Long, Yan (2016) *Essays on robust mechanism design.* PhD thesis.

Essays on Robust Mechanism Design

Yan Long

University of Glasgow

This dissertation is submitted for the degree of

PhD in Economics

College of Social Science June, 2016

I would like to dedicate this thesis to my loving family.

Declaration

I declare that except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Signature:
Printed name:

Yan Long
June, 2016

Acknowledgements

My deepest gratitude is to my supervisor Prof. Herve Moulin. I have learned a great deal from not only his expertise and insight, which is, one of the best in Theory, but also his attitude towards research and academics, the curiosity and passion he possesses. I would also like to thank the many good economists, senior or junior, I have met during my PhD studies, in various seminars and conferences. It is such a comfort to know that one is not alone in this journey. And I hope one day I myself would become a source of comfort to young students.

I would like to acknowledge the generosity of the two universities in which I carried out PhD studies and research: Rice University and Glasgow University. In particular, I would like to thank the administrative staff in both universities who have made the practical side of my life as easy as possible.

I would like to thank the friendship I obtained these years, one of the most cherishable parts of doing PhD and becoming a researcher. And I would not have the strength and courage needed without the support and love of my family: I am so pound to be able to dedicate this thesis to you.

Preface

The dissertation brings together three papers that were written during my PhD studies, each dealing with a specific robust mechanism design problem. The objective of this preface is to discuss briefly the term "robust mechanism design" used in literature, and to outline the framework of the three papers.

The mechanism design literature has been a great success during the last thirty years; however since the very beginning of the theoretic development, it has been argued the the mechanisms working in theory are not "robust"–i.e. they are too sensitive to fine details that will not be available to the designer in practice. Robust mechanism design responds to the concerns from various aspects, and to different degrees. In all of the papers collected here, a completely prior-free approach is adopted, that is, both prior-free solution concepts and prior-free objective functions are imposed. The mechanisms obtained are hence immune to uncertainties about private information held by each agent.

In an environment where common value is assumed (Chapter 1), ex-post implementability is imposed as the solution concept; in the other environments where private value is assumed (Chapter 2 and 3), dominant-strategy implementability, a.k.a. strategy-proof, is imposed. The objective set up in the first two chapters is worst-case guarantee, and in the last chapter Pareto efficiency.

Due to the fundamental work of Bergemann and Morris[9], it is known that the two strong solution concepts are exactly what are required for a well-defined "robustness to private information" in each setting respectively. The worst-case guarantee, in which a ratio that compares the absolute gain (or loss) to some reference point is concerned, is imposed in a seemly "ad hoc" way. This "maxmin-ratio" objective, however, is a well established and widely used criterion. (See Chapter 1 and 2 for more discussion). While a decision-theory-based justification is still absent and would be very welcomed, its merit as an objective lies largely in its intrinsic simplicity and desirability.

Abstract

Chapter 1: Under the average common value function, we select almost uniquely the mechanism that gives the seller the largest portion of the true value in the worst situation among all the direct mechanisms that are feasible, ex-post implementable and individually rational.

Chapter 2: Strategy-proof, budget balanced, anonymous, envy-free *linear* mechanisms assign p identical objects to n agents. The *efficiency loss* is the largest ratio of surplus loss to efficient surplus, over all profiles of non-negative valuations. The *smallest* efficiency loss $\frac{n-p}{n^2-n}$ is uniquely achieved by the following simple allocation rule: assigns one object to each of the $p-1$ agents with the highest valuation, a large probability to the agent with the pth highest valuation, and the remaining probability to the agent with the $(p+1)$th highest valuation. When "envy freeness" is replaced by the weaker condition "voluntary participation", the optimal mechanism differs only when p is much less than n.

Chapter 3: One group is to be selected among a set of agents. Agents have preferences over the size of the group if they are selected; and preferences over size as well as the "stand-outside" option are single-peaked. We take a mechanism design approach and search for group selection mechanisms that are efficient, strategy-proof and individually rational. Two classes of such mechanisms are presented. The *proposing mechanism* allows agents to either maintain or shrink the group size following a fixed priority, and is characterized by group strategy-proofness. The *voting mechanism* enlarges the group size in each voting round, and achieves at least half of the maximum group size compatible with individual rationality.

Contents

Chapter 1

Maxmin mechanism in a simple common value auction

1.1 Introduction

Auction mechanisms have been widely used in selling goods and services; one of the most commonly used formats is the sealed bid first price auction. In this format, bidders simultaneously submit bids, and the one with the highest bids wins and pays her bid. When the object being sold is of common value to all the bidders, there is a well-documented phenomenon called "winner's curse": the winner is tend to be the one who over-estimates the true value. Though "winner's curse" is not supposed to happen in the equilibrium, the literature on common value auctions found that bidders have difficulty learning to avoid it (see [38] for a comprehensive discussions).

"Winner's curse" can turn out to be a curse for the auctioneer. Consider a government auctioning oil fields. It is quite possible that the winning buyer, suffering from the winner's curse, will eventually go bankrupt and leave a huge mess for the government. Instead of maximizing the revenue from the auction, the government may be more interested in avoiding the risk caused by winner' curse, while obtaining a reasonably high share of the value generated by the oil fields.

One auction format that many organizations have been using to achieve these goals is the average bidding auction. In its most standard format, the winner is the one who bids closest to the average bid and the price is the average bid. The naive argument is that the average bid is likely to be close to the true value. However, there is no guarantee that agents will truthfully submit their estimation. In fact, bidders have incentive to all submit identical bids, and a continuum of Nash equilibria in which all bidder submit the same low-enough bid exists.

While the effectiveness of the average bidding auction are still being debated (see [2], [19], [21], [25]), we think it is worthwhile carrying out a mechanism design directly. The objective of the auctioneer is to maximize the *worst share* of value, under the restriction that the mechanism is truth-reporting and bidders never lose money. More specifically, the true value of the object for each bidder is v, unknown to everybody ex-ante. Each agent $i \in \{1,2,\cdots,n\}$ gets a private signal s_i and $v = \frac{\sum s_i}{n}$. The worst share is defined to be the smallest ratio of the revenue to the true value of the object, over all the non-negative signal vectors. Note that by using the worst share, we avoid distribution assumptions of the signal profiles, and make sure that the auctioneer gets something (as long as the worst share is positive) under *any* signal profiles.

We show that the seller can achieve the largest worst share using the simple *Maxmin mechanism*. The Maxmin mechanism randomly assigns the object to one agent and charges the agent i who gets the object a price $\frac{\sum_{j\neq i} s_j}{n}$. It is easy to see that this mechanism is ex-post implementable and individually rational; and the seller will get $\frac{n-1}{n}$ of the true value v under any signal profiles. In fact, $\frac{n-1}{n}$ is the largest worst share that can be achieved by any ex-post implementable and individually rational mechanism, and is achieved (almost) uniquely by the Maxmin mechanism.

One disadvantage of the Maxmin mechanism is the lack of strict incentive for truth-telling: agents get the same outcomes irrespective of their own reports. An easy way to modify it is to randomly assign the object to an agent with the first mth highest signals, where m is an integer between 1 and n; and the worst share under such mechanism is $\frac{m-1}{m}$. We have to trade worst share for incentive.

literature review

The literature on (pure) common value auction largely focus on the Mineral Right Model [64]. In this model the true value of the object is a *random variable* V. Given a realization v of V, each agent i gets private signal s_i generated from the distribution $F(s_i|v)$, and the signals are usually assumed to be mutually independent. [46] shows that it is possible for the auctioneer to extract almost all the surplus in this setting. A simple and special case is the following: suppose the bidders can construct unbiased estimated $z(s_i)$ of v from their signal s_i and s_i is independent of s_j, given v, then the mechanism assigning the object to i and charging her $z(j)$ extracts all the rents. A more detailed review on auctions with interdependent value can be found in [42].

Our paper takes the true value to be the average of all the signals, as in several other papers ([10], [27], [14], [40], [34]). We offer some justification for the average value function form. Suppose that for each bidder the true value of the object is v, a *parameter* that is unknown to everybody. Each agent $i \in \{1, \cdots, n\}$ gets a private signal s_i that is a realization of a random variable S_i with $\mathrm{E}(S_i) = v$. There are different ways to estimate v, depending on the specific assumptions on the distributions of $\{S_i : i = 1, 2, \cdots, n\}$. We argue that a focal point is the Ordinary Least Square (OLS) estimation, i.e., $\hat{v} = \frac{1}{n}\sum_i^n s_i$. Hence the assumption under the average value function is that all the agents, as well as the seller, agree on that true value is nothing more than the most reasonable estimation, and take OLS estimation to be that reasonable estimation. Note that the data generating process here is modeled based on the classical approach in statistical inference, while the Mineral Right Model follows Bayesian inference.

We use the worst case analysis to find desirable mechanisms. This approach has been successfully applied in mechanism design with private value. (See [41], [29], [50], [28].) However, its application in common value case is not known to us. Common value assumption makes it possible to guarantee the seller a fixed portion of the true value of the object, which is not possible in the private value setting no matter how many copies of the objects are being sold (see [28]).

We could also set the objective to be the well-established revenue maximization, as in [54]. If we assume that $\{S_i : i = 1, 2, \cdots, n\}$ are mutually independent and identically distributed according to some regular distribution function $G(\cdot)$, then the mechanism that maximize the seller's expected revenue is found in [54]: if no bid is higher than the reservation price, then the seller keeps the object; otherwise the bidder with the highest bid wins and pays either the reservation price or $\frac{s_{(2)}+\sum_{k\geq 2} s_{(k)}}{n}$, whichever is higher, where $s_{(k)}$ is the kth highest signal. The worst share is zero in this mechanism.

This paper is organized as the following. In section 1.2 we set up the model and give the axioms we impose on the mechanisms. In section 1.3 we select the Maxmin mechanism, the mechanism that has the best performance in the worse case when revenue share is concerned. Section 1.4 concludes.

1.2 Setting and Axioms

An indivisible object is to be sold to at most one of n agents, where n is fixed; the value of the object is the same among agents. Denote the set of agents by $I = \{1, 2, \cdots, n\}$; agents are risk neutral. Each agent $i \in I$ has a private signal (individual estimation) $s_i \in S_i = \mathbb{R}_+$ about the value of the object. Let $s \in S = \mathbb{R}_+^n$ be a signal profile. The notation s_{-i} stands for the vector obtained form s by deleting s_i, and $(\tilde{s}_i; s_{-i})$ stands for the signal profile obtained from s by replacing the signal s_i with $\tilde{s}_i$. For any s, the true value of the object for each agent is the average of s, i.e., $v(s) = \frac{1}{n}\sum_{i=1}^{n} s_i$. The object has no value for the seller if it is not sold. Note that (1) we only restrict signals to be non-negative and do not specify probability distribution of the signal space; (2) No efficiency issue occurs in this common value setting as long as the object is sold.

We focus on direct mechanisms, that is, agents report their own signals and the allocation only depends on the reported profile of signals.

Definition 1.1. A direct mechanism μ is a pair (a,t) such that

$$a : S \to [0,1]^n, \quad t : S \to \mathbb{R}^n;$$

for each $s \in S$, $a_i(s)$ is the probability that agent i gets the object and $t_i(s)$ is her expected payment.

Assuming risk neutral preference, agent i's expected utility under μ is $a_i(s) \cdot v(s) - t_i(s)$ for any $s \in S$. We define some standard axioms for direct mechanisms.

Definition 1.2. A direct mechanism $\mu = (a,t)$ is feasible if

$$\sum_{i \in I} a_i(s) \leq 1, \quad \forall s \in S.$$

This is simply saying that the seller has only one object to sell.

Definition 1.3. A direct mechanism $\mu = (a,t)$ is ex-post implementable if

$$a(s) \cdot v(s) - t_i(s) \geq a_i(\tilde{s}_i; s_{-i}) \cdot v(s) - t_i(\tilde{s}_i; s_{-i}),$$

$\forall i \in I, \forall s \in S, \forall \tilde{s}_i \in S_i$.

A direct mechanism is ex-post implementable if each agent has no incentive to misreport his or her signal when she is assured that other agents report truthfully.

Definition 1.4. A direct mechanism $\mu = (a,t)$ is individually rational if

$$a_i(s) \cdot v(s) - t_i(s) \geq 0, \quad \forall i \in I, \forall s \in S.$$

Let $\mathscr{M}$ be the class of direct mechanisms that are feasible, ex-post implementable and individually rational. We give a standard characterization of $\mathscr{M}$ by the following lemma.

Lemma 1.1. *A direct mechanism $\mu = (a,t)$ is in $\mathscr{M}$ iff for all $i \in I$, for all $s \in S$, $\sum_{i \in I} a_i(s_i) \leq 1$; for all $i \in I$, for all $s_{-i} \in \mathbb{R}_+^{n-1}$, $a_i(s_i; s_{-i})$ is non-decreasing in s_i; and*

$$t_i(s_i;s_{-i}) = a_i(s_i;s_{-i}) \cdot v(s) - \frac{1}{n}\int_0^{s_i} a_i(x_i;s_{-i})\mathrm{d}x_i + h_i(s_{-i}),$$

where $h_i : \mathbb{R}_+^{n-1} \to \mathbb{R}$ *and* $h_i(s_{-i}) \leq 0$ *for all* $i \in I$*, all* $s_{-i} \in \mathbb{R}_+^{n-1}$.

Proof. We omit the proof, which is an easy modification of the proof of Lemma 2 in [54]. □

Note 1. (1) any seller who likes more revenue than less will simply set $h_i(s_{-i}) = 0$ for all $i \in I$, all $s_{-i} \in \mathbb{R}_+^{n-1}$, so we will ignore the $h_i(\cdot)$ part of any $\mu \in \mathscr{M}$ from now on; (2) $t_i(s_i;s_{-i}) \geq 0$ if $h_i(s_{-i}) = 0$, that is, the seller will get non-negative expected revenue from each agent and hence non-negative expected revenue under μ.

1.3 Maxmin mechanism

Our goal in this section is to find the mechanisms within $\mathscr{M}$ that give the seller the largest portion of the true value under the worst situation, when expected revenue is concerned.

First we introduce some notation. For any mechanism $\mu \in \mathscr{M}$, for any $s \in S$, let $\mathscr{A}_\mu(s)$ be the revenue the seller gets under μ when the signal profile is s. Note that for a give s, $\mathscr{A}_\mu(s)$ is a random variable whose outcomes are not specified if μ is a randomized mechanism; however, the expected revenue is known for sure: $\mathrm{E}[\mathscr{A}_\mu(s)] = \sum_{i=1}^n t_i(s)$. For any mechanism $\mu \in \mathscr{M}$, we define worst share $Q(\mu)$ to be the minimum portion of the true value that the seller gets among all signal profiles, that is, $Q(\mu) = \inf_{s\in S_+} \mathrm{E}[\mathscr{A}_\mu(s)]/v(s)$, where $S_+ = S\backslash\{\mathbf{0}\}$. Note that in order to define $Q(\mu)$, we take out the zero profile, in which case any mechanism in $\mathscr{M}$ gives the seller zero revenue.

Using lemma 1.1, our problem can be written as the following:

$$\sup_{a(\cdot)} \inf_{s\in S_+} \frac{\mathrm{E}[\mathscr{A}_\mu(s)]}{v(s)} \tag{1.1}$$

$$s.t. \begin{cases} a_i(s) \geq 0 & \forall i, \forall s \\ \sum_{i\in I} a_i(s) \leq 1 & \forall s \\ a_i(s_i;s_{-i}) \quad \mathrm{non-decreasing\, w.r.t.}\, s_i & \forall i, \forall s_{-i} \end{cases}$$

where

$$\mathrm{E}[\mathscr{A}_\mu(s)] = \sum_{i\in I} a_i(s)\cdot v(s) - \frac{1}{n}\sum_{i\in I}\int_0^{s_i} a_i(x_i;s_{-i})\mathrm{d}x_i.$$

First Note that for any $\mu \in \mathscr{M}$, $0 \leq \mathrm{E}[\mathscr{A}_\mu(s)] \leq \sum_{i\in I} a_i(s)\cdot v(s) \leq v(s)$ for all $s \in S_+$, hence $0 \leq Q(\mu) \leq 1$.

Lemma 1.2. *For any $\mu = (a,t) \in \mathscr{M}$,*

$$Q(\mu) \leq \sup_{s\in S_+} \sum_{i=1}^{n} a_i(s) - \sup_{i\in I, s\in S_+} a_i(s);$$

if $\sum_{i\in I} a_i(s) = 1$ for all $s \in S_+$, then $Q(\mu) = 1 - \sup_{i\in I, s\in S_+} a_i(s)$.

Proof. See Appendix. □

Corollary 1.1. *For any mechanism $\mu \in \mathscr{M}$, $Q(\mu) \leq \frac{n-1}{n}$.*

Proof. For any mechanism $\mu = (a,t) \in \mathscr{M}$, if $a_i(s) > \frac{1}{n}$ for some s and i, since $\sup_{s\in S_+} \sum_{i=1}^{n} a_i(s) \leq 1$, we have that

$$\sup_{s\in S_+} \sum_{i=1}^{n} a_i(s) - \sup_{i\in I, s\in S_+} a_i(s) < \frac{n-1}{n};$$

if $a_i(s) \leq \frac{1}{n}$ for all s and i, we have $\sup_{s\in S_+} \sum_{i=1}^{n} a_i(s) \leq \frac{1}{n}$. Then

$$\begin{aligned} Q(\mu) &\leq \sup_{s\in S_+} \sum_{i=1}^{n} a_i(s) - \sup_{i\in I, s\in S_+} a_i(s) \\ &\leq n \cdot \sup_{i\in I, s\in S_+} a_i(s) - \sup_{i\in I, s\in S_+} a_i(s) \\ &= (n-1)\cdot \sup_{i\in I, s\in S_+} a_i(s) \leq \frac{n-1}{n}. \end{aligned}$$

Therefore $Q(\mu) \leq \frac{n-1}{n}$. □

Now we give the definition of the Maxmin mechanism.

Definition 1.5. The Maxmin mechanism $\mu^* = (a^*, t^*)$ is the following:

$$a_i^*(s) = \frac{1}{n}, t_i^* = \frac{1}{n^2}\sum_{j\neq i} s_j, \quad \forall i \in I, \forall s \in S.$$

Note that $\mu^* \in \mathscr{M}$. As stated by Theorem 1.1, it is the "almost" unique solution of problem 1.1, i.e., it is the "almost" unique mechanism in $\mathscr{M}$ that guarantees the seller $\frac{n-1}{n}$ of the true value in any signal profile, which justifies its name.

Theorem 1.1. *$Q(\mu^*) = \frac{n-1}{n}$; for any $\mu = (a,t) \in \mathscr{M}$ such that $a(s) \neq a^*(s)$ for some $s \in S'_+$, where $S'_+ = \{s \in S_+ : \exists i, j \in I\, s.t.\, i \neq j\, and\, s_i \cdot s_j > 0\}$, we have $Q(\mu) < Q(\mu^*)$,* *Proof.* See Appendix. □

Note 2. (1) For any $\mu \in \mathscr{M}$, $Q(\mu) = \frac{n-1}{n}$ only if μ coincides with μ^* in almost every signal profiles, more precisely, in every signal profile with at least two positive arguments. We allow some flexibility in the allocation rule when the signal profile has only one positive argument; see the proof for detailed description. However since that is highly knife edge case, we conclude that Maxmin mechanism is the "almost" unique solution of problem 1.1.

(2) For the second part of the Theorem, if we assume that $\sum_{i\in I} a_i(s) = 1$ for any $s \in S_+$, then it is much easier to prove since we know in any solution $a_i(s) \leq \frac{1}{n}$ for all $i \in I$, $s \in S_+$. However, it is not clear at all why $\sum_{i\in I} a_i(s) = 1$ for any $s \in S_+$ must hold in any solution. By induction on the number of zero arguments in signal profiles, we show a slightly more general result first, which gives both upper bounds and lower bounds of $a_i(s)$ for each $i \in I$, $s \in S_+$. These bounds converge to $\frac{1}{n}$ when $Q(\mu)$ approaches to $\frac{n-1}{n}$; hence our Theorem is proved. However, the lower bounds reduce to non-negative restrictions quickly when decreasing $Q(\mu)$ from $\frac{n-1}{n}$ and hence we do not know much about what a mechanism in $\mathscr{M}$ with $Q(\mu)$ less than $\frac{n-1}{n}$ must look like. On the other hand, if the mechanism always sells the object, we have an explicit expression of $Q(\mu)$, which is the second part of Lemma 1.2.

1.4 Discussion and conclusion

Now consider the outcomes of the Maxmin mechanism. Since $t_i^*(\cdot)$ is the average payment for agent i, there are many ways to assign payments for outcomes in which he or she wins or loses respectively. A natural way is to let agent i pay $\frac{1}{n}\sum_{j\neq i} s_j$ when he or she wins, and pay 0 when loses. As a result, if the worst outcome of the mechanism for the seller is realized, that is, agent with the highest signal gets the object, the seller will end up getting $\frac{1}{n}\sum_{i=2}^{n} s_{(i)}$, where $s_{(1)}, s_{(2)}, \cdots, s_{(n)}$ is a reordering of $s_1, s_2, \cdots, s_n$ with $s_{(1)} \geq s_{(2)} \geq \cdots \geq s_{(n)}$. Fixed $s_{(2)}, \cdots s_{(n)}$, $\frac{1}{n}\sum_{i=2}^{n} s_{(i)}/v(s)$ goes to zero as $s_{(1)}$ goes to infinity; hence the Maxmin mechanism with this natural payment schedule does not guarantee the seller a (positive) fixed portion of the true value in all realizations of the lottery in assigning the object and for all $s \in S_+$. A second thought may yield a payment schedule that works: since the expected payment of agent i is $\frac{1}{n^2}\sum_{j\neq i} s_j$ for all $i \in I$, we could ask each agent to pay this same amount no matter he or she wins or loses; as a result, the seller will get $\frac{n-1}{n} \cdot v(s)$ for sure. The drawback with this payment schedule is that agents will get negative utilities if they lose, which is of high probability. In fact we could further avoid this undesirability. Since the expected utility of agent i is $\frac{1}{n^2} \cdot s_i$, we could ask each agent to pay (or receive) certain amounts in different outcomes so that he or she gets this same utility no matter he or she wins or loses; as a result, the seller will get $\frac{n-1}{n} \cdot v(s)$ for sure. This works because we are in the common value world, i.e., no matter which agent wins, the total surplus to be divided between the seller and agents is the same. Therefore if each agent gets the same utility in each outcome of the lottery, the seller will also get the same revenue in each outcome of the lottery.

We conclude our paper with possible extensions. First note that the procurement model in which a buyer is buying an object from a set of bidders can not be treated similarly when worst case analysis is applied. The difference is largely due to the signal space $\mathbb{R}^n_+$, which is bounded below but not above. In an auction, the lower bound makes it possible to define the Maxmin mechanism, and the lack of upper bound makes the situation that one agent gets a very, very high signal the biggest concern. In a procurement, however, we cannot define a parallel mechanism due to the lack of upper bound, and the lower bound restrict the exploitation of the extreme case.

Second, other forms of value functions also arise naturally in reality. One possible extension is the weighted average, with the weight of each agent a common prior, i.e., $v(s) = \alpha_1 s_1 + \cdots + \alpha_n s_n$, where $\alpha_i > 0$ for all $i \in I$ and $\sum_{i\in I}\alpha_i = 1$. This does not change our analysis since we could ask each agent $i \in I$ to report his or her adjusted signal $a_i s_i$; another extension also takes the form of weighted average, with the weights a common prior; however the weights are given to the order statistics, that is, $v(s) = \beta_1 s_{(1)} + \cdots + \beta_n s_{(n)}$, where $s_{(1)}, \cdots, s_{(n)}$ is a reordering of $s_1, \cdots, s_n$ with $s_{(1)} \geq \cdots \geq s_{(n)}$, $\beta_i \geq 0$ for all $i \in I$ and $\sum_{i\in I}\beta_i = 1$. Functions like Max, Min, Median, Average are all special cases. It is interesting to see whether the random allocation rule is still uniquely optimal in our worst case analysis with such function forms.

1.5 Appendix

1.5.1 Proof of Lemma 1.2.

Proof. The first part:

For any $\mu = (a,t) \in \mathscr{M}$, suppose for the sake of contradiction, there exists $i \in I$ such that $a_i(s) = \frac{1}{m} > \sup_{z\in S_+}\sum_{i=1}^n a_i(z) - Q(\mu)$ for some $s = (s_i; s_{-i})$. Since (a,t) is ex-post implementable, $a_i(z_i; s_{-i}) \geq \frac{1}{m}\ \forall z_i \geq s_i$. Let $\tilde{s} = (\tilde{s}_i; s_{-i}) = (\lambda \cdot s_i; s_{-i})$, where $\lambda > 1$, then

$$\frac{\sum_{i=1}^n \int_0^{\tilde{s}_i} a_i(x_i;\tilde{s}_{-i})\mathrm{d}x_i}{\sum_{i=1}^n \tilde{s}_i} \geq \frac{\frac{1}{m}(\lambda-1)\cdot s_i + \sum_{j\neq i}\int_0^{\tilde{s}_j} a_j(x_j;\tilde{s}_{-j})\mathrm{d}x_j}{\lambda\cdot s_i + \sum_{-i} s_j}.$$

The last expression goes to $\frac{1}{m}$ when λ goes to infinity. Therefore, there exist $\tilde{s} = (\lambda \cdot s_i; s_{-i})$ with λ large enough that

$$\frac{\sum_{i=1}^n \int_0^{\tilde{s}_i} a_i(x_i;\tilde{s}_{-i})\mathrm{d}x_i}{\sum_{i=1}^n \tilde{s}_i} > \sup_{z\in S_+}\sum_{i=1}^n a_i(z) - Q(\mu).$$

That is

$$\frac{1}{n}\sum_{i=1}^n \int_0^{\tilde{s}_i} a_i(x_i;\tilde{s}_{-i})\mathrm{d}x_i > (\sup_{z\in S_+}\sum_{i=1}^n a_i(z))\cdot v(\tilde{s}) - Q(\mu)\cdot v(\tilde{s}).$$

Therefore

$$\begin{aligned}
\mathrm{E}[\mathscr{A}(\tilde{s})] &= \sum_{i=1}^n a_i(\tilde{s})\cdot v(\tilde{s}) - \sum_{i=1}^n \frac{1}{n}\int_0^{\tilde{s}_i} a_i(x_i;s_{-i})\mathrm{d}x_i \\
&\leq (\sup_{z\in S_+}\sum_{i=1}^n a_i(z))\cdot v(\tilde{s}) - \sum_{i=1}^n \frac{1}{n}\int_0^{\tilde{s}_i} a_i(x_i;s_{-i})\mathrm{d}x_i \\
&< Q(\mu)\cdot v(\tilde{s}).
\end{aligned}$$

We get $Q(\mu) > \mathrm{E}[\mathscr{A}(\tilde{s})]/v(\tilde{s})$, contradicting with the definition of $Q(\mu)$.

The second part:

If $\sum_{i\in I} a_i(s) = 1$ for any $s \in S_+$, we know from the first statement that $Q(\mu) \leq 1 - \sup_{i\in I, s\in S_+} a_i(s)$. We show that $Q(\mu) \geq 1 - \sup_{i\in I,s\in S_+} a_i(s)$. For all $s \in S_+$,

$$\begin{aligned}
\mathrm{E}[\mathscr{A}(s)] &= \sum_{i=1}^n a_i(s)\cdot v(s) - \sum_{i=1}^n \frac{1}{n}\int_0^{s_i} a_i(x_i;s_{-i})\mathrm{d}x_i \\
&= v(s) - \sum_{i=1}^n \frac{1}{n}\int_0^{s_i} a_i(x_i;s_{-i})\mathrm{d}x_i \\
&\geq v(s) - \frac{1}{n}\sum_{i\in I} s_i \cdot \sup_{i\in I,s\in S_+} a_i(s) \\
&= v(s)\cdot(1 - \sup_{i\in I,s\in S_+} a_i(s))
\end{aligned}$$

Therefore we have $Q(\mu) \geq 1 - \sup_{i\in I,s\in S_+} a_i(s)$. □

1.5.2 Proof of Theorem 1.1.

Proof. $Q(\mu^*) = \frac{n-1}{n}$ is easy to see. We show the second part by proving a slightly more general result.

For any $s \in S_+$, let $|s| = \#\{i \in I : s_i > 0\}$, i.e., $|s|$ is the number of positive arguments in s. Let $S_k = \{s \in S_+ : |s| = k\}$ for $k = 1, 2, \cdots, n$, i.e., S_k is the set of signal profiles that have k positive arguments. For any $\mu \in \mathscr{M}$, let $X_k = \{a_i(s) : i \in I, s \in S_k\}$ for $k = 1, 2, \cdots, n$. For any $c \in \mathbb{R}$, We write $X_k \geq c$ if each element in X_k is no less than c; $X_k \leq c$ if each element in

X_k is no greater than c.

We show that for any $k = 2, 3, \cdots, n$,

$$k \cdot b + \frac{Q(\mu)}{n-1} \leq X_k \leq 1 - Q(\mu),$$

where $b = Q(\mu) - (n-1) \cdot (1 - Q(\mu))$. Note that if $Q(\mu) = \frac{n-1}{n}$, then $b = 0$ and $\frac{1}{n} = \frac{Q(\mu)}{n-1} \leq X_k \leq 1 - Q(\mu) = \frac{1}{n}$ for $k = 2, \cdots, n$, which is the second part of Theorem 2;

First by Lemma 2, $X_k \leq 1 - Q(\mu)$ for any $k = 1, 2, \cdots, n$.

Consider profile $s = (0, \cdots, 0, s_n)$, where $s_n \geq 0$. Then $\mathrm{E}[\mathscr{A}(s)]/v(s)$ equals

$$\sum_{1 \leq i \leq n-1} a_i(0, \cdots, 0, s_n) + a_n(0, \cdots, 0, s_n) - \frac{1}{n} \cdot \int_0^{s_n} a_n(0, \cdots, 0, x_n) \mathrm{d}x_n / s_n$$

Since $\{a_n(0, \cdots, 0, s_n) - \int_0^{s_n} a_n(0, \cdots, 0; x_n) \mathrm{d}x_n / s_n\} \to 0$ as $s_n \to 0$, under the condition "equal treatments of equal signals[1]", we have

$$a_i(0, \cdots, 0, s_n) \geq \frac{Q(\mu)}{n} \quad as\, s_n \to 0, \quad \forall i \neq n. \tag{1.2}$$

Now consider profile $s = (0, \cdots, 0, s_{n-1}, s_n)$, where s_{n-1}, $s_n > 0$. Fix any $s_{n-1} > 0$, by the non-decreasing property of a_i, $a_{n-1}(0, \cdots, 0, s_{n-1}, s_n) \geq \frac{Q(\mu)}{n}$ as $s_n \to 0$.

Then $\mathrm{E}[\mathscr{A}(s)]/v(s)$ is

$$\sum_{i \in I} a_i(0, \cdots, 0, s_{n-1}, s_n) - \frac{\int_0^{s_{n-1}} a_{n-1}(0, \cdots, 0, x_{n-1}, s_n) \mathrm{d}x_{n-1} + \int_0^{s_n} a_n(0, \cdots, 0, s_{n-1}, x_n) \mathrm{d}x_n}{s_{n-1} + s_n}.$$

As $s_n \to 0$,

$$\frac{\int_0^{s_{n-1}} a_{n-1}(0, \cdots, 0, x_{n-1}, s_n) \mathrm{d}x_{n-1} + \int_0^{s_n} a_n(0, \cdots, 0, s_{n-1}, x_n) \mathrm{d}x_n}{s_{n-1} + s_n} \geq \frac{Q(\mu)}{n}.$$

Since $\mathrm{E}[\mathscr{A}(s)]/v(s) \geq Q(\mu)$ and $X_2 \leq 1 - Q(\mu)$, we have for all $i \in I$,

$$a_i(0, \cdots, 0, s_{n-1}, s_n) \geq Q(\mu) - (n-1) \cdot (1 - Q(\mu)) + \frac{Q(\mu)}{n}, as\, s_n \to 0.$$

Especially, $a_n(0, \cdots, 0, s_{n-1}, s_n) \geq Q(\mu) - (n-1) \cdot (1 - Q(\mu)) + \frac{Q(\mu)}{n} = b + \frac{Q(\mu)}{n}$ as $s_n \to 0$.

[1]This symmetry condition is needed for 1.2. However, if $Q(\mu) = \frac{n-1}{n}$, this condition is not needed for the conclusion that $a_i(0, \cdot, 0, s_n) \geq \frac{1}{n}$ for all $i \neq n$ since we must have $a_i(0, \cdot, 0, s_n) \leq \frac{1}{n}$ by Lemma 2.

Therefore by the non-decreasing property of a_n, we have

$$a_n(0,\cdots,0,s_{n-1},s_n) \geq b + \frac{Q(\mu)}{n} \quad \forall s_n > 0.$$

By varying s_{n-1}, we get that

$$a_n(0,\cdots,0,s_{n-1},s_n) \geq b + \frac{Q(\mu)}{n} \quad \forall s_{n-1}, s_n > 0.$$

Changing the rule of $n-1$ and n, we get that

$$a_{n-1}(0,\cdots,0,s_{n-1},s_n) \geq b + \frac{Q(\mu)}{n} \quad \forall s_{n-1}, s_n > 0.$$

Now reconsider profile $s = (0,\cdots,0,s_{n-1},s_n)$, where s_{n-1}, $s_n > 0$. We have

$$\frac{\int_0^{s_{n-1}} a_{n-1}(0,\cdots,0,x_{n-1},s_n)\mathrm{d}x_{n-1} + \int_0^{s_n} a_n(0,\cdots,0,s_{n-1},x_n)\mathrm{d}x_n}{s_{n-1}+s_n} \geq b + \frac{Q(\mu)}{n}.$$

Therefore for all $i \in I$,

$$a_i(0,\cdots,0,s_{n-1},s_n) \geq Q(\mu) - (n-1)\cdot(1-Q(\mu)) + (b + \frac{Q(\mu)}{n}) = 2b + \frac{Q(\mu)}{n}.$$

By the same argument, $a_i(s) \geq 2b + \frac{Q(\mu)}{n}$ for all $i \in I$, $s \in S_2$. Hence $X_2 \geq 2b + \frac{Q(\mu)}{n}$.

Now consider profile $s = (0,\cdots,0,s_{n-2},s_{n-1},s_n)$ where s_{n-2}, s_{n-1}, $s_n > 0$. By the non-decreasing property of a_i, we have $a_i(0,\cdots,0,s_{n-2},s_{n-1},s_n) \geq 2b + \frac{Q(\mu)}{n}$ for $i = n-2$, $n-1$ and n. Using the same logic, we must have for all $i \in I$,

$$a_i(0,\cdots,s_{n-2},s_{n-1},s_n) \geq Q(\mu) - (n-1)\cdot(1-Q(\mu)) + (2b + \frac{Q(\mu)}{n}) = 3b + \frac{Q(\mu)}{n}.$$

Hence $X_3 \geq 3b + \frac{Q(\mu)}{n}$.

Continue with signal profile from $X_4,\cdots,X_5$ and using the same argument, we could show that $X_k \geq k\cdot b + \frac{Q(\mu)}{n-1}$ for any $k = 2,\cdots,n$. □

Note 3. (1) There is some flexibility in the allocation rule for $s \in S_1$. Consider $s = (0,\cdots,0,s_n)$ with $s_n > 0$. If we let $a_i(s) = \frac{1}{n}$ for all $i \neq n$, since $\sum_{i=1}^{n-1} t_i(s) = \frac{n-1}{n}\cdot v(s)$, $a_n(\cdot)$ could be any function that is non-decreasing w.r.t. s_n and no greater than $\frac{1}{n}$; however, if $t_n(s) > 0$ in at least some s_n, we could reduce $a_i(s)$ from $\frac{1}{n}$ slightly for $i \neq n$ and still achieve $\frac{n-1}{n}$ of $v(s)$. This is why we do not have a simple necessary condition for the solution of problem (3.1) on S_1. However, when expected revenue is concerned, the solution that cannot be improved for each $s \in S_1$ is to set $a_i(s) = \frac{1}{n}$ for all $i \neq n$ and $a_n(s) = \frac{1}{n}$ for not all but some $s_n > 0$.

(2) $b < 0$ if $Q(\mu) < \frac{n-1}{n}$. Hence as k increases, the lower bound on X_k gets smaller. In fact, if $Q(\mu) < \frac{2(n-1)}{2n+\frac{1}{n-1}}$, all these inequalities on X_k for $k = 2, \cdots, n$ simply reduce to non-negative constraints.

Chapter 2

Budget balanced and almost efficient assignment of multiple objects

2.1 Introduction

Uniform randomization (lottery) is common practice in assignment. When everyone has equal claim, lottery is appealing in that it is simple and fair. However, it is also notably inefficient, since people usually appreciate the same thing to different extent. If quasi-linear preferences are assumed and money transfers are allowed, there is an easy way to improve on efficiency while keeping the favorable aspects of the lottery: suppose one object is to be assigned to n agents and $v_i \in \mathbb{R}_+$ is agent i's private valuation; write v_{*k} to be the kth highest valuation. Let agent with v_{*1} get the object with probability $\frac{n-1}{n}$ and pay $\frac{n-2}{n}v_{*2}$, agent with v_{*2} get the object with probability $\frac{1}{n}$ and pay nothing, all other agents receive $\frac{1}{n}v_{*2}$. It is easy to check that this mechanism is strategy-proof, budget balanced, anonymous, envy-free, and each agent i gets at least $\frac{1}{n}v_i$: nice properties apparently shared by the lottery. When (ex-ante) surplus is concerned, the current mechanism loses $\frac{1}{n}v_{*1}$ at most; the lottery, however, loses $\frac{n-1}{n}v_{*1}$ in the worst case.

We define worse efficiency loss or simply efficiency loss to be the largest ratio of surplus loss to the efficient surplus, over all profiles of non-negative valuations[1]. We show that the above mechanism is the unique one to achieve the smallest efficiency loss among all strategy-proof, budget balanced, anonymous, envy-free *linear* mechanisms. By linear mechanism, we mean mechanism whose allocation rule can be represented by a vector $(a_1, a_2, \cdots, a_n)$, where a_k is the probability that agent with the kth highest valuation gets an object.

[1]The worst case analysis has been applied to different settings, see the introduction of [49] for a brief survey.

In general, if p identical mechanisms are to be assigned to n agents ($p < n$) and each agents demands at most one, then the optimal mechanism gives one object to each of the $p-1$ agents with the highest evaluations, a large probability to the agent with the pth highest valuation, and the remaining probability to the agent with the $(p+1)$th highest valuation. The optimal efficiency loss is $\frac{n-p}{n^2-n}$. Note that for $p = 1$, the optimal mechanism is the same as "randomly picking one agent to be the residual claimer of the pivotal mechanism run among the other agents (proposed in [26])"; for $p > 1$, the two are different.

We also find out the optimal efficiency loss when "envy freeness" is replaced by "voluntary participation", a weaker requirement in our setting. It turns out the above mechanism remains optimal when p is larger than $\bar{p}(n)$, a threshold less than $\frac{n}{2}$. For p no greater than $\bar{p}(n)$, an optimal voluntary mechanism will assign an equally small probability to a group of agents (rather than just one agent) following the agent with the pth highest valuation, and the optimal efficiency loss is bounded away from $\frac{n-p}{n^2-n}$.

To solve for the optimal efficiency loss, we first characterize the whole class of anonymous, strategy-proof, and budget balanced linear mechanisms. This class, to our best knowledge, appears the first time in literature. It turns out that in order to be budge balanced, the coordinates of the allocation vector only need to satisfy one linear equality. As showed in the paper, when budget balance is a hard constraint, this class is a nice starting point for further exploration.

literature review

When allocating objects among a group of agents, allocation efficiency requires to assign the objects to the agents with the highest valuations; however it is well know that no truth-telling mechanism can be allocation efficient and budget balanced at the same time. Different approaches exist in encountering the impossibility.

Several papers enforce allocation efficiency and explore the idea of redistributing VCG payments while respecting incentives. (See [15], [49] [31]). [49] and [31] find independently the VCG mechanisms that minimize (relative) budget surplus. In particular, [49] points out that if participation is voluntary, no VCG mechanism guarantees that budget surplus remains relatively small when p is close to n; and in the most dramatic $p = n-1$ case, we may lose all the welfare if all budget surplus has to be burned.

In [30], [51], [23], etc., welfare loss is defined to be the sum of efficient allocation loss and budget surplus. [30] compute the optimal (relative) efficiency loss in the restricted class of linear mechanisms. [51] explores the trade-offs between k-fairness (a concept of equity first proposed by [58]) and the optimal efficiency loss in the general class of (deterministic) strategy-proof mechanism. [23] deals with a similar problem as in [30], in a Bayesian setting, for two agents. Interestingly, they find out that under a mild distribution assumption, an optimal deterministic mechanism will be budget balanced.

Note that in all the papers mentioned above (except [23]), efficiency loss is defined relative to the efficient surplus. Even if the worst efficiency loss, which is a ratio, is close to zero, it is still possible that a huge amount of money will be burn if agents have high enough valuations. While uniform randomization (lottery) is wildly used, one rarely observes in reality allocations accompanied by burning money: mis-allocation seems to be more easily tolerated for whatever reason. Our paper, inspired by the lottery, enforce budget balance as a hard constraint and search for mechanisms that behave well in efficiency. Building on the linear mechanism proposed in [30], we characterize the class of all budget balanced linear mechanisms, and solve our optimization problems analytically. The optimal mechanism we obtain share all the good aspects of the lottery (except for the "money-free" property), and is hoped to be an easy alternative. Additionally, to compare with [51], we are able to bound the efficiency loss of optimal voluntary mechanism by switching to a different problem,

[32] enforces budget balance in a pubic good provision setting. Budget balance is achieved based on the idea proposed in [26]. That is, one agent is excluded at random and is made the residual claimant of the payments collected by a VCG mechanism in the market with only the remaining agents. In a public good provision setting with exclusion, [36] studies the efficiency and fairness properties of the equal cost sharing with maximal participation mechanism (the mechanism is budget balanced) and find out conditions in which the mechanism is optimal. In a cost-sharing setting, [53], [37] discuss the trade-off between budget balance and allocation efficiency for (group) strategyproof cost sharing mechanisms. In a bilateral trading setting, two classic papers ([55], [33]) also enforce budget balance: the former characterizes mechanisms that are Bayesian-Nash incentive compatible, interim individually rational and budget balanced; the latter characterizes mechanisms that are dominant strategy incentive compatible, ex-post individually rational and budget balanced.

2.2 Setting

p identical objects are to be assigned to n agents; and $1 \leq p < n$. Each agent $i \in N = \{1, \cdots, n\}$ demands at most one object and has a private valuation $v_i \in \mathbb{R}_+$ for the object. For any profile of valuations $v \in \mathbb{R}_+^N$, for any $i \in N$, v_{-i} stands for the vector obtained from v by deleting v_i, and (v', v_{-i}) stands for the valuation profile obtained from v by replacing v_i with v_i'.

We consider direct mechanisms, that is, agents report their own valuations and the allocation only depend on the reported valuation profile.

Definition 2.1. A direct mechanism is a pair (σ,t) such that

$$\sigma : \mathbb{R}_+^N \to [0,1]^n, \quad t : \mathbb{R}_+^N \to \mathbb{R}^n;$$

for any $v \in \mathbb{R}_+^N$, for any $i \in N$, $\sigma_i(v)$ is the probability that i gets an object and $t_i(v)$ is her payment.

Assuming risk neutral preference, agent i's utility under (σ,t) is $u_i(v) = \sigma_i(v) \cdot v_i - t_i(v)$ for each v.

A direct mechanism (σ,t) is *feasible* if $\sum_i \sigma_i(v) \leq p$ for any v; is *strategy-proof* if for any i, for any v_i, v_i', v_{-i}, $u_i(v_i,v_{-i}) \geq u_i(v_i',v_{-i})$; is *anonymous* if $\sigma(\cdot)$ and $t(\cdot)$ are both symmetric in all its variables; is *budget balanced* if $\sum_i t_i(v) = 0$ for any v; is *envy-free* if $\sigma_i(v) \cdot v_i - t_i(v) \geq \sigma_j(v) \cdot v_i - t_j(v)$ for any i, j, and for any v; is *voluntary* (or satisfies the constraint of *voluntary participation)* if for any i, for any v, $u_i(v) \geq 0$.

We further restrict our attention to *linear* mechanism whose allocation rule can be represented by a vector $(a_1, a_2, \cdots, a_n) \in [0,1]^n$, where a_l is the probability of getting an object for the agent with the lth highest valuation. The formal definition is given below.

For any $v \in \mathbb{R}_+^n$, let $r_i(v) = \{l \in \mathbb{N} : |\{j \in N : v_j > v_i\}| < l \leq n - |\{j \in \mathbb{N} : v_j > v_i\}|\}$. That is, $r_i(v)$ is agent i's set of rankings.

Definition 2.2. A direct mechanism (σ,t) is a linear mechanism if there exists $\{a_l\}_{l=1}^n$, with $a_l \in [0,1]$ constant for each l, such that $\sigma_i(v) = \frac{\sum_{l \in r_i(v)} a_l}{|r_i(v)|}$ for any i, any v.

Now we write a linear mechanism as $(\{a_l\};t)$. Let $\mathscr{M}$ be the set of linear mechanisms that are feasible, strategy-proof and anonymous.

Lemma 2.1. *A linear mechanism $(\{a_l\};t)$ is in $\mathscr{M}$ if and only if*

$$a_n \leq a_{n-1} \leq \cdots \leq a_1;$$

$\sum_{l=1}^n a_l \leq p$; and for any i, any v,

$$t_i(v) = \sigma_i(v) v_i - \int_0^{v_i} \sigma_i(x_i, v_{-i}) \mathrm{d}x_i - h(v_{-i})$$

where $\sigma_i(v) = \frac{\sum_{l \in r_i(v)} a_l}{|r_i(v)|}$ and $h(\cdot)$ is a symmetric function from $\mathbb{R}_+^{n-1}$ to $\mathbb{R}$.

Proof. See [30] for the proof of the $a_n \leq a_{n-1} \leq \cdots \leq a_1$ part ; the proof of the remaining part is omitted. $\square$

Now we use $(\{a_l\};h)$ to refer to a mechanism in $\mathscr{M}$. Let $\mathscr{V} = \{v \in \mathbb{R}^n_+ : v_1 \geq v_2 \geq \cdots \geq v_n \,\text{and}\, v_1 > 0\}$. For any $x \in \mathbb{R}^{n-1}_+$, let $x_{*1}, x_{*2}, \cdots, x_{*(n-1)}$ be a reordering of $x_1, x_2, \cdots, x_{n-1}$ with $x_{*1} \geq x_{*2} \geq \cdots \geq x_{*(n-1)}$.

Lemma 2.2. *A mechanism $(\{a_l\};h) \in \mathscr{M}$ is budget balanced iff for all $x \in \mathbb{R}^{n-1}_+$, $h(x) = \sum_{k=1}^{n-1} \delta_k x_{*k}$ for some $\{\delta_k\}_{k=1}^{n-1}$ that satisfy the following system of linear equations:*

$$k\delta_k + (n-k-1) \cdot \delta_{k+1} = k(a_k - a_{k+1}) \tag{2.1}$$

for all $0 \leq k \leq n-1$, where $\delta_0 = \delta_n = a_0 = 0$.

Proof. By Lemma 2.1, we have that for all $v \in \mathscr{V}$,

$$\begin{aligned} \sum_{i=1}^{n} t_i(v) &= \sum_{i=1}^{n} a_i v_i - \sum_{i=1}^{n}\sum_{j=i}^{n} (v_j - v_{j+1}) a_j - \sum_{i=1}^{n} h(v_{-i}) \\ &= \sum_{i=1}^{n-1} i \cdot (a_i - a_{i+1}) v_{i+1} - \sum_{i=1}^{n} h(v_{-i}) \end{aligned} . \tag{2.2}$$

We show in the appendix (Lemma 2.6) that $\sum_i t_i(v) = 0$ for all $v \in \mathscr{V}$ implies that $h(\cdot)$ takes a "linear" form. That is, for any $x \in \mathbb{R}^{n-1}_+$,

$$h(x) = \delta_1 x_{*1} + \delta_2 x_{*2} + \cdots + \delta_{n-1} x_{*(n-1)}.$$

Since for any $v \in \mathscr{V}$,

$$\sum_{i=1}^{n} h(v_{-i}) = \sum_{i=1}^{n} ((i-1)\delta_{i-1} + (n-i)\delta_i) v_i \quad ,$$

with $\delta_0 = \delta_n = 0$, it is easy to see that $\sum_i t_i(v) = 0$ for all $v \in \mathscr{V}$ if and only if the system of linear equations (2.1) holds. □

Remark 2.1. The system of linear equations (2.1) have $n-1$ unknowns $\{\delta_k\}_{k=1}^{n-1}$ and n equations. By straightforward calculation (see Lemma 2.7 in the appendix for details) the linear system has a solution if and only if the following constraint on $\{a_l\}_{l=1}^{n}$ holds:

$$\sum_{l=1}^{n} (-1)^{l-1} \cdot \gamma_l \cdot a_l = 0, \tag{2.3}$$

where $\gamma_l = \mathscr{C}_{l-2}^{n-2} + \mathscr{C}_{l-1}^{n-2}$ for $n > 2$ and $l = 2, \cdots, n-1$, and $\gamma_1 = \gamma_n = 1$.

Let $\mathscr{M}^b$ be the set of budget balanced mechanisms in $\mathscr{M}$. Now we use $(\{a_l\};\{\delta_k\})$ to refer to a mechanism in $\mathscr{M}^b$. The *worst efficiency loss,* or simply the *efficiency loss* for any mechanism $(\{a_l\};\{\delta_k\}) \in \mathscr{M}^b$ is the largest ratio of surplus loss to efficient surplus, over all profiles of non-negative valuations:

$$\max_{v\in\mathscr{V}}\{1-\frac{\sum_{l=1}^{n} a_l\cdot v_l}{\sum_{l=1}^{p} v_l}\} \tag{2.4}$$

Lemma 2.3. *For any $a_1 \geq a_2 \geq \cdots \geq a_n \geq 0$,*

$$\frac{\sum_{l=1}^{p} a_l}{p} = \min_{v\in\mathscr{V}} \frac{\sum_{l=1}^{n} a_l v_l}{\sum_{l=1}^{p} v_l}. \tag{2.5}$$

Proof. For any $v \in \mathscr{V}$, for any i, j, we have

$$(a_i - a_j)(v_i - v_j) \geq 0.$$

That is

$$a_i v_i + a_j v_j \geq a_i v_j + a_j v_i.$$

Adding up all pairs of (i, j), we have

$$p(a_1v_1 + a_2v_2 + \cdots + a_pv_p) \geq (a_1 + a_2 + \cdots + a_p)(v_1 + v_2 + \cdots + v_p).$$

Therefore for any v, we have

$$\frac{\sum_{l=1}^{n} a_l v_l}{\sum_{l\leq p} v_l} \geq \frac{\sum_{l\leq p} a_l v_l}{\sum_{l\leq p} v_l} \geq \frac{\sum_{l\leq p} a_l}{p}. \tag{2.6}$$

If we take $v_1 = v_2 = \cdots = v_p > 0$, and $v_{p+1} = v_{p+2} = \cdots = v_n = 0$, we have that (2.6) hold with equality. Therefore we have (2.5). □

Remark 2.2. By Lemma 2.3 and the definition of efficiency loss, we see that minimizing efficiency loss within $\mathscr{M}^b$ is equivalent to maximizing $\sum_{l=1}^{p} a_l$ of $(\{a_l\},\{\delta_k\}) \in \mathscr{M}^b$.

Lemma 2.4. *A mechanism $(\{a_l\};\{\delta_k\}) \in \mathscr{M}^b$ is envy-free iff $\delta_k \geq 0$ for all $1 \leq k \leq n-1$.*

Proof. First we show the "only if" part. Suppose $v_{l+1} > v_l > v_{l+1} > v_{l+2}$, then by Lemma 2.1 and Lemma 2.2, we have

$$t_l - t_{l+1} = (a_l - a_{l+1})v_{l+1} - \delta_l(v_{l+1} - v_l) = \delta_l v_l + (a_l - a_{l+1} - \delta_l)v_{l+1}.$$

Envy-freeness says that $a_{l+1}v_{l+1} - t_{l+1} \geq a_l v_{l+1} - t_l$. And $a_l \geq a_{l+1}$. Therefore we have

$$t_l - t_{l+1} \geq (a_l - a_{l+1})v_{l+1} \geq 0.$$

Therefore we must have $\delta_l \geq 0$; otherwise $t_l - t_{l+1} \geq 0$ will fail for $v_l >> v_{l+1}$.

Now we show the "if" part. Suppose $\delta_k \geq 0$ for all $1 \leq k \leq n-1$, then according to (2.1),

$$a_k - a_{k+1} - \delta_k = \frac{n-k-1}{k}\delta_{k+1} \geq 0.$$

Suppose $v_{i+1} > v_i > v_j > v_{j+1}$ (hence $i < j$)[2], then

$$\begin{aligned} t_i - t_j &= \sum_{l=i}^{j-1}(a_l - a_{l+1})v_{l+1} + \delta_i v_i - \sum_{l=i}^{j-2}(\delta_l - \delta_{l+1})v_{l+1} - \delta_{j-1}v_j \\ &= \delta_i v_i + \sum_{l=i}^{j-2}(a_l - a_{l+1} - \delta_l + \delta_{l+1})v_{l+1} + (a_{j-1} - a_j - \delta_{j-1})v_j \end{aligned}.$$

Since $a_l - a_{l+1} - \delta_l + \delta_{l+1} \geq a_l - a_{l+1} - \delta_l \geq 0$ for any $i \leq l \leq j-2$ and $a_{j-1} - a_j - \delta_{j-1} \geq 0$ for any j, we have that

$$t_i - t_j \leq \delta_i v_i + \sum_{l=i}^{j-2}(a_l - a_{l+1} - \delta_l + \delta_{l+1})v_i + (a_{j-1} - a_j - \delta_{j-1})v_i = (a_i - a_j)v_i,$$

and

$$t_i - t_j \geq \delta_i v_j + \sum_{l=i}^{j-2}(a_l - a_{l+1} - \delta_l + \delta_{l+1})v_j + (a_{j-1} - a_j - \delta_{j-1})v_j = (a_i - a_j)v_j.$$

Therefore no envy occurs between i and j. □

Lemma 2.5. *A mechanism* $(\{a_l\};\{\delta_k\}) \in \mathscr{M}^b$ *is voluntary iff for all* $1 \leq k \leq n-1$,

$$\sum_{j=1}^{k}\delta_j \geq 0. \tag{2.7}$$

Proof. First note that a mechanism $(\{a_l\};h) \in \mathscr{M}$ is voluntary iff $h(x) \geq 0$ for all $x \in \mathbb{R}_+^{n-1}$. Hence a mechanism $(\{a_l\};\{\delta_k\}) \in \mathscr{M}^b$ is voluntary iff $\sum_{k=1}^{n-1}\delta_k x_{*k} \geq 0$ for all $x \in \mathbb{R}_+^{n-1}$. This condition holds iff $\sum_{j=1}^{k}\delta_j \geq 0$ for all $1 \leq k \leq n-1$. □

Remark 2.3. A mechanism is envy-free implies that it is voluntary, but not visa verse.

[2] We omit in the proof the case when ties exist.

2.3 Optimal envy-free mechanism

Let $\mathscr{M}_e^b$ be the set of envy-free mechanisms in $\mathscr{M}^b$.

Definition 2.3. The Optimal Envy-free (OE) mechanism $(\{a_l^\star\};\{\delta_k^\star\})$ is the following:
$a_l^\star = 1$ for all $l < p$, $a_{p+1}^\star = 1 - a_p^\star = \frac{(n-p)p}{n^2-n}$, $a_l^\star = 0$ for all $l > p+1$;
$\delta_k^\star = 0$ for all $k < p$ and $k > p+1$, $\delta_p^\star = \frac{(p-1)p}{n^2-n}$, $\delta_{p+1}^\star = \frac{(n-p)p}{n^2-n}$.

It is easy to check that the OE mechanism is in $\mathscr{M}_e^b$. Note that if $p = 1$, we have that $a_1^\star = \frac{n-1}{n}$ and $a_2^\star = \frac{1}{n}$, the mechanism pointed out in the introduction.

Proposition 2.1. *For any $v \in \mathbb{R}_+^N$, the OE mechanism guarantees at least $\frac{p}{n} \cdot v_i$ for each agent i.*

Proof. Without loss of generality assume $v_1 \geq v_2 \geq \cdots \geq v_n$. Note that we have $a_p^\star - a_{p+1}^\star - \delta_p^\star = \frac{n-p-1}{p}\delta_{p+1}^\star > 0$ by 2.1. Since

$$u_l^\star = v_l - (1-a_p^\star)v_p - (a_p^\star - a_{p+1}^\star - \delta_p^\star)v_{p+1} \geq [1-(1-a_p^\star)-(a_p^\star - a_{p+1}^\star - \delta_p^\star)]v_l = \frac{p}{n}\cdot v_l, \forall l < p,$$
$$u_p^\star = a_p^\star v_p - (a_p^\star - a_{p+1}^\star - \delta_p^\star)v_{p+1} \geq [a_p^\star - (a_p^\star - a_{p+1}^\star - \delta_p^\star)]v_p = \frac{p}{n}\cdot v_p,$$
$$u_{p+1}^\star = a_{p+1}^\star v_{p+1} + \delta_p^\star v_p \geq (a_{p+1}^\star + \delta_p^\star)v_{p+1} = \frac{p}{n}\cdot v_{p+1},$$
$$u_l^\star = \delta_p^\star v_p + \delta_{p+1}^\star v_{p+1} \geq (\delta_p^\star + \delta_{p+1}^\star)v_l = \frac{p}{n}\cdot v_l, \forall l > p+1,$$

we are done. □

Theorem 2.1. *The OE mechanism is the unique mechanism that achieves the smallest efficiency loss in $\mathscr{M}_e^b$.*

Proof. Suppose there exists another mechanism $(\{a_l\};\{\delta_k\}) \in \mathscr{M}_e^b$ that achieves the same or smaller worst efficiency loss, then by Lemma 2.3 we must have $a_p \geq a_p^\star$, $a_{p+1} \leq a_{p+1}^\star$, and $a_p - a_{p+1} > a_p^\star - a_{p+1}^\star$.

Since $(p+1)\delta_{p+1} + (n-p-2)\delta_{p+2} = (p+1)(a_{p+1} - a_{p+2})$, and $\delta_{p+2} \geq 0 = \delta_{p+2}^\star$, $a_{p+1} \leq a_{p+1}^\star$, $a_{p+2} \geq 0 = a_{p+2}^\star$, we must have $\delta_{p+1} \leq \delta_{p+1}^\star$.

Since $(p-1)\delta_{p-1} + (n-p)\delta_p = (p-1)(a_{p-1} - a_p)$, and $\delta_{p-1} \geq 0 = \delta_{p-1}^\star$, $a_{p-1} \leq 1 = a_{p-1}^\star$, $a_p \geq a_p^\star$, we have that $\delta_p \leq \delta_p^\star$. However since $p\delta_p + (n-p-1)\delta_{p+1} = p(a_p - a_{p-1})$, and $\delta_p \leq \delta_p^\star$, $a_p - a_{p+1} > a_p^\star - a_{p+1}^\star$, we must have $\delta_{p+1} > \delta_{p+1}^\star$. Contradiction! □

Therefore the smallest efficiency loss $L^e(n,p)$ in $\mathscr{M}_e^b$ is $\frac{n-p}{n^2-n}$.

Example 2.1. The OE mechanism for $n = 4$, $p = 2$.

$a_1^\star = 1$, $a_2^\star = \frac{2}{3}$, $a_3^\star = \frac{1}{3}$, $a_4^\star = 0$; since $\delta_1^\star = 0$, $\delta_2^\star = \frac{1}{6}$, $\delta_3^\star = \frac{1}{3}$, the payment functions (under the assumption that $v_1 > v_2 > v_3 > v_4$) are

$$\begin{aligned}
t_1 &= (\frac{1}{3}v_2 + \frac{1}{3}v_3 + \frac{1}{3}v_4) - (\frac{1}{6}v_3 + \frac{1}{3}v_4) && = \frac{1}{3}v_2 + \frac{1}{6}v_3 \\
t_2 &= \frac{1}{3}v_3 + \frac{1}{3}v_4 - (\frac{1}{6}v_3 + \frac{1}{3}v_4) && = \frac{1}{6}v_3 \\
t_3 &= \frac{1}{3}v_4 - (\frac{1}{6}v_2 + \frac{1}{3}v_4) && = -\frac{1}{6}v_2 \\
t_4 &= && -\frac{1}{6}v_2 - \frac{1}{3}v_3
\end{aligned}.$$

One can check that $\sum_{i=1}^4 t_i = 0$ for all v. The smallest efficiency loss L^e if $\frac{1}{6}$.

2.4 Optimal voluntary mechanism

Let $\mathscr{M}_v^b$ be the set of voluntary mechanisms in $\mathscr{M}^b$. Let $\phi(n,p,s) = \frac{\mathscr{C}_{p+2s}^{n-2} + \mathscr{C}_{p-1}^{n-2}}{2s+1}$. For $p \leq \lfloor \frac{n}{2} \rfloor$, let $s^\star(n,p) = \min\{\tilde{s} : \tilde{s} \in \arg\max_{s \in \{0,1,\cdots,\lfloor \frac{n-2-p}{2} \rfloor\}} \phi(n,p,s)\}$. Let $\bar{p}(n) = \max\{p \leq \lfloor \frac{n}{2} \rfloor : s^\star(n,p) \neq 0\}$ if the set is non-empty; $\bar{p}(n) = 0$ otherwise.

Definition 2.4. The Optimal Voluntary (OV) mechanism $(\{a_l^o\}; \{\delta_k^o\})$ is the following:

if $p > \bar{p}(n)$, then $(\{a_l^o\}; \{\delta_k^o\}) = (\{a_l^\star\}; \{\delta_k^\star\})$;

if $p \leq \bar{p}(n)$, then

$$a_1^o = \cdots = a_{p-1}^o = 1,\ a_{p+1}^o = \cdots = a_{p+1+2s^\star}^o > 0,\ a_{p+2+2s^\star}^o = \cdots = a_n^0 = 0,$$

$a_p = 1 - (1+2s^\star)a_{p+1} = \frac{\phi(n,p,s^\star) + \mathscr{C}_{p-2}^{n-2}}{\phi(n,p,s^\star) + \mathscr{C}_{p-2}^{n-2} + \mathscr{C}_{p-1}^{n-2}}$, and $\{\delta_k^o\}$ is the solution of (2.1).

It is easy to check that $\{a_l^o\}$ satisfies the budget balance constraint (2.3). The following proposition shows that the OV mechanism differs with the OE mechanism if and only if $p \leq \bar{p}(n)$.

Proposition 2.2. *(1) Suppose $s^\star(n,p) > 0$ for some p, then $s^\star(n,p') > 0$ for all $p' < p$.*

(2) $\frac{n-8}{3} \leq \bar{p}(n) < \lfloor \frac{n}{2} \rfloor$.

Proof. See appendix. □

Example 2.2. The OV mechanism for $n = 15$, $p = 3$. (Note that $s^\star(15,3) = 1$.)

$a_1^o = a_2^o = 1$, $a_3^o = \frac{6}{7}$, $a_4^o = a_5^o = a_6^o = \frac{1}{21}$, $a_7^o = \cdots = a_{15}^o = 0$; and since $\delta_1^o = \delta_2^o = 0$, $\delta_7^o = \delta_8^o = \cdots = \delta_{14}^o = 0$, $\delta_3^o = \frac{1}{42}$, $\delta_4^o = \frac{3}{14}$, $\delta_5^o = -\frac{3}{35}$, $\delta_6^o = \frac{1}{21}$, the payment functions (under the assumption that $v_1 > v_2 > \cdots > v_{15}$) are

$$t_1 = \frac{1}{7}v_3 + \frac{11}{14}v_4 - \frac{3}{14}v_5 + \frac{3}{35}v_6$$
$$t_2 = \frac{1}{7}v_3 + \frac{11}{14}v_4 - \frac{3}{14}v_5 + \frac{3}{35}v_6$$
$$t_3 = \frac{11}{14}v_4 - \frac{3}{14}v_5 + \frac{3}{35}v_6$$
$$t_4 = -\frac{1}{42}v_3 - \frac{3}{14}v_5 + \frac{3}{35}v_6$$
$$t_5 = -\frac{1}{42}v_3 - \frac{3}{14}v_4 + \frac{3}{35}v_6$$
$$t_6 = -\frac{1}{42}v_3 - \frac{3}{14}v_4 + \frac{3}{35}v_5$$
$$t_7 = \cdots = t_{15} = -\frac{1}{42}v_3 - \frac{3}{14}v_4 + \frac{3}{35}v_5 - \frac{1}{21}v_6$$

One can check that $\sum t_i = 0$ for any v. Since $\sum_{l=1}^{k} \delta_l^o \geq 0$ for all $1 \leq k \leq n-1$, the mechanism is voluntary. And the efficiency loss is $\frac{1}{21} \approx 0.0476$. We show below that the the OV mechanism is voluntary for any $p < n$, and it achieves the smallest efficiency loss among mechanisms in $\mathscr{M}_v^b$.

Proposition 2.3. *The OV mechanism is voluntary.*

Proof. We only need to show it for $p \leq \bar{p}(n)$. According to (2.1), we have $\delta_1^o = \cdots = \delta_{p-1}^o = 0, \delta_{p+2+2s^\star(p)}^o = \cdots = \delta_{n-1}^o = 0$; and $(\delta_p^o, \delta_{p+1}^o, \cdots, \delta_{p+1+2s^\star}^o)$ are of the signs $(+,+,-,+,-,\cdots,+,-,+)$. Furthermore, since $p+1+2s^\star(p) \leq \lfloor \frac{n}{2} \rfloor + 1$, we have $|\delta_p^o| > |\delta_{p+1}^o| > \cdots > |\delta_{p+1+2s^\star}^o|$, according to (2.1). Hence $\sum_l^k \delta_l^o \geq 0$ for all $1 \leq k \leq n-1$. □

Theorem 2.2. *The OV mechanism achieves the optimal efficiency loss in $\mathscr{M}_v^b$.*

Proof. The very long proof in the appendix is divided into three parts: first we find out the mechanism that achieve the smallest efficiency loss in $\mathscr{M}^b$, by very careful perturbations. Second we show that the optimal mechanism in $\mathscr{M}^b$ is voluntary iff $p < n - \bar{n}(p)$ (hence it is the same as the OV mechanism for $p < n - \bar{n}(p)$). Finally we show that for $p \geq n - \bar{n}(p)$, the OV mechanism is optimal in $\mathscr{M}_e^b$, using the Duality Theorem. □

Now we look at the smallest efficiency loss $L^v(n,p)$ in $\mathscr{M}_v^b$.

For $p > \bar{p}(n)$

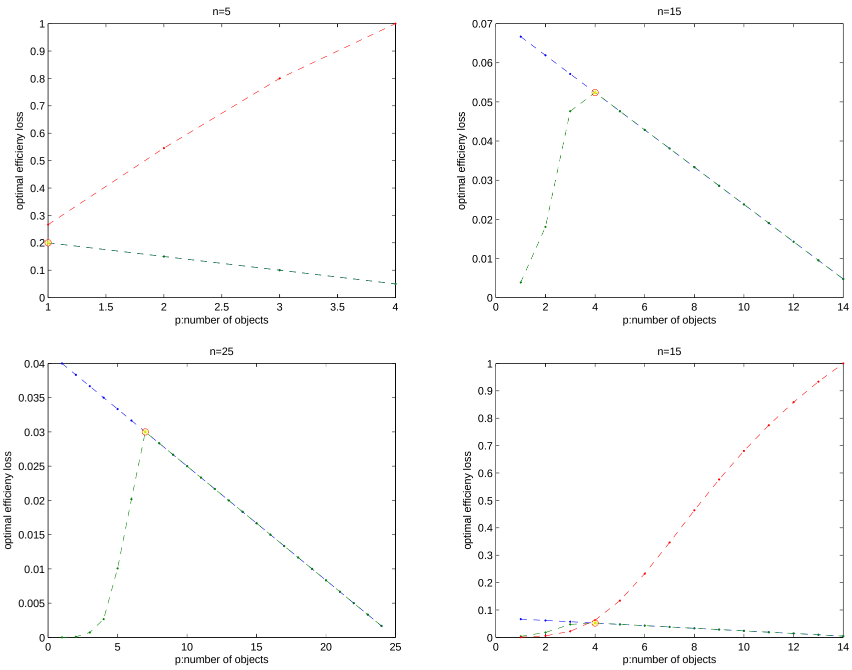
$$L^v(n,p) = \frac{(n-p)}{n^2-n} \quad ;$$

for $p \leq \bar{p}(n)$

$$L^v(n,p) = \frac{1}{p} \cdot \frac{\mathscr{C}_{p-1}^{n-2}}{\mathscr{C}_{p-2}^{n-2} + \mathscr{C}_{p-1}^{n-2} + \frac{\mathscr{C}_{p-1}^{n-2} + \mathscr{C}_{p+2s^\star}^{n-2}}{1+2s^\star}}.$$

Fix n, if $p > \bar{p}(n)$, then $L^v(n,p)$ is decreasing with respect to p. However, the monotone properties of $L^v(n,p)$ otherwise remains an open question, though we do see a pattern in the examples below.

Example 2.3. Optimal worst efficiency for $n = 5$, 15, 25, and $p = 1, \cdots, n-1$.



In each graph the green line is $L^v(n,p)$, and the blue line is $L^e(n,p)$. (When $n = 5$, the two lines coincide.) And the yellow point is the largest $L^v(n,p)$ as p changes from 1 to $n-1$, and it is also the point where $p = \bar{p}(n) + 1$.

In the first and the fourth graphs, the red line represents the smallest efficiency loss with voluntary participation in [49]. For $n = 5$, it is above L^v. In fact, one can check that this domination holds for all $n \leq 5$. For $n > 5$, however, no domination occurs. As showed in the fourth graph, the OV mechanism does worse when p is much smaller than n, but much better when p is close to n.

2.5 Appendix

2.5.1 Lemma 2.6 and its proof

Lemma 2.6. *Suppose a symmetric function $h:\mathbb{R}_+^{n-1}\to\mathbb{R}$ satisfies the following functional equation: for any $(x_1,x_2,\cdots,x_n)\in\mathscr{V}$,*

$$\sum_i h(x_{-i})=\sum_{i=2}^n \lambda_i x_i,$$

where $\lambda_i\geq 0$ for all $i\geq 2$, then $h(z)=\delta_1 z_{*1}+\delta_2 z_{*2}+\cdots+\delta_{n-1}z_{*(n-1)}$ for some $(\delta_1,\cdots,\delta_{n-1})\in\mathbb{R}^{n-1}$, where $(z_{*1},z_{*2},\cdots,z_{*n-1})$ is a reordering of $(z_1,z_2,\cdots,z_n)$ such that $z_{*1}\geq z_{*2}\geq\cdots\geq z_{*(n-1)}$.

Proof. We show it by induction. Since

$$\begin{aligned}\lambda_2x_2+\lambda_3x_3+\cdots+\lambda_nx_n = {} & h(x_2,x_3,\cdots,x_n)\\ & +h(x_1,x_3,\cdots,x_n)\\ & +\cdots\\ & +h(x_1,x_2,\cdots,x_{n-1})\end{aligned},$$

taking $x_1=\cdots=x_n=0$, we have $h(0,0,\cdots,0)=0$. Taking $x_1\geq x_2=\cdots=x_n=0$, we have $h(x_1,0,\cdots,0)=0\triangleq\delta_1x_1$. Now suppose $h(x_1,x_2,\cdots,x_k,0,\cdots,0)=\delta_1'x_1+\delta_2'x_2+\cdots+\delta_k'x_k$ for $x_1\geq x_2\geq\cdots\geq x_k\geq 0$, we show that $h(x_1,x_2,\cdots,x_k,x_{k+1}0,\cdots,0)=\delta_1''x_1+\delta_2''x_2+\cdots+\delta_k''x_k+\delta_{k+1}''x_{k+1}$.

Since

$$\begin{aligned}\lambda_2x_2+\lambda_3x_3+\cdots+\lambda_kx_k+\lambda_{k+1}x_{k+1} = {} & h(x_2,x_3,\cdots,x_{k+1},0,\cdots,0)\\ & +h(x_1,x_3,\cdots,x_{k+1},0,\cdots,0)\\ & +\cdots\\ & +h(x_1,x_2,\cdots,x_k,0,\cdots,0)\\ & +(n-k-1)h(x_1,x_2,\cdots,x_k,x_{k+1},0,\cdots,0)\end{aligned},$$

It is easy to see that $h(x_1,x_2,\cdots,x_k,x_{k+1},0,...,0)=\delta_1''x_1+\delta_2''x_2+\cdots\delta_k''x_k+\delta_{k+1}''x_{k+1}$. □

2.5.2 Lemma 2.7 and its proof

Lemma 2.7. *The system of linear equations* (2.1) *of* $\{\delta_1, \delta_2, \cdots, \delta_{n-1}\}$ *has a solution if and only if the constraint* (2.3) *on* $\{a_i\}$ *holds.*

Proof. The system of linear equations (2.1) have $n-1$ unknowns and n equations.

We show the "only if" part. According to the first $n-1$ equations of the system, we have

$$\begin{aligned}\delta_{n-1} = &(-1)^{n-1} \cdot \prod_{i=1}^{n-2} \lambda_i \cdot a_1 + (-1)^{n-2} \cdot (\prod_{i=1}^{n-2} \lambda_i + \prod_{i=1}^{n-3} \lambda_i) \cdot a_2 + (-1)^{n-3} \cdot (\prod_{i=1}^{n-3} \lambda_i + \prod_{i=1}^{n-4} \lambda_i) \cdot a_3 \\ &+ \cdots + (-1)^2 \cdot (\prod_{i=1}^{2} \lambda_i + \lambda_1) \cdot a_{n-2} + (-1) \cdot \lambda_1 \cdot a_{n-1}\end{aligned} \tag{2.8}$$

where $\lambda_i = \frac{n-i-1}{i}$. Note that $\prod_{i=1}^{m} \lambda_i = \mathscr{C}_m^{n-2} = \mathscr{C}_{n-2-m}^{n-2} = \prod_{i=1}^{n-2-m} \lambda_i$.

According to the last equation of the system, we have □

$$\delta_{n-1} = a_{n-1} - a_n. \tag{2.9}$$

Combining (2.8) and (2.9), we have (2.3).

Now the "if" part is easy to see.

2.5.3 Proof for Proposition 2.2

Proof. (1) Suppose that $s^\star(n,p) > 0$, we show that $s^\star(n,p') > 0$ for $p' = p-1$.

Since $s^\star > 0$, we have that $\phi(n,p,s^\star) > \phi(n,p,0)$, that is, $\mathscr{C}_{p+2s^\star}^{n-2} - \mathscr{C}_p^{n-2} > 2s(\mathscr{C}_p^{n-2} + \mathscr{C}_{p-1}^{n-2})$. We show below that $\phi(n,p-1,s^\star) > \phi(n,p-1,0)$, that is, $\mathscr{C}_{p-1+2s^\star}^{n-2} - \mathscr{C}_{p-1}^{n-2} > 2s(\mathscr{C}_{p-1}^{n-2} + \mathscr{C}_{p-2}^{n-2})$. Note that

$$\begin{aligned}\mathscr{C}_{p-1+2s^\star}^{n-2} - \mathscr{C}_{p-1}^{n-2} &= \frac{p+2s^\star}{n-2-p-2s^\star+1}\mathscr{C}_{p+2s^\star}^{n-2} - \frac{p}{n-2-p+1}\mathscr{C}_p^{n-2} \\ &> \frac{p+2s^\star}{n-2-p-2s^\star+1}(\mathscr{C}_{p+2s^\star}^{n-2} - \mathscr{C}_p^{n-2})\end{aligned};$$

and

$$\begin{aligned}\mathscr{C}_{p-1}^{n-2} + \mathscr{C}_{p-2}^{n-2} &= \frac{p}{n-2-p+1}\mathscr{C}_p^{n-2} + \frac{p-1}{n-2-(p-1)+1}\mathscr{C}_{p-1}^{n-2} \\ &< \frac{p+2s^\star}{n-2-p-2s^\star+1}(\mathscr{C}_p^{n-2} + \mathscr{C}_{p-1}^{n-2})\end{aligned}.$$

We are done.

(2) First note that

$$\phi(n,p,s)=\frac{(\mathscr{C}_{p+2s}^{n-2}-\mathscr{C}_{p+2(s-1)}^{n-2})+(\mathscr{C}_{p+2(s-1)}^{n-2}-\mathscr{C}_{p+2(s-2)}^{n-2})+\cdots+(\mathscr{C}_{p+2}^{n-2}-\mathscr{C}_{p}^{n-2})+(\mathscr{C}_{p}^{n-2}+\mathscr{C}_{p-1}^{n-2})}{2s+1}.$$

Let $\bar{s}(n,p)=\min\{s:s\in\arg\max_{s\in\{1,2,\cdots,\lfloor\frac{n-2-p}{2}\rfloor\}}(\mathscr{C}_{p+2s}^{n-2}-\mathscr{C}_{p+2(s-1)}^{n-2})\}$. Suppose $(\mathscr{C}_{p+2}^{n-2}-\mathscr{C}_{p}^{n-2})/2>\mathscr{C}_{p}^{n-2}+\mathscr{C}_{p-1}^{n-2}$, that is $p<\frac{n-5}{3}$, then $s^\star(n,p)=\bar{s}(n,p)$; suppose $(\mathscr{C}_{p+2}^{n-2}-\mathscr{C}_{p}^{n-2})/2\leq\mathscr{C}_{p}^{n-2}+\mathscr{C}_{p-1}^{n-2}$, that is, $p\geq\frac{n-5}{3}$, then $s^\star=0$ or $\bar{s}(n,p)$. Therefore we have that $\bar{p}(n)\geq\frac{n-5}{3}-1=\frac{n-8}{3}$. It is easy to see that $\bar{p}(n)<\lfloor\frac{n}{2}\rfloor$.

□

2.5.4 Proof of Theorem 2.2

2.5.4.1 Optimal mechanism in $\mathscr{M}^b$

We first solve the following problem:

$$\max_{\{a_l\}_{l=1}^n}\sum_{l=1}^{p}a_l \tag{2.10}$$

such that

$$\textstyle\sum_{l=1}^n a_l\leq p,\quad 0\leq a_n\leq\cdots\leq a_2\leq a_1\leq 1\ ;$$

and (2.3).

For $p>\lfloor\frac{n}{2}\rfloor$, let $s^\star(n,p)=s^\star(n,n-p)$.

Definition 2.5. $\{a_i\}_{i=1}^n$ is called the Optimal Budget balanced (OB) solution if:

(1) for $p\leq\lfloor\frac{n}{2}\rfloor$,

$$a_1=\cdots=a_{p-1}=1,\, a_{p+1}=\cdots=a_{p+1+2s^\star}>0,\, a_{p+2+2s^\star}=\cdots=a_n=0,$$

$a_p=1-(1+2s^\star)a_{p+1}=\frac{\phi(n,p,s^\star)+\mathscr{C}_{p-2}^{n-2}}{\phi(n,p,s^\star)+\mathscr{C}_{p-2}^{n-2}+\mathscr{C}_{p-1}^{n-2}}.$

(2) for $p>\lfloor\frac{n}{2}\rfloor$,

$$a_1=\cdots=a_{p-1-2s^\star}=1,\, a_{p-2s^\star}=\cdots=a_p>0,\, a_{p+2}=\cdots=a_n=0,$$

$a_{p+1} = (1+2s^\star)(1-a_p) = \frac{\mathscr{C}_{p-1}^{n-2}}{\phi(n,n-p,s^\star)+\mathscr{C}_p^{n-2}+\mathscr{C}_{p-1}^{n-2}}$

We write the OB solution as $\{a_i^\phi\}_{i=1}^n$. Note that the OB solution coincides with the allocation part of the OV mechanism $\{a_l^o\}$ for $p < n-\bar{n}(p)$.

Lemma 2.8. *$\{a_i^\phi\}_{i=1}^n$ is an optimal solution of the problem* (2.10).[3]

Proof. The very long proof is left in the next section 2.5.4.2. □

Remark 2.4. Since the optimal solution $\{a_i^\phi\}$ is the allocation part of the optimal mechanism in $\mathscr{M}^b$, we look at it more carefully.

When $p \le \lfloor \frac{n}{2} \rfloor$, the OB solution assigns one object to each of the $(p-1)$ agents with the highest valuations, and a substantial portion to the agent with the pth highest valuation. The remaining portion will be distributed to a group of agent immediately behind, and each receives an equal slice. The size of this "equal-group" depends on p and n.

When $p > \lfloor \frac{n}{2} \rfloor$, the OB solution assigns a slight portion to the agent with $(p+1)$th highest valuation, and an equally substantial portion to each agent in a group immediately before her, and one object to each of the agent before the group. Again the size of the "equal-group" depends on p and n.

Interestingly, the solution of the $p > \lfloor \frac{n}{2} \rfloor$ case can be constructed directly using the $p \le \lfloor \frac{n}{2} \rfloor$ case: first rank agents from the lowest valuation to the highest valuation and "distribute" $n-p \le \lfloor \frac{n}{2} \rfloor$ "null objects" to agents according to the OB solution (Hence agents with the lower valuations are the ones receiving "null objects"); then assign each agent 1 minus the portion of the "null object". Formally, let $\{b_i^\phi\}_{i=1}^n$ be the OB solution of the problem (2.10) with p replaced by $n-p$. Then we have for $i = 1,\cdots,n$, $a_i^\phi = 1 - b_{n+1-i}^\phi$. Intuitively, allocating p objects as efficiently as possible is equivalent to allocating $n-p$ "null-objects" as efficiently as possible.

Mathematically, the above Lemma reduces the original complicated maximization problem (2.10) to a small one: choosing s to maximize $\phi(n,p,s)$. Since $1+2s$ is the size of the "equal-group", so by choosing s we are determining the optimal size of the "equal-group". We know from Proposition 2.2 that when $\bar{p}(n) < p < n-\bar{p}(n)$, $s^\star(n,p) = 0$; that is, the "equal-group" reduces to a singleton.

2.5.4.2 Proof for Lemma 2.8

Proof. We could write an optimal solution as:

[3] If $S^\star(n,p)$ is a singleton, then $\{a_i^\phi\}_{i=1}^n$ is the optimal solution. If $S^\star(n,p)$ contains more than one elements, then the optimal solutions all have the same structure and differ only in the size of the "equal-group".

$$(a_1, a_2, \cdots, a_n) = (1, 1, \cdots, 1, a_{l+1}, \cdots, a_{n-m}, 0, \cdots, 0)$$

with $1 > a_{l+1} \geq a_{n-m} > 0$. Note that l is the number of agents who get one object; m is the number of agents who get zero object; and $n-m-l$ is the number of agents who get part of an object. We know that $l \in \{1, 2, \cdots, p-1\}$ (since first best is impossible), $m \in \{1, 2, \cdots, n-1\}$ (since zero vector is obviously not an optimal solution) and $l+m \leq n$. Let $k = \lfloor \frac{n}{2} \rfloor$.

Case 1: $p \leq k$**.**

Step 1: We show that in any optimal solution,$\sum_{i=1}^n a_i = p$.

Suppose not, increase a_{l+1},a_{l+2} by ε_{l+1} and ε_{l+2} respectively, where $\varepsilon_{l+2} = \frac{\gamma_{l+1}}{\gamma_{l+2}} \cdot \varepsilon_{l+1} < \varepsilon_{l+1}$. (Since $l+1 \leq p \leq k$, we have $\gamma_{l+1} < \gamma_{l+2}$). We can take ε_{l+1} small enough so that all constraints still hold and the value of the objective function increases.

Step 2: We show that $a_{k+2} = a_{k+3} = \cdots = a_n = 0$.

Suppose not, consider a_{n-m} and a_{n-m-1}. Decrease a_{n-m} and a_{n-m-1} by ε_{n-m} and ε_{n-m-1}, where $\varepsilon_{n-m} = \frac{\gamma_{n-m-1}}{\gamma_{n-m}} \cdot \varepsilon_{n-m-1} > \varepsilon_{n-m-1}$. (Since $n-m \geq k+2$, We have $\gamma_{n-m-1} > \gamma_{n-m}$.) What we get is still an optimal solution since all the constraints still hold and the the value of the objective function does not change. However we now have $\sum_{i=1}^n a_i < p$, contradicting step 1.

Step 3: We show that $n-m-p$ is odd, that is, the size of the tail group is an odd number.

First we know that $a_{p+1} \neq 0$. Otherwise we will have $\sum_{i=1}^n a_i = \sum_{i=1}^p a_i < p$, contradicting step 1. Now suppose $a_{p+1} \geq a_{p+2} > 0$ and $n-m-p$ is even. We could decrease $a_{p+1}, \cdots, a_{n-m}$ each by ε_{n-m}, increase a_{n-m+1} by ε_{n-m+1} and increase $a_{l+1}, \cdots, a_p$ each by ε_p such that $\varepsilon_p = (\varepsilon_{n-m}(\gamma_{n-m} - \cdots + \gamma_{p+1}) - \varepsilon_{n-m+1}\gamma_{n-m+1})/(\gamma_p - \gamma_{p-1} + (-1)^{p-l}\gamma_{l+1})$. Note that we can choose ε_{n-m} and ε_{n-m+1} properly so that $\varepsilon_p < [\varepsilon_{n-m}(n-m-p) - \varepsilon_{n-m+1}]/(p-l)$. (Take $\varepsilon_{n-m+1} = 0$, if $\varepsilon_p < \varepsilon_{n-m}(n-m-p)/(p-l)$, then we are done; if not, since $\varepsilon_p < 0$ if we take $\varepsilon_{n-m+1} = \varepsilon_{n-m}$, there exists some $\varepsilon_{n-m+1} < \varepsilon_{n-m}$ such that $0 < \varepsilon_p < \varepsilon_{n-m}(n-m-p)/(p-l)$ by continuity.) Hence the reduction is greater than the increase. And all the constraints still hold and the objective function increases.

Step 4: We show that $a_p > a_{p+1}$.

Since $a'_1 = a'_2 = \cdots = a'_{p-1} = 1$, $a'_p = 1 - \frac{(n-p)p}{n^2-n}$, $a'_{p+1} = \frac{(n-p)p}{n^2-n}$ (it is easy to check that $a'_{p+1} < a'_p$), $a'_{p+2} = a'_{p+3} \cdots = a'_n = 0$ is a feasible solution, any optimal solution $\{a_i\}$ that is different has to increase a'_p , and decrease $a'_1 + \cdots + a'_{p-1}$ by a less or equal margin. This cannot be achieved if $a_p = a_{p+1}$ since we then have to increase a'_p and a'_{p+1} simultaneously.

Step 5: We show that if $[\gamma_{n-m} - \gamma_{n-m-1} + \cdots - \gamma_{p+2} + \gamma_{p+1}]/(n-m-p) < \gamma_{p+1}$, then $n-m-p = 1$.

Suppose $[\gamma_{n-m}-\gamma_{n-m-1}+\cdots-\gamma_{p+2}+\gamma_{p+1}]/(n-m-p)<\gamma_{p+1}$, if $a_{p+2}\geq a_{p+3}\geq\cdots\geq a_{n-m}>0$, we could increase a_{p+1} (since $a_p>a_{p+1}$) by ε_{p+1}, decrease $a_{p+2},\cdots,a_{n-m}$ each by ε_{n-m}, where $\varepsilon_{n-m}=\gamma_{p+1}\varepsilon_{p+1}/[\gamma_{n-m}-\gamma_{n-m-1}+\cdots-\gamma_{p+2}]$, and keep all constraints hold. Now we have that $\varepsilon_{n-m}\cdot(n-m-p-1)>\varepsilon_{p+1}$, that is, reduction is larger than increase. However we have now $\sum_{i=1}^{n}a_i<p$, contradicting step 1; hence we must have $n-m-p=1$.

<u>Step 6</u>: We show that if $[\gamma_{n-m}-\gamma_{n-m-1}+\cdots-\gamma_{p+2}+\gamma_{p+1}]/(n-m-p)>\gamma_{p+1}$, then $n-m-p=1+2s'(n,p)$ for some $s'(n,p)\in S^{\star}(n,p)$.

Now suppose $n-m-p>1+2s'(n,p)$ for all $s'(n,p)\in S^{\star}(n,p)$We could decrease $a_{p+1+2s'(n,p)+1}$, $a_{p+1+2s'(n,p)+2}$, $\cdots$, a_{n-m} each by ε_{n-m} and increase a_{p+1}, a_{p+2}, a_{p+3},$\cdots$, $a_{p+1+2s'(n,p)}$ each by ε_{p+1},where $\varepsilon_{p+1}=\{[n-m-p-1-2s'(n,p)][\gamma_{n-m}-\gamma_{n-m-1}+\cdots+\gamma_{p+1+2s'(n,p)+2}-\gamma_{p+1+s'(n,p)+1}]\varepsilon_{n-m}\}/\{[1+2s'(n,p)][\gamma_{p+1+2s'(n,p)}-\gamma_{p+1+2s'(n,p)}+\cdots-\gamma_{p+2}+\gamma_{p+1}]\}$, and keep all constraints hold. Now we have $\varepsilon_{p+1}\cdot[1+2s'(n,p)]<\varepsilon_{n+m}\cdot[n-m-p-1-2s'(n,p)]$, that is, reduction is larger than increase. Therefore we have now $\sum_{i=1}^{n}a_i<p$, contradicting step 1; hence we must have $n-m-p\leq 1+2s'(n,p)$. Using similar argument, we can conclude that $n-m-p\geq 1+2s'(n,p)$. Therefore we have $n-m-p=1+2s'(n,p)$.

<u>Step 7</u>: We show that if $[\gamma_{n-m}-\gamma_{n-m-1}+\cdots-\gamma_{p+2}+\gamma_{p+1}]/(n-m-p)>\gamma_{p+1}$, and $s'(n,p)=(n-m-p-1)/2$, then $a_{p+1}=a_{p+2}=\cdots=a_{p+1+2s'(n,p)}$.

7a) First we show that $a_{p+2s}=a_{p+2s+1}$ for all $1\leq s\leq s'(n,p)$.

If not, we reduce a_{p+2s} by ε and reduce a_{n-m} by ε', where $\varepsilon'=\gamma_{p+2s}\varepsilon/\gamma_{n-m}<\varepsilon$. We can make ε small enough so that all the constraints still hold and the value of the objective function does not change. We have now that $\sum_{i=1}^{n}a_i<p$, contradicting step 1.

7b) Then we show that $a_{p+2}=a_{p+3}\cdots=a_{p+1+s'(n,p)}$.

Now suppose that $a_{p+1+2\tilde{s}+1}=a_{p+1+2\tilde{s}+2}>a_{p+1+2\tilde{s}+3}=a_{p+1+2\tilde{s}+4}$ for some $\tilde{s}\leq s'(n,p)-2$, then we could reduce both $a_{p+1+2\tilde{s}+1}$ and $a_{p+1+2\tilde{s}+2}$ by ε, increase both $a_{p+1+2\tilde{s}+3}$ and $a_{p+1+2\tilde{s}+4}$ by ε', where $\varepsilon=[\gamma_{p+1+2\tilde{s}+4}-\gamma_{p+1+2\tilde{s}+3}]\varepsilon'/[\gamma_{p+1+2\tilde{s}+2}-\gamma_{p+1+2\tilde{s}+1}]>\varepsilon'$. (Note that $\gamma_{p+1+2\tilde{s}+4}-\gamma_{p+1+2\tilde{s}+3}>\gamma_{p+1+2\tilde{s}+2}-\gamma_{p+1+2\tilde{s}+1}$ since $\tilde{s}\leq s'(n,p)-2$.) All constraint still hold and we now have $\sum_{i=1}^{n}a_i<p$, contradicting step 1; hence we must have $a_{p+2}=a_{p+3}\cdots=a_{p+1+2s'(n,p)}$.

7c) Finally we show that $a_{p+1}=a_{p+2}$.

If $a_{p+1}>a_{p+2}$, we could reduce a_{p+1} by ε_{p+1} and increase $a_{p+2},\cdots,a_{n-m}$ each by ε_{n-m}, where $\varepsilon_{n-m}=\gamma_{p+1}\varepsilon_{p+1}/[\gamma_{n-m}-\gamma_{n-m-1}+\cdots-\gamma_{p+2}]$, and keep all constraints hold. Now we have that $\varepsilon_{n-m}\cdot(n-m-p-1)<\varepsilon_{p+1}$, that is, reduction is larger than increase. Therefore we have now $\sum_{i=1}^{n}a_i<p$, contradicting step 1; Hence we have $a_{p+1}=a_{p+2}$.

<u>Step 8</u>: We show that $l=p-1$, i.e. $a_1=a_2=\cdots=a_{p-1}=1$.

Using the results from the above steps, we can write (2.3) as

$$C_1a_1 + C_2a_2 + \cdots + C_pa_p = pC, \tag{2.11}$$

where $C = (\gamma_{n-m} - \gamma_{n-m-1} + \cdots - \gamma_{p+2} + \gamma_{p+1})$ and $C_i = C + (-1)^{p-i}\gamma_i$ for $i = 1, \cdots, p$. Note that $C_i < C_p$ for all $i < p$. And notice we have $\sum_{i=1}^{p} a_i < p$.

Suppose $l < p-1$. We could increase a_{l+1} by ε_{l+1} and decrease a_p by ε_p (since $a_p > 0$) such that $\varepsilon_p = \frac{C_{l+1}}{C_p} < \varepsilon_{l+1}$ and $\sum_{i=1}^{p} a_i < p$ still holds. Now all the constraints still hold and the objective function increases. Hence we must have $a_1 = \cdots = a_{p-1} = 1$ in any optimal solution.

Case 2: $p > k$**.**

Step 1: We show that $\sum_{i=1}^{n} a_i = p$

If $l < k-1$, the same argument in Step 1 of Case 1 applies. We get $\sum_{i=1}^{n} a_i = p$. Now we assume $l \geq k-1$.

1a) We first show that $a_{p+2} = \cdots = a_n = 0$; and $p+1-l$ is even.

Suppose $a_{p+2} > 0$. If $n-m-l$ is even, then we decrease both a_{n-m} and a_{n-m-1} (note that $l+1 \leq p$) by ε_{n-m}, and increase a_{l+1} by $\frac{\gamma_{n-m-1}-\gamma_{n-m}}{\gamma_{l+1}} \cdot \varepsilon_{n-m} < \varepsilon_{n-m}$ (since $\gamma_{l+1} \geq \gamma_{n-m-1} > \gamma_{n-m}$) ; if $n-m-l$ is odd, then we decrease a_{n-m} by ε_{n-m}, and increase a_{l+1} by $\frac{\gamma_{n-m}}{\gamma_{l+1}} \cdot \varepsilon_{n-m} < \varepsilon_{n-m}$ (since $\gamma_{l+1} > \gamma_{n-m}$). All the constraints still hold but the objective function increases. So we must have $a_{p+2} = \cdots = a_n = 0$; and if $p+1-l$ is odd, we must also have $a_{p+1} = 0$.

1b) We then show that $p+1-l$ is even.

If for the sake of contradiction $p+1-l$ is odd, since $a_{p+1} = a_{p+2} = \cdots = a_n = 0$, we have $\sum_{i=1}^{n} a_p = \sum_{i=1}^{p} < p$. We first show that $a_{p-1} = a_p$, $a_{p-3} = a_{p-2}$, $\cdots$, $a_{l+1} = a_{l+2}$. Suppose $a_{l-1+2s} > a_{l+2s}$ for some $1 \leq s \leq (p-l)/2$, then we can increase a_{l+1}, a_{l+2},$\cdots$, a_{l-1+2s} by ε, and increase a_{l+2s} by $\varepsilon' = (\gamma_{l+1} - \gamma_{l+2} + \cdots + \gamma_{l-1+2s})(2s-1)\varepsilon/\gamma_{l+2s}$, where ε is small enough such all the constraints still hold. However the value of the objective function increases. Contradiction!

Now the budget balanced constraint becomes $(\gamma_{l+1} - \gamma_{l+2})a_{l+1} + (\gamma_{l+3} - \gamma_{l+4})a_{l+3} + \cdots + (\gamma_{p-1} - \gamma_p)a_{p-1} = \gamma_l - \gamma_{l-1} + \gamma_{l-2} - \gamma_{l-3} + \cdots + (-1)^{l-1}\gamma_1$. The left side is less than $\mathscr{C}_l^{n-2} - \mathscr{C}_{p-2}^{n-2}$ (take $a_{l+1} = a_{l+3} = \cdots = a_{p-1} = 1$), and less than the right side, which is $\mathscr{C}_{l-1}^{n-2}$. Contradiction! Therefore $p+1-l$ is even.

1c) We now show that $\sum_{i=1}^{n} a_i = p$. First note that in any optimal solution we have $a_p > a_{p+1}$, which can be proved using the same argument in Step 5 of Case 1. Suppose $\sum_{i=1}^{p} a_i < p$, we can increase $a_{l+1}, \cdots, a_p$ each by ε and increase a_{p+1} by $\frac{\gamma_{l+1}-\gamma_{l+2}+\cdots-\gamma_p}{\gamma_{p+1}} \cdot \varepsilon > \varepsilon$, where ε is small enough such that all constraints still hold. But the value of the objective function increases. Contradiction!

Step 2: We finish the proof for case 2 by showing that solving for some $p > k$ is essentially the same as solving for $n - p$.

Let $b_i = 1 - a_{n+1-i}$, our problem can be written as

$$max_{\{b_1,b_2,\cdots,b_n\}}(b_1 + b_2 + \cdots + b_{n-p}) \tag{2.12}$$

s.t.

$$0 \leq b_n \leq b_{n-1} \leq \cdots \leq b_1 \leq 1$$

$$\sum_{i=1}^{n} b_i = n - p$$

$$\gamma_1 b_n - \gamma_2 b_{n-1} + \cdots + (-1)^{n-1}\gamma_n b_1 = 0$$

Note that $\gamma_n = \gamma_1$, $\gamma_{n-1} = \gamma_2$, $\cdots$; the budget constraint can be written as

$$\gamma_1 b_1 - \gamma_2 b_2 + \cdots + (-1)^{n-1}\gamma_n b_n = 0$$

Now we have the same problem as in Case 1 since $n - p < k$. This symmetry comes from the fact that if we are going to allocate p objects as efficient as possible, it is the same as allocating $n - p$ "non-objects" as efficient as possible. And notice we must show $\sum_{i=1}^{n} a_i = p$ in both cases to conclude that the two cases are essentially the same. Suppose $(b_1, b_2, \cdots, b_n)$ is an optimal solution as described in Case 1 for the problem (2.12), then $(1 - b_n, 1 - b_{n-1}, \cdots, 1 - b_1)$ is an optimal solution of the original problem (2.10).

□

2.5.4.3 When voluntary participation holds

Now we look at the problem with voluntary participation:

$$\max_{(\{a_l\};\{\delta_k\})} \sum_{l=1}^{p} a_l \tag{2.13}$$

such that

$$\sum_{l=1}^{n} a_l \leq p, \quad 0 \leq a_n \leq \cdots \leq a_2 \leq a_1 \leq 1 \ ;$$

and (2.1), (2.7).

Proposition 2.4. *OB solution is part of a feasible solution to the problem* (2.13) *if and only if* $p < n - \bar{p}(n)$.

Proof. Let $\{\delta_k\}$ be the solution of (2.1) for $\{a_l\} = \{a_l^\phi\}$. We only need to show that $\sum_l^k \delta_l \geq 0$ for all $k = 1, 2, \cdots, n-1$ does not hold if $p \geq n - \bar{p}(n)$.

According to (2.1), we have $\delta_1 = \cdots = \delta_{p-1-2t^\star(p)} = 0, \delta_{p+2} = \cdots = \delta_n = 0$; and $(\delta_{p-2t^\star(p)}, \cdots, \delta_p, \delta_{p+1})$ are of the signs $(+, -, +, -, \cdots, +, -, +, +)$. And

$$|\delta_{p-2t^\star(p)}| < |\delta_{p+1-2t^\star(p)}| < |\delta_{p+2-2t^\star(p)}| < \cdots < |\delta_p| < |\delta_{p+1}|.$$

Since $|\delta_{p-2t^\star(p)}| < |\delta_{p+1-2t^\star(p)}|$, we know that $\sum_l^k \delta_l \geq 0$ for all $k = 1, 2, \cdots, n-1$ does not hold.

□

2.5.4.4 Optimal voluntary mechanism for $p \geq n - \bar{p}(n)$.

We show that for $p \geq \bar{p}(n)$, the OV mechanism (solution) is an optimal solution for the problem (2.13).

We first prove the following lemma.

Lemma 2.9. *For* $p > \bar{p}(n)$*, there exists an optimal solution* $\{a_1, \cdots, a_n; \delta_1, \cdots, \delta_{n-1}\}$ *of problem* (2.13) *such that* $a_{p+2} = a_{p+3} = \cdots = a_n = 0$*, and* $\delta_{p+2} = \delta_{p+3} = \cdots = \delta_{n-1} = 0$.

Proof. If $p = n-1$, then the statement holds obviously. Now assume $p < n-1$. Suppose we have an optimal solution $\{a_i\}$, where $j = \max\{i : a_i > 0\} \geq p+2$. Then we reduce both a_{j-1} and a_j by ε_{j-1} and ε_j respectively, with $\varepsilon_j = a_j$ and $\varepsilon_{j-1} = \frac{\gamma_j}{\gamma_{j-1}}\varepsilon_j < \varepsilon_j$. Then all the constraints of problem (2.10) still hold and the value of objective function remains the same. We check that VP constraint (2.7) still holds. Since $a_1, \cdots, a_{j-2}$ remain the same, so does $\delta_2, \cdots, \delta_{j-2}$. And $\delta_{j+1}, \cdots, \delta_{n-1}$ also remain to be 0. We only need to pay attention to the changes of δ_{j-1} and δ_j . Since

$$\begin{aligned} (j-2)\delta_{j-2} + (n-j+1)\delta_{j-1} &= (j-2)(a_{j-2} - a_{j-1}) \\ (j-1)\delta_{j-1} + (n-j)\delta_j &= (j-1)(a_{j-1} - a_j) \quad , \\ j\delta_j &= ja_j \end{aligned}$$

it is easy to see that 1) δ_j decreases by $\varepsilon_j = a_j$ and becomes zero; 2) δ_{j-1} increases. Since VP constraint (2.7) holds originally, it still holds under such changes of δ_j and δ_{j-1}. Hence we get another optimal solution in which $a_j = 0$. We can use the same trick to reduce all a_{p+2}, a_{p+3}, $\cdots$, a_{j-1} to zero and get an optimal solution as required. □

By the above Lemma, it is enough to show that the truncated OV solution, which is $(\{a_l^o\}_{l=1}^{p+1}, \{\delta_k^o\}_{k=1}^{p+1})$, is an optimal solution for the following truncated problem:

$$\max_{(a_1,a_2,\cdots,a_{p+1};\delta_1,\delta_2,\cdots,\delta_p,\delta_{p+1})} (a_1 + a_2 + \cdots + a_p) \tag{2.14}$$

s.t.

$$0 \leq a_{p+1} \leq \cdots \leq a_2 \leq a_1 \leq 1$$

$$\sum_{i=1}^{p+1} a_i \leq p$$

$$\begin{aligned}
(n-1) \cdot \delta_1 &= 0 \\
\delta_1 + (n-2) \cdot \delta_2 &= (a_1 - a_2) \\
2\delta_2 + (n-3) \cdot \delta_3 &= 2(a_2 - a_3) \\
&\vdots \\
p \cdot \delta_p + (n-p-1) \cdot \delta_{p+1} &= p \cdot (a_p - a_{p-1}) \\
(p+1) \cdot \delta_{p+1} &= (p+1) \cdot a_{p+1}
\end{aligned}$$

$$\begin{gathered}
\delta_1 \geq 0 \\
\delta_1 + \delta_2 \geq 0 \\
\delta_1 + \delta_2 + \delta_3 \geq 0 \\
\vdots \\
\delta_1 + \delta_2 + \cdots + \delta_{p+1} \geq 0
\end{gathered}$$

We rewrite the problem:

Let

$$\begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ \vdots \\ t_{p-1} \\ t_p \end{bmatrix} = \begin{bmatrix} \delta_2 \\ \delta_2+\delta_3 \\ \delta_2+\delta_3+\delta_4 \\ \vdots \\ \delta_2+\delta_3+\cdots+\delta_p \\ \delta_2+\delta_3+\cdots+\delta_p+\delta_{p+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & 1 & 0 \\ 1 & 1 & 1 & \cdots & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} \delta_2 \\ \delta_3 \\ \delta_4 \\ \vdots \\ \delta_p \\ \delta_{p+1} \end{bmatrix}$$

and

$$\begin{bmatrix} x_1 \\ \frac{x_2}{2} \\ \frac{x_3}{3} \\ \vdots \\ \frac{x_p}{p} \\ \frac{x_{p+1}}{p+1} \end{bmatrix} = \begin{bmatrix} a_1-a_2 \\ a_2-a_3 \\ a_3-a_4 \\ \vdots \\ a_p-a_{p+1} \\ a_{p+1} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_p \\ a_{p+1} \end{bmatrix}$$

Hence

$$a_1+a_2+\cdots+a_p = x_1+x_2+\cdots+x_p+\frac{p}{p+1}x_{p+1}$$

and

$$a_1+a_2+\cdots+a_p+a_{p+1} = x_1+x_2+\cdots+x_p+x_{p+1}.$$

First we take care of the constraint $0 \le a_{p+1} \le a_p \le \cdots \le a_1 \le 1$, which can be written as the following:

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & 0 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_p \\ a_{p+1} \end{bmatrix} \le \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}.$$

This is equivalent to

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & 0 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} x_1 \\ \frac{x_2}{2} \\ \frac{x_3}{3} \\ \vdots \\ \frac{x_p}{p} \\ \frac{x_{p+1}}{p+1} \end{bmatrix} \leq \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}.$$

That is

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \frac{x_2}{2} \\ \frac{x_3}{3} \\ \vdots \\ \frac{x_p}{p} \\ \frac{x_{p+1}}{p+1} \end{bmatrix} \leq \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}.$$

Next we take care of the BB constraint:
Since

$$\begin{bmatrix} n-2 & 0 & 0 & \cdots & 0 & 0 \\ 2 & n-3 & 0 & \cdots & 0 & 0 \\ 0 & 3 & n-4 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & p & n-p-1 \\ 0 & 0 & 0 & \cdots & 0 & p+1 \end{bmatrix} \cdot \begin{bmatrix} \delta_2 \\ \delta_3 \\ \delta_4 \\ \vdots \\ \delta_p \\ \delta_{p+1} \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_p \\ x_{p+1} \end{bmatrix},$$

we have

$$\begin{bmatrix} n-2 & 0 & 0 & \cdots & 0 & 0 \\ 2 & n-3 & 0 & \cdots & 0 & 0 \\ 0 & 3 & n-4 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & p & n-p-1 \\ 0 & 0 & 0 & \cdots & 0 & p+1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & 1 & 0 \\ 1 & 1 & 1 & \cdots & 1 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ \vdots \\ t_{p-1} \\ t_p \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_p \\ x_{p+1} \end{bmatrix}.$$

That is

$$\begin{bmatrix} n-2 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 5-n & n-3 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -3 & 7-n & n-4 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -4 & 9-n & n-5 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -p & 2p+1-n & n-p-1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -(p+1) & p+1 \end{bmatrix} \cdot \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ \vdots \\ t_{p-1} \\ t_p \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_p \\ x_{p+1} \end{bmatrix}. \tag{2.15}$$

Therefore, our problem (2.14) can be written as the following:

$$\max_{(t_1,t_2,\cdots,t_p;x_1,x_2,\cdots,x_{p+1})} (x_1+x_2+\cdots+x_p+\frac{p}{p+1}x_{p+1}) \tag{2.16}$$

s.t.

$$\begin{aligned} t_j \geq 0, & \qquad \forall j=1,\cdots,p \\ x_i \geq 0, & \qquad \forall i=1,\cdots,p+1 \\ x_1+x_2+\cdots+x_p+x_{p+1} \leq p & \\ x_1+\tfrac{x_2}{2}+\cdots+\tfrac{x_p}{p}+\tfrac{x_{p+1}}{p+1} \leq 1 & \end{aligned}$$

and (2.15).

Note that the feasible solution of this problem which corresponds to the OV solution of the problem (2.13) is the following:

$$\begin{gathered} x_1^0=\cdots=x_{p-2}^0=0 \\ x_{p-1}^0=\frac{p(p-1)(n-p)}{n(n-1)}, x_p^0=p[1-\frac{2p(n-p)}{n(n-1)}], x_{p+1}^0=\frac{p(p+1)(n-p)}{n(n-1)}. \\ t_1^0=\cdots=t_{p-2}^0=0, t_{p-1}^0=\frac{p(p-1)}{n(n-1)}, t_p^0=\frac{p(n-1)}{n(n-1)} \end{gathered} \tag{2.17}$$

Let $\mathbf{A}=\begin{bmatrix} A^1, & \mathbf{A^2} \end{bmatrix}$, where

$$
\mathbf{A^1} = \left[\begin{array}{cccccc|cc}
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
n-2 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
5-n & n-3 & 0 & 0 & \cdots & 0 & 0 & 0 \\
-3 & 7-n & n-4 & 0 & \cdots & 0 & 0 & 0 \\
0 & -4 & 9-n & n-5 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 2p-1-n & n-p & 0 \\
0 & 0 & 0 & 0 & \cdots & -p & 2p+1-n & n-p-1 \\
0 & 0 & 0 & 0 & \cdots & 0 & -(p+1) & p+1
\end{array}\right]
$$

$$
\mathbf{A^2} = \left[\begin{array}{cccc|ccc}
1 & 1 & 1 & \cdots & 1 & 1 & 1 \\
1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{p-1} & \frac{1}{p} & \frac{1}{p+1} \\
-1 & 0 & 0 & \cdots & 0 & 0 & 0 \\
0 & -1 & 0 & \cdots & 0 & 0 & 0 \\
0 & 0 & -1 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & -1 & 0 & 0 \\
0 & 0 & 0 & \cdots & 0 & -1 & 0 \\
0 & 0 & 0 & \cdots & 0 & 0 & -1
\end{array}\right]
$$

$$
\mathbf{b} = \begin{bmatrix} p & 1 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}^T
$$

$$
\mathbf{c} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & 1 & 1 \cdots & 1 & 1 & \frac{p}{p+1} \end{bmatrix}^T .
$$

We can write the dual problem as

$$
\min_{\mathbf{y}=(y_1, y_2, \cdots, y_{p+3})} \mathbf{b}^T \cdot \mathbf{y} \tag{2.18}
$$

s.t.

$$
\mathbf{y}^T \cdot \mathbf{A} \geq \mathbf{c}^T
$$

$$
y_1 \geq 0, y_2 \geq 0
$$

Now we show that (2.17) is an optimal solution of the original problem (2.16) by duality theorem. That is, we will find a feasible vector y of the dual problem (2.18) such that $\mathbf{b}^T \cdot \mathbf{y} = x_1^0 + x_2^0 \cdots + x_p^0 + \frac{p}{p+1} x_{p+1}^0$.

Since $t_{p-1}^0, t_p^0, x_{p-1}^0, x_p^0, x_{p+1}^0 > 0$, by the Complementary Slackness Theorem, if such a y exists, we must have

$$\mathbf{y}^{\mathrm{T}} \cdot (\mathbf{A^1}_{(\mathrm{p}-1)}, \mathbf{A^1}_{\mathrm{p}}, \mathbf{A^2}_{(\mathrm{p}-1)}, \mathbf{A^2}_{\mathrm{p}}, \mathbf{A^2}_{\mathrm{p}+1}) = (\mathbf{c}^{\mathrm{T}}_{(\mathrm{p}-1)}, \mathbf{c}^{\mathrm{T}}_{\mathrm{p}}, \mathbf{c}^{\mathrm{T}}_{(2\mathrm{p}-1)}, \mathbf{c}^{\mathrm{T}}_{2\mathrm{p}}, \mathbf{c}^{\mathrm{T}}_{2\mathrm{p}+1},),$$

where M_k is the kth column of the matrix M, for $\mathrm{M} = \mathrm{A}^1, \mathrm{A}^2, \mathrm{c}^T$.

That is

$$\begin{bmatrix} 0 & 0 & n-p & 2p+1-n & -(p+1) \\ 0 & 0 & 0 & n-p-1 & p+1 \\ 1 & \frac{1}{p-1} & -1 & 0 & 0 \\ 1 & \frac{1}{p} & 0 & -1 & 0 \\ 1 & \frac{1}{p+1} & 0 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ y_{p+1} \\ y_{p+2} \\ y_{p+3} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ \frac{p}{p+1} \end{bmatrix}.$$

Solving it we get

$$\begin{bmatrix} y_1 \\ y_2 \\ y_{p+1} \\ y_{p+2} \\ y_{p+3} \end{bmatrix} = \begin{bmatrix} \frac{n-p}{n} \\ \frac{p(p-1)}{n-1} \\ \frac{p}{n(n-1)} \\ -\frac{n-p}{n(n-1)} \\ \frac{(n-p-1)(n-p)}{(p+1)n(n-1)} \end{bmatrix}.$$

Now we check that $\mathbf{b}^T \cdot \mathbf{y} = x_1^0 + x_2^0 \cdots + x_p^0 + \frac{p}{p+1} x_{p+1}^0$. That is,

$$py_1 + y_2 = x_1^0 + \cdots + x_p^0 + \frac{p}{p+1} x_{p+1}^0 = \frac{p(n-p)}{n} + \frac{p(p-1)}{n-1}.$$

The remaining work is to show that the above $(y_1, y_2, y_{p+1}, y_{p+2}, y_{p+3})$ is part of a feasible solution of the problem (2.18). Since $y_1, y_2 \geq 0$, we only need to show that given the value of $y_1, y_2, y_{p+1}, y_{p+2}, y_{p+3}$, $\mathbf{y^T A} \geq \mathbf{c}^T$ has a feasible solution. That is

$$\begin{bmatrix} y_3 \\ y_4 \\ y_5 \\ \cdots \\ y_{p-2} \\ y_{p-1} \\ y_p \end{bmatrix} \leq \begin{bmatrix} y_1 + y_2 - 1 \\ y_1 + \frac{y_2}{2} - 1 \\ \vdots \\ y_1 + \frac{y_2}{p-5} - 1 \\ y_1 + \frac{y_2}{p-4} - 1 \\ y_1 + \frac{y_2}{p-3} - 1 \\ y_1 + \frac{y_2}{p-2} - \frac{p}{p+1} \end{bmatrix},$$

and

$$\begin{bmatrix} y_3 \\ y_4 \\ y_5 \\ \cdots \\ y_{p-2} \\ y_{p-1} \\ y_p \end{bmatrix}^T \cdot \begin{bmatrix} n-2 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 5-n & n-3 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -3 & 7-n & n-4 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -4 & 9-n & n-5 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 2p-5-n & n-p+2 & 0 \\ 0 & 0 & 0 & 0 & \cdots & -(p-2) & 2p-3-n & n-p+1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -(p-1) & 2p-1-n \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & -p \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix}$$

$$\geq \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ (p-1)y_{p+1} \\ py_{p+2} - (2p-1-n)y_{p+1} \end{bmatrix}^T.$$

Let $y_3 = y_4 = \cdots = y_{p-3} = y_1 + \frac{y_2}{p-5} - 1$. Note that since $y_1 + \frac{y_2}{p} = 1$ and $y_1, y_2 > 0$, we have $y_1 + y_2 - 1 > y_1 + \frac{y_1}{2} - 1 > \cdots > y_1 + \frac{y_2}{p-5} - 1 > 0$. We only need to find y_{p-2}, y_{p-1}, y_p such that

$$\begin{bmatrix} y_{p-2} \\ y_{p-1} \\ y_p \end{bmatrix} \leq \begin{bmatrix} y_1 + \frac{y_2}{p-4} - 1 \\ y_1 + \frac{y_2}{p-3} - 1 \\ y_1 + \frac{y_2}{p-2} - \frac{p}{p+1} \end{bmatrix}$$

and

$$(n-p+2)y_{p-4}+(2p-3-n)y_{p-3}+(-p+1)y_{p-2}\geq(p-1)y_{p+1}$$

$$(n-p+1)y_{p-3}+(2p-1-n)y_{p-2}+(-p)y_{p-1}\geq py_{p+2}-(2p-1-n)y_{p+1}$$

Taking $y_p << y_{p-1} << y_{p-2} << 0$, the above inequality system is satisfied. Hence $(y_1,y_2,y_{p+1},y_{p+2},y_{p+3})$ is part of a feasible solution. The proof is completed.

Chapter 3

Group selection under single-peaked preference: a mechanism design approach

3.1 Introduction

Human beings are social animals, yet to varying degrees: some more gregarious, some more solitary. It is therefore no surprise that they have different ideas about the optimal group size when sharing a facility or taking a common task. Our paper is about how to select the group among people with different preferences over group size.

This issue of selecting a group is barely new in real life; however, the current literature on group formation focus mainly on the strategic interactions among agents and the decentralized formation of coalitions.[1] Our paper takes a mechanism design approach, which has been successfully adopted in matching and school choice, and works in a novel setting that is becoming increasingly relevant with the rise of the "open source project"[2], in which a set of volunteers are attracted to contribute. One remarkable example is the "Wikiproject"[3] of Wikipedia, that is, a group of contributors who work together as a team to assess articles' importance and quality in a specific topic area. Having a group assembling mechanism that takes the size preference seriously is of special importance.

In general, when people share a facility, such as students sharing a study room, guests

[1] See [11] for a recent survey, see [22] for more exhaustive surveys and see [60] for a recent monograph on coalition formation.

[2] See https://en.wikipedia.org/wiki/Open-source_movement for more information. Prime examples of open-source products include Mozilla Firefox, Google Chromium, Android.

[3] See https://en.wikipedia.org/wiki/WikiProject for more information.

sharing a swimming pool, size effect mainly comes from two sources: comradeship effect on the positive side, congestion effect on the negative side. The trade-off is the same for everyone; the optimum differs. When people take a shared task, such as volunteers developing a software, coauthors working on a joint paper, an additional source of size effect appears: individual contribution and individual cost. Both decrease as the group grows and again the balance differs. Though identities of companions often matter, size effect may dominate in situations when agents in the group do not interact with others much, or when cooperation and interaction happen among strangers, which becomes common in societies more and more connected by smart-phones.

The above situations share the following formal structure: a group is to be selected among a set N of agents; each agent $i \in N$ has strict preference R_i over the size of the group if i is a member of the group, and over the "stand-outside" option. We say that i is in a group of size *zero* in the latter case. Furthermore preferences are assumed to be single-peaked over the set $S = \{0, 1, \cdots, n\}$ of sizes.[4]

We impose three standard designing constraints on group selection mechanisms. First, efficiency, i.e., the allocation can not be Pareto improved for each preference profile. Second, individual rationality, also known as voluntary participation, acknowledging the fact that agents cannot be forced to stay in the group. Third, the mechanisms must be robust to manipulations: we require strategy-proof mechanisms.[5]

In our model, these three criteria are compatible: two classes of desirable (direct) mechanisms are proposed, each with additional nice properties.

In the *"Proposing in turn"* mechanism, agents are ordered in a queue to make choices one by one. Observing the current group size, which is determined by the agents before her, an agent either walks away or joins the group; if she joins the group, she either ends the process (hence excluding all agents behind her), or continues and waits for new agents to join. The mechanism makes sure that the final group size does not exceed her peak. For instance, the first agent joins the group with peak 4, the second agent joins the group with peak 3, the third agent walks away, and the fourth agent joins with a peak larger than 3 (the specific number is irrelevant), then the process ends.

Our first result reads as follows: the "proposing in turn" mechanism is efficient, individually rational and group-strategyproof; conversely, if we generalize the mechanism slightly by allowing for more complicated priority, in which agents could have different successors depending on whether or not she joins the group, then the generalized class is characterized

[4]See [4, 12, 45] for single-peaked preferences over the group size. However, size 0, which did not appear in their papers, plays an important role in ours. See appendix for more discussions.

[5]See [48] and [5] for general introductions to the axioms. For applications on specific settings, see [1], [61], [3], [44], [39], [43], [52] etc.

by efficiency, individual rationality and group strategy-proofness.

The mechanism can also be interpreted as follows: starting with the grand coalition and following the priority, each agent has the power to either maintain or shrink the current group size. Therefore the resulting group size may be much smaller than the *maximum group size* compatible with individual rationality, which is a legitimate concern in many situations. Our second mechanism, though more complicated, brings us closer to the maximum group size.

To describe this *"Voting on ascending-size"* mechanism, we start with an intuitive yet flawed procedure, and revise it only later. The intuitive procedure mimics the usual ascending auction, in which agents "bid" on group size. In each round k ($k \geq 1$), agents are asked to vote on group size k and only agents who say "yes" continue to vote in any future rounds. If the number q_k of agents who vote "yes" is strictly larger than k, round $k+1$ begins; otherwise the procedure stops and the final outcome is the following: if $q_k = k$, then a group with size k forms, containing all the q_k agents who vote "yes" to size k; if $q_k < k$, then a group with size $k-1$ forms, selecting among the q_{k-1}(note that $q_{k-1} > k-1$) agents who vote "yes" to size $k-1$ by a fixed ordering σ of agents.

If agents say "yes" to size k when size k is preferred to size 0, the above procedure will lead to a group with the maximum group size. The problem is that agents have incentive to misreport, which can be easily illustrated by an example with two agents. Suppose there are two agents $\{a,b\}$ and $\sigma = (a,b)$. Suppose both agents vote "yes" in round 1. In round 2, if both vote "yes" for group size 2, then $\{a,b\}$ is the final group; otherwise $\{a\}$ is the final group. When size 1 is better than size 2, agent a will vote "no" for size 2 even if she prefers size 2 to size 0.

Our "voting on ascending-size" mechanism offers the right incentive by ending the "auction" *earlier* in certain situations. In each round k, agents are asked to vote for size k, and then to compare size $k-1$ to size k. Eve if $q_k \geq k$, a group of size $k-1$ will form when a *coalition* of agents preferring size $k-1$ to size k can *collectively* misreport at round k (vote "no" to size k instead of the true answer "yes") and join a size $k-1$ group; otherwise a group of size k will form if $q_k = k$, and round $k+1$ will start if $q_k > k$.

In the above two agents example, whenever a can benefit by voting "no" to size 2 instead of her honest answer "yes", she can get the same benefit by truthfully answering the two questions (vote "yes" for size 2 and then say size 1 is preferred to size 2); thus the incentive problem in the original procedure is removed. This is not an accident, but a key feature of the mechanism that provides agents with the right incentive. The reader will find in the paper a more detailed description of the incentive problem, which becomes thornier as the number of agents increases, and will have a better understanding of this particular design (say, why consider a coalition?), whose validity relies crucially on the single-peaked preference assumption. The mechanism is efficient, individually rational, and weakly group strategy-proof; it enlarges the group size in each round and guarantees at least one-half of the maximum group size.

The remaining paper is organized as follows. We end the introduction by a literature review. In section 3.2 we introduce the setting and the main criteria for group selection mechanisms. In section 3.3 we first describe the "proposing in turn" mechanism and show that it is efficient, strategy-proof and individually rational; then we generalize the priority, define the proposing mechanism associated with a priority tree, and characterize the mechanism by group strategy-proofness. Section 3.4 focuses on the "voting on ascending-size" mechanism: we first show that it is efficient, strategy-proof and individually rational, and then discuss its properties on group size. Section 3.5 concludes. The appendix contains a discussion on extensions of preference domain and all omitted proofs.

3.1.1 Relation to the literature

Our paper is related to [35] and [45]. In all three paper, a club (group) is to be formed among a set of agents, and agents care about the club size. The difference is, in their paper an alternative consumed collectively by the club has to be decided along with the membership, while in our paper we assume hedonic club[6] and focus solely on the effect of club size. In [35], agents have single-peaked preference over the alternatives (position on an interval); the size effect is assumed to be either pure "cost-sharing" (larger size is always preferred to smaller size), or pure congestion (smaller size is always preferred to larger size), or that there exists common optimal size among agents. A fundamental result is that in all three cases efficient and strategy-proof mechanisms must fix the club size and not allow it to vary with agents' preferences; and if there is heterogeneity in agents' preference over the optimal size, then only dictatorial mechanisms satisfy strategy-proofness, Pareto efficiency and outsider independence (a partial non-bossiness condition). In [45], agents have arbitrary preference over the alternatives, and for each alternative, they always prefer the set of users becomes larger. They show that no efficient and stable (no agents can be forced to be a user and no agent who wants to be a user can be excluded) mechanisms can be Nash-implemented, and propose instead a two-stage sequential mechanism whose unique sub-game perfect equilibrium outcome is efficient and stable. Our paper, on the contrary, yields positive results by looking solely at the effect of group size and assuming

[6]We borrow the term "hedonic" from the literature on "hedonic games", see [12] for the definition. Just as hedonic game is a hedonic reduction of more general cooperative games, in which an alternative is consumed by each coalition, our setting is a hedonic reduction of the more general one-club formation problem discussed in these two paper.

single-peaked preferences over sizes.

Our paper is also related to [20], in which they propose an ascending auction-like mechanism for selecting the club and sharing among the club members the cost of the club good. If we fix the cost sharing mechanism to be the average cost sharing mechanism and do not allow money transfer among agents, then the preferences of agents can be described as over group sizes, and their mechanism reduces to a group selection mechanism (see appendix for more discussions). Though largely complicated by the potential manipulations induced by tie-breaking, our "voting on ascending size" mechanism is partly inspired by their auction-like mechanism and still bears similar spirit.

It is interesting to note that although single-peaked preference is essential for our positive results, the way it works is quite novel. As a result, our mechanisms are very different from the well known mechanisms associated with single-peaked preferences. That is, the generalized median voting mechanism and its variations in the public choice setting (see [47], [6], [7], etc.) and the uniform mechanism in the private goods division setting (see [62], [18], [39]). A recent paper by [52] treated a general collective decision problem where preferences are single-peaked over one-dimensional allocation space, including the above two settings as two special cases. The mechanism proposed there equalizes in the leximin sense individual gains from an arbitrary benchmark allocation. Again our mechanisms work differently. Since our setting is neither purely public nor purely private, our results further confirm the salience of single-peaked preferences in strategy-proof mechanism design.[7]

3.2 Setting

Given a set of agents N, with $|N| = n \geq 2$, a subset G of N will be selected. If agent i belongs to G, i is in a group of size $|G|$. If agent i does not belong to G, for convenience we say that i is in a group of size zero. Each agent only cares about the size of the group she belongs to and the preference over sizes is strict. That is, the preference R_i of agent i is a reflexive, transitive, complete and anti-symmetric binary relation[8] over the set $S = \{0, 1, \cdots, n\}$. The associated strict relation is denoted by P_i; we write kP_il if kR_il and $k \neq l$. Let $R = \{R_i\}_{i \in N}$ be a preference profile; R_{-i} be a preference profile of all the agents except for i; and for any $M \subseteq N$, let R_M be a preference profile of all agents in M, and R_{-M} be a preference profile for all the agents in $N \backslash M$. let $\mathscr{U}$ be the set of all preference profiles.

For each R_i, denote by $T(R_i)$ the upper contour set of size zero, i.e., $T(R_i) = \{k \in S :$

[7]For the salience of single-peaked preferences in public choice setting, see [16], [17].

[8]Reflexive: for all $k \in S$, kR_ik; transitive: for all $k, l, h \in S$, kR_il and lR_ih imply kR_ih; complete: for all $k, l \in S$, kR_il or lR_ik; and anti-symmetric: for all $k, l \in S$, kR_il and lR_ik imply $k = l$.

$kR_i0\}$; denote by $t(R_i)$ the largest element in $T(R_i)$, i.e. the largest size that is not worse than size zero; denote by $p(R_i)$ the most preferred group size. A preference R_i is *single peaked* if for any $k,l \in S$, $k < l < p(R_i)$ implies lP_ik, and $k > l > p(R_i)$ implies lP_ik. Note that for any single-peaked preference R_i, $T(R_i)$ is a integer interval, i.e, $T(R_i) = \{0, \cdots, k\}$ for some $k = t(R_i) \geq 0$. Let $\mathscr{P}$ be the set of all preference profiles $\{R_i\}_{i \in N}$ such that R_i is single-peaked for each i.

A group selection (direct) mechanism f is a set-valued functional $f : \mathscr{R} \to 2^N$, where $\mathscr{R} \subseteq \mathscr{U}$ is the set of all admissible preference profiles. Note that $f(R) = \{\emptyset\}$ is allowed.[9] We write $f_i(R) = 0$ if $i \notin f(R)$, and $f_i(R) = |f(R)|$ if $i \in f(R)$. Unless stated otherwise, we take $\mathscr{R} = \mathscr{P}$.

Definition 3.1. (Efficiency)

A group selection mechanism f is efficient (Eff) if for any R, there dose not exist $G \subset N$ such that G as a chosen group Pareto dominates $f(R)$, i.e., for all $i \in G$, $|G|R_if_i(R)$, for all $i \notin G$, $0R_if_i(R)$, and for some $j \in G$, $|G|P_jf_j(R)$ or for some $j \notin G$, $0P_jf_j(R)$.

Definition 3.2. (Individual rationality)

A group selection mechanism f is individually rational (IR) if for all R, $i \in N$, $f_i(R)R_i0$.

Definition 3.3. (Strategy-proofness)

A group selection mechanism f is strategy-proof (SP) if for all R, $i \in N$, and R'_i, $f_i(R)R_if_i(R'_i, R_{-i})$.

We present below two simple priority mechanisms; each fails a criterion.

Let $\Sigma(N)$ be the set of permutations of N, and $\sigma = (\sigma_1, \cdots, \sigma_n) \in \Sigma(N)$ be a permutation (ordering) of agents, in which σ_i is the ith ranked agent.

Example 3.1. (A priority mechanism)

Fix an ordering σ of agents. For each R, if $0R_i1$ for all i, then $f(R) = \{\emptyset\}$; otherwise $f(R) = \sigma_k$, where σ_k is the first agent preferring size 1 to size 0 in the ordering σ.

This mechanism is SP, IR, but not Eff, since it ignores any benefit agents may get from accompanying each other.

Example 3.2. (Another priority mechanism)

Fix an ordering σ of agents. For each R, if $0R_i1$ for all i, then $f(R) = \{\emptyset\}$; otherwise let σ_k be the first agent preferring size 1 to size 0 in the ordering σ. let $p^\star = p(R_{\sigma_k})$. If $p^\star \geq n-k+1$, then $f(R) = \{\sigma_s\}_{s \geq k}$. If $p^\star < n-k+1$, select the first $p^\star$ agents who prefer

[9] $f(R) = \{\emptyset\}$ iff $T(R_i) = \{0\}$ for each i. If a task has to been done, i.e., $f(R)$ cannot be $\{\emptyset\}$, and agents are symmetric in the sense that the set of admissible preference is the same for each agent, then we can model the situation simply by getting rid of all preference profiles in which $T(R_i) = \{0\}$ for some i.

size $p^\star$ to size 0 by the priority σ; if there are less than $p^\star$ agents who prefer size $p^\star$ to size 0, select the remaining agents to be forced into the group by the priority $(\sigma_n, \sigma_{n-1}, \cdots, \sigma_1)$.[10]

This mechanism is the usual serial dictatorship, in which agents, one by one in the priority list, are guaranteed their most preferred outcome that is still available. It is Eff, SP, but not IR.

From now on we only consider group selection mechanisms that are individually rational. Hence it is enough for us to focus on $T(R_i)$ (the upper contour set of size zero) and single-peaked preferences over $T(R_i)$. We use an ordered list of sizes in $T(R_i)$, such as $(3,2,0)$, to indicate a preferences from better to worse; in particular, (0) represents a (class of) preference R_i such that $T(R_i) = \{0\}$.

3.3 "Proposing in turn"

We start with the two agents case and check the restrictions Eff, SP and IR impose on group selection mechanisms.

Example 3.3. (Eff, SP and IR mechanisms for $n = 2$)

Let $N = \{a,b\}$; let f be an Eff, SP and IR mechanism. We first figure out how f behaves in the following preference profiles.

(1) $R_a = (1,0)$, $R_b = (1,0)$

(2) $R_a = (1,0)$, $R_b = (1,2,0)$

(3) $R_a = (1,2,0)$, $R_b = (1,2,0)$

(4) $R_a = (1,2,0)$, $R_b = (2,1,0)$

(5) $R_a = (2,1,0)$, $R_b = (1,2,0)$

We use $R^{(k)}$ to represent the preference profile (k). First by Eff and IR, $f(R^{(1)}) = \{a\}$ or $f(R^{(1)}) = \{b\}$. Assume $f(R^{(1)}) = \{a\}$. Then by SP, $f(R^{(2)}) \neq \{b\}$; by IR, $f(R^{(2)}) \neq \{a,b\}$; and by Eff, $f(R^{(2)}) \neq \{\emptyset\}$. Hence $f(R^{(2)}) = \{a\}$. Then by SP, $f(R^{(3)}) = \{a\}$. Again by SP, $f(R^{(4)}) \neq \{b\}$ and $f(R^{(4)}) \neq \{a,b\}$; and by Eff $f(R^{(4)}) \neq \{\emptyset\}$. Hence $f(R^{(4)}) = \{a\}$. Then by SP, $f(R^{(5)}) \neq \{b\}$. By Eff, $f(R^{(5)}) \neq \{a\}$ and $f(R^{(5)}) \neq \{\emptyset\}$. Hence $f(R^{(5)}) = \{a,b\}$.

[10] To be precise, for each $l > k$, $\sigma_l \in f(R)$ iff (1) $p^\star R_{\sigma_l} 0$ and $|\{s \in \{k,\cdots,l\} : p^\star R_{\sigma_s} 0\}| \leq p^\star$, or (2) $0 R_{\sigma_l} p^\star$ and $p^\star - |\{s \in \{k,\cdots,l\} : p^\star R_{\sigma_s} 0\}| \geq n - l + 1$.

It is easy now to describe f fully:

$$f(R)=\begin{cases}\{a\} & \text{(i)}\,p(R_a)=1\,\text{or}\,\text{(ii)}\,p(R_a)=2\,\&\,0R_b2;\\ \{a,b\} & p(R_a)=2\,\&\,2R_b0;\\ \{b\} & R_a=(0)\,\&\,1R_b0;\\ \{\emptyset\} & R_a=R_b=(0).\end{cases}$$

Note first that f does not satisfy "free entry".[11] In fact, for each R, any group with a size less than $m^\star(R) := \max\{m \in S : |\{i \in N : mR_i0\}| \geq m\}$ does not satisfy free entry. On the other hand, any group with a size larger than $m^\star(R)$ must violate IR; therefore we call $m^\star$ the *maximum group size* compatible with individual rationality.

Note also that agent a has all the privilege restricted only by the individual rationality of agent b. If we change the role of agents a and b, that is, assume $f(R^1) = \{b\}$ in the above example, we get another mechanism in which agent b has all the privilege. These are all the direct mechanisms that are Eff, SP and IR for $n = 2$. For arbitrary n, we propose below a class of mechanisms that satisfies Eff, SP, and IR, and is very easy to understand.

3.3.1 "Proposing in turn"

Definition 3.4. ("proposing in turn" mechanism)

Fix an ordering σ of agent. For any announced preference profile R, we find the group as follows:

In each step $1 \leq k \leq n$, if the algorithm has not yet ended, according to the announced R_i, agent $i = \sigma_k$ *joins the group (J)* if the size of the current group after her join is preferred to size zero, or *walks away (W)* otherwise; if she joins, then she *ends the process (E)* if the size of the current group after her join is no less than her peak, or *continues the process (C)* otherwise. The algorithm ends when one of the following cases happens:

Case (i): an agent chooses *"J and E";*

Case (ii): the number of the agents choosing *"J and C"* equals the minimum of the the peaks of all the agents choosing J ;

Case (iii): all agents have made their choices.

Proposition 3.1. *For any ordering σ, the "proposing in turn" mechanism ψ^σ is Pareto efficiency, individually rational, and strategy proof.*

[11] Free entry means that agents can join the group freely. Hence a group selection mechanism satisfying free entry has the property that agents outside the final group always like to stay outside than to join the group. The stability concept in [45] is equivalent to individual rationality plus free entry. Hence the $n = 2$ case shows that in our setting stability is incompatible with efficiency and strategy-proofness.

Proof. Individual rationality is easy to see since (1) an agent walks away if the size of the current group after her join is worse than size zero and (2) the final size of the group will not exceed her peak if she joins the group.

Now we show Pareto efficiency. Suppose an agent is selected into the original group, then any other outcome in which she is not selected is worse for her. Hence we only need to consider outcomes that contain the original selected group. Suppose the process ends under either Case (i) or Case (ii), then enlarging the group is either impossible or will hurt someone originally selected; suppose the process ends under case (iii), then enlarging the group is either impossible, or will hurt someone that is added to the original group. Therefore no Pareto improvement is possible.

Now we show strategy-proof. Suppose for the sake of contradiction there exists R_i, R_i' and R_{-i} such that $\psi_i^\sigma(R_i', R_{-i}) P_i \psi_i^\sigma(R_i, R_{-i})$. Since $\psi_i^\sigma(R_i', R_{-i}) P_i \psi_i^\sigma(R_i, R_{-i})$, we have that $\psi_i^\sigma(R_i', R_{-i}) = |\psi^\sigma(R_i', R_{-i})| > 0$. Consider the following cases:

Case (1). Suppose $|\phi^\sigma(R_i', R_{-i})| < p(R_i')$.

If $p(R_i) > |\phi^\sigma(R_i', R_{-i})|$, then $\psi_i^\sigma(R_i', R_{-i}) = \psi_i^\sigma(R_i, R_{-i})$; if $p(R_i) < |\phi^\sigma(R_i', R_{-i})| < p(R_i')$, then $\psi_i^\sigma(R_i, R_{-i}) = p(R_i)$.

Case (2). Suppose $p(R') \leq |\phi^\sigma(R_i', R_{-i})| \leq t(R_i')$.

If $p(R_i) \leq |\phi^\sigma(R_i', R_{-i})| \leq t(R_i)$, then $\psi_i^\sigma(R_i', R_{-i}) = \psi_i^\sigma(R_i, R_{-i})$; if $t(R_i) < |\phi^\sigma(R_i', R_{-i})|$, then $\psi_i^\sigma(R_i, R_{-i}) = 0$; if $p(R_i) > |\phi^\sigma(R_i', R_{-i})|$, then $\psi_i^\sigma(R_i', R_{-i}) \leq \psi_i^\sigma(R_i, R_{-i}) \leq p(R_i)$.

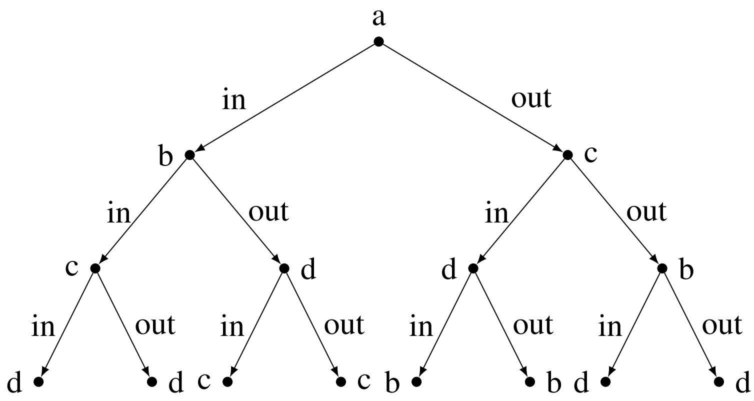
There is no case in which $\psi_i^\sigma(R_i', R_{-i}) P_i \psi_i^\sigma(R_i, R_{-i})$. Contradiction! □

Remark 3.1. Note that for any R_i, R_i' such that $t(R_i) = t(R_i')$ and $p(R_i) = p(R_i')$, for any R_{-i}, we have $\psi^\sigma(R_i, R_{-i}) = \psi^\sigma(R_i', R_{-i})$. That is, the information needed to carry out the mechanism is $p(R_i)$ and $t(R_i)$ for each R_i.

3.3.2 Generalizing the priority

In the “proposing in turn” mechanism the priority of agents is specified by a fixed ordering, or say, a queue. It is easy to generalize the priority queue into a priority tree, in which agents may have different successors depending on whether she is in or out of the group, and run the mechanism based on the priority tree. The structure of a priority tree for $n = 4$ is illustrated in the following example.

Example 3.4. (a priority tree for $N = \{a, b, c, d\}$)



Each node has two child nodes except the terminal nodes. And each node is labeled by an agent, which gives messages of who is going to choose at a particular stage. Arcs are labeled either "in" or "out", which together with the labels of the nodes, give a way to track the history and decide what the current stage is.

The formal definition provides a precise description of the tree structure illustrated above.[12]

Definition 3.5. (Priority tree)

A *priority tree* $\Gamma = (V, Q; \mathscr{L}, \mathscr{H})$ is a *rooted tree* with labels, where V is the set of vertices, and $Q \subset V \times V$ is the set of arcs: if $(v_i, v_j) \in Q$ for $v_i, v_j \in V$, then there is an arc from v_i to v_j; for each $v \in V$, $\mathscr{L}(v)$ is the label of v, and for each $(v_i, v_j) \in Q$, $\mathscr{H}(v_i, v_j)$ is the label of (v_i, v_j). A *Q-path* from v_{s_1} to v_{s_r} is a sequence $\{v_{s_j}\}_{j=1}^{r}$, where $r \geq 2$, such that for all $j = 1, \cdots, r-1$, $(v_{s_j}, v_{s_{j+1}}) \in Q$. The *length* of a Q-path is the number of the connecting arcs. Since Γ is a rooted tree, Q is *acyclic:* if there is a Q-path from v_i to v_j, then $(v_j, v_i) \notin Q$, i.e., there are no cycles. Furthermore, for all $v_i, v_j \in V$, there is at most one Q-path from v_i to v_j. Thus, if $\{v_{s_j}\}_{j=1}^{r}$ is a Q-path, the *distance* of v_{s_r} from v_{s_1} is unambiguously defined by the length of the Q-path: $d(v_{s_1}, v_{s_r}) = r - 1$. Finally Γ has a unique *root* $v_1 \in V$, that is, v_1 is the only vertex such that there is no $v \in V$ with $(v, v_1) \in Q$.

The following properties define the structure and dimensions of the priority tree:

(**A**.1) $\max_{v \in V} d(v_1, v) = n - 1$.

(**A**.2) The number of arcs starting from v_1 is 2.

(**A**.3) For all $v \in V$ such that there is a Q-path from v_1 to v, with $d(v_1, v) = r < n - 1$, the number of arcs starting from v is 2.

The following properties concern the labeling of vertices:

(**B**.1) All vertices are labeled by agents: for all $v \in V$, $\mathscr{L}(v) \in N$.

[12] The language used in the definition is borrowed from [56], to which we are very grateful. Since our problem is quite different from the assignment problem studied there, our priority tree differs in many aspect with the inheritance trees defined there.

(**B**.2) Every vertex of a Q-path represents a different agent: for all $v_i, v_j \in V$ such that there is a Q-path from v_i to v_j, we have $\mathscr{L}(v_i) \neq \mathscr{L}(v_j)$.

The following properties concerns the labeling of arcs:

(**C**.1) All arcs are labeled either "*in*" or "*out*": for all $(v_i, v_j) \in Q$, $\mathscr{H}(v_i, v_j) \in \{in, out\}$.

(**C**.2) For any vertex that starts two arcs, one arc is labeled "*in*", the other is labeled "*out*": for all $v_i, v_j, v_l \in V$ such that $(v_i, v_j) \in Q$ and $(v_i, v_l) \in Q$ and $j \neq l$, we have $\mathscr{H}(v_i, v_j) \cup \mathscr{H}(v_i, v_l) = \{in, out\}$

Let $\mathscr{T}$ be the set of priority trees. Then for each priority tree Γ, a proposing mechanism associated with Γ can be thought of as carrying on a "proposing in turn" mechanism along a particular path of Γ. In the above example, if agent a chooses W, then agent c is the next to make choice; if agent c chooses "J and C", then agent d is the next to act, etc.

In general, we say that a group selection mechanism f is a proposing mechanism if there exists $\Gamma \in \mathscr{T}$ such that f is the proposing mechanism associated with Γ. In this case we will also say that Γ is an underlying priority tree for f. The formal definition of the proposing mechanism is given below.

For any vertex v_i that starts two arcs, we use $\overrightarrow{v_i}^{in}$ to represent the vertex v_j such that $(v_i, v_j) \in Q$ and $\mathscr{H}(v_i, v_j) = in$, and $\overrightarrow{v_i}^{out}$ to represent the vertex v_l such that $(v_i, v_l) \in Q$ and $\mathscr{H}(v_i, v_l) = out$.

Definition 3.6. (proposing mechanism associated with priority tree Γ)

Given a priority tree Γ.

Step **1**: denote the peak of agent $\sigma_1 = \mathscr{L}(v_1)$ by p_1. If $p_1 = 0$, let $G_1 = \emptyset$, $w_1 = v_1$ and move to step **2(a)**; if $p_1 = 1$, the algorithm ends and the final outcome is $\{\sigma_1\}$; if $p_1 > 1$, let $G_1 = \{\sigma_1\}$, $w_1 = v_1$ and move to step **2(b)**.

...

Step **k(a)**: denote the peak of agent $\sigma_k = \mathscr{L}(\overrightarrow{w_{k-1}}^{out})$ by p_k. If $p_k < |G_{k-1}| + 1$, then check whether she prefers $|G_{k-1}| + 1$ to 0, if yes, the algorithm ends and the final outcome is $G_{k-1} \cup \{\sigma_k\}$, if no, let $G_k = G_{k-1}$, $w_k = \overrightarrow{w_{k-1}}^{out}$ and move to step **(k+1)(a)**; If $\min\{p_j : j \in G_{k-1} \cup \{\sigma_k\}\} = |G_{k-1}| + 1$, the algorithm ends and the final outcome is $G_{k-1} \cup \{\sigma_k\}$; If $\min\{p_j : j \in G_{k-1} \cup \{\sigma_k\}\} > |G_{k-1}| + 1$, let $G_k = G_{k-1} \cup \{\sigma_k\}$, $w_k = \overrightarrow{w_{k-1}}^{out}$ and move to step **(k+1)(b)**.

Step **k(b)**: denote the peak of agent $\sigma_k = \mathscr{L}(\overrightarrow{w_{k-1}}^{in})$ by p_k. If $p_k < |G_{k-1}| + 1$, then check whether she prefers $|G_{k-1}| + 1$ to 0, if yes, the algorithm ends and the final outcome is $G_{k-1} \cup \{\sigma_k\}$, if no, let $G_k = G_{k-1}$, $w_k = \overrightarrow{w_{k-1}}^{in}$ and move to step **(k+1)(a)**; If $\min\{p_j : j \in G_{k-1} \cup \{\sigma_k\}\} = |G_{k-1}| + 1$, the algorithm ends and the final outcome is $G_{k-1} \cup \{\sigma_k\}$; If $\min\{p_j : j \in G_{k-1} \cup \{\sigma_k\}\} > |G_{k-1}| + 1$, let $G_k = G_{k-1} \cup \{\sigma_k\}$, $w_k = \overrightarrow{w_{k-1}}^{in}$ and move to step **(k+1)(b)**.

...

Step **n**: check whether agent $\sigma_n = \mathscr{L}(\overrightarrow{w_{n-1}}^{in})$ (note that $\mathscr{L}(\overrightarrow{w_{n-1}}^{in}) = \mathscr{L}(\overrightarrow{w_{n-1}}^{out})$) prefers $|G_{n-1}|+1$ to 0, if yes, the algorithm ends and the final outcome is $G_{n-1} \cup \{\sigma_n\}$, if no, the algorithm ends and the final outcome is G_{n-1}.

3.3.3 Characterization of the proposing mechanism

In this subsection we present a characterization of the proposing mechanism, in which non-bossiness plays a key role.

Definition 3.7. (Non-bossiness)

A group selection mechanism f is non-bossy (NB) if for all R, $i \in N$, and R'_i, $f_i(R) = f_i(R'_i, R_{-i})$ implies $f(R) = f(R'_i, R_{-i})$.

In our setting non-bossiness is nicely related to group strategy-proofness as in several rather different models[13].

Definition 3.8. (Group strategy-proofness)

A group selection mechanism f is group strategy-proof (GSP) if for all R, there does not exists $M \subset N$ and R'_M such that for all $i \in M$, $f_i(R'_M, R_{-M}) R_i f_i(R)$ and for some $j \in M$, $f_j(R'_M, R_{-M}) P_j f_j(R)$.

Lemma 3.1. *A group selection mechanism f is group strategy proof if and only if it is strategy-proof and non-bossy.*

Proof. We modify the proof in [56] to accommodate the single-peakedness of the preferences. See appendix for details. □

Theorem 3.1. *A proposing mechanism is efficient, group strategy-proof and individually rational; furthermore, any group selection mechanism that is efficient, group strategy-proof and individually rational is a proposing mechanism.*

Proof. We omit the proof for the first half of the theorem since a large part of it is similar to the proof of Proposition 3.1 and non-bossiness of the mechanism is easy to see. For the second half, we show here any Eff, GSP and IR mechanisms are proposing mechanisms for $n = 3$; the general arguments are presented in the appendix with the help of heavier notation. The $n = 3$ case, though much simpler, does provide us the logic framework that will be developed further in proving the general case.

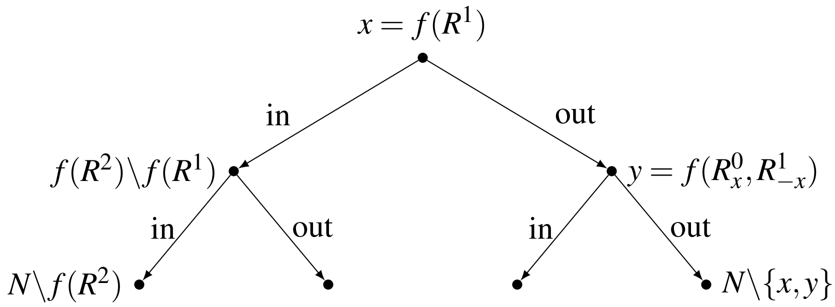
[13] See the appendix of [63] for a list of models in which non-bossiness is (or is not) related to group strategy-proofness.

Suppose $N = \{a,b,c\}$. Fix any Eff, GSP and IR mechanism f. For simplicity we write $R_i^k = (k, k-1, \cdots 1, 0)$, that is, $p(R_i^k) = t(R_i^k) = k$; $R_S^k = \{R_i^k\}_{i \in S}$ for each $S \subseteq N$, and $R^k = \{R_i^k\}_{i \in N}$. We will prove the following two statements.

(i) If $f(R^1) = \{a\}$, then $\{a\} \subset f(R^2)$.

(ii) If $f(R^1) = \{a\}$, then $f(R) = \{a\}$ for any R such that $p(R_a) = 1$; furthermore if $f(R^2) = \{a,b\}$, then $f(R) = \{a,b\}$ for any R such that $p(R_a) \geq 2$, $p(R_b) \leq 2$, and $2R_b0$.

Using the first statement, we can find out a priority tree associated with f as follows:



The second statement then shows that the outcome of f coincides with "the proposing in turn" mechanism running on the priority tree.

<u>Now we show (i):</u> First by efficiency we have $|f(R^k)| = k$ for any $k \in S$. Suppose for the sake of contradiction, $f(R^2) = \{b,c\}$. Since $f(R_a^1, R_b^2, R_c^2) \neq \{a\}$ by SP, we have $f(R_a^1, R_b^2, R_c^2) = \{b,c\}$ by Eff and IR. Since $f(R^1) = \{a\}$, $f(R_a^1, R_b^1, R_c^2) \neq \{c\}$ by SP, and $|f(R_a^1, R_b^1, R_c^2)| \neq 2$ by IR, we have $f(R_a^1, R_b^1, R_c^2) = \{a\}$ by NB of agent c. Hence $f(R_a^1, R_b^1, \tilde{R}_c^2) = \{a\}$ for $\tilde{R}_c^2 = (1,2,0)$ by similar argument. And by symmetry of agent b and c, we have $f(R_a^1, \tilde{R}_b^2, R_c^1) = \{a\}$ for $\tilde{R}_b^2 = (1,2,0)$. Hence $f(R_a^1, \tilde{R}_b^2, \tilde{R}_c^2) \neq \{b\}$ or $\{c\}$ by SP and NB. Furthermore, since $f(R_a^1, R_b^2, R_c^2) = \{b,c\}$, by GSP we have $f(R_a^1, \tilde{R}_b^2, \tilde{R}_c^2) = \{b,c\}$.

However using the same argument in the two agent case (example 3.3), we know that $f(R_a^0, \tilde{R}_b^2, \tilde{R}_c^2) \neq \{b,c\}$. Therefore agent a is bossy. Contradiction!

<u>We show (ii):</u> by the above argument, we know that $f(R_a^1, R_b^2, R_c^2) \neq \{b,c\}$. Hence $f(R_a^1, R_b^2, R_c^2) = \{a\}$. Hence by SP and NB, we have that $f(R_a^1, R'_{-a})$ for any R'_{-a}. And by SP, we have $f(R) = \{a\}$ for any R such that $p(R_a) = 1$.

Since $f(R^2) = \{a,b\}$, we have that $f(R_a^3, R_b^2, R_c^3) = \{a,b\}$ by SP and NB. Then $f(R_a^3, \tilde{R}_b^3, R_c^3) = \{a,b\}$ for $\tilde{R}_b^3 = (2,3,1,0)$ by SP. Therefore $f(R_a^3, \tilde{R}_b, R_c^3) = \{a,b\}$ for $\tilde{R}_b$ such that $p(\tilde{R}_b) \leq 2$ and $2R_b0$. Therefore $f(R) = \{a,b\}$ for any R such that $p(R_a) \geq 2$, $p(R_b) \leq 2$, and $2R_b0$. □

Remark 3.2. The characterization is tight. In example 3.1 and example 3.2, we have mechanisms that fail Eff and IR respectively (both are non-bossy); and we will present a class of mechanisms in the next section which satisfies Eff, SP and IR, but fails non-bossiness.

An obvious feature of this mechanism is that the final group size could be much smaller than the maximum group size $m^\star$. Though we know from the two agent case (example 3.3) that any Efficient, SP and IR mechanism will yield group size smaller than $m^\star$ at some states, it is still a significant fact that the final group size could be 1 while the maximum group size be n.[14] The mechanisms proposed in the next section will do much better in terms of group size.

3.4 “Voting on ascending-size”

We present in this section another class of mechanisms that are Pareto efficient, individually rational, and strategy-proof; furthermore, the final group size generated by each mechanism in the class will be at least one half of the maximum group size.

We start with a intuitive procedure (to achieve the maximum group size $m^\star$): agents are asked to approve for a group size in each round according to their announced preferences; only agents votes “yes” remain active in any future rounds. Starting with size 1, in each round, if more than enough agents are demanding that size (agents who vote “yes”), increase the size by one and proceed to the next round. Since the demand is non-increasing with the size, at some point, the group size being voted will equal or exceed the demand and we find the maximum group size. Form a group with the maximum size, breaking ties when necessary. However we already know from the two agents case (example 3.3) that no matter what tie-breaking method we use, this procedure cannot be strategy-proof.

To see more clearly the incentive problem caused by tie-breaking, imagine the following scenario. Suppose three agents a, b, c vote “yes” for group size 2; if all three agents vote “yes” for group size 3, then a group of size 3 will form; if only two of them vote “yes” for group size 3, then a group of size 2 will form among the three agents, using a tie-breaking mechanism. Suppose ties are broken using a fixed ordering $\sigma = (a,b,c)$, then agent a would rather report $(2,1,0)$ if her true preference is $(2,3,1,0)$. If we modify the ties-breaking mechanism in certain occasions to punish potential dishonesty, say, whenever two vote “yes” and one votes “no”, the two “yes” will be selected, then agent c would rather report $(2,3,1,0)$ if her true preference is $(2,1,0)$, when agent a votes “no” and agent b votes “yes” to size 3.

Instead of modifying the fixed tie-breaking mechanism, we modify the procedure to respect the priorities some agents have in tie-breaking, as showed in the following example.

Remember that for any m, for any R, $\Delta_m(R)$ is the set of agents who prefer size m to size 0. For any l such that $|\Delta_l(R)| \geq l$, for any ordering σ, let $\tilde{\Delta}_l^\sigma(R)$ be the set containing l

[14] Say, if $R_{\sigma_1} = (1,2,\cdots,n,0)$, $R_{\sigma_i} = (n,n-1,\cdots,1,0)$ for all $i = 2,\cdots,n$. Then $|f(R)| = 1$ while $m^\star(R) = n$.

agents selected from $\Delta_l(R)$ by the priority σ. (We suppress R in the notation if no confusion is caused.)

Example 3.5. (A direct mechanism for $n = 3$.)

Suppose $N = \{a,b,c\}$. Fixed an order of agent $\sigma = (a,b,c)$. For any announced preference profile R, the group is selected as follows.

Step 1: If $|\Delta_1| \leq 1$, then we select Δ_1 and the algorithm ends. If $|\Delta_1| > 1$, we move to step 2.

Step 2: If $|\Delta_2| < 2$, then select $\tilde{\Delta}_1^\sigma$ and the algorithm ends. If $|\Delta_2| = 2$, and there exists an agent in $\tilde{\Delta}_1^\sigma \cap \Delta_2$ who prefers size 1 to size 2, then select $\tilde{\Delta}_1^\sigma$ and the algorithm ends; otherwise select Δ_2 and the algorithm ends. If $|\Delta_2| > 2$, move to step 3.

Step 3: If $|\Delta_3| < 3$, then select $\tilde{\Delta}_2^\sigma$ and the algorithm ends. If $|\Delta_3| = 3$, and there exists an agents in $\tilde{\Delta}_2^\sigma$ who prefers size 2 to size 3, then select $\tilde{\Delta}_2^\sigma$ and the algorithm ends; otherwise select Δ_3 and the algorithm ends.

When $|\Delta_2| = 2$ (or $|\Delta_3| = 3$), some agents are *pivotal* in the sense that her report will determine whether she ends up in a size-two group or a size-one group (a size-three group or a size-two group). The above mechanism allows these pivotal agents to choose the smaller size whenever any of them wants to. We will show later that for $n = 3$ this is enough for truthful report. For $n > 3$, however, manipulation continues. To illustrate, let $N = \{a,b,c,d\}$, and $\sigma = (a,b,c,d)$. We already know that when exactly three agents are supposed to vote for size 3, we have to form a group with size 2 if any of the three agents who has priority to enter the size-two group prefers size 2 to size 3. Once we do that, however, we trigger a "manipulation-restoration-further manipulation" chain. Suppose now all four agents are supposed to vote "yes" for group size 3. If agent a untruthfully says "no" to size 3, she makes agent b pivotal. And then she could join a group of size 2 if agent b also prefers size 2 to size 3. To maintain truthful report under this scenario, we cannot always form a group of size 3 or size 4 even if all four agent vote "yes" for size 3; instead, we compare size 2 with size 3, and form group $\{a,b\}$ if both a and b prefer size 2 to size 3. With more than four agents, the manipulation chain gets even longer.

To cut off this evil chain and restore truthful reports, the procedure stops when a *coalition* of agents preferring size $k-1$ to size k can *collectively* misreport at round k (say "no" to size k instead of the true answer "yes") and join a size $k-1$ group, as showed in the following.

Definition 3.9. ("voting on ascending-size" or voting mechanism)

Fixed an ordering σ of agents. For any announced preference profile R, the group is selected as follows.

Step 1: If $|\Delta_1| \leq 1$, then we select Δ_1 and the algorithm ends. If $|\Delta_1| > 1$, we move to step 2.

Step k $(2 \leq k \leq n)$:

If $|\Delta_k| < k$, then select $\tilde{\Delta}^{\sigma}_{k-1}$ and the algorithm ends.

If $|\Delta_k| = k$, and at least one agent in $\tilde{\Delta}^{\sigma}_{k-1} \cap \Delta_k$ prefers size $k-1$ to size k”, then select $\tilde{\Delta}^{\sigma}_{k-1}$ and the algorithm ends; otherwise select Δ_k and the algorithm ends.

If $|\Delta_k| > k$, and at least $|\Delta_k| - k + 1$ of agents in $\tilde{\Delta}^{\sigma}_{k-1} \cap \Delta_k$ prefer size $k-1$ to size k”, then select $\tilde{\Delta}^{\sigma}_{k-1}$ and the algorithm ends; otherwise move to step $k+1$.

Remark 3.3. (1) Note that $\tilde{\Delta}^{\sigma}_{k-1} \cap \Delta_k$ is the set of active agents in round k who also has the priority to enter the group with size $k-1$. If enough of them prefer size $k-1$ to size k, the process has to end. We can write the mechanism in a more concise way.

For any $R \in \mathscr{R}$, for any $m \in S$, let $\Lambda_m(R) = \{i \in \Delta_m : p(R_i) < m\}$; let $\kappa^{\sigma}(R) = \max\{m \in S : |\Delta_m| - |\Lambda_m \cap \tilde{\Delta}^{\sigma}_{m-1}| \geq m\}$. (We suppress R and σ in the notation when no confusion is caused.) Then the selected group $\phi^{\sigma}(R) = \tilde{\Delta}^{\sigma}_{\kappa}$. It is easy to check that $|\Delta_m| - |\Lambda_m \cap \tilde{\Delta}^{\sigma}_{m-1}|$ is non-increasing with m (due to the single-peakedness of preferences); hence κ is achieved under the above mechanism.

(2) Note that if $|\Delta_2| > 2$, then $|\Delta_2| - |\Lambda_2 \cap \tilde{\Delta}^{\sigma}_1| \geq 2$, and hence $\kappa \geq 2$. So the “voting on ascending-size” mechanism for $n = 3$ is reduced to what is presented in the example 3.5.

Proposition 3.2. *For any ordering σ, the voting mechanism ϕ^{σ} is Pareto efficient, individually rational, and strategy-proof.*

Proof. Since the final group is $\tilde{\Delta}^{\sigma}_{\kappa}$, individual rationality is clear. To show Pareto efficiency, we only need to consider outcomes that contain the original selected group. If $\kappa = m^{\star}$, then we either cannot enlarge the group or enlarging the group will violate individual rationality, if $\kappa < m^{\star}$,then enlarging the group will hurt someone who is in the original group.

The formal proof of strategy-proofness is left in the appendix. We give some intuition here. Consider an agent i with preference R_i such that $(m-1)R_i m R_i 0$. And suppose she has the priority to join a group with size $m-1$, that is, $i \in \tilde{\Delta}^{\sigma}_{m-1}$. Note that this is the situation that causes trouble in the native procedure discussed at the beginning of the section. If she misreports, saying that $(m-1)R_i 0 R_i m$, then both $|\Delta_m|$ and $|\Lambda_m \cap \tilde{\Delta}^{\sigma}_{m-1}|$ decrease by 1; therefore $|\Delta_m| - |\Lambda_m \cap \tilde{\Delta}^{\sigma}_{m-1}|$ does not change. Hence by misreporting she will not change the outcome even if $|\Delta_m| - |\Lambda_m \cap \tilde{\Delta}^{\sigma}_{m-1}| = m$. □

Remark 3.4. Note that for any R_i, R'_i such that $t(R_i) = t(R'_i)$ and $p(R_i) = p(R'_i)$, for any R_{-i}, we have $\phi^{\sigma}(R_i, R_{-i}) = \phi^{\sigma}(R'_i, R_{-i})$. Again the information needed to run the voting on ascending-size mechanism is $p(R_i)$ and $t(R_i)$ for each R_i.

Proposition 3.3. *(Group size)*
For any σ, for any R, $|\phi^{\sigma}(R)| \geq \frac{1}{2} \cdot m^{\star}(R)$.

Proof. Suppose that $|\phi^{\sigma}(R)| = k$, then by definition of ϕ^{σ}, $k = \max\{m \in S : |\Delta_m| \geq m + |\Lambda_m \cap \tilde{\Delta}^{\sigma}_{m-1}|\}$. We show that $|\Delta_{k+1}(R)| \leq 2k$. Suppose for the sake of contradiction $|\Delta_{k+1}(R)| > 2k$, then we must have $|\Delta_{k+1}| \geq k + 1 + |\Lambda_{k+1} \cap \tilde{\Delta}^{\sigma}_k|$ since $|\tilde{\Delta}^{\sigma}_k| = k$. Contradiction!

If $m^{\star}(R) > k$, then $m^{\star}(R) \leq |\Delta_{k+1}(R)| \leq 2k$, we have $|\phi^{\sigma}(R)| \geq \frac{1}{2} \cdot m^{\star}(R)$. □

We can see from the $n = 3$ case (example 3.5) that ϕ^{σ} does not satisfy non-bossiness for any ordering σ: say, $\sigma = (a,b,c)$, then $\phi(R) = \{b,c\}$ for R such that $R_a = (1,0)$, $R_b = R_c = (1,2)$, and $\phi(R') = \{b\}$ for R' such that $R'_a = (0)$, $R_b = R_c = (1,2)$. Agent a is bossy. It is hence not group strategy-proof; however it satisfies weak group stratety-proofness, defined below.

Definition 3.10. A group selection mechanism f is weakly group strategy-proof (w-GSP) if for all R, there does not exists $M \subset N$ and R'_M such that for all $i \in M$, $f_i(R'_M, R_{-M}) P_i f_i(R)$.

Proposition 3.4. *For any ordering σ, ϕ^{σ} is weakly group strategy-proof.*

Proof. We show it by first showing that ϕ^{σ} satisfied certain partial non-bossiness properties. See appendix for details. □

We end the section by an open question: can we achieve even larger group size by any other Eff, IR and SP mechanisms? For $n = 3$ the answer is no, as showed in the following proposition. However, the question remains open for $n > 3$.

Proposition 3.5. *(Maximal group size) For $n = 3$ and for any fixed ordering σ, there is no Eff, SP and IR mechanism g such that $|g(R)| \geq |\phi^{\sigma}(R)|$ for all R and $|g(R')| > |\phi^{\sigma}(R')|$ for some R'.*

Proof. See Appendix. □

3.5 Conclusion

This paper studies the group selection problem under single-peaked preferences, and proposes two classes of mechanisms that are efficient, strategy-proof and individually rational. There is a natural way to combine them and generate a class of hybrid mechanisms: partition the set N into two subsets N_1 and N_2. First run the “proposing in turn” mechanism in N_1 and let O_1 be the group chosen. If the mechanism ends under case (i) or (ii) (see Definition 3.4), then O_1 is final group; otherwise run the “voting on ascending-size” mechanism on

N_2, starting with group size $|O_1|+1$, ending at most at size $p = \min_{i \in O_1} p_i$, keeping in mind that the number of positions available for agents in N_2 in each round is the group size minus $|O_1|$. Let O_2 be the chosen group within N_2. The final group is then $O_1 \cup O_2$.

This hybrid mechanism is Eff, SP and IR. Note that if we take $N_1 = N$, then we get the "proposing in turn" mechanism; if we take $N_1 = \emptyset$, then we get the "voting on ascending-size" mechanism; if we take $N_1 = \{i\}$ for some $i \in N$, then we get the mechanism that favors agent i mostly among all hybrid mechanisms.

We end the paper by a discussion of open questions and possible extensions:

In settings where private objects are to be allocated, strategy-proofness is usually not enough to characterize a clearly-cut classes of mechanisms; more axioms (e.g. non-bossiness, population monotonicity, resource monotonicity, etc.) are added to offer more structural restrictions (See [8], [24, 56, 57]). Our setting is similar to the private good setting in that there are two different outcomes for agents; and bossiness is hence possible. The following example offers a glimpse of mechanisms that are strategy-proof and bossy.

Example 3.6. $N = \{a,b,c\}$. Let $\sigma = (a,b,c)$. Modify the tie-breaking rule in the mechanism presented in example 3.5 as follows: if $|\Delta_2| = 3$, then $\tilde{\Delta}_2 = \{a,b\}$ if $3R_a0$, and $\tilde{\Delta}_2 = \{a,c\}$ if $0R_a3$.

Adding non-bossiness, we are able to characterize the proposing mechanism. Interestingly, we find another class of mechanisms that is intuitively appealing and weakly group strategy-proof. Weak group strategy-proofness, however, is not helpful in characterization due to the fact that it cuts off an open set of mechanisms. In fact it is easy to check that the mechanism in the above example is w-GSP. Since axioms like population monotonicity and resouce monotonicity are largely irrelevant in our setting, how to characterize the voting mechanism is a challenging question. Furthermore, our understandings towards the structural restrictions imposed by w-GSP is far from complete, both in our setting and in more general terms. (see [59] and [63].) Needless to say, our model could serve as a motivation for further theoretic inspection.

Finally, there are many situations in which more than one group could or should be formed; it is both interesting and challenging to study group selection mechanism in those settings.[15]

[15] For example, see [13], in which a fixed number ($k = 1,2,\cdots,n$) of facility have to be installed.

3.6 Appendix

3.6.1 Universal preference domain: impossibility

In this subsection we take $\mathscr{R} = \mathscr{U}$ and present an impossibility result.

Theorem 3.2. *There is no group selection mechanism that is efficient, strategy-proof and individually rational.*

Proof. We prove the result for $n = 2$. The proof will be identical for $n > 2$: just endow the two agents with the same preferences that we consider here, and let the preferences of all others be (0).

Let $N = \{a, b\}$. Suppose for the sake of contradiction f is an Eff, SP and IR mechanism. Consider the following preference profiles:

(1) $R_a = (1,0)$, $R_b = (2,0)$

(2) $R_a = (1,2,0)$, $R_b = (2,0)$

(3) $R_a = (1,2,0)$, $R_b = (2,1,0)$

(4) $R_a = (1,2,0)$, $R_b = (1,2,0)$

We use $R^{(k)}$ to represent the above profile (k). First by Eff and IR, $f(R^{(1)}) = \{a\}$. By SP, $f(R^{(2)}) = \{a\}$. By SP, $f(R^{(3)}) \neq \{a,b\}$. And by Eff, $f(R^{(3)}) \neq \{b\}$ and $f(R^{(3)}) \neq \{\emptyset\}$. Therefore $f(R^{(3)}) = \{a\}$. By SP, $f(R^{(4)}) \neq \{b\}$ and $f(R^{(4)}) \neq \{a,b\}$. By Eff, $f(R^{(4)}) \neq \{\emptyset\}$. Therefore $f(R^{(4)}) = \{a\}$. However, if we start with $R_a = (2,0)$, $R_b = (1,0)$, by similar argument, we get that $f(R^{(4)}) = \{b\}$. Contradiction! $\square$

The above result mechanisms out positive results in an interesting sub-domain $\mathscr{P}' \subset \mathscr{U}$, where for each i, for each R_i, the *strict* upper contour set of size 0 is an integer interval and preference restricted to the interval is single-peaked. The reduced preferences profiles in [20] belongs to $\mathscr{P}'$, and in fact forms a sub-domain of $\mathscr{P}'$ in which larger group size is always preferred to smaller size. And their auction-like mechanism is Eff, IR and SP (within the sub-domain).

Note that our single-peaked preference domain $\mathscr{P}$, by requiring any nonempty strict upper contour set of size 0 for R_i be an integer interval *starting with size* 1, is another sub-domain of $\mathscr{P}'$. And it is a maximum domain where efficient, individually rational and strategy-proof group selection mechanisms exist (to see this, consider the $n = 2$ case in the proof of Theorem 3.2).

3.6.2 Proof

3.6.2.1 Proof for Lemma 3.1

Proof. Group strategy-proofness implies strategy-proofness and non-bossiness is obvious. We prove the converse. Let f be strategy-proof and non-bossy. Let $M \subset N$, R and R'_M be such that for all $i \in M$, $f_i(R'_M, R_{-M}) R_i f_i(R)$. For all $i \in M$, let $\hat{R}_i$ be the preference such that $p(\hat{R}_i) = f_i(R'_M, R_{-M}) = p^\star$ and for any $b, c \in S$ such that $p^\star R_i b$ and $p^\star R_i c$ we have $b \hat{R}_i c$ iff $b R_i c$. (We leave for the reader to check that such $\hat{R}_i$ always exists.) We first show that $f_i(R) = f_i(\hat{R}_i, R_{-i})$. Since $p^\star R_i f_i(R)$, and $f_i(R) R_i f(\hat{R}_i, R_{-i})$ by strategy-proofness, we have $p^\star R_i f_i(\hat{R}_i, R_{-i})$. Since $p^\star R_i f_i(R)$ and $p^\star R_i f_i(\hat{R}_i, R_{-i})$, we have $f_i(R) R_i f_i(\hat{R}_i, R_{-i})$ iff $f_i(R) \hat{R}_i f_i(\hat{R}_i, R_{-i})$. By strategy-proof, we must have $f_i(R) = f_i(\hat{R}_i, R_{-i})$. Then $f(R) = f(\hat{R}_i, R_{-i})$ by non-bossiness. Repeating the same argument for individuals in M one by one, we have $f(\hat{R}_M, R_{-M}) = f(R)$. On the other hand, we have $f(\hat{R}_M, R_{-M}) = f(R'_M, R_{-M})$ by strategy-proof and non-bossiness. Thus $f(R) = f(R'_M, R_{-M})$, which implies f is group strategy-proof. □

3.6.2.2 Proof for Theorem 3.1

First we introduce group non-bossiness, which is implied by group strategy-proofness and will play a key role in the proof.

Definition 3.11. (Group non-bossiness)

A group selection mechanism f is group non-bossiness (GNB) if for all R, $M \subset N$, and R'_M, $f_M(R) = f_M(R'_M, R_{-M})$ implies $f(R) = f(R'_M, R_{-M})$.

First, fix an Eff, GSP and IR mechanism f. For simplicity we write $R_i^k = (k, k-1, \cdots, 1, 0)$ and $R^k = \{R_i^k\}_{i \in N}$ for each $k \in S$; and for any $M \subsetneq N$, we write $f(R_M)$ instead of $f(R_M, R^0_{-M})$.

Lemma 3.2. *For any $k \in S$, $|f(R^k)| = k$; and $f(R) = f(R^k)$ for all R such that $p(R_i) \geq k$ for all $i \in f(R^k)$, and $|\{i \in N : (k+1) R_i 0\}| < k+1$.*

Proof. By Eff we have $|f(R^k)| = k$ for any $k \in S$. To show the second half, first note that $|f(R)| < k+1$ by IR. For $j \notin f(R^k)$, by SP and NB we have $f(R_j, R^k_{-j}) = f(R^k)$. Therefore changing R_j^k to R_j for each $j \notin f(R^k)$ one by one, and then changing R_j^k to R_j for each $j \in f(R^k)$ one by one, we have $f(R) = f(R^k)$ by SP and NB in each step. □

Lemma 3.3. $f(R_M^m) \subset f(R_M^{m+1})$ *for any* $0 \leq m \leq n-1$*, any* $M \subseteq N$.

Proof. The statement holds for all cases such that $|M| \leq m$, in which cases $f(R_M^m) = f(R_M^{m+1}) = M$. Hence we focus on cases such that $|M| > m$.

Suppose for the sake of contradiction the statement is not true. That is, there exists $M \subseteq N$, and $m < |M|$ such that $f(R_M^m) \backslash f(R_M^{m+1}) \neq \emptyset$. Let $A = f(R_M^m)$, $B = f(R_M^{m+1})$, $C = f(R_M^m) \backslash f(R_M^{m+1})$, $D = f(R_M^m) \cup f(R_M^{m+1})$. Then $f(R_B^{m+1}) = f(R_M^{m+1}) = B$; and by Lemma 3.2 $f(R_D^m) = f(R_M^m) = A$. Since $f(R_B^m) \subset B$ and $|f(R_B^m)| = |f(R_D^m)| = m$, there exists $a_l \in f(R_B^m) \backslash f(R_D^m)$, and $a_s \in f(R_B^{m+1}) \backslash f(R_B^m)$. Let $B' = B \backslash \{a_l\}$, then $f(R_{B'}^{m+1}, R_{a_l}^m) = f(R_B^m)$ by Lemma 3.2. Let $\tilde{R}_i^{m+1} = (m, m+1, m-1, \cdots, 0)$ for $i = a_l$, then $f(R_{B'}^{m+1}, \tilde{R}_{a_l}^{m+1}) = f(R_B^m)$ by SP.

However since $D = B \cup C \subseteq M$, $f(R_C^m, R_B^m) = A$, and $f(R_C^{m+1}, R_B^{m+1}) = B$, we have that $f(R_C^m, R_{B'}^{m+1}, \tilde{R}_{a_l}^{m+1}) = f(R_B^{m+1}) \neq f(R_B^m)$. This violates group non-bossiness of C. □

With the help of the above lemma, we can construct the priority tree. Given a priority tree $\Gamma = (V, Q; \mathscr{L}, \mathscr{H})$, we write a Q-path $\{v_s\}_{s=1}^r$ with length $r-1$ as $v_1 \overset{\star_1}{\to} v_2 \overset{\star_2}{\to} v_3 \overset{\star_3}{\to} \cdots \overset{\star_{r-1}}{\to} v_r$, where $\star_i = \mathscr{H}(v_i, v_{i+1})$ for $1 \leq i \leq r-1$.

Lemma 3.4. *(Construction of the priority tree for f)*

Fix a rooted tree (V, Q) with properties **A**.1, **A**.2 *and* **A**.3, *and with arcs labeled according to* **C**.1 *and* **C**.2. *Denote by $\mathscr{H}(v_i, v_j)$ the label of each $(v_i, v_j) \in Q$. We label the vertices by induction on the length of Q-path starting with v_1:*

First,$\mathscr{L}(v_1) = f(R^1)$; $\mathscr{L}(\overrightarrow{v_1}^{in}) = f(R^2) \backslash \{i\}$ and $\mathscr{L}(\overrightarrow{v_1}^{out}) = f(R^1_{N \backslash \{i\}})$, where $i = \mathscr{L}(v_1)$.

For any $k \in \{2, \cdots, n-1\}$, suppose there is a Q-path starting with v_1, with length $k-1$ and all vertices already labeled: $v_1 = v_{p_1} \overset{\star_1}{\to} v_{p_2} \overset{\star_2}{\to} v_{p_3} \overset{\star_3}{\to} \cdots \overset{\star_{k-1}}{\to} v_{p_k}$, where $\star_j \in \{in, out\}$ for all $1 \leq j \leq k-1$. Let $I = \{\mathscr{L}(v_{p_j}) : 0 \leq j \leq k-1, \star_j = in\}$ and $t = |I|$; $O = \{\mathscr{L}(v_{p_j}) : 0 \leq j \leq k-1, \star_j = out\}$. Then

$\mathscr{L}(\overrightarrow{v_{p_k}}^{in}) = f(R_M^{t+2}) \backslash (I \cup \{v_{p_k}\})$, where $M = N \backslash O$;

$\mathscr{L}(\overrightarrow{v_{p_k}}^{out}) = f(R_{M'}^{t+1}) \backslash I$, where $M' = N \backslash (O \cup \{v_{p_k}\})$.

Then $\Gamma = (V, Q; \mathscr{L}, \mathscr{H})$ is a priority tree.

Proof. First by Lemma 3.3, $\mathscr{L}(\cdot)$ is well defined. That is, $|f(R_M^{t+2}) \backslash (I \cup \{v_{p_k}\})| = 1$, and $|f(R_{M'}^{t+1}) \backslash I| = 1$. To show $\Gamma = (V, Q; \mathscr{L}, \mathscr{H})$ is a priority tree, we only need to show that $\mathscr{L}(\cdot)$ satisfies properties $B.1$ and $B.2$. First it is easy to see that $B.1$ is satisfied, that is, for all $v \in V$, $\mathscr{L}(v) \in N$. Second, since $f(R_M^{t+2}) \backslash (I \cup \{v_{p_k}\}) \notin I \cup O$ and $f(R_{M'}^{t+1}) \backslash I \notin I \cup O$, property $B.2$ is satisfied, that is, every vertex of a Q-path represents a different agent. □

Lemma 3.5. *For any $M \subseteq N$, for any $m < |M|$, $f(R) = f(R_M^m)$ for all R such that $p(R_s) = t(R_s) = m$ for some $s \in f(R_M^m)$, $p(R_i) = |M|$ for all $i \in f(R_M^m) \backslash \{s\}$,and $R_i = (0)$ for $i \notin M$.*

Proof. First we must have $|f(R)| \geq |f(R^m_M)|$ for any R as described in the statement. Suppose $|f(R)| = k < |f(R^m_M)|$. Let $A = f(R)\backslash f(R^m_M)$. By GSP we have that $f(R^k_A, R_{-A}) = f(R)$. Now let $B = f(R^m_M)$, then by NB of the group $N\backslash(A\cup B)$, we have that $f(R^k_A, R_B) = f(R)$. However, we know from Lemma 3.2 that $f(R^k_A, R^m_B) = B$. Hence $f(R^k_A, R_B) = f(R)$ is impossible by GSP. Suppose $|f(R)| = |f(R^m_M)|$, then by similar argument we must have $f(R) = f(R^m_M)$.

Hence we assume for the sake of contradiction that there exist R such that $l = |f(R)| > |f(R^m_M)| = m$. Then $s \notin f(R)$ by IR. Let $A = f(R)\cup f(R^m_M)$. Then by SP $m < |f(R_{A\backslash s}, \tilde{R}^l_s)| = k \leq l$, where $\tilde{R}^l_s = (m, m+1, \cdots, l-1, l, m-1, m-2, \cdots, 1, 0)$. And $f(R^m_M) \subset f(R_{A\backslash s}, \tilde{R}^l_s)$; in particular, $s \in f(R_{A\backslash s}, \tilde{R}^l_s)$. By GSP we have that $f(R^k_{A\backslash s}, \tilde{R}^l_s) = f(R_{A\backslash s}, \tilde{R}^l_s)$. Let $B = A\backslash f(R^k_{A\backslash s}, \tilde{R}^l_{a_m})$. Note that $B \neq \emptyset$.

Since $s \in f(R^{k-1}_{A\backslash B})$, by Lemma 3.2 we have $|f(R^k_{A\backslash(B\cup s)}, R^{k-1}_s)| = k-1$ and $s \in f(R^k_{A\backslash(B\cup s)}, R^{k-1}_s)$. Therefore by SP, we have $|f(R^k_{A\backslash(B\cup s)}, \tilde{R}^l_s)| \leq k-1$ and $s \in f(R^k_{A\backslash(B\cup s)}, \tilde{R}^l_s)$. Since $|f(R^k_{A\backslash s}, \tilde{R}^l_s)| = k$, group non-bossiness of B is violated. □

Lemma 3.6. *For any $M \subseteq N$, for any $m < |M|$, let $f(R^{m-1}_M) = \{a_1, a_2, \cdots, a_{m-1}\}$, and $a_m = f(R^m_M)\backslash f(R^{m-1}_M)$; then $f(R) = f(R^m_M)$ for all R such that $p(R_i) = |M|$ for all $i \in \{a_1, \cdots, a_{m-1}\}$, $p(R_i) < m$, $t(R_i) = m$ for $i = a_m$, and $R_i = (0)$ for $i \notin M$.*

Proof. We omit the proof since it is very similar to the proof of Lemma 3.5. □

Lemma 3.7. *For any $M \subseteq N$ with $|M| = m$, let $\{v_{p_s}\}^n_{s=1}$ be the Q-path of Γ such that $\{\mathscr{L}(v_{p_j}) : \mathscr{H}(v_{p_j}, v_{p_{j+1}}) = out\} = N\backslash M$. Let $\sigma_s = \mathscr{L}(v_{p_s})$ for $1 \leq s \leq n$. Let $s^\star = \max\{s : \sigma_s \in M\}$, $\bar{M} = \{\sigma_s : 1 \leq s \leq s^\star\}$; and let $l_s = |\{\sigma_t \in M : t \leq s\}|$ for $s = \{1, \cdots, n\}$. Then $f(R) = M$ for any R such that (1) $p(R_{\sigma_s}) \geq m$ for all $\sigma_s \in M$, and $0R(l_s+1)$ for all $\sigma_s \notin M$, or (2) $p(R_{\sigma_s}) \geq m$ for all $\sigma_s \in M$, $p(R_{\sigma_w}) = m$ for some $\sigma_w \in M$, and $0R_{\sigma_s}(l_s+1)$ for all $\sigma_s \in \bar{M}\backslash M$, or (3) $p(R_{\sigma_s}) \geq m$ for all $\sigma_s \in M\backslash\{\sigma_{s^\star}\}$, $p(\sigma_{s^\star})R_{\sigma_{s^\star}}mR_{\sigma_{s^\star}}0$, and $0R_{\sigma_s}(l_s+1)$ for all $\sigma_s \in \bar{M}\backslash M$.*

Proof. Suppose R satisfies condition (1). First we show that if $\sigma_s \notin M$, then $\sigma_s \notin f(R)$. Suppose for the sake of contradiction there exists $\sigma_s \notin M$ and $\sigma_s \in f(R)$, since $0R_{\sigma_s}l_s$, by IR we have $|f(R)| < (l_s+1)$. Since $l_s = |\{\sigma_t \in M : t \leq s\}|$, there exists $k < s$ such that $\sigma_k \in M$ and $\sigma_k \notin f(R)$. Let $A = f(R)\cup\{\sigma_k\}$; let $h = |f(R)|$. Since $k < s$, we have that $\sigma_k \in f(R^h_A)$. Since $p(R_{\sigma_k}) \geq m \geq l_s > h$, by SP, we must have $\sigma_k \in f(R_{\sigma_k}, R^h_{A\backslash\{\sigma_k\}})$. And by SP and NB, we must have that $\sigma_k \in f(R)$. Contradiction! Therefore if $\sigma_s \notin M$, then $\sigma_s \notin f(R)$. Since $p(R_{\sigma_s}) \geq m$ for all $\sigma_s \in M$, by Eff, we have $f(R) = M$.

Suppose R satisfies condition (2) or (3). Using similar argument, we can show that if $\sigma_s \in \bar{M}\backslash M$, then $\sigma_s \notin f(R)$. Then by NB $f(R_{N\backslash(\bar{M}\backslash M)}) = f(R)$. And by Lemma 3.5

and Lemma 3.6 (more accurately, a slight extension of the two Lemma), we have that $f(R_{N\backslash(\bar{M}\backslash M)}) = M$. Therefore $f(R) = M$. $\square$

Now we show that f coincides with the proposing mechanism associated with Γ. For any R, let M be the outcome generated by the proposing mechanism associated with Γ. Let $\{v_{p_s}\}_{s=1}^{n}$ be the Q-path of Γ such that $\{\mathscr{L}(v_{p_j}) : \mathscr{H}(v_{p_j}, v_{p_{j+1}}) = out\} = N\backslash M$. Let $\sigma_s = \mathscr{L}(v_{p_s})$ for $1 \leq s \leq n$. Let $s^\star = \max\{s : \sigma_s \in M\}$, $\bar{M} = \{\sigma_s : 1 \leq s \leq s^\star\}$; and let $l_s = |\{\sigma_t \in M : t \leq s\}|$ for $s = \{1, \cdots, n\}$. We know from the definition of the proposing mechanism that R satisfies condition (1) or condition (2) or condition (3) in Lemma 3.7. And by Lemma 3.7, we have $f(R) = M$, as desired.

3.6.2.3 Proof for Proposition 3.2

Proof. We show strategy-proofness of ϕ^σ. For simplicity we suppress σ in all the notation below. Suppose for the sake of contradiction there exists R_i, R'_i and R_{-i} such that $\phi_i(R'_i, R_{-i}) P_i \phi_i(R_i, R_{-i})$. Write $\kappa = \kappa(R_i, R_{-i})$ and $\kappa' = \kappa(R'_{-i}, R_{-i})$. Consider the following cases:

case (1): $\kappa < p(R_i)$.

Suppose for the sake of contradiction we have $\kappa' > \kappa$. Note that $\Delta_{\kappa+1}(R'_i, R_{-i}) \subseteq \Delta_{\kappa+1}(R_i, R_{-i})$. And since $i \notin \Lambda_{\kappa+1}(R_i, R_{-i})$, we have that $\Lambda_{\kappa+1}(R_i, R_{-i}) \cap \tilde{\Delta}_\kappa(R_i, R_{-i}) \subseteq \Lambda_{\kappa+1}(R'_i, R_{-i}) \cap \tilde{\Delta}_\kappa(R'_i, R_{-i})$. Therefore $|\Delta_{\kappa+1}(R', R_{-i})| - |\Lambda_{\kappa+1}(R'_i, R_{-i}) \cap \tilde{\Delta}_\kappa(R'_i, R_{-i})| \leq |\Delta_{\kappa+1}(R, R_{-i})| - |\Lambda_{\kappa+1}(R_i, R_{-i}) \cap \tilde{\Delta}_\kappa(R_i, R_{-i})| < \kappa + 1$; hence $\kappa' < \kappa + 1$. Either $i \in \tilde{\Delta}_\kappa(R_i, R_{-i})$ or $i \notin \tilde{\Delta}_\kappa(R_i, R_{-i})$, $\phi_i(R'_i, R_{-i}) P_i \phi_i(R_i, R_{-i})$ is impossible.

case (2): $\kappa > p(R_i)$ and $i \in \tilde{\Delta}_\kappa(R_i, R_{-i})$.

Suppose for the sake of contradiction that $\kappa' < \kappa$ and $i \in \tilde{\Delta}_{\kappa'}(R'_i, R_{-i})$. We first show that $\kappa' \geq \kappa - 1$. Since $\kappa' \geq \kappa - 2$ (by changing from R_i to R'_i, we can reduce $|\Delta_\kappa(R, R_{-i})| - |\Lambda_\kappa(R_i, R_{-i}) \cap \tilde{\Delta}_{\kappa-1}(R_i, R_{-i})|$ by 2 at most), we have $i \in \tilde{\Delta}_{\kappa-2}(R'_i, R_{-i}) = \tilde{\Delta}_{\kappa-2}(R_i, R_{-i})$. Now if $i \in \Delta_{\kappa-1}(R'_i, R_{-i})$, then $\Delta_{\kappa-1}(R'_i, R_{-i}) = \Delta_{\kappa-1}(R_i, R_{-i})$ and $\Lambda_{\kappa-1}(R_i, R_{-i}) \cap \tilde{\Delta}_{\kappa-2}(R_i, R_{-i}) \supseteq \Lambda_{\kappa-1}(R'_i, R_{-i}) \cap \tilde{\Delta}_{\kappa-2}(R'_i, R_{-i})$. Hence $|\Delta_{\kappa-1}(R', R_{-i})| - |\Lambda_{\kappa-1}(R'_i, R_{-i}) \cap \tilde{\Delta}_{\kappa-2}(R'_i, R_{-i})| \geq |\Delta_{\kappa-1}(R, R_{-i})| - |\Lambda_{\kappa-1}(R_i, R_{-i}) \cap \tilde{\Delta}_{\kappa-2}(R_i, R_{-i})| > \kappa - 1$; hence $\kappa' \geq \kappa - 1$.

If $i \notin \Delta_{\kappa-1}(R'_i, R_{-i})$, then $\Delta_{\kappa-1}(R'_i, R_{-i})| = \Delta_{\kappa-1}(R_i, R_{-i}) \backslash \{i\}$, and $\Lambda_{\kappa-1}(R'_i, R_{-i}) = \Lambda_{\kappa-1}(R_i, R_{-i}) \backslash \{i\}$. Therefore $|\Delta_{\kappa-1}(R', R_{-i})| - |\Lambda_{\kappa-1}(R'_i, R_{-i}) \cap \tilde{\Delta}_{\kappa-2}(R'_i, R_{-i})| = |\Delta_{\kappa-1}(R, R_{-i})| - |\Lambda_{\kappa-1}(R_i, R_{-i}) \cap \tilde{\Delta}_{\kappa-2}(R_i, R_{-i})| > \kappa - 1$. hence $\kappa' \geq \kappa - 1$.

Therefor $\kappa' \geq \kappa - 1$ and we have that $i \in \tilde{\Delta}_{\kappa-1}(R_i, R_{-i}) = \tilde{\Delta}_{\kappa-1}(R'_i, R_{-i})$.

Now if $i \in \Delta_\kappa(R'_i, R_{-i})$, then $\Delta_\kappa(R'_i, R_{-i}) = \Delta_\kappa(R_i, R_{-i})$ and $\Lambda_\kappa(R_i, R_{-i}) \cap \tilde{\Delta}_{\kappa-1}(R_i, R_{-i}) \supseteq \Lambda_\kappa(R'_i, R_{-i}) \cap \tilde{\Delta}_{\kappa-1}(R'_i, R_{-i})$. Hence $|\Delta_\kappa(R', R_{-i})| - |\Lambda_\kappa(R'_i, R_{-i}) \cap \tilde{\Delta}_{\kappa-1}(R'_i, R_{-i})| \geq |\Delta_\kappa(R, R_{-i})| - |\Lambda_\kappa(R_i, R_{-i}) \cap \tilde{\Delta}_{\kappa-1}(R_i, R_{-i})| \geq \kappa$; hence $\kappa' \geq \kappa$.

If $i \notin \Delta_\kappa(R'_i, R_{-i})$, then $\Delta_\kappa(R'_i, R_{-i})| = \Delta_\kappa(R_i, R_{-i}) \backslash \{i\}$, and $\Lambda_\kappa(R'_i, R_{-i}) = \Lambda_\kappa(R_i, R_{-i}) \backslash \{i\}$. Therefore $|\Delta_\kappa(R', R_{-i})| - |\Lambda_\kappa(R'_i, R_{-i}) \cap \tilde{\Delta}_{\kappa-1}(R'_i, R_{-i})| = |\Delta_\kappa(R, R_{-i})| - |\Lambda_\kappa(R_i, R_{-i}) \cap \tilde{\Delta}_{\kappa-1}(R_i, R_{-i})| \geq \kappa$. hence $\kappa' = \kappa$.

Case (3): $\kappa > p(R_i)$, $t(R_i) > \kappa$ and $i \notin \tilde{\Delta}_\kappa(R_i, R_{-i})$.

Suppose for the sake of contradiction we have $\kappa' > \kappa$ and $i \in \tilde{\Delta}_{\kappa'}(R'_i, R_{-i})$. Note that again we have $\Delta_{\kappa+1}(R'_i, R_{-i}) \subseteq \Delta_{\kappa+1}(R_i, R_{-i})$. Since $i \notin \tilde{\Delta}_\kappa(R_i, R_{-i}) = \tilde{\Delta}_\kappa(R'_i, R_{-i})$, we have that $\Lambda_{\kappa+1}(R_i, R_{-i}) \cap \tilde{\Delta}_\kappa(R_i, R_{-i}) = \Lambda_{\kappa+1}(R'_i, R_{-i}) \cap \tilde{\Delta}_\kappa(R'_i, R_{-i})$. Therefore $|\Delta_{\kappa+1}(R', R_{-i})| - |\Lambda_{\kappa+1}(R'_i, R_{-i}) \cap \tilde{\Delta}_\kappa(R'_i, R_{-i})| \leq |\Delta_{\kappa+1}(R, R_{-i})| - |\Lambda_{\kappa+1}(R_i, R_{-i}) \cap \tilde{\Delta}_\kappa(R_i, R_{-i})| < \kappa + 1$; hence $\kappa' < \kappa + 1$.

Case (4): $\kappa > p(R_i)$, $t(R_i) \leq \kappa$ and $i \notin \tilde{\Delta}_\kappa(R_i, R_{-i})$.

Then $\phi_i(R'_i, R_{-i}) P_i \phi_i(R_i, R_{-i})$ is impossible. □

3.6.2.4 Proof for Proposition 3.4

It is easy to check that ϕ^σ satisfies non-bossiness of members, defined below; and it satisfies the following weak non-bossiness of non-members condition[16].

Definition 3.12. A group selection mechanism f is non-bossy for members (NBM) if for all R,$i \in N$, and R'_i, $f_i(R) = f_i(R'_i, R_{-i}) \neq 0$ implies that $f(R) = f(R'_i, R_{-i})$.

Definition 3.13. A group selection mechanism f is preference-restricted non-bossy of non-members (PRNBN) if for all R, $i \in N$, R'_i such that $T(R'_i) = T(R)$, $f_i(R) = f_i(R'_i, R_{-i}) = 0$ implies that $f(R) = f(R'_i, R_{-i})$.

Proof. We modify the proof for Lemma 1 to show that a SP and IR group selection mechanism is weakly group strategy-proof if it satisfies NBM and PRNBN.

Let f be a mechanism that is SP, IR and satisfies NBM and PRNBN. Suppose for the sake of contradiction, there exists $M \subset N$, R and R'_M be such that for all $i \in M$, $f_i(R'_M, R_{-M}) P_i f_i(R)$. Since f is individual rational, we must have $p^\star := f_i(R'_M, R_{-M}) > 0$ and $p^\star \in T(R_i)$ for all $i \in M$.

For all $i \in M$, let $\hat{R}_i$ be the preference such that $p(\hat{R}_i) = p^\star$, $T(\hat{R}) = T(R)$, and for any $b, c \in S$ such that $p^\star R_i b$ and $p^\star R_i c$ we have $b \hat{R}_i c$ iff $b R_i c$. (We leave for the reader to check that such $\hat{R}_i$ always exists.) We first show that $f_i(R) = f_i(\hat{R}_i, R_{-i})$. Since $p^\star R_i f_i(R)$, and $f_i(R) R_i f(\hat{R}_i, R_{-i})$ by strategy-proofness, we have $p^\star R_i f_i(\hat{R}_i, R_{-i})$**.** Since $p^\star R_i f_i(R)$ and $p^\star R_i f_i(\hat{R}_i, R_{-i})$, we have $f_i(R) R_i f_i(\hat{R}_i, R_{-i})$ iff $f_i(R) \hat{R}_i f_i(\hat{R}_i, R_{-i})$. By strategy-proof, we must have $f_i(R) = f_i(\hat{R}_i, R_{-i})$. If $f_i(R) > 0$, then $f(R) = f(\hat{R}_i, R_{-i})$ by non-bossiness of

[16] See the weak non-bossiness definition in [59]. Inspired by that paper, we propose this condition to show in a more convenient way that our mechanism is weakly group strategy-proof.

members; if $f_i(R) = 0$, then $f(R) = f(\hat{R}_i, R_{-i})$ by PRNBN. Repeating the same argument for individuals in M one by one, we have $f(\hat{R}_M, R_{-M}) = f(R)$. On the other hand, we have $f(\hat{R}_M, R_{-M}) = f(R'_M, R_{-M})$ by strategy-proof and non-bossiness for members. Thus $f(R) = f(R'_M, R_{-M})$. Contradiction! $\square$

3.6.2.5 Proof for Proposition 3.5

Proof. Let $N = \{a,b,c\}$ and $\sigma = (a,b,c)$. Suppose for the sake of contradiction, g is such a mechanism. Then $|g(R)| \geq |\phi(R)|$ for all R, and $|g(R')| > |\phi(R')|$ for some R'. We consider the following two cases:

Case 1: $|g(R')| = 2$, $|\phi(R')| = 1$. Since $|g(R')| = 2$, we know that $|\{i : 2R_i0\}| \geq 2$ by IR. Since $|\phi(R')| = 1$, we know that $|\{i : 2R_i0\}| \leq 2$ by the definition of ϕ (see (2) of Remark 3.3). Hence $|\{i : 2R_i0\}| = 2$.

Suppose $g(R') = \{b,c\}$. Then for $|\phi(R')| = 1$, we must have $R'_a = (0)$, $p(R'_b) = 1$. However, we know already that any Eff, SP and IR mechanism will yield $\{b\}$ as a outcome. So $g(R') \neq \{b,c\}$.

Suppose $g(R') = \{a,b\}$. Then for $|\phi(R')| = 1$, we must have $p(R'_a) = 1$. Then for $|g(R')| = 2$, we must have $R'_c \neq (0)$. Hence $R'_c = (1,0)$. By SP and IR, we have that $g(\tilde{R}_a, R'_b, R'_c) \in \{b,c\}$, where $\tilde{R}_a = (1,0)$. And $g(\tilde{R}_a, R'_b, R'_c) \neq \{c\}$, otherwise $g(\tilde{R}_a, R'_b, \tilde{R}_c) = \{c\}$ for $\tilde{R}_c = (1,2,0)$ (then $|g| = 1 < |\phi|$). Therefore $g(\tilde{R}_a, R'_b, R'_c) = \{b\}$. Hence $g(\tilde{R}_a, \tilde{R}_b R'_c) = \{b\}$ for $\tilde{R}_b = (1,0)$ by SP. Then $g(\hat{R}_a, \tilde{R}_b, R'_c) \in \{b,c\}$ for $\hat{R}_a = (2,1,0)$ by SP and IR. However both $\{b\}$ and $\{c\}$ are impossible outcomes (suppose $g(\hat{R}_a, \tilde{R}_b, R'_c) = \{b\}$, then $g(\hat{R}_a, \hat{R}_b, R'_c) = \{b\}$ for $\hat{R}_b = (1,2,0)$. Then $|g| = 1 < |\phi|$).

Suppose $g(R') = \{a,c\}$. Using similar arguments as in the above, we get contradiction.

Case 2: $|g(R')| = 3$, $|\phi(R')| = 2$. Then we must have that $\phi(R') = \{a,b\}$ and either $2R_a3$ or $2R_b3$ or both.

Suppose $2R_a3$. First we have $g(R'_a, \tilde{R}_b, \tilde{R}_c) = \{a,b,c\}$ for $\tilde{R}_b = \tilde{R}_c = (3,2,1,0)$ by SP. Then $g(\hat{R}_a, \tilde{R}_b, \tilde{R}_c) = \{b,c\}$ for $\hat{R}_a = (2,1,0)$ by SP. Since $\phi(\tilde{R}_a, \tilde{R}_b, \bar{R}_c) = \{a,b,c\}$ for $\tilde{R}_a = (3,2,1,0)$ and $\bar{R}_c = (2,3,1,0)$, we must have $g(\tilde{R}_a, \tilde{R}_b, \bar{R}_c) = \{a,b,c\}$. Hence $g(\tilde{R}_a, \tilde{R}_b, \hat{R}_c) = \{a,b\}$ for $\hat{R}_c = (2,1,0)$ by SP. So we must have $g(\hat{R}_a, \tilde{R}_b, \hat{R}_c) = \{a,c\}$ by SP. Then we must have $g(\hat{R}_a, \hat{R}_b, \hat{R}_c) = \{a,c\}$ for $\hat{R}_b = (2,1,0)$. Since $g(\tilde{R}_a, \hat{R}_b, \hat{R}_c) = \{a,b\}$ by SP, we must have $c \notin g(\tilde{R}_a, \hat{R}_b, \tilde{R}_c)$ by SP; similarly since $g(\hat{R}_a, \hat{R}_b, \tilde{R}_c) = \{b,c\}$ by SP, we must have $a \notin g(\tilde{R}_a, \hat{R}_b, \tilde{R}_c)$ by SP. This is impossible since $|g(\tilde{R}_a, \hat{R}_b, \tilde{R}_c)| = 2$ by Eff and IR.

Suppose $2R_b3$. Using similar arguments as in the above, we get contradiction. $\square$

References

[1] Atila Abdulkadiroglu and Tayfun Sonmez, T.nmez. House Allocation with Existing Tenants. *Journal of Economic Theory*, 88(2):233–260, October 1999.

[2] M. Bianchi Albano G. L. and G. Spagnolo. Bid average methods in procurement. *Rivista di Politica Economica*, 96 (1):41–62, 2006.

[3] Jose Alcalde and Salvador Barbera. Top dominance and the possibility of strategy-proof stable solutions to matching problems. *Economic Theory*, 4(3):417–435, May 1994.

[4] Suryapratim Banerjee, Hideo Konishi, and Tayfun Sonmez. Core in a simple coalition formation game. *Social Choice and Welfare*, 18(1):135–153, January 2001.

[5] Salvador Barbera. An introduction to strategy-proof social choice functions. *Social Choice and Welfare*, 18(4):619–653, October 2001.

[6] Salvador Barbera, Faruk Gul, and Ennio Stacchetti. Generalized Median Voter Schemes and Committees. *Journal of Economic Theory*, 61(2):262–289, December 1993.

[7] Salvador Barbera and Matthew Jackson. A characterization of strategy-proof social choice functions for economies with pure public goods. *Social Choice and Welfare*, 11(3):241–252, July 1994.

[8] Salvador Barbera, Matthew O. Jackson, and Alejandro Neme. Strategy-Proof Allotment Rules. *Games and Economic Behavior*, 18(1):1–21, January 1997.

[9] Dirk Bergemann and Stephen Morris. *Robust Mechanism Design: The Role of Private Information and Higher Order Beliefs*. World Scientific, 2012.

[10] Sushil Bikhchandani and John G Riley. Equilibria in open common value auctions. *Journal of Economic Theory*, 53:101–130, 1991.

[11] Francis Bloch and Bhaskar Dutta. Formation of networks and coalitions. *Handbook of social economics. North Holland, Amsterdam*, 2011.

[12] Anna Bogomolnaia and Matthew O. Jackson. The Stability of Hedonic Coalition Structures. *Games and Economic Behavior*, 38(2):201–230, February 2002.

[13] Anna Bogomolnaia and Antonio Nicolo. Stable Assignment of Public Facilities under Congestion. *Journal of Public Economic Theory*, 7(1):65–91, February 2005.

[14] Jeremy Bulow, Ming Huang, and Paul Klemperer. Toeholds and takeovers. *Journal of Political Economy*, 107:427–454, 1999.

[15] Ruggiero Cavallo. Optimal decision-making with minimal waste: Strategyproof redistributionofVCGpayments. In *InProceedings of the International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS*, pages 882–889, 2006.

[16] Shurojit Chatterji, Remzi Sanver, and Arunava Sen. On domains that admit well-behaved strategy-proof social choice functions. *Journal of Economic Theory*, 148(3):1050–1073, May 2013.

[17] Shurojit Chatterji, Arunava Sen, and Huaxia Zeng. A Characterization of Single-Peaked Preferences via Random Social Choice Functions. 2014.

[18] Stephen Ching. An alternative characterization of the uniform rule. *Social Choice and Welfare*, 11(2):131–136, April 1994.

[19] Timothy G. Conley and Francesco Decarolis. Detecting bidders groups in collusive auctions. Technical report, Working Paper, University of Wisconsin, Madison. 16, 2011.

[20] Rajat Deb and Laura Razzolini. Auction-Like Mechanisms for Pricing Excludable Public Goods. *Journal of Economic Theory*, 88(2):340–368, October 1999.

[21] Francesco Decarolis. When the highest bidder loses the auction: theory and evidence from public procurement. Temi di discussione (Economic working papers) 717, Bank of Italy, Economic Research and International Relations Area, 2009.

[22] Gabrielle Demange and Myrna Wooders. *Group Formation in Economics: Networks, Clubs, and Coalitions*. Cambridge University Press, January 2005.

[23] Moritz Drexl and Andreas Kleiner. Optimal private good allocation: The case for a balanced budget. *Games and Economic Behavior*, 94:169–181, November 2015.

[24] Lars Ehlers and Bettina Klaus. Coalitional strategy-proof and resource-monotonic solutions for multiple assignment problems. *Social Choice and Welfare*, 21(2):265–280, October 2003.

[25] Andreas R. Engel, J.-J. Ganuza, Esther Hauk, and Achim Wambach. Managing risky bids. In *Handbook of Procurement*. Cambridge University Press, 2006.

[26] Boi Faltings. A Budget-Balanced, Incentive-Compatible Scheme for Social Choice. In Peyman Faratin and Juan A. RodrÃguez-Aguilar, editors, *Agent-Mediated Electronic Commerce VI. Theories for and Engineering of Distributed Mechanisms and Systems*, number 3435 in Lecture Notes in Computer Science, pages 30–43. Springer Berlin Heidelberg, January 2005.

[27] Jacob K. Goeree and Theo Offerman. Efficiency in auctions with private and common values: An experimental study. *The American Economic Review*, 92:625–643, 2002.

[28] Andrew V. Goldberg, Jason D. Hartline, Anna R. Karlin, Michael Saks, and Andrew Wright. Competitive auctions. *Games and Economic Behavior*, 55:242–269, 2006.

[29] Mingyu Guo and Vincent Conitzer. Worst-case optimal redistribution of VCG payments. In *Proceedings of the 8th ACM conference on Electronic commerce*, 2007.

[30] Mingyu Guo and Vincent Conitzer. Better redistribution with inefficient allocation in multi-unit auctions with unit demand. In *Proceedings of the 9th ACM conference on Electronic commerce*, EC '08, pages 210–219, New York, NY, USA, 2008. ACM.

[31] Mingyu Guo and Vincent Conitzer. Worst-case optimal redistribution of VCG payments in multi-unit auctions. *Games and Economic Behavior*, 67(1):69–98, September 2009.

[32] Mingyu Guo, Victor Naroditskiy, Vincent Conitzer, Amy Greenwald, and Nicholas R. Jennings. Budget-Balanced and Nearly Efficient Randomized Mechanisms: Public Goods and beyond. In Ning Chen, Edith Elkind, and Elias Koutsoupias, editors, *Internet and Network Economics*, number 7090 in Lecture Notes in Computer Science, pages 158–169. Springer Berlin Heidelberg, January 2011.

[33] Kathleen M Hagerty and William P Rogerson. Robust trading mechanisms. *Journal of Economic Theory*, 42(1):94–107, June 1987.

[34] Charles A. Holt and Roger Sherman. Risk aversion and the winner's curse. *Unpublished manuscript, April*, 2000.

[35] Matthew O. Jackson and Antonio Nicolo. The strategy-proof provision of public goods under congestion and crowding preferences. *Journal of Economic Theory*, 115(2):278–308, April 2004.

[36] Masso Jordi, Nicolo Antonio, Sen Arunava, Sharma Tridib, and Ulku Levent. On cost sharing in the provision of a binary and excludable public good. *Journal of Economic Theory*, 155:30–49, January 2015.

[37] Ruben Juarez. The worst absolute surplus loss in the problem of commons: random priority versus average cost. *Economic Theory*, 34(1):69–84, December 2006.

[38] John Henry Kagel and Dan Levin. *Common Value Auctions and the Winner's Curse*. Princeton University Press, Princeton, N.J, July 2002.

[39] Bettina Klaus, Hans Peters, and Ton Storcken. Strategy-proof division with single-peaked preferences and individual endowments. *Social Choice and Welfare*, 15(2):297–311, February 1998.

[40] Paul Klemperer. Auctions with almost common values: TheWallet game'and its applications. *European Economic Review*, 42:757–769, 1998.

[41] Elias Koutsoupias and Christos Papadimitriou. Worst-case equilibria. *Computer science review*, 3:65–79, 2009.

[42] Vijay Krishna. *Auction theory*. Academic press, 2009.

[43] Jin Li and Jingyi Xue. Egalitarian division under Leontief Preferences. *Economic Theory*, 54(3):597–622, October 2012.

[44] Jinpeng Ma. Strategy-proofness and the strict core in a market with indivisibilities. *International Journal of Game Theory*, 23(1):75–83, March 1994.

[45] Jordi Masso and Antonio Nicolo. Efficient and stable collective choices under gregarious preferences. *Games and Economic Behavior*, 64(2):591–611, November 2008.

[46] R. Preston McAfee, John McMillan, and Philip J. Reny. Extracting the surplus in the common-value auction. *Econometrica*, 57:1451–1459, 1989.

[47] H. Moulin. On strategy-proofness and single peakedness. *Public Choice*, 35(4):437–455, January 1980.

[48] Herve Moulin. *Axioms of Cooperative Decision Making*. Cambridge University Press, Cambridge, reprint edition edition, July 1991.

[49] Herve Moulin. Almost budget-balanced VCG mechanisms to assign multiple objects. *Journal of Economic Theory*, 144(1):96–119, January 2009.

[50] Herve Moulin. Auctioning or assigning an object: some remarkable VCG mechanisms. *Social Choice and Welfare*, 34:193–216, 2009.

[51] Herve Moulin. Auctioning or assigning an object: some remarkable VCG mechanisms. *Social Choice and Welfare*, 34(2):193–216, February 2010.

[52] Herve Moulin. One dimensional mechanism design. 2015.

[53] Herve Moulin and Scott Shenker. Strategyproof sharing of submodular costs:budget balance versus efficiency. *Economic Theory*, 18(3):511–533, November 2001.

[54] Roger B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6:58–73, 1981.

[55] Roger B. Myerson and Mark A. Satterthwaite. Efficient mechanisms for bilateral trading. *Journal of economic theory*, 29(2):265–281, 1983.

[56] Szilvia Papai. Strategyproof Assignment by Hierarchical Exchange. *Econometrica*, 68(6):1403–1433, November 2000.

[57] Szilvia Papai. Strategyproof and Nonbossy Multiple Assignments. *Journal of Public Economic Theory*, 3(3):257–271, July 2001.

[58] Ryan Porter, Yoav Shoham, and Moshe Tennenholtz. Fair imposition. *Journal of Economic Theory*, 118(2):209–228, October 2004.

[59] TCA Madhav Raghavan. Sufficient Conditions for Weak Group-Strategy-Proofness. 2014.

[60] Debraj Ray. *A Game-Theoretic Perspective on Coalition Formation*. Oxford University Press, November 2007.

[61] Tayfun Sonmez. Strategy-proofness and Essentially Single-valued Cores. *Econometrica*, 67(3):677–689, May 1999.

[62] Yves Sprumont. The Division Problem with Single-Peaked Preferences: A Characterization of the Uniform Allocation Rule. *Econometrica*, 59(2):509–519, March 1991.

[63] William Thomson. Non-bossiness. Technical report, University of Rochester-Center for Economic Research (RCER), 2014.

[64] Robert Wilson. A bidding model of perfect competition. *The Review of Economic Studies*, 44:511–518, 1977.