect on the outcome. Even though which mechanism will be applied to which group is chosen randomly, the sessions might have some effect. For Efficiency, MST, and Inequality, we run three different OLS Regression where errors are clustered by session. The model is as follows:

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \epsilon_i \tag{5.1}$$

Where Y_i is the binary variable that takes the value 1 if the outcome is Efficient (1), the binary variable that takes the value 1 if the outcome satisfies MST (2), and Inequality between matched participants. (3). X_1, X_2, and X_3 are dummy

variables for ASL, CRk, and GV, respectively. ϵ is clustered by session. Table 5.22 summarizes the OLS Regression results.

Table 5.22 OLS regression results for PO, MST, and inequality

	Dependent variable:		
	(1)	(2)	(3)
ASL	0.042^{**}	-0.031	6.787^{***}
	(0.021)	(0.021)	(0.021)
CRk	-0.042	-0.079^{***}	-1.119^{***}
	(0.027)	(0.027)	(0.027)
GV	-0.099^{***}	-0.071^{***}	-20.120^{***}
	(0.016)	(0.016)	(0.016)
Constant	0.861^{***}	0.833^{***}	97.084^{***}
	(0.015)	(0.015)	(0.015)
Note:			*p<0.1; **p<0.05; ***p<0.01

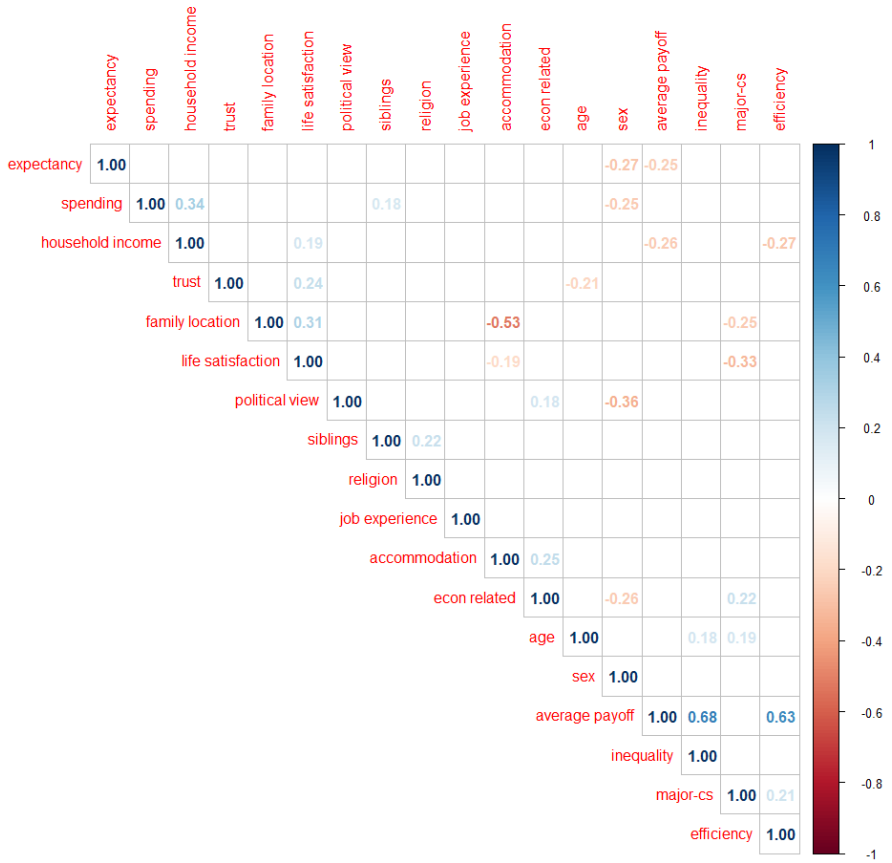
Previous results are consistent with these regression results. For the efficiency, ASL has a significantly higher average, and CRk follows it, but the coefficient is insignificant. GV is significantly the most inefficient alternative. For MST, ASL is the relatively successful one, and CRk is significantly the least successful in this regression. For Inequality, all coefficients are significant. ASL is the most unequal, and GV is the most equal. All these results are the same as the previous sections.

5.12 Participant Analysis

In the en-of-the survey questionnaire, the participants asked for several personal information. Then, we calculated each subject's average success in reaching an efficient outcome, their average payoff, and the inequality they create between matched players. Figure 5.1 represents the results with 0.1 significance level.

We used the variables life expectancy, monthly spending, household income, trust toward others, family location, life satisfaction, political view, number of siblings, the existence of a religious belief, job experience, accommodation type, whether they have an ECON major/minor, age, sex, and whether their major is Computer

Figure 5.1 Correlation matrix of participant variables



Science (CS).

There are not many, but a couple of significant relationships. First, the average payoff decreases as life expectancy increases, and the coefficient is -0.25. There is a similar relationship between household income and reaching an efficient outcome. As household income increases, their success in reaching PO outcomes decreases. In contrast, higher ages have a positive relationship with reaching efficient outcomes. And finally, being a Computer Science student is positively correlated with reaching efficient outcomes as well.

6. CONCLUSION

We have taken two sequential mechanisms, the Alternate Shortlisting Mechanism (ASL) and the Compromise Rule of k Names, mentioned in the arbitrator selection literature, which are theoretically claimed to yield better results in certain respects than the Shortlisting mechanism. In addition, since the search for alternative simultaneous mechanisms to sequential mechanisms is also included in the literature, we have added a recently defined simultaneous veto mechanism, called the Gradual Veto (GV) mechanism to the study. We made comparisons and analyses on Truthfulness, Efficiency, Minimal Satisfaction Test (MST), Fallback Bargaining, Inequality, First Mover Advantage, Equal Loss Principle, Average Payoffs, Fairness, and Reciprocity in a between-subject design with 290 participants from Sabancı University.

First, we looked at truthful behavior and strategizing. The GV mechanism has the most truthful behavior patterns. ASL and CRk include strategic behavior more than SL. Even though GV is a nearly-strategy-proof mechanism, it falls behind other mechanisms in terms of efficiency. Overall, ASL and SL are doing better jobs of reaching efficient outcomes and satisfying MST. Regarding inequality, GV is creating a significantly more equal environment. Even though ASL is the most efficient, it also brings higher inequality between matched players. Among sequential ones, SL has a clear first-mover advantage. Average payoff statistics follow a similar path as Efficiency. We repeat these analyses with three OLS regressions, where errors are clustered by session to see whether this performance is related to the session. The results do not change and stay significant. ASL is the most efficient and satisfying MST and the most unequal. And finally, we define a punishment pattern and analyze whether such a desire to punish is common among participants. SL allows such direct punishment more, and there is no such behavior in CRk.

Which mechanism should be adapted to daily life problems is a question that needs to be considered in detail. In this thesis, we have evaluated the proposed mechanisms from many aspects. If the priority is to achieve efficient outcomes, we have shown empirically that different mechanisms do this as well as SL and do not give a player

a direct first-mover advantage. On the other hand, if the priority is inequality or strategy-proofness, it can be observed that the simultaneous veto mechanism creates better results.

BIBLIOGRAPHY

Afacan, Mustafa Oğuz, Nejat Anbarci, and Özgür Kıbrıs. 2022. "Arbiter Assignment.".

Anbarci, Nejat. 2006. "Finite Alternating-move Arbitration Schemes and the Equal Area Solution." *Theory and Decision* 61: 21–50.

Barberà, Salvador, and Danilo Coelho. 2010. "On the Rule of k Names." *Games and Economic Behavior* 70: 44–61.

Barberà, Salvador, and Danilo Coelho. 2018. "On the advantages and disadvantages of being the first mover under rules of k names." *International Journal of Economic Theory* 14(1): 51–60.

Barberà, Salvador, and Danilo Coelho. 2022. "Compromising on Compromise Rules." *The RAND Journal of Economics* 53(1): 95–112.

Bloom, David E., and Christopher L. Cavanagh. 1986. "An Analysis of the Selection of Arbitrators." *NBER Working Paper Series* 1938.

Bol, D., J. F. Laslier, and M. Núñez. 2022. "Two Person Bargaining Mechanisms: A Laboratory Experiment." *Group Decision and Negotiation* 31(6): 1145–1177.

Brams, S. J., and D. M. Kilgour. 2001. "Fallback Bargaining." *Group Decision and Negotiation* 10: 287–316.

Chun, Y. 1988. "The Equal-Loss Principle for Bargaining Problems." *Economics Letters* 26(2): 103–106.

Davis, J. A., and T. W. Smith. 1991. *The NORC General Social Survey: A user s guide*. SAGE.

De Clippel, G., and K. Eliaz. 2012. "Reason-based Choice: A Bargaining Rationale for the Attraction and Compromise Effects." *Theoretical Economics* 7(1): 125–162.

De Clippel, Geoffroy, Kfir Eliaz, and Brian Knight. 2014. "On the Selection of Arbitrators." *American Economic Review* 104(11): 3434–3458.

Fischbacher, U. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics* 10(2): 171–178.

Hurwicz, L., and D. Schmeidler. 1978. "Construction of Outcome Functions Guaranteeing Existence and Pareto Optimality of Nash Equilibria." *Econometrica: Journal of the Econometric Society* pp. 1447–1474.

Hurwicz, L., and M. R. Sertel. 1999. "Designing Mechanisms, in particular for Electoral Systems: the Majoritarian Compromise." *Contemporary Economic Issues: Economic Behaviour and Design* pp. 69–88.

Kıbrıs, Ö., and M. R. Sertel. 2007. "Bargaining Over a Finite Set of Alternatives." *Social Choice and Welfare* 28(3): 421–437.

Lambert, R. D. 1983. "Question design, response set and the measurement of left/right thinking in survey research." *Canadian Journal of Political Science/Revue canadienne de science politique* 16(1): 135–144.

Laslier, J. F., M. Nunez, and M. R. Sanver. 2021. "A Solution to the Two-person Implementation Problem." *Journal of Economic Theory* 194.

Maskin, E. 1999. "Nash Equilibrium and Welfare Optimality." *The Review of Economic Studies* 66(1): 22–38.

Naef, M. 2009. "Measuring Trust: Experiments and Surveys in Contrast and Combination.".

Van der Linden, M. 2017. "Impossibilities for Strategy-proof Committee Selection Mechanisms with Vetoers." *Journal of Mathematical Economics* 73: 111–121.

APPENDIX A

Ethics Committee Approvals for the Study

Evrak Tarih ve Sayısı: 29.12.2022-7469

Sabancı Üniversitesi

Sabancı University Research Ethics Council (SUREC)

Date: December 2022

To: Özgür Kıbrıs (PI) Sonnur Baş (Co-I)
From: Director of Research and Graduate Policy, Chairman of the Ethics Committee

Protocol Number: FASS-2022-63
Protocol Name: An Experiment on Two-Person Social Choice

Subject: SUREC Approval
Official Approval Date: Dec 29th, 2022

Sabancı University Research Ethics Council has approved the above named and numbered protocol through expedited review. You are responsible for promptly reporting to the SUREC:

- any severe adverse effects
- any unanticipated problems involving risks to subjects or others;
- any proposed changes in the research activity

Enclosed you can find the below noted approved documents.

☒ Protocol Application

☒ Informed Consent Form

If you have any questions please feel free to contact me via phone at 216-483 9666 or via e-mail at RGP Director.

Best Regards,

Director of Research and Graduate Policy,
Chairman of the Ethics Committee

Mevcut Elektronik İmzalar

Prof. Dr. Mehmet Zafer Gedik - Fizik (Öğretim Üyesi) - 29.12.2022
Dr. Öğr. Üyesi Nedim Nami Nomer - Siyaset Bilimi (Öğretim Üyesi) - 29.12.2022
Prof. Dr. Syeda Arzu Wasti - Organizasyon (Öğretim Üyesi) - 29.12.2022
Doç. Dr. Sabahat Çiğdem Hemşinlioğlu - Psikoloji (Öğretim Üyesi) - 29.12.2022
Dr. Öğr. Üyesi Ogün Adebali - Moleküler Biyoloji, Genetik ve Biyomühendislik (Öğretim Üyesi) - 29.12.2022
Prof. Dr. Mehmet Yıldız - Rektör Yardımcılığı (Araştırma) (Rektör Yardımcısı) - 29.12.2022

FRG-A410-01-03

Research and Graduate Policies Directorate
Phone: 216-483 9099 Fax:216-4839118

Ön izleme Bu evrak elektronik imzalıdır.
ikinci satır https://turkiye.gov.tr/ebd?eK=6673&eD=BSPBKEAF9&eS=7469

For SUREC Use Only

Protocol No: **FASS-2022-63** Approval Date: **Dec 29th, 2022**
Modification Requested Date: Modification Approval Date:

Title: An Experiment on Two-Person Social Choice
Principal Investigator: Prof. Özgür Kıbrıs / FASS
Co-Investigator: Sonnur Baş / FAS Research Assistant

The SUREC approval is valid for 3 years after the given approval date.

THIS SPACE FOR SUREC USE ONLY

☐ The protocol has been determined to be exempt from SUREC review in accordance with Sabancı University Research Ethics Council procedure.
☒ The protocol has been approved through expedited review in accordance with Sabancı University Research Ethics Council procedure.
☐ The Institutional Review Board has been approved the protocol through full review review in accordance with Sabancı University Research Ethics Council procedure.

APPROVED BY THE SABANCI UNIVERSITY RESEARCH ETHICS COUNCIL

PARTICIPANTS:

SUREC Chair : Prof. Dr. Mehmet Yıldız
SUREC Member : Prof. Dr. Arzu S. Wasti
SUREC Member : Asst. Prof. S. Çiğdem Bağcı
SUREC Member : Asst. Prof. Nedim Nomer
SUREC Member : Asst. Prof. Ogün Adebali
SUREC Member : Prof. Dr. Zafer Gedik

Mevcut Elektronik İmzalar

Prof. Dr. Mehmet Zafer Gedik - Fizik (Öğretim Üyesi) - 29.12.2022
Dr. Öğr. Üyesi Nedim Nami Nomer - Siyaset Bilimi (Öğretim Üyesi) - 29.12.2022
Prof. Dr. Syeda Arzu Wasti - Organizasyon (Öğretim Üyesi) - 29.12.2022
Doç. Dr. Sabahat Çiğdem Hemşinlioğlu - Psikoloji (Öğretim Üyesi) - 29.12.2022
Dr. Öğr. Üyesi Ogün Adebali - Moleküler Biyoloji, Genetik ve Biyomühendislik (Öğretim Üyesi) - 29.12.2022
Prof. Dr. Mehmet Yıldız - Rektör Yardımcılığı (Araştırma) (Rektör Yardımcısı) - 29.12.2022

FRG-A410-01-03-V01

Instructions

This study is about the mechanisms used for joint decision-making between two individuals. By participating in this study, you will earn Experimental Money, which will later be converted into research credits. Every 100 EM will be equivalent to 0.5 research credits. Your total earnings will consist of your participation fee of 300 EM plus an additional (bonus) amount that will vary between 0 EM and 400 EM depending on your decisions during the study. Therefore, you will complete the

work with an amount between 300 EM and 700 EM.

The experiment consists of 40 rounds. In each round, the computer will randomly match you to another participant and both of you will play the game described below. The outcome of this game is a selection of a single option from a list of five, named as {a,b,c,d,e}

At the beginning of each round, you will be informed in a table similar to the example below about how many EM were assigned to which option for you and how many EM were assigned to each option for the participant you matched. Similarly, the participant you are matched with will see the values assigned to him and you. The 40 rounds you will play are divided into 4 stages, and each stage consists of 10 rounds. (That is, stage 1 consists of rounds 1-10, stage 2 consists of rounds 11-20, stage 3 consists of rounds 21-30, and stage 4 consists of rounds 31-40.) In each stage, half of the participants will be randomly assigned to Group A and the other half to Group B. Participants in the same group will have the same payoffs throughout the stage. In the 10 rounds of the same stage, participants in one group will be randomly matched with participants in the other group. During a stage, the payoff table for you and your matched participant will not change.

Payoff	You	Other Participant
400 EM	d	c
300 EM	b	a
200 EM	a	e
100 EM	c	b
0 EM	e	d

At the beginning of each stage, you will see a message about switching to a new chapter. This message will warn you that the payoff table for you and your matched participant in the next stage may be different from the previous one. You will also be able to see these payoff table on your screen every round.

In each round, you will play a game with your matched participant to choose one of the alternatives {a, b, c, d, e.} At the end of each round, the selected alternative and the winnings you will get by choosing this alternative will be displayed on the screen. At the end of the experiment, we will randomly select one of the 40 rounds played and calculate your final payoff. The total amount of EM you will earn will be equal to the sum of the 300 EM you will earn for your participation and the amount of EM you have earned in the randomly selected round.

Questionnaire

Date of birth?

__ ...

Place of birth?

__ ...

What is your assigned-at-birth-sex?

__ male
__ female
__ other.
__ do not want to say.

What is your gender identification?

__ male
__ female
__ non-binary
__ transsexual
__ other.
__ do not want to say.

Where does your family live?

__ ...

Where do you live right now?

__ dormitory
__ shared flat
__ with parents
__ other
__ do not want to say

What is your major?

__ ...

What is your double major and/or minor (if any)?

__ ...

Which ECON courses have you taken so far?

__ ...

What is your monthly spending?

__ ...

Do you have any work experience?

__ yes
__ no

What is your monthly household income?

__ ...

How many siblings do you have?

__ ...

What is your religion? (Optional)

__ ...

Do you expect your life next year to be better, worse, or more or less the same as this year?

__ better
__ worse
__ more or less the same

How satisfied are you with your life? (1: Not at all satisfied, 10: Completely satisfied)

__ ...

Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people? (1: I never trust, 10: I usually trust)

__ ...

Some people think of political attitudes as being on the "left" or "right". Here is a scale stretching from left to right, with "1" as being the most to the left and "10" being the most to the right. When you think of your political opinions, where would you put yourself on this scale?

__ ...

APPENDIX B

Truthfulness: Early Decisions

Here is the list of decision types that are classified as *Early* and *Truthful* in our analysis in Section 4.1

- **ASL**
  - The first mover proposes her best $3, 4$, or 5 alternatives.
  - The second-mover accepts the set if the set exactly consists of her best $3, 4$, or 5 alternatives.
  - If the second mover decides to propose, she proposes her best 4 or 5 alternatives.
- **CRk**
  - If the second mover decides to propose, she proposes her best k alternatives.
  - If the second mover decides to choose, the first mover proposes her best k alternatives.
- **GV**
  - Players veto their worst-ranked alternative from the biggest set, a, b, c, d, e
- **SL**
  - The first mover proposes her best 3 alternatives.