Incentives and Institutions: Essays in Mechanism Design and Game Theory with Applications

Thesis by

Serkan Küçükşenel

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2009

(Defended 05.15.2009)

To my parents, Olcay and Murat Küçükşenel,

and

to my fiancee, Meltem Küçükşahin

Acknowledgements

I am very grateful to my main advisor, John Ledyard, for being a role model with his passion for research. I have learned a great deal about auction and mechanism design from him. I thank him for his continuous support, funding, and encouragement. I am also indebted to the other members of my dissertation committee, Federico Echenique, Thomas Palfrey, and Kim Border, for their insightful discussions and assistance. I have benefited greatly from Federico's expertise in economic theory and his research guidance in matching theory. Tom has also been a wonderful advisor with his useful suggestions and comments.

I would like to extend my sincerest thanks to Mehmet Baç, Özgür Kıbrıs, Tayfun Sönmez, and Utku Ünver who made my admission to Caltech possible, hence this dissertation. I am grateful to my fellow graduate students, Alan, Christoph, Guido, Michael, Min, Morgan, Noah, and Sera, for being such good friends and colleagues. I also thank all HSS division officers and staff for always helping me out.

Finally, I wish to thank my family, Murat, Olcay Yeter, and Serpil, whose loving support from Turkey I constantly felt. Meltem, without your love and marvelous support, this dissertation would never have been completed.

All possible errors are my own responsibility.

Abstract

In the first part of this dissertation we study the problem of designing *desirable* mechanisms for economic environments with different types of informational and consumption externalities. We first study the mechanism design problem for the class of Bayesian environments where preferences of individuals depend not only on their allocations but also on the welfare of other individuals. For these environments, we fully characterize interim efficient mechanisms and examine their properties. This set of mechanisms is compelling, since interim efficient mechanisms are the best in the sense that there is no other mechanism which generates unanimous improvement. For public good environments, we show that these mechanisms produce public goods closer to the efficient level of production as the degree of altruism in the preferences increases. For private good environments, we show that altruistic agents trade more often than selfish agents.

We next consider a mechanism design problem for matching markets where externalities are present. We present mechanisms that implement the core correspondence of many-to-one matching markets, such as college admissions problems, where the students have preferences regarding the other students who would attend the same college. With an unrestricted domain of preferences the non-emptiness of the core is

not guaranteed. We present a sequential mechanism implementing the core without any restrictions on the preferences. We also show that simple two-stage mechanisms cannot be used to implement the core correspondence in subgame perfect Nash equilibrium without strong assumptions on agents' preferences.

In the final part of the dissertation we focus on another matching market, one-to-one assignment games with money. We present an alternative way to characterize the core as the fixed points of a certain mapping. We also introduce the first algorithm that finds all core outcomes in assignment games. The lattice property of the stable payoffs, as well as its non-emptiness, are proved using Tarski's fixed point theorem. We show that there is a polarization of interests in the core by using our formulation.

Contents

Acknowledgements **iv**

Abstract **v**

1 Introduction **1**

2 Behavioral Mechanism Design **5**

2.1 Introduction 6

2.2 The Model 9

2.3 Results 13

2.3.1 Preliminaries to the Main Results 14

2.3.2 Interim Efficient Mechanisms 16

2.3.3 Modified Virtual Valuations 22

2.3.4 A Sufficient Condition for Regularity 23

2.4 Applications 25

2.4.1 Public Goods 26

2.4.2 Bargaining: One Buyer and One Seller 31

2.5 Concluding Remarks 38

2.6 Appendix 40

3 Implementation of the Core in College Admissions Problems When Colleagues Matter **47**

3.1 Introduction 48

3.2 Preliminaries 50

3.2.1 The Model 50

3.2.2 Subgame Perfect Implementation 53

3.3 The Matching Mechanism 53

3.3.1 Implementing the Core 56

3.4 The Colleges Propose and Students Choose Mechanism 59

3.5 The Students Propose and Colleges Choose Mechanism 64

3.6 Final Remarks 67

4 A Theory of Stability in Assignment Games **69**

4.1 Introduction 70

4.2 Assignment Games with Money 72

4.3 Preliminary Definitions 75

4.4 The Core as a Set of Fixed Points 76

4.5 The Lattice Structure of Stable Payoffs 83

4.6 The Algorithm 85

4.7 Formulation with Core Outcomes Is Not Possible 91

4.8 Final Remarks 94

Bibliography **95**

Chapter 1

Introduction

Mechanism design is a general way of thinking about institutions. An institution or mechanism takes into it messages from agents and responds with an outcome. In this framework, a group of individuals must choose an alternative from the set of possible alternatives and must decide how to arrange monetary transfers. Initially, each agent obtains private information about each of the possible alternative outcomes. An agent's utility for a given alternative depends not only on her own material utility but also on the welfare of other agents. This implies agents are unselfish or altruistic. We first show how the existence of agents with this type of preference change the mechanism design problem. Then, we propose the set of mechanisms that we can observe in practice. These mechanisms correspond to the most efficient (or interim efficient) mechanisms within the mechanisms that satisfy incentive and feasibility constraints. If a mechanism is interim efficient, it can never be common knowledge that there is another mechanism which makes some types of agents better off without hurting other types of agents. This implies that any other mechanisms should not be observed in practice. We also provide applications for both public and private goods environments. In our applications, efficient decisions are independent of the social concerns of the agents. However, we show that interim efficient decisions depend on social concerns. In these mechanisms, inefficiencies in public good production decreases as the agents care more about the welfare of the other agents in the society. For bilateral trade environments, we show that altruistic agents trade more often than selfish agents. On the other hand, altruistic agents do not trade when it is not optimal to trade.

We next move to the sequential mechanism design problem for matching markets. We introduce simple sequential mechanisms that implement the core correspondence in college admissions problems when students do care about who else goes to the same college. This is an important problem because it provides the noncooperative foundations of the core and it might also help to design new institutions for the real-life college admissions problem. In this matching market, there are two finite disjoint sets of agents, the set of colleges and the set of students. Each college has a preference relation over groups of students. Each student has a preferences relation over colleges and groups of students (or classmates). A matching will be a particular assignment of students to colleges. The solution concept, core, specifies the set of matchings we might observe in practice. In college admissions problems, it has been shown that two-stage simple mechanisms, such as the "students propose and colleges choose" mechanism and the "colleges propose and students choose" mechanism, implement the core. We show that only extension of the colleges propose and students choose mechanism can implement the core under the restrictions that guarantee the nonemptiness of the core. Therefore, the symmetry between these two mechanisms does not hold when students also care about their classmates. We also provide a multi-stage mechanism that implements the core without any restrictions on the preferences. This chapter also shows that we should take into consideration whether students care about their classmates or not when designing institutions for the real-life matching problem.

We analyze another matching market, a one-to-one assignment game with money, in the final chapter. An assignment game is a two-sided matching market with monetary transfers. In this market, there are two exogenously specified disjoint sets of agents, say firms and workers. The agents engage in bilateral transactions (if worker i works for firm j then firm j employs worker i) and make monetary transfers. Each firm can employ no more than one worker and each worker can not work for more than one firm. A natural solution concept for such markets is the core. The core outcomes specify which partnership we can expect to observe and how the agents will divide their gains. The assignment game has traditionally been studied in terms of its linear programming formulations. We propose an alternative way to characterize the core of assignment games as the fixed points of a certain mapping. Moreover, our characterization gives an algorithm for finding the core outcomes. The characterization is useful because it allows us to construct a simple algorithm to find all core outcomes and it provides a very simple proof for the lattice structure of the core and for the polarization of interests in the core.

Chapter 2

Behavioral Mechanism Design

2.1 Introduction

This paper studies the problem of designing mechanisms for the class of Bayesian environments with interdependent preferences. A group of individuals must choose an alternative from the set of possible alternatives and must decide how to arrange monetary transfers. Initially, each agent obtains private information about each possible alternative. An agent's utility for a given alternative depends not only on her own material utility but also on the welfare of other agents. This implies agents are unselfish or altruistic. In this framework, we characterize the most efficient mechanisms within the mechanisms that satisfy incentive and feasibility constraints.

The assumption of self interest is problematic. Self-interest hypothesis states that preferences among allocations depend only on an agent's own material well being. Experimental results suggest that people often do care for well being of others and have other preferences. For example, there is more contribution to public goods than purely selfish maximization can lead us to expect. Moreover, people should not vote in elections or contribute to public television if they are purely self interested. See Ledyard (1995) for a survey on public goods which documents these and several other anomalies. Similar anomalous results are also observed in private goods environments. For example, in an ultimatum game one player has a strictly dominant strategy if the player is self-interested but he or she does not choose this selfish strategy. See also Fehr and Schmidt (2006) for more experimental evidence on unselfish preferences. Given these observations, I ask a basic question: How does the existence of agents exhibiting interdependent preferences change the mechanism design problem?

There exists an extensive literature on mechanism design. We refer the reader to Jackson (2003) for a survey on mechanism design literature. In those studies the main focus is on either (the impossibility of) efficient or optimal mechanism design with selfish agents. In contrast to previous literature, we are interested in characterizing interim efficient mechanisms with unselfish agents. Interim is used to denote the informational time frame: agents select their messages after receiving their signals but before learning the signals of others. We assume that all decisions, including whether to change the mechanism, are made at the interim stage. Interim efficiency is a natural extension of efficiency to incomplete information environments. If a mechanism is interim efficient, then it can never be common knowledge that there is another feasible mechanism which makes some types of agents better off without hurting other types of agents. This implies that any other mechanism would be unanimously rejected by all agents and should thus not be observed in practice. We show that these mechanisms correspond to decision rules based on *modified* virtual cost-benefit criterion, together with the appropriate incentive taxes. Moreover, we show that interim efficient decisions depend on the social concerns of the agents even though classical efficient decisions do not depend on the social concerns of the agents. There are a few papers that explore the properties of interim efficient allocation rules for standard mechanism design environments with selfish agents. See Wilson (1985) and Gresik (1991) for a characterization of ex ante efficient mechanisms for bilateral trade environments (double auctions), and Ledyard and Palfrey (2007) for a characterization of interim efficient mechanisms for public good environments.

Our characterization is general and can be applied to different economic settings. We provide applications for both public and private goods environments. In our applications, efficient decisions are independent of the social concerns of the agents. However, we show that interim efficient mechanisms produce public goods more often as the degree of altruism in the preferences goes up. That is, inefficiencies in public good production decreases as the agents care more about the welfare of the other agents in the society. For bilateral trade environments, we show that these mechanisms give the good to the agent with the highest positive *modified* virtual valuations. If there is no buyer with a positive modified virtual valuation which is higher than the virtual valuation of the seller, then the seller keeps the good. We also show that altruistic agents trade more often than selfish agents. This means that there are some information states of the economy where it is optimal to trade but selfish agents will not trade and altruistic agents will trade. Moreover, altruistic agents do not trade when it is not optimal to trade.

The remainder of the paper is organized as follows. In the next section, we describe the environment and introduce the basic notation. In Section 2.3 we formulate the set of constraints and provide necessary and sufficient conditions for incentive compatibility and individual rationality. The tools of mechanism design are used to provide these necessary and sufficient conditions. Then, we present the characterization results and proofs. Section 2.4 provides applications of our characterization for both public good and bilateral trade environments. Finally, we summarize the findings of the paper and make some concluding remarks in Section 2.5. The proofs are delegated to the Appendix, Section 2.6.

2.2 The Model

Consider a Bayesian mechanism design framework with n agents. The set of agents is denoted by $N = \{1, ..., n\}$. Each agent has a type θ^i which is her private information. We assume that each agent knows her own type and does not know the types of the other agents. Each θ^i is independently drawn from cumulative distribution function $F^i(.)$ on $\Theta^i = [\underline{\theta}^i, \overline{\theta}^i]$ with $0 \leq \underline{\theta}^i \leq \overline{\theta}^i < \infty$. Types are drawn independently across agents; that is, the θ^is are independent random variables. We denote a generic profile of agent types by $\theta = (\theta^1, ..., \theta^n) \in \Theta \equiv \times_{i=1}^{N} \Theta^i$. For any $\theta \in \Theta$, we adopt the standard notation so that $\theta^{-i} = (\theta^1, ..., \theta^{i-1}, \theta^{i+1}, ..., \theta^n)$, and $\theta = (\theta^i, \theta^{-i})$ where $f(\theta) = \prod_{i=1}^{N} f^i(\theta^i)$. Let X be a finite set of possible nonmonetary decisions, or allocations (e.g., X could be a subset of an Euclidean space and represent the set of possible allocations of private and public goods).

Let $\Delta(X)$ be the set of probability distributions on X. A mechanism $\zeta = (y, t)$ consists of an allocation rule y and a payment rule t. Let $y^x(\theta)$ denote the probability of choosing $x \in X$, given the profile of types $\theta \in \Theta$. A feasible allocation rule (or social choice function) $y : \Theta \to \Delta(X)$ is a function from agents' reported types to a probability distribution over allocations such that $\sum_{x \in X} y^x(\theta) = 1$ and $y^x(\theta) \geq 0$ for all $\theta \in \Theta$. We allow allocation rules to randomize over feasible allocations. Let Y be the set of all possible allocation rules and $\Omega \subseteq Y$ be the set of all feasible allocation rules. The payment rule $t : \Theta \to \mathbb{R}^N$ is a map from the agents' reported types to monetary compensations where $\sum_{i=1}^{N} t^i(\theta) \geq 0$. This condition (ex-post budget balance) requires that there is no outside source to finance the compensations. Therefore, a mechanism cannot run a deficit.

The individual payoff function (or material utility) of an agent i given an allocation rule y, and her monetary payment t^i is

$$\Pi^i(y, t^i, \theta^i) = \sum_{x \in X} y^x(\theta) v^i(x, \theta^i) - t^i$$

where $v^i(x, \theta^i)$ is agent i's valuation of allocation x which depends on her private information. We assume that $v^i(x, \theta^i)$ is differentiable, monotone increasing, and convex in θ^i for all i and $x \in X$.

Beyond her individual payoff, agent i cares about the payoffs of others:

$$\begin{aligned} u^i(y, t, \theta) &= \rho^i \Pi^i + (1 - \rho^i)\overline{\Pi} \\ &= \sum_{x \in X} y^x(\theta) V^i(x, \theta, \rho^i) + \rho^i(\frac{\sum_{j \in N} t^j}{N} - t^i) - \frac{\sum_{j \in N} t^j}{N} \end{aligned}$$

where $\overline{\Pi} = \frac{\sum_{j \in N} \Pi^j}{N}$ is the average payoff in the population and

$$V^i(x, \theta, \rho^i) = \rho^i v^i(x, \theta^i) + (1 - \rho^i)\frac{\sum_{j \in N} v^j(x, \theta^j)}{N}$$

is the total value of allocation x for agent i . The constant $\rho^i \in [0, 1]$ is an agent-specific weighting factor showing each agent's social concerns. If $\rho^i = 1$, the agent has selfish preferences which do not directly depend on the well being of others. If

$\rho^i < 1$, the agent has altruistic preferences which are increasing in the well being of others. Note that as ρ^i increases the degree of altruism in preferences goes down. If all agents are identical in their social concerns ($\rho^i = \rho^j = \rho$), and $\rho = 0$, the model is a common value setting where full social preferences are in action and the society is homogeneous. If $\rho = 1$, the model is equivalent to the standard mechanism design environment with selfish agents. We assume that agents have identical social concerns to simplify the analysis for the rest of the paper ($\rho^i = \rho^j = \rho$ for all $i, j \in N$). The model can also be extended to environments where agents are spiteful ($\rho > 1$).

We only consider direct mechanisms in which the set of reported types is equal to the set of possible types in the rest of the paper. By the revelation principle, any allocation rule that results from equilibrium in any mechanism is also an equilibrium allocation rule of an incentive compatible, direct mechanism. Therefore, there is no loss of generality in restricting our attention to these simple type of mechanisms.

Let $U^i(\zeta, \theta^i, s^i)$ be the interim expected utility of agent i when he reports $s^i \neq \theta^i$, assuming all other agents truthfully report their type. That is

$$U^i(\zeta, \theta^i, s^i) = E_{\theta^{-i}}[u^i(y(s^i, \theta^{-i}), t(s^i, \theta^{-i}), \theta)].$$

Denote $U^i(\zeta, \theta^i) \equiv U^i(\zeta, \theta^i, \theta^i)$. The ex-ante utility of agent i is

$$U^i(\zeta) = E_\theta[u^i(y(\theta), t(\theta), \theta)].$$

Define also the conditional expected payment function $a^i : \Theta^i \to \mathbb{R}$ such that

$$a^i(\theta^i) = E_{\theta^{-i}}[t^i(\theta)].$$

A mechanism is interim incentive compatible (IIC) if honest reporting of types defines a Bayesian-Nash equilibrium. That is ζ is IIC if and only if $U^i(\zeta, \theta^i) \geq U^i(\zeta, \theta^i, s^i)$ for all i, s^i, θ^i. We call a mechanism interim individual rational (IIR) if every agent wants to participate in the mechanism: $U^i(\zeta, \theta^i) \geq 0$ for all i, θ^i. A mechanism is ex ante budget balanced (EABB) if a mechanism designer does not expect to pay subsidies to the agents, e.g., $E_\theta(\sum_{i=1}^N t^i(\theta)) \geq 0$. A mechanism is feasible if it satisfies IIC, IIR, and EABB.

A mechanism ζ is interim efficient (IE) if it is feasible and there is no other feasible mechanism $\widehat{\zeta}$ such that $U^i(\widehat{\zeta}, \theta^i) \geq U^i(\zeta, \theta^i)$ for all i, θ^i and $U^i(\widehat{\zeta}, \theta^i) > U^i(\zeta, \theta^i)$ for some i and for all $\theta^i \in \widetilde{\Theta}^i \subset \Theta^i$, where $\widetilde{\Theta}^i$ has strictly positive measure relative to Θ^i. IE is an extension of efficiency to the environments with private information. A mechanism is IE if there does not exist an alternative feasible mechanism that interim Pareto-dominates it. Note that the idea of Pareto-domination is applied to the expected utilities after the agents have learned their types. IE mechanisms can also be represented as the solutions to a set of maximization problems. A mechanism ζ is an IE mechanism if and only if there exists $\lambda = \{\lambda^i : \Theta^i \to \mathbb{R}^+\}_{i=1}^N$ with $\int_{\underline{\theta}^i}^{\overline{\theta}^i} \lambda^i(\theta^i) dF^i(\theta^i) > 0$ for some i, such that ζ maximizes $\sum_{i=1}^N \int_{\underline{\theta}^i}^{\overline{\theta}^i} \lambda^i(\theta^i) U^i(\zeta, \theta^i) dF^i(\theta^i)$ subject to ζ is feasible. Note that the weight attached to an agent i can vary with her type. See Holmstrom and Myerson (1983) for more on this. Thus, an IE mechanism maximizes weighted sum of agents' utilities subject to IIC, IIR, and EABB constraints. We could also use ex post budget balance condition. It turns out that EABB is equivalent to ex post budget balance condition in our setting.

2.3 Results

Given welfare weights $\lambda > 0$, our main problem can now be stated as finding mechanisms that maximize

$$\sum_{i=1}^{N} \int_{\underline{\theta}^i}^{\overline{\theta}^i} \lambda^i(\theta^i) U^i(\zeta, \theta^i) dF^i(\theta^i)$$

subject to IIC, IIR, EABB, and obvious quantity constraints.

We now proceed to characterize the complete set of interim efficient mechanisms. We first start to reformulate the constraint set such that we can provide necessary and sufficient conditions for IIC and IIR. The second step in the characterization involves a general solution to the maximization problem with the constraints rewritten as described below. The constraints for IIC correspond to the first-order and second-order conditions of an individual optimization problem. Following the same idea in Myerson (1981), I find the solution to the case where the second-order IIC condition is not binding (regular problems). Then, I provide a sufficient condition in which the solution to the regular problem coincides with the solution to the original problem. The standard tools of mechanism design are used to get the following preliminary results.

2.3.1 Preliminaries to the Main Results

IIC requires that it is a Bayesian equilibrium for each agent to report her type truthfully, i.e., none of the agents can obtain strictly higher payoffs by deviating individually. Let $S^i : \Theta^i \rightarrow \mathbb{R}$ be agent i's expected surplus function. Then given a mechanism ζ surplus function is

$$S^i(\theta^i) := sup\{U^i(\zeta, \theta^i, s^i) | s^i \in \Theta^i\}.$$

This optimization problem determines the agent i's optimal report. It is easy to see that a mechanism is IIC if and only if $S^i(\theta^i) = U^i(\zeta, \theta^i)$ for all i, θ^i. In our framework incentive compatibility can also be characterized by means of an envelope and a monotonicity condition as in standard mechanism design problems. The following result, which is useful in our characterization, states that the derivative of the expected marginal total value of type θ^i under the mechanism should be nondecreasing and the expected utility function of that type is uniquely determined by the expected utility of the lowest type and the allocation rule. The proof is similar to the selfish preferences environment. A similar result is also proved by Rochet (1987) for linear environments with selfish preferences.

Lemma 1 *A mechanism is IIC if and only if*

$$(s^i - \theta^i) \times (Q^i(s^i, \rho) - Q^i(\theta^i, \rho)) \geq 0 \quad for\ all\ s^i, \theta^i \in \Theta^i \tag{2.1}$$

$$U^i(\zeta,\theta^i) = U^i(\zeta,\underline{\theta}^i) + \int_{\underline{\theta}^i}^{\theta^i} Q^i(s,\rho)ds \tag{2.2}$$

where

$$Q^i(\theta^i,\rho) \equiv \int_{\Theta^{-i}} \sum_{x\in X} y^x(\theta)\frac{\partial V^i(x,\theta,\rho)}{\partial\theta^i} dF^{-i}(\theta^{-i}).$$

The first condition is the monotonicity condition which states that the expected marginal total value of agent i in her own type, $Q^i(\theta^i,\rho)$, should be monotone increasing in her own private information. This implies that $\frac{\partial Q^i(\theta^i,\rho)}{\partial\theta^i} \geq 0$. The second condition is the envelope condition. The monotonicity condition has implications only for allocation rules. Notice that the expected payment function a^i is completely determined by a constant $a^i(\underline{\theta}^i)$ and the allocation rule y. The constant of integration, $U^i(\zeta,\underline{\theta}^i)$ is uniquely determined by N constants $a(\underline{\theta})$ and y for all agents.

Now we can write expected budget surplus in an IIC mechanism using the result above.

$$B(\zeta) \equiv \sum_{i=1}^{N}\int_{\Theta} t^i(\theta)dF(\theta) = \sum_{i=1}^{N}\int_{\Theta}(\rho t^i(\theta) + (1-\rho)\frac{\sum_j t^j(\theta)}{N})dF(\theta)$$

$$= \sum_{i=1}^{N}\left(\int_{\Theta}\sum_{x\in X} y^x(\theta)V^i(x,\theta,\rho)dF(\theta) - U^i(\zeta,\underline{\theta}^i) - \int_{\Theta^i}[\int_{\underline{\theta}^i}^{\theta^i} Q^i(s,\rho)ds]dF^i(\theta^i)\right)$$

Using integration by parts,

$$= \sum_{i=1}^{N}\int_{\Theta}\sum_{x\in X} y^x(\theta)\left(V^i(x,\theta,\rho) - \frac{\partial V^i(x,\theta,\rho)}{\partial\theta^i}\frac{1-F^i(\theta^i)}{f^i(\theta^i)}\right)dF(\theta) - \sum_{i=1}^{N} U^i(\zeta,\underline{\theta}^i).$$

Let $\Phi(\zeta,\rho) \equiv \sum_{i=1}^{N} \int_{\Theta} \sum_{x\in X} y^x(\theta) \left(V^i(x,\theta,\rho) - \frac{\partial V^i(x,\theta,\rho)}{\partial \theta^i} \frac{1-F^i(\theta^i)}{f^i(\theta^i)} \right) dF(\theta)$. Notice that if ζ is EABB then $B(\zeta) \geq 0$. This also implies the mechanism designer does not expect to pay subsidies to the agents.

IIR requires that each type of each agent must be at least as well off by participating as they would be by not participating at the interim stage. We assume that outside options are exogenously given and without loss of generality normalized to zero. We next combine IIR and IIC to get a useful result for later.

Lemma 2 *An IIC mechanism ζ is IIR if and only if for all $i \in N$, $U^i(\zeta, \underline{\theta}^i) \geq 0$.*

2.3.2 Interim Efficient Mechanisms

Welfare weights play an important role in my analysis. Before stating the main characterization, the following definition and lemma will be useful in reformulating the original problem.

Definition 1 *If $\lambda^{0i} \equiv \int_{\underline{\theta}^i}^{\overline{\theta}^i} \lambda^i(\theta^i) dF^i(\theta^i) > 0$, let $\Lambda^i(\theta^i) = \frac{1}{\lambda^{0i}} \int_{\underline{\theta}^i}^{\theta^i} \lambda^i(s) dF^i(s)$. If $\lambda^{0i} = 0$, let $\Lambda^i(\theta^i) = 0$.*

λ^{0i} is agent i's ex ante welfare weight relative to other agents. $\Lambda^i(\theta^i)$ is a relative weight of agent i's lower types given her private information.

Lemma 3

$$\int_{\underline{\theta}^i}^{\overline{\theta}^i} \lambda^i(\theta^i) \left(U^i(\zeta, \underline{\theta}^i) + \int_{\underline{\theta}^i}^{\theta^i} Q^i(s,\rho) ds \right) dF^i(\theta^i)$$

$$=$$

$$\lambda^{0i}\left[U^i(\zeta,\underline{\theta}^i)+\int_{\underline{\theta}^i}^{\overline{\theta}^i}(\frac{1-\Lambda^i(\theta^i)}{f^i(\theta^i)})Q^i(\theta^i,\rho)dF^i(\theta^i)\right].$$

Now we can provide our first characterization by using the previous lemmas. This result implies that the objective function is just a function of utilities of the lowest types $U^i(\zeta,\underline{\theta}^i)$ and the allocation rule y. It does not depend on the transfers anymore.

Theorem 1 *A mechanism $\zeta=(y,a)$ is IE if and only if there exists non-negative type-dependent welfare weights, $\{\lambda^i\}_{i=1}^N$, where $\sum_{i\in N}\lambda^{0i}>0$, and N constants, $\{c^i(\underline{\theta}^i)\}_{i=1}^N$, such that $(y,\{c^i(\underline{\theta}^i)\}_{i=1}^N)$ solves,*

$$max_{y\in\Omega}\sum_{i=1}^N\lambda^{0i}\left[U^i(\zeta,\underline{\theta}^i)+\int_{\underline{\theta}^i}^{\overline{\theta}^i}(\frac{1-\Lambda^i(\theta^i)}{f^i(\theta^i)})Q^i(\theta^i,\rho)dF^i(\theta^i)\right] \tag{2.3}$$

subject to

$$\Phi(\zeta,\rho)-\sum_{i=1}^N U^i(\zeta,\underline{\theta}^i)\geq 0 \tag{2.4}$$

$$U^i(\zeta,\underline{\theta}^i)=\int_{\Theta^{-i}}\sum_{x\in X}y^x(\underline{\theta}^i,\theta^{-i})V^i(x,\underline{\theta}^i,\theta^{-i},\rho)dF^{-i}(\theta^{-i})-c^i(\underline{\theta}^i)\geq 0 \tag{2.5}$$

$$Q^i(\theta^i,\rho)\ \ monotone\ increasing\ for\ all\ i,\theta^i. \tag{2.6}$$

Following the same idea in Myerson (1981) we characterize the solution to the problem in Theorem 1 for the case where monotonicity constraint is not binding. In this case solution can be obtained by pointwise maximizing the integrand in the objective function (2.3). Then we provide conditions under which the solutions to this reduced problem satisfies the monotonicity constraint. When solutions to the original problem and the reduced problem coincide, we refer to the problem as regular.

Given non-negative type-dependent welfare weights, $\{\lambda^i\}_{i=1}^N$, we can define the Lagrangian function as

$$\mathcal{L}(y,(c^i(\underline{\theta}^i))_{i=1}^N,\gamma,(\mu^i)_{i=1}^N)=\sum_{i=1}^N\lambda^{0i}\left[U^i(\zeta,\underline{\theta}^i)+\int_{\underline{\theta}^i}^{\overline{\theta}^i}(\frac{1-\Lambda^i(\theta^i)}{f^i(\theta^i)})Q^i(\theta^i,\rho)dF^i(\theta^i)\right]$$

$$+\gamma\left[\Phi(\zeta,\rho)-\sum_{i=1}^N U^i(\zeta,\underline{\theta}^i)\right]$$

$$+\sum_{i=1}^N\mu^i\left(\int_{\Theta^{-i}}\sum_{x\in X}y^x(\underline{\theta}^i,\theta^{-i})V^i(x,\underline{\theta}^i,\theta^{-i},\rho)dF^{-i}(\theta^{-i})-c^i(\underline{\theta}^i)\right).$$

The first-order conditions with respect to γ (EABB multiplier) and with respect to μ^i (IIR multiplier) imply that

$$\gamma\geq 0, B(\zeta)\geq 0\ \ and\ \ \gamma B(\zeta)=0,$$

$$\mu^i\geq 0, U^i(\zeta,\underline{\theta}^i)\geq 0\ \ and\ \ \mu^i U^i(\zeta,\underline{\theta}^i)=0\ \ for\ all\ i\in N.$$

The first-order condition with respect to $a^i(\underline{\theta}^i)$ yields $-\lambda^{0i}+\gamma-\mu^i=0$. Then $\gamma\geq\lambda^{0i}$ for all $i\in N$. This implies the EABB constraint is always binding ($\gamma>0$) since there is $i\in N$ such that $\lambda^{0i}>0$ and $\mu^i\succeq 0$ for all $i\in N$. The intuition of this result is the following. We assumed that contributions in excess are not socially valued. If the EABB is not binding, a redistribution of budget surplus to the agents would result in an interim Pareto improvement. If we assume that there is a seller (or collector) who keeps the excess surplus, then excess contributions are not lost for everybody. This implies the constraint might not be binding depending on the welfare weight of the seller.

Let $\overline{\lambda} = max_{i \in N}\{\lambda^{0i}\}$. Define also $K = \{k \mid \overline{\lambda} = \lambda^{0k},\ \forall\, k \in N\}$, the set of agents who have the highest ex ante welfare weight, and $M = \{m \mid \overline{\lambda} > \lambda^{0m},\ \forall\, m \in N\}$, the set of agents whose welfare weights are lower than the highest ex ante welfare weight, where $N = K \cup M$. There are two possible cases:

Case 1: $\gamma > \overline{\lambda}$. This implies for all $i \in N$, $\mu^i > 0 \Rightarrow U^i(\zeta, \underline{\theta}^i) = 0 \Rightarrow$ IIR constraints are binding for all agents' lowest types.

Case 2: $\gamma = \overline{\lambda}$. This implies for each $k \in K$, $\gamma = \overline{\lambda} = \lambda^{0k} \Rightarrow \mu^k = 0 \Rightarrow U^k(\zeta, \underline{\theta}^k) \geq 0$ and for each $m \in M$, $\gamma = \overline{\lambda} > \lambda^{0m} \Rightarrow \mu^m > 0 \Rightarrow U^m(\zeta, \underline{\theta}^m) = 0 \Rightarrow$ IIR constraints are binding for all agents' lowest types in M and the constraints are not binding for all agents in K.

If IIR constraints for the lowest types of agents with non maximal expected welfare weight is not binding, redistribution of wealth to agents with maximal expected welfare weight would increase the weighted welfare function. The following lemma summarizes the discussion above.

Lemma 4

$$\begin{aligned}\sum_{i=1}^{N} \lambda^{0i} U^i(\zeta, \underline{\theta}^i) &= \gamma \sum_{i=1}^{N} U^i(\zeta, \underline{\theta}^i) \\ &= \gamma \Phi(\zeta, \rho).\end{aligned}$$

Using the above result, the objective function (2.3) can be written as follows

$$\sum_{i=1}^{N}\left[\int_{\Theta}\sum_{x\in X} y^{x}(\theta)\left(V^{i}(x,\theta,\rho)-\frac{\partial V^{i}(x,\theta,\rho)}{\partial\theta^{i}}\frac{1-F^{i}(\theta^{i})}{f^{i}(\theta^{i})}\right)dF(\theta)\right.$$
$$\left.+\frac{\lambda^{0i}}{\gamma}\int_{\Theta^{i}}(\frac{1-\Lambda^{i}(\theta^{i})}{f^{i}(\theta^{i})})Q^{i}(\theta^{i},\rho)dF^{i}(\theta^{i})\right]=$$

$$\sum_{i=1}^{N}\int_{\Theta}\sum_{x\in X} y^{x}(\theta)\left[V^{i}(x,\theta,\rho)+\frac{\partial V^{i}(x,\theta,\rho)}{\partial\theta^{i}}\left(\frac{F^{i}(\theta^{i})-1}{f^{i}(\theta^{i})}+\frac{\lambda^{0i}}{\gamma}\frac{1-\Lambda^{i}(\theta^{i})}{f^{i}(\theta^{i})}\right)\right]dF(\theta).$$

Suppose we are in Case 1. In this case both constraints are binding. This implies,

$$\Phi(\zeta,\rho)=0$$

from EABB and

$$\int_{\Theta^{-i}}\sum_{x\in X} y^{x}(\underline{\theta}^{i},\theta^{-i})V^{i}(x,\underline{\theta}^{i},\theta^{-i},\rho)dF^{-i}(\theta^{-i})=c^{i}(\underline{\theta}^{i})=\rho a^{i}(\underline{\theta}^{i})$$

from IIR. Therefore, we can uniquely solve for the set of expected payments of all agents' minimum types.

Suppose now we are in Case 2. The argument for each $i\in M$ is similar to Case 1. On the other hand, for each $i\in K$ the IIR constraint may not be binding. Therefore, we need $|K|$ constants to solve for the payment function. Note that the agents with the maximal expected welfare weight share the remaining surplus (or cost) to make the EABB constraint binding. This implies agents in set K will be residual claimants.

When we combine both cases and rearrange terms, we get the following result.[1] This result is simplified reformulation of Theorem 1 for regular problems. The utilities of lowest types $U^i(\zeta, \underline{\theta}^i)$ are explicitly entered into the objective function and the constraints of the maximization problem reduce to two constraints representing the EABB constraint for each of the two cases as discussed above.

Theorem 2 *Suppose the type-dependent welfare weights are such that* $\lambda^{0i} = \overline{\lambda} > \lambda^{0j}$ *for all* $j \in N \backslash \{i\}$ *and* $\gamma \geq \overline{\lambda}$. *Then, for regular problems, a mechanism* ζ *is IE if and only if* $y \in \Delta(X)$ *simultaneously solves the following inequalities*

$$\max_{y \in \Omega} \sum_{i=1}^{N} \int_{\Theta} \sum_{x \in X} y^x(\theta) \Big[V^i(x,\theta,\rho) + \frac{\partial V^i(x,\theta,\rho)}{\partial \theta^i} \left(\frac{F^i(\theta^i) - 1}{f^i(\theta^i)} + \frac{\lambda^{0i}}{\gamma} \frac{1 - \Lambda^i(\theta^i)}{f^i(\theta^i)} \right) \Big] dF \tag{2.7}$$

$$0 \leq \Phi(\zeta, \rho) \tag{2.8}$$

$$0 = (\gamma - \overline{\lambda}) \Phi(\zeta, \rho). \tag{2.9}$$

The payment function (if $\rho \neq 0$*) is given by*[2]*:*

$$\forall\ j \neq i, \quad a^j(\theta^j) = \frac{\int_{\Theta^{-j}} \sum_{x \in X} y^x(\theta) V^j(x,\theta,\rho) dF^{-j}(\theta^{-j}) - \int_{\underline{\theta}^j}^{\theta^j} Q^j(s,\rho) ds}{\rho}, \tag{2.10}$$

$$a^i(\theta^i) = \frac{\int_{\Theta^{-i}} \sum_{x \in X} y^x(\theta) V^i(x,\theta,\rho) dF^{-i}(\theta^{-i}) - \Phi(\zeta,\rho) - \int_{\underline{\theta}^i}^{\theta^i} Q^i(s,\rho) ds}{\rho}. \tag{2.11}$$

[1] For this result, we assume that $|K| = 1$. It is easy to generalize this result to the cases where $|K| \neq 1$.

[2] If $\rho = 0$, any payment scheme that adds up to zero will work since in this case agents only care about the average payment and we know that budget is always balanced.

2.3.3 Modified Virtual Valuations

In this section we show the effects of interdependent preferences in our setting. Let

$$W^i(x,\theta,\rho,\lambda^i) = V^i(x,\theta,\rho) + \frac{\partial V^i(x,\theta,\rho)}{\partial\theta^i}\left(\frac{F^i(\theta^i)-1}{f^i(\theta^i)} + \frac{\lambda^{0i}}{\gamma}\frac{1-\Lambda^i(\theta^i)}{f^i(\theta^i)}\right). \qquad (2.12)$$

We call $W^i(x,\theta,\rho,\lambda^i)$ as the (modified) virtual valuation of agent i for allocation x following Myerson (1981). Rather than directly working with total valuations, interim efficient mechanisms use the agents' total valuations suitably adjusted. The virtual valuation for a given allocation is equal to the agent's total valuation for the allocation, $V^i(x,\theta,\rho)$, with two adjustments that depend on the distribution of types, welfare weights, and social concerns of the agents. The first one, $\frac{\partial V^i(x,\theta,\rho)}{\partial\theta^i}\frac{F^i(\theta^i)-1}{f^i(\theta^i)}$, is due to the informational rent to be given for truthful revelation of the agent's private information. The second one is due to distortions arising from redistribution of income $(\frac{\partial V^i(x,\theta,\rho)}{\partial\theta^i}\frac{\lambda^{0i}}{\gamma}\frac{1-\Lambda^i(\theta^i)}{f^i(\theta^i)})$. Note that these adjustments are weighted with the marginal total valuation of agent i for a given allocation and hence virtual valuations also depend on the allocation.

The modified virtual valuations reduce to those given in Ledyard and Palfrey (2007) if $\rho = 1$ and $v^i(x^i,\theta) = x^i\theta^i$, which implies that valuations are independent. For that case these authors show that

$$W^i(\theta^i,\lambda^i) = \theta^i + \frac{F^i(\theta^i)-1}{f^i(\theta^i)} + \frac{\lambda^{0i}}{\gamma}\frac{1-\Lambda^i(\theta^i)}{f^i(\theta^i)}.$$

If valuations are linear in the allocation and $\rho = 1$, then virtual valuations are also linear in x^i and agents are selfish. Then we can redefine modified virtual valuations such that they are independent of the allocation.

2.3.4 A Sufficient Condition for Regularity

In this section we provide sufficient conditions under which the solution to the regular problem coincide with the solution to the original problem in Theorem 1.

Substituting (2.12) into (2.7) gives us:

$$\max_{y \in F} \sum_{i=1}^{N} \int_{\Theta} \sum_{x \in X} y^x(\theta) W^i(x, \theta, \rho, \lambda^i) dF(\theta). \tag{2.13}$$

The regular problem to find IE mechanisms can now be stated as (y, a) is an interim efficient mechanism if and only if the allocation rule $y \in \Delta(X)$ simultaneously solves (2.13), (2.8), and (2.9), and the payment function a is given by (2.10) and (2.11).

The problem stated above has a simple solution defined by[3]

$$y^x(\theta, \lambda) = \begin{cases} 1 & \text{if } x = argmax_{m \in X} \sum_{i=1}^{N} W^i(m, \theta, \rho, \lambda^i) \\ 0 & \text{otherwise.} \end{cases} \tag{2.14}$$

This implies an IE mechanism assigns probability one to an allocation with the highest sum of modified virtual valuations. Note that to find the interim efficient mechanism we use the minimum possible $\gamma \geq \bar{\lambda}$ such that (2.8) and (2.9) are satisfied.

This solution also provides an algorithm to find the interim efficient mechanisms.

[3]For simplicity, we assume that there are no allocations $x, y, x \neq y$ such that $\sum_{i=1}^{N} W^i(x, \theta, \rho, \lambda^i) = \sum_{i=1}^{N} W^i(y, \theta, \rho, \lambda^i)$. We can also use a random tie-breaking rule.

Firstly, given welfare weights, set $\gamma = \bar{\lambda}$ and find the allocation rule $y^x(\theta, \lambda)$ for each $\theta \in \Theta$. If this solution satisfies (2.8) and (2.9) then the expected transfer functions $a(\theta)$ are given by (2.10) and (2.11). Then, (y, a) is the solution. If the solution does not satisfy the constraints, then for each $\gamma > \bar{\lambda}$ find the allocation rule. Then, find the minimum value of γ such that the allocation rule y_γ satisfies the constraints. Given the allocation rule, calculate the expected transfer functions a_γ as before. Then, (y_γ, a_γ) is the solution.

We now provide a condition under which the solution (2.14) and the condition imply that the monotonicity constraint is satisfied.

Assumption 1 *(a)* $W^i(x, \theta, \rho, \lambda^i)$ *is non decreasing in* θ^i *for all* $i \in N$, $x \in X$ *and all* $\theta \in \Theta$, *and (b)* $\frac{\partial V^i(x,\theta,\rho)}{\partial \theta^i}$ *is non decreasing in* x *for all* $i \in N$, *all* $\theta \in \Theta$, *and all* $x \in X$.

Note that we did not make any assumption about how the total valuations depend on the allocation up to now. Assumption 1 basically restricts the set of admissible valuation functions and welfare weights such that the solution to the reduced problem satisfies the monotonicity condition (2.1). Assumption 1(a) reduces to a joint condition on priors F^i, welfare weights and the curvature of the total valuation functions. We already know by the initial assumption that $\frac{\partial V^i(x,\theta^i,\rho)}{\partial \theta^i}$ increases (decreases) when θ^i increases (decreases). However, the allocation might also change as a result of increase in an agent's signal. Note that the allocation can not decrease due to Assumption 1(a). Assumption 1(b) guarantees that the derivative of the total valuation functions $\frac{\partial V^i(x,\theta,\rho)}{\partial \theta^i}$ are also non decreasing in x. For example, if priors are uniform on

$[0,1]$, $V^i(x,\theta,\rho) = x(\rho\theta^i + (1-\rho)\frac{\sum_{j\in N}\theta^j}{N})$, and $\rho = 0$ then

$$W^i(.) = \frac{\sum_{j\in N}\theta^j}{N} + \frac{1}{N}(\theta^i - 1 + \frac{\lambda^{0i}}{\gamma}(1-\Lambda^i(\theta^i)).$$

So Assumption $1(a)$ requires $\frac{\partial W^i}{\partial \theta^i} \geq 0$. This is true if and only if $\lambda^i(\theta^i) \leq 2\gamma$ for all $i \in N$ and all $\theta^i \in \Theta^i$. We know that $\gamma \geq \overline{\lambda}$ and welfare weights are always non-negative. Therefore, this condition is satisfied for all possible welfare weights. For general priors and social concerns, this assumption requires

$$\gamma \geq \frac{\lambda^i(\theta^i)}{2f^i(\theta^i) - \frac{\frac{\partial f^i(\theta^i)}{\partial \theta^i}(F^i(\theta^i)-\Lambda^i(\theta^i))}{f^i(\theta^i)}}.$$

This implies Assumption $1(a)$ may not be satisfied for all welfare weights with arbitrary priors. We showed that with uniform priors the assumption can be satisfied for all welfare weights and thus it is satisfied by all incentive efficient mechanisms. Note also that the total valuation function satisfies Assumption $1(b)$.

Theorem 3 *If each* $W^i(.)$ *and* $V^i(.)$ *satisfies Assumption* 1*, then the solution* (2.14) *satisfies all constraints in Theorem* 1.

2.4 Applications

Our characterization is general and can be applied to different economic settings. In this section, we present the main intuition of the characterization by providing simplified applications for both public and private goods environments where the private valuations are linear in types and allocation.

2.4.1 Public Goods

There are N people, i=1,...,N, who must decide on the level of a public good which is produced according to constant returns to scale. In addition, they must decide how to distribute the production costs. Let $X = \{0, 1\}$ denote the possible values of the public good. The cost of producing $x \in X$ is equal to Kx. In our main model, we assumed that social allocation is costless but it is easy to incorporate the cost of social allocation to our model.

The individual payoff function of agent i, given a decision rule y and her monetary payment t^i, is

$$\Pi^i(y, t^i, \theta^i) = \sum_{x \in X} y^x(\theta) v^i(x, \theta^i) - t^i.$$

We had started with the assumption that agents not only care about their individual payoffs but also they care about the payoffs of other agents,

$$u^i(y, t, \theta) = \rho \Pi^i + (1 - \rho)\overline{\Pi}.$$

For this application we assume that the total valuation functions have the following form:

$$V^i(x, \theta, \rho) = x\left(\rho\theta^i + (1 - \rho)\frac{\sum_{j \in N} \theta^j}{N}\right).^4 \tag{2.15}$$

[4]The private valuation of agent i is $v^i(x, \theta^i) = x\theta^i$.

This implies the utility function of agent i is

$$u^i(y,t,\theta) = \sum_{x\in X} y^x(\theta)V^i(x,\theta,\rho) + \rho(\frac{\sum_{j\in N} t^j}{N} - t^i) - \frac{\sum_{j\in N} t^j}{N}, \tag{2.16}$$

where $\rho \in [0,1]$ is the measure of social concerns. If $\rho = 1$, the model is equivalent to the standard public good environment where agents are selfish. If $\rho = 0$, the model is a common values setting where full social preferences are in action and the society is homogeneous (every agent has the same valuation for the public good). If $1 > \rho \geq 0$, agents have interdependent preferences. Note that as ρ decreases the degree of altruism in preferences goes up and the model converges to the full social preferences setting.

For the regular case, given welfare weights $\lambda^i : \Theta^i \to \mathbb{R}^+$, an IE mechanism satisfies:

$$\max_{y\in\Omega} \sum_{i=1}^{N} \int_{\Theta} \left(\sum_{x\in X} y^x(\theta)(W^i(x,\theta,\rho,\lambda^i) - \frac{K}{N}x) \right) dF(\theta) \tag{2.17}$$

$$0 \leq \Phi(\zeta,\rho) - \int_{\Theta} K \sum_{x\in X} y^x(\theta) x dF(\theta) \tag{2.18}$$

$$0 = (\gamma - \overline{\lambda}) \left[\Phi(\zeta,\rho) - \int_{\Theta} K \sum_{x\in X} y^x(\theta) x dF(\theta) \right]. \tag{2.19}$$

Suppose IIR was not required. It is much easier to solve the problem without IIR constraints. First-order conditions imply $\lambda^{0i} = \gamma = \overline{\lambda}$ for all $i \in N$. Hence the ex-ante welfare weights must all be equal. Otherwise, the solution does not exist, since it is always possible to improve welfare by making arbitrarily large transfers between agents with different welfare weights. The problem stated above has a simple solution:

$$y^x(\theta, \lambda) = \begin{cases} 1 & \text{if } x = argmax_{m \in X} \sum_{i=1}^N W^i(m, \theta, \rho, \lambda^i) - Km \\ 0 & \text{otherwise,} \end{cases} \tag{2.20}$$

and the payment functions can be found using Theorem 2 after subtracting the expected cost of the public good from the constraint on the sum of the expected payment of the lowest types of each agent.

The public good is produced if

$$\sum_{i=1}^N \theta^i + (\rho + \frac{1-\rho}{N}) \sum_{i=1}^N \frac{F^i(\theta^i) - \Lambda^i(\theta^i)}{f^i(\theta^i)} \geq K.$$

The first best decision is to produce the public good when

$$\sum_{i=1}^N V^i(1, \theta, \rho) = \sum_{i=1}^N V^i(1, \theta, \rho = 1) = \sum_{i=1}^N v^i(1, \theta) = \sum_{i=1}^N \theta^i \geq K.$$

Let $\Theta^e = \{\theta | \sum_{i=1}^N \theta^i \geq K\}$ and $\Theta^\rho = \{\theta | \sum_{i=1}^N W^i(.) \geq K\}$. Efficiency dictates that the probability of producing the public good does not depend on whether agents are selfish or unselfish. In interim efficient mechanisms, there are distortions from the first best due to informational rents, the type-dependent welfare weights and the measure of social concerns ρ. Note that even though the sum of the valuations for the public good is independent of the measure of social concerns, the production of the public good depends on the interdependence among preferences.

Suppose $\lambda^i(\theta^i)$ is decreasing for all i and θ^i (lower types are weighted more heavily).

This implies the aim of the planner is that agents valuing the public good more should bear a larger share of the costs. Then, $\sum_{i=1}^{N} \frac{F^i(\theta^i)-\Lambda^i(\theta^i)}{f^i(\theta^i)} < 0$. This implies there is less production than the ex post efficient mechanisms for all ρ since sum of the modified virtual valuations is less than the sum of the valuations. Moreover, as ρ decreases, or the degree of altruism in the preferences goes up, the public good is produced more often. That is, when higher types are less heavily weighted than lower types, underproduction is a more efficient way to relax incentive compatibility constraints than transfers. However, incentive compatibility constraints are less binding as we converge to the full social preferences environment (ρ decreases) and there is no need to relax the incentive compatibility constraints. This leads to a relative increase in the production of the public good.

Now, suppose $\lambda^i(\theta^i)$ is increasing for all i and θ^i (higher types are weighted more heavily). Then, $\sum_{i=1}^{N} \frac{F^i(\theta^i)-\Lambda^i(\theta^i)}{f^i(\theta^i)} > 0$. This implies there is more production than the ex post efficient mechanisms for all ρ since sum of the modified virtual valuations is more than the sum of the valuations. Moreover, as ρ decreases, the public good is produced less often. It is also easy to see that if $\lambda^i(\theta^i) = c \in \mathbb{R}_+$ for all i and θ^i, ex ante efficient public decisions which are also interim efficient correspond to the classical first best decision (or ex-post efficient allocation). The following comparative statistics result directly follows from the above discussion.

Proposition 1 *If the welfare weights are decreasing in type, the public good is produced more often as the degree of altruism in preferences goes up.*

Now suppose that IIR constraints are required. The main question is then whether the individual rationality constraint will be binding or not for the agent who is assigned the highest welfare weight. We know that individual rationality constraints will be binding for all other agents since $\gamma \geq \bar{\lambda}$. Note that γ is found using the algorithm in Section 1.3.4. Suppose $\gamma > \bar{\lambda}$. This implies individual rationality constraints are binding for all agents. For this case, virtual valuations are equivalent to[5]

$$W^i(\theta, \rho, \lambda^i) = \left(\rho\theta^i + (1-\rho)\frac{\sum_{j\in N}\theta^j}{N} + (\rho + \frac{1-\rho}{N})\left(\frac{F^i(\theta^i)-1}{f^i(\theta^i)} + \frac{\lambda^{0i}}{\gamma}\frac{1-\Lambda^i(\theta^i)}{f^i(\theta^i)}\right)\right). \tag{2.21}$$

The public good is produced if

$$\sum_{i=1}^{N}\theta^i + (\rho + \frac{1-\rho}{N})\left(\sum_{i=1}^{N}\frac{F^i(\theta^i)-1}{f^i(\theta^i)} + \frac{\lambda^{0i}}{\gamma}\frac{1-\Lambda^i(\theta^i)}{f^i(\theta^i)}\right) \geq K. \tag{2.22}$$

With IIR constraints, virtual valuations are lower for all agents. Hence interim efficient choice of the public good is always lower with the constraints than without. This implies that in some cases it might be efficient to produce the public good but there might not be enough surplus to cover the incentive costs without violating individual rationality constraints. Note that the adjustment term, $(\rho + \frac{1-\rho}{N})\left(\sum_{i=1}^{N}\frac{F^i(\theta^i)-1}{f^i(\theta^i)} + \frac{\lambda^{0i}}{\gamma}\frac{1-\Lambda^i(\theta^i)}{f^i(\theta^i)}\right)$, is always negative. If interim efficient trade occurs, then $\sum_{i=1}^{N}\theta^i > K$. This is because classical (or ex post) efficiency, interim incentive compatibility, and interim individual rationality are incompatible in our environment. IE mechanisms do not produce the public good when it is not optimal

[5] In this application the virtual valuations are linear in the allocation x since the valuation functions are linear in x. We can redefine modified virtual valuations such that they will not depend on the level of public good.

or not classically efficient ($\sum_{i=1}^{N} \theta^i < K$) to produce the public good, and IE mechanisms may not produce the public good when it is optimal to produce ($\sum_{i=1}^{N} \theta^i \geq K$). However, as degree of altruism in the preferences goes up the adjustment term becomes smaller. This implies agents earn less informational rents, the budget balance constraint is relaxed and the constrained optimum is getting more efficient, since it is easier to satisfy individual rationality constraints with relatively more unselfish agents. That is, the inefficient provisioning of the public good will decrease as ρ decreases, and there will be fewer information states of the economy in which it might be optimal to produce the public good but the public good is not produced. Moreover, inefficiency in public good production is the smallest when full social preferences are in action ($\rho = 0$). These observations lead to the following result.

Proposition 2 *Inefficiencies in interim efficient public good provision decreases as the degree of altruism in preferences goes up* ($\Theta^e \supseteq \Theta^{\rho'} \supseteq \Theta^{\rho}$ *for all* $\rho, \rho' \in [0,1]$ *such that* $\rho > \rho'$).

2.4.2 Bargaining: One Buyer and One Seller

There is a risk-neutral seller who wants to sell an indivisible object that she owns and a risk-neutral buyer who wants to buy the object. The seller's type is $\theta^s \in [\underline{\theta}^s, \overline{\theta}^s]$, and the buyer's type is $\theta^b \in [\underline{\theta}^b, \overline{\theta}^b]$. We assume that $[\underline{\theta}^s, \overline{\theta}^s] \cap [\underline{\theta}^b, \overline{\theta}^b] \neq \emptyset$. That is, there are gains from trade for some information states of the economy. A nonmonetary decision may be represented by a vector $x = (x^s, x^b)$, where $x^s = 1$ if the seller keeps the good, $x^s = 0$ if the seller sells the good, $x^b = 1$ if the buyer gets the good, and

$x^b = 0$ if the buyer does not get the good. The set of possible allocations is then $X = \{(1,0),(0,1)\}$. Agent i's individual payoff depends on the decision rule y, her private information θ^i and her monetary transfer t^i,

$$\Pi^i(y, t^i, \theta^i) = \sum_{x \in X} y^x(\theta) v^i(x^i, \theta^i) - t^i.$$

Beyond her individual payoff, agent $i \in \{b, s\}$ cares about the payoff of the other agent,

$$u^i(y, t, \theta) = \sum_{x \in X} y^x(\theta) V^i(x, \theta, \rho) - \rho t^i - (1-\rho)\frac{t^s + t^b}{2}.$$

For this application we assume that total valuation functions have the following form:[6]

$$V^i(x, \theta, \rho) = \left(\rho x^i \theta^i + (1-\rho)\frac{x^s\theta^s + x^b\theta^b}{2}\right).$$

If the parties do not reach an agreement, they get their outside options. The seller's outside option is $U^{0s}(\theta^s) = \frac{1+\rho}{2}\theta^s$ and $U^{0b}(\theta^b) = \int_{\Theta^s} \frac{1-\rho}{2}\theta^s dF^s(\theta^s)$ for the buyer, the buyer's expected value when he does not make any payments to the seller and the seller keeps the good. Note that the interpretation of outside options is not standard in our model. We could also set $U^{0s}(\theta^s) = \theta^s$ and $U^{0b}(\theta^b) = 0$. This would imply interdependence among preferences are not observed if there is no trade.

IIR requires an agent's net utility given incentive taxes to be non-negative for all

[6] If $\rho = 1$, the model is equivalent to the original Myerson and Satterthwaite (1983) bargaining problem with selfish agents.

of that agent's types:

$$U^i(\zeta,\theta^i) - U^{0i}(\theta^i) = U^i(\zeta,\underline{\theta}^i) + \int_{\underline{\theta}^i}^{\theta^i} Q^i(s,\rho)ds - U^{0i}(\theta^i) \geq 0 \;\; for\ all\ i \in M, \theta^i \in \Theta^i.$$

This is only true if

$$U^i(\zeta,\underline{\theta}^i) + min_{\theta^i}\left[\int_{\underline{\theta}^i}^{\theta^i} Q^i(s)ds - U^{0i}(\theta^i)\right] \geq 0.$$

It is easy to see that IIR constraint is binding for the lowest possible type of the buyer and for the highest possible type of the seller. Note that for the buyer individual rationality is satisfied if and only if $U^b(\zeta,\underline{\theta}^b) \geq \int_{\Theta^s} \frac{1-\rho}{2}\theta^s dF^s(\theta^s)$. For the seller, individual rationality requires $U^s(\zeta,\underline{\theta}^s) - \int_{\underline{\theta}^s}^{\overline{\theta}^s} Q^s(a,\rho)da - \frac{1+\rho}{2}\overline{\theta}^s \geq 0$. The expected surplus using our formulation in the paper can be written as

$$B(\zeta) \equiv \sum_{i\in\{b,s\}} \int_\Theta \sum_{x\in X} y^x(\theta)\left(V^i(x,\theta,\rho) - \frac{\partial V^i(x,\theta,\rho)}{\partial\theta^i}\frac{1-F^i(\theta^i)}{f^i(\theta^i)}\right)dF(\theta) +$$

$$-U^b(\zeta,\underline{\theta}^b) - U^s(\zeta,\underline{\theta}^s).$$

Repeating the same arguments, an interim efficient mechanism maximizes

$$\max_{y\in\Omega} \sum_{i\in\{b,s\}} \int_\Theta \Bigg[\sum_{x\in X} y^x(\theta)(V^i(x,\theta,\rho) + \frac{\partial V^i(x,\theta,\rho)}{\partial\theta^i}\left(\frac{F^i(\theta^i)-1}{f^i(\theta^i)} + \frac{\lambda^{0i}}{\gamma}\frac{1-\Lambda^i(\theta^i)}{f^i(\theta^i)} + (1-\frac{\lambda^{0i}}{\gamma})\frac{I^i(\theta^i)}{f^i(\theta^i)}\right)\Bigg] dF(\theta) \tag{2.23}$$

$$0 \leq \Psi(\zeta,\rho) \tag{2.24}$$

$$0 \leq \gamma - \overline{\lambda} \tag{2.25}$$

$$0 = (\gamma - \overline{\lambda})\Psi(\zeta, \rho) \tag{2.26}$$

where

$$I^i(\theta^i) = \begin{cases} 1 & \text{if } \theta^i < argmin_{\theta^i} \left[U^i(\zeta, \underline{\theta}^i) + \int_{\underline{\theta}^i}^{\theta^i} Q^i(s,\rho)ds - U^{0i}(\theta^i) \right]; \\ 0 & \text{if } \theta^i \geq argmin_{\theta^i} \left[U^i(\zeta, \underline{\theta}^i) + \int_{\underline{\theta}^i}^{\theta^i} Q^i(s,\rho)ds - U^{0i}(\theta^i) \right] \end{cases}$$

and

$$\begin{aligned} \Psi(\zeta,\rho) = \sum_{i \in \{b,s\}} \Bigg[\int_\Theta \sum_{x \in X} y^x(\theta) \left(V^i(x,\theta,\rho) + \frac{\partial V^i(x,\theta,\rho)}{\partial \theta^i} \; \frac{F^i(\theta^i) - 1}{f^i(\theta^i)} \right) dF(\theta) \; + \\ \min_{\theta^i} \left(\int_{\underline{\theta}^i}^{\theta^i} Q^i(s,\rho)ds - U^{0i}(\theta^i) \right) \Bigg] \\ = \int_\Theta y^{(0,1)}(\theta) \left[\theta^b - \theta^s + \frac{1+\rho}{2} \left(\frac{F^b(\theta^b) - 1}{f^b(\theta^b)} - \frac{F^s(\theta^s)}{f^s(\theta^s)} \right) \right] dF(\theta). \end{aligned}$$

The modified virtual valuation of the buyer is:

$$W^b(x,\theta,\rho,\lambda^b) = V^b(x,\theta,\rho) + \frac{\partial V^b(x,\theta,\rho)}{\partial \theta^b} \left(\frac{F^b(\theta^b) - 1}{f^b(\theta^b)} + \frac{\lambda^{0b}}{\gamma} \frac{1 - \Lambda^b(\theta^b)}{f^b(\theta^b)} \right),$$

and the modified virtual valuation of the seller is

$$W^s(x,\theta,\rho,\lambda^s) = V^s(x,\theta,\rho) + \frac{\partial V^s(x,\theta,\rho)}{\partial \theta^s} \left(\frac{F^s(\theta^s)}{f^s(\theta^s)} - \frac{\lambda^{0s}}{\gamma} \frac{\Lambda^s(\theta^s)}{f^s(\theta^s)} \right).$$

An IE mechanism gives the good to the agent with the highest positive modified virtual valuations. Trade will take place whenever the seller's modified virtual valuation is below the buyer's modified virtual valuation. This implies trade occurs if

$$\theta^b - \theta^s + \frac{1+\rho}{2}\left(\frac{F^b(\theta^b)-1}{f^b(\theta^b)} + \frac{\lambda^{0b}}{\gamma}\frac{1-\Lambda^b(\theta^b)}{f^b(\theta^b)} - \frac{F^s(\theta^s)}{f^s(\theta^s)} + \frac{\lambda^{0s}}{\gamma}\frac{\Lambda^s(\theta^s)}{f^s(\theta^s)}\right) \geq 0.$$

The first best decision is to trade the good when $\theta^b \geq \theta^s$. Efficient trade cannot occur with probability one in interim efficient mechanisms. There are distortions from the first best due to informational rents, redistribution of income, and our behavioral assumption. Note that the sum of total valuations of agents is independent of the degree of altruism in preferences. However, interim efficient trade depends on the interdependence among preferences.

With IIR constraints, modified virtual valuation of the buyer is lower and modified virtual valuation of the seller is higher than the case without IIR constraints. Hence trade occurs less often with the constraints than without. It might be efficient to trade in some cases but there might not be enough surplus to cover incentive costs without violating individual rationality constraints.

Suppose the priors are uniform on $[0,1]$. Then trade occurs if and only if

$$\theta^b - \theta^s \geq \frac{1+\rho}{3+\rho}\left(1 - \frac{\lambda^{0b}}{\gamma}(1-\Lambda^b(\theta^b)) - \frac{\lambda^{0s}}{\gamma}\Lambda^s(\theta^s)\right).$$

In the ex ante efficient mechanism, which is also interim efficient, $\lambda^s = \lambda^b = 1$, trade occurs if and only if

$$\theta^b - \theta^s \geq \frac{(1+\rho)(\gamma-1)}{(3+\rho)\gamma-(1+\rho)}.$$

The probability of trade in an interim efficient mechanism can be higher or lower than the ex ante efficient mechanism depending on welfare weights. Note that the set of ex ante efficient mechanisms is a subset of the set of interim efficient mechanisms. In the ex ante efficient mechanism the seller adjusts her total valuation upward and the buyer adjusts her total valuation downward. They are willing not to trade even if trade is beneficial to both parties to get more favorable total payoffs. This may not be the case in an interim efficient mechanism depending on welfare weights.

If we apply our algorithm from Section 1.3.4, we see that γ is positively correlated with ρ. The following table summarizes the relationship among the resource feasibility Lagrangian multiplier (γ), the degree of altruism ρ, and information state of the economy $\theta = (\theta^b, \theta^s)$ for which trade occurs.

ρ	γ	$\theta^b - \theta^s \geq$
0	1	0
0.3	1.15	0.08
0.6	1.3	0.17
1	1.45	0.25

Table 2.1: Relationship between ex ante efficient trade and the degree of altruism in preferences.

Note that the probability of ex ante efficient trade is equal to the probability of efficient trade when $\rho = 0$. In this case agents only care about the total valuations but do not care about the transfers, so the problem is equivalent to finding efficient mechanisms. This will not be true for all interim efficient mechanisms since welfare weights might be type dependent. We can conclude from the table that the probability of trade decreases as ρ increases since there will be less (θ^s, θ^b) for which trade occurs as ρ increases. This implies altruistic agents trade more often than selfish agents. Moreover, selfish agents are more willing to risk losing beneficial trades to get a more favorable payment than unselfish agents. The following result states that this observation can easily be extended to all interim efficient mechanisms.

Proposition 3 *Trade occurs more often as the degree of altruism in preferences goes up* ($\Theta^e \supseteq \Theta^{\rho'} \supseteq \Theta^{\rho}$ *for all* $\rho, \rho' \in [0,1]$ *such that* $\rho > \rho'$).

The above result implies that agents will trade more often as ρ decreases. That is, there will be more information states of the economy ($\theta \in \Theta$) where it is interim efficient to trade. Moreover, agents do not trade when it is not optimal (or not classically efficient) to trade ($\theta^b - \theta^s < 0$) and they may not trade when it is optimal to trade ($\theta^b - \theta^s \geq 0$). However, there will be fewer information states of the economy where trade does not occur but trade is optimal as the degree of altruism in the preferences increases.

These results and characterization of interim efficient mechanisms can be applied to markets with many buyers and many sellers and to auctions with one seller and many buyers. For competitive environments such as markets and auctions, we do not have any evidence to support unselfish preferences. We assume that self-interest assumption provides a good description for most people's behavior for these applications ($\rho = 1$).

2.5 Concluding Remarks

In this paper, we have characterized interim efficient mechanisms for Bayesian environments with interdependent preferences. We showed that interim efficient allocation rules assign probability one to an allocation that maximizes sum of agents' modified virtual valuations that are carefully defined for these environments. We mostly concentrated on regular problems where we assumed that monotonicity constraint is not binding and provided a sufficient condition for regular problems. The extension of characterization to irregular problems remains open. We also provided some applications of this characterization for both public and private goods environments.

Our initial intention was to extend our analysis to the case where the individuals share prior claims to the objects (dissolving a partnership). In that case individual rationality constraints are type specific and the determination of buyers and sellers is endogenous. Moreover, individual rationality constraints bind in the interior and this interior point (or region) is also endogenous. This creates difficulties in separating virtual valuations from the allocation rule, and hence virtual valuations are also endogenous. Then, our formulation does not work for this case. Our conjecture is that the set of initial shares for which efficient dissolution is possible extends as the degree of altruism in preferences goes up. Extending the formulation to this problem is an open question. We did not have this problem in our formulation because individual rationality constraints are binding for the lowest or highest types of agents for all incentive-compatible mechanisms.

One possibility for future research is to consider a model where the social concerns of agents are also private information. Then, types are multidimensional. This extension appears to be a difficult open question since the problems with multidimensional analysis are well known in mechanism design literature. Another possibility for future research is to consider a mechanism design problem with Fehr and Schmidt (1999) type of preferences where individuals are inequity-averse. This complicates the mechanism design problem since this type of preferences introduces discontinuities.

The extension of our characterization to the Bayesian environments with inequity averse agents and to the environments with private social concerns will be a subject of our future research.

2.6 Appendix

Proof of Lemma 1. ($\Rightarrow$) Let $s^i > \theta^i$. IIC implies $U^i(\zeta, \theta^i) \geq U^i(\zeta, \theta^i, s^i)$ and $U^i(\zeta, s^i) \geq U^i(\zeta, s^i, \theta^i)$ where

$$U^i(\zeta, \theta^i, s^i) = U^i(\zeta, s^i) - \int_{\Theta^{-i}} \sum_{x \in X} y^x(s^i, \theta^{-i}) V^i(x(s^i, \theta^{-i}), s^i, \theta^{-i}, \rho) dF^{-i}(\theta^{-i})$$

$$+ \int_{\Theta^{-i}} \sum_{x \in X} y^x(s^i, \theta^{-i}) V^i(x(s^i, \theta^{-i}), \theta, \rho) dF^{-i}(\theta^{-i})$$

and

$$U^i(\zeta, s^i, \theta^i) = U^i(\zeta, \theta^i) - \int_{\Theta^{-i}} \sum_{x \in X} y^x(\theta) V^i(x(\theta), \theta, \rho) dF^{-i}(\theta^{-i})$$

$$+ \int_{\Theta^{-i}} \sum_{x \in X} y^x(\theta) V^i(x(\theta), s^i, \theta^{-i}, \rho) dF^{-i}(\theta^{-i}).$$

This implies $Q^i(s^i, \rho) \geq \frac{U^i(\zeta, s^i) - U^i(\zeta, \theta^i)}{s^i - \theta^i} \geq Q^i(\theta^i, \rho)$ and hence $Q^i(\theta^i, \rho)$ is nondecreasing. Letting $s^i \to \theta^i$ implies $\frac{\partial U^i(\zeta, \theta^i)}{\partial \theta^i} = Q^i(\theta^i, \rho)$. Then $U^i(\zeta, \theta^i) = U^i(\zeta, \underline{\theta}^i) + \int_{\underline{\theta}^i}^{\theta^i} Q^i(s, \rho) ds$.

($\Leftarrow$) Now suppose 1 and 2 hold. Then $U^i(\zeta, s^i) - U^i(\zeta, \theta^i) = \int_{\theta^i}^{s^i} Q^i(s, \rho) ds \geq (s^i - \theta^i) Q^i(\theta^i, \rho)$. This implies, repeating the construction backwardly in the necessary part, $U^i(\zeta, \theta^i) \geq U^i(\zeta, \theta^i, s^i)$ and $U^i(\zeta, s^i) \geq U^i(\zeta, s^i, \theta^i)$. □

Proof of Lemma 2. IIR is satisfied if and only if $U^i(\zeta, \theta^i) \geq 0$ for all i, θ^i. By IIC

$$U^i(\zeta, \theta^i) = U^i(\zeta, \underline{\theta}^i) + \int_{\underline{\theta}^i}^{\theta^i} Q^i(s, \rho) ds \geq 0 \quad \textit{for all } i \in N, \theta^i \in \Theta^i.$$

That is, it requires

$$min_{\theta^i \in \Theta^i} \left[U^i(\zeta, \underline{\theta}^i) + \int_{\underline{\theta}^i}^{\theta^i} Q^i(s, \rho) ds \right] \geq 0,$$

$$\Leftrightarrow$$

$$U^i(\zeta, \underline{\theta}^i) + min_{\theta^i \in \Theta^i} \left[\int_{\underline{\theta}^i}^{\theta^i} Q^i(s, \rho) ds \right] \geq 0 \Leftrightarrow U^i(\zeta, \underline{\theta}^i) \geq 0 \;\; for \; all \; i \in N.$$

The other direction is trivial since $V^i(x, \theta^i)$ is monotone increasing in θ^i. □

Proof of Lemma 3. By changing the order of integration we get:

$$\begin{aligned} LHS &= \lambda^{0i} U^i(\zeta, \underline{\theta}^i) + \int_{\underline{\theta}^i}^{\overline{\theta}^i} Q^i(s, \rho) \left[\int_s^{\overline{\theta}^i} \lambda^i(\theta^i) dF^i(\theta^i) \right] ds \\ &= \lambda^{0i} \left[U^i(\zeta, \underline{\theta}^i) + \int_{\underline{\theta}^i}^{\overline{\theta}^i} Q^i(s, \rho)(1 - \Lambda^i(s)) ds \right] \\ &= \lambda^{0i} \left[U^i(\zeta, \underline{\theta}^i) + \int_{\underline{\theta}^i}^{\overline{\theta}^i} (\frac{1 - \Lambda^i(\theta^i)}{f^i(\theta^i)}) Q^i(\theta^i, \rho) dF^i(\theta^i) \right]. \end{aligned}$$

□

Proof of Theorem 1**.** Directly follows from Lemmas 1, 2, and 3. (2.4) is EABB, (2.5) is IIR, and (2.6) is the first part of IIC. □

Proof of Lemma 4**.** Let $\overline{\lambda} = max_{i \in N}\{\lambda^{0i}\}$. Define also $K = \{k \mid \overline{\lambda} = \lambda^{0k}, \; \forall \; k \in N\}$, the set of agents who have the highest ex ante welfare weight, and $M = \{m \mid \overline{\lambda} > \lambda^{0m}, \; \forall \; m \in N\}$, the set of agents whose welfare weights are lower than the highest ex ante welfare weight, where $N = K \cup M$. There are two possible cases:

Case 1: $\gamma > \overline{\lambda}$. This implies for all $i \in N$, $\mu^i > 0 \Rightarrow U^i(\zeta, \underline{\theta}^i) = 0 \Rightarrow$ IIR constraints are binding for all agents' lowest types.

Case 2: $\gamma = \overline{\lambda}$. This implies for each $k \in K$, $\gamma = \overline{\lambda} = \lambda^{0k} \Rightarrow \mu^k = 0 \Rightarrow U^k(\zeta, \underline{\theta}^k) \geq 0$ and for each $m \in M$, $\gamma = \overline{\lambda} > \lambda^{0m} \Rightarrow \mu^m > 0 \Rightarrow U^m(\zeta, \underline{\theta}^m) = 0 \Rightarrow$ IIR constraints are binding for all agents' lowest types in M and the constraints are not binding for all agents in K.

From Case 1 and 2, if $U^i(\zeta, \underline{\theta}^i) \neq 0$ for some $i \in M \subseteq N$ then for all $i \in M$ ex ante welfare weights are equal to $\gamma = \overline{\lambda} = \lambda^{0i}$. This implies $\sum_{i=1}^{N} \lambda^{0i} U^i(\zeta, \underline{\theta}^i) = \gamma \sum_{i=1}^{N} U^i(\zeta, \underline{\theta}^i) = \gamma \Phi(\zeta, \rho)$. The second equality follows by EABB constraint which is always binding. $\square$

Proof of Theorem 2. It follows from Lemma 4 and the discussion for Case 1 and 2 as stated above. We know that IIR constraint is binding for the lowest types of all $-i$. This implies $\sum_{l \in N} U^l(\zeta, \underline{\theta}^l) = U^i(\zeta, \underline{\theta}^i) = \Phi(\zeta, \rho)$ since EABB is always binding. Then, $\rho a^i(\theta^i) = \int_{\Theta^{-i}} \sum_{x \in X} y^x(\theta) V^i(x, \theta, \rho) dF^{-i}(\theta^{-i}) - U^i(\zeta, \theta^i)$. By using the envelope condition, we get the payment function of the agent with maximal welfare weight. This implies agent i is the residual claimant. $\square$

Proof of Theorem 3. The solution is constructed such that all constraints other than monotonicity are satisfied. We only need to show that the solution satisfies the monotonicity constraint. Suppose $\theta^i \geq s^i$, $x = argmax_{m \in X} \sum_{i=1}^{N} W^i(m, \theta, \rho, \lambda^i)$ and

$y = argmax_{m\in X} \sum_{i=1}^{N} W^i(m, s^i, \theta^{-i}, \rho, \lambda^i)$. This implies $x \geq y$ by Assumption 1(a). By Assumption 1(b),

$$Q^i(\theta^i,\rho) = \int_{\Theta^{-i}} \sum_{x\in X} y^x(\theta^i,\theta^{-i}) \frac{\partial V^i(x,\theta,\rho)}{\partial \theta^i} dF^{-i}(\theta^{-i}) = \int_{\Theta^{-i}} \frac{\partial V^i(x,\theta,\rho)}{\partial \theta^i} dF^{-i}(\theta^{-i})$$

$$\geq Q^i(s^i,\rho) = \int_{\Theta^{-i}} \frac{\partial V^i(y,s^i,\theta^{-i},\rho)}{\partial s^i} dF^{-i}(\theta^{-i}).$$

This implies $Q^i(\theta^i,\rho)$ is monotone increasing. Note that if $x = y$, $Q^i(\theta^i,\rho)$ is obviously monotone increasing since we initially assumed that the valuation functions are monotone increasing in type for all agents. Therefore, the solution (2.14) satisfies all constraints in Theorem 1. □

Proof of Proposition 1. Suppose the welfare weights are decreasing in type. This implies $\sum_{i=1}^{N} \frac{F^i(\theta^i)-\Lambda^i(\theta^i)}{f^i(\theta^i)} < 0$. We know that the probability of public good production (the ratio of type profiles in which public good is produced) is equal to $Prob(y_\rho^{x=1} > 0) = Prob(\theta | \sum_{i=1}^{N} \theta^i + (\rho + \frac{1-\rho}{N}) \sum_{i=1}^{N} \frac{F^i(\theta^i)-\Lambda^i(\theta^i)}{f^i(\theta^i)} \geq K)$. Now consider $\hat{\rho} > \rho$ and θ where $y^{x=1}(\rho,\theta) = 1$. It is easy to see that $Prob(y_\rho^{x=1} > 0) \geq Prob(y_{\hat{\rho}}^{x=1} > 0)$ since there is $\hat{\rho}$ such that $\sum_{i=1}^{N} \theta^i + (\rho + \frac{1-\rho}{N}) \sum_{i=1}^{N} \frac{F^i(\theta^i)-\Lambda^i(\theta^i)}{f^i(\theta^i)} \geq K \geq \sum_{i=1}^{N} \theta^i + (\hat{\rho} + \frac{1-\hat{\rho}}{N}) \sum_{i=1}^{N} \frac{F^i(\theta^i)-\Lambda^i(\theta^i)}{f^i(\theta^i)}$. Note also that if $y^{x=1}(\rho,\theta) = 0$ then $y^{x=1}(\hat{\rho},\theta) = 0$. Next consider $\rho > \tilde{\rho}$ and θ where $y^{x=1}(\rho,\theta) = 0$. Then $Prob(y_{\tilde{\rho}}^{x=1} > 0) \geq Prob(y_\rho^{x=1} > 0)$ since there is $\tilde{\rho}$ such that $\sum_{i=1}^{N} \theta^i + (\tilde{\rho} + \frac{1-\tilde{\rho}}{N}) \sum_{i=1}^{N} \frac{F^i(\theta^i)-\Lambda^i(\theta^i)}{f^i(\theta^i)} \geq K \geq \sum_{i=1}^{N} \theta^i + (\rho + \frac{1-\rho}{N}) \sum_{i=1}^{N} \frac{F^i(\theta^i)-\Lambda^i(\theta^i)}{f^i(\theta^i)}$. Note also that if $y^{x=1}(\rho,\theta) = 1$ then $y^{x=1}(\tilde{\rho},\theta) = 1$. The proof for the case of increasing welfare weights is also similar. □

Proof of Proposition 2. Given welfare weights and priors, let $\Theta^{\rho} = \{\theta \mid \sum_j W^j(\theta, \rho, \lambda^i) \geq K\}$ be the set of types where public good is produced by an IE mechanism ζ, given (ρ, γ). Let $\Theta^{\rho'} = \{\theta | \sum_j W^j(\theta, \rho', \lambda^i) \geq K\}$ be the set of types where public good is produced by ζ given (ρ', γ'). Note that for all $\rho^* \in [0, 1]$ and all $\theta \in \Theta^{\rho^*}$, $\sum_{i=1}^N \theta^i \geq K$ since efficiency, interim incentive compatibility, and interim individual rationality are incompatible. This implies the adjustment term in modified virtual valuations is always negative. Suppose without loss of generality $\rho' < \rho$. We want to show that there are more information states of the economy where the public good is produced as the degree of altruism in preferences goes up, $\Theta^{\rho'} \supseteq \Theta^{\rho}$. Suppose on the contrary there is θ such that $\theta \in \Theta^{\rho}$ and $\theta \notin \Theta^{\rho'}$. Then

$$\sum_{i=1}^N \theta^i + (\rho + \frac{1-\rho}{N}) \left(\sum_{i=1}^N \frac{F^i(\theta^i) - 1}{f^i(\theta^i)} + \frac{\lambda^{0i}}{\gamma} \frac{1 - \Lambda^i(\theta^i)}{f^i(\theta^i)} \right) \geq K$$

and

$$\sum_{i=1}^N \theta^i + (\rho' + \frac{1-\rho'}{N}) \left(\sum_{i=1}^N \frac{F^i(\theta^i) - 1}{f^i(\theta^i)} + \frac{\lambda^{0i}}{\gamma'} \frac{1 - \Lambda^i(\theta^i)}{f^i(\theta^i)} \right) < K.$$

This is only possible if $\gamma' > \gamma$. We also know that $\gamma \geq \overline{\lambda}$ from first-order conditions. This implies

$$\Psi(\zeta, \rho) - \int_{\Theta^{\rho}} K dF(\theta) \geq \Psi(\zeta, \rho') - \int_{\Theta^{\rho'}} K dF(\theta) = 0$$

where

$$\Psi(\zeta,\rho') = \int_{\Theta^{\rho'}} \left(\sum_{i=1}^{N} \theta^i + (\rho' + \frac{1-\rho'}{N}) \sum_{i=1}^{N} \frac{F^i(\theta^i) - 1}{f^i(\theta^i)} \right) dF(\theta).$$

Let, without loss of generality, $\Theta^{\rho} = \{\theta | \sum_j \theta^j > A(\theta, K)\}$ and $\Theta^{\rho'} = \{\theta | \sum_j \theta^j > B(\theta, K)\}$. Since $(\rho' + \frac{1-\rho'}{N}) \sum_{i=1}^{N} \frac{F^i(\theta^i)-1}{f^i(\theta^i)} > (\rho + \frac{1-\rho}{N}) \sum_{i=1}^{N} \frac{F^i(\theta^i)-1}{f^i(\theta^i)}$ for all $\theta \in \Theta$, $B(\theta, K) > A(\theta, K)$. This implies if $\theta \in \Theta^{\rho}$ then $\theta \in \Theta^{\rho'}$, contradicting our initial assumption. □

Proof of Proposition 3. Given welfare weights and priors, let Θ^{ρ} be the set of types where trade occurs given (ρ, γ) and $\Theta^{\rho'}$ be the set of types where trade occurs given (ρ', γ'). Let, without loss of generality, $\rho' < \rho$ (altruism in the preferences increases). We want to show that $\Theta^{\rho'} \supseteq \Theta^{\rho}$. Suppose on the contrary there exists θ such that $\theta \in \Theta^{\rho}$ but $\theta \notin \Theta^{\rho'}$. (We know that either $\Theta^{\rho'} \supseteq \Theta^{\rho}$ or $\Theta^{\rho'} \subseteq \Theta^{\rho}$.) This implies

$$\theta^b - \theta^s + \frac{1+\rho}{2} \left(\frac{F^b(\theta^b) - 1}{f^b(\theta^b)} + \frac{\lambda^{0b}}{\gamma} \frac{1 - \Lambda^b(\theta^b)}{f^b(\theta^b)} - \frac{F^s(\theta^s)}{f^s(\theta^s)} + \frac{\lambda^{0s}}{\gamma} \frac{\Lambda^s(\theta^s)}{f^s(\theta^s)} \right) \geq 0$$

and

$$\theta^b - \theta^s + \frac{1+\rho'}{2} \left(\frac{F^b(\theta^b) - 1}{f^b(\theta^b)} + \frac{\lambda^{0b}}{\gamma'} \frac{1 - \Lambda^b(\theta^b)}{f^b(\theta^b)} - \frac{F^s(\theta^s)}{f^s(\theta^s)} + \frac{\lambda^{0s}}{\gamma'} \frac{\Lambda^s(\theta^s)}{f^s(\theta^s)} \right) \leq 0.$$

This is only possible if $\gamma' > \gamma \geq \overline{\lambda}$. This implies $\Psi(\zeta, \rho') = 0$. Since $\rho > \rho'$ and trade does not occur in ρ', we have

$$\theta^b - \theta^s + \frac{1+\rho}{2} (\frac{F^b(\theta^b) - 1}{f^b(\theta^b)} - \frac{F^s(\theta^s)}{f^s(\theta^s)}) \leq \theta^b - \theta^s + \frac{1+\rho'}{2} (\frac{F^b(\theta^b) - 1}{f^b(\theta^b)} - \frac{F^s(\theta^s)}{f^s(\theta^s)}) \leq 0.$$

We know that trade occurs for θ in ρ. Hence,

$$\Psi(\zeta, \rho) = \int_{\Theta^\rho} \left[\theta^b - \theta^s + \frac{1+\rho}{2} \left(\frac{F^b(\theta^b) - 1}{f^b(\theta^b)} - \frac{F^s(\theta^s)}{f^s(\theta^s)} \right) \right] dF(\theta) \leq \Psi(\zeta, \rho') = 0.$$

This is a contradiction since γ is chosen by the algorithm such that $\Psi(\zeta, \rho) \geq 0$. $\square$

Chapter 3

Implementation of the Core in College Admissions Problems When Colleagues Matter

3.1 Introduction

This paper presents simple sequential mechanisms that implement the core correspondence in college admissions problems when students do care about who else goes to the same college. In this matching market, there are two finite disjoint sets of agents, the set of colleges C and the set of students S. Each college has a preference relation over groups of students. Each student has a preferences relation over colleges and groups of students (or classmates). A matching will be a particular assignment of students to colleges. The solution concept, core, specifies the set of matchings we might observe in practice.

It is well known that the core of this matching market can be empty. Dutta and Massó (1997) present conditions under which the core is non-empty. They show that if the students' preferences are college-lexicographic and colleges' preferences satisfy substitutability, the core is non-empty. There are also other papers in which other conditions are presented for non-emptiness of the core. See, for example, Revilla (2004) and Pycia (2007). We first introduce a multi-stage mechanism that implements the core in Subgame Perfect Nash Equilibrium on an unrestricted domain. We also show that two-stage mechanisms cannot implement the core without these restrictions.

In college admissions problems, Alcalde and Romero-Medina (2000) show that two-stage simple mechanisms, such as the "students propose and colleges choose mechanism" and the "colleges propose and students choose" mechanism, implement the core. We show that only extension of the colleges propose and students choose mechanism can implement the core under the restrictions that guarantee the non-emptiness of the core (students' preferences are college-lexicographic and colleges' preferences satisfy substitutability). Therefore, the symmetry between these two mechanisms does not hold when students also care about their classmates.

Implementation of pairwise-stable outcomes in other matching markets has been studied by several authors. Roth (1984), Gale and Sotomayor (1985), Alcalde (1996), and Sotomayor (2003) deal with implementation in one-to-one matching markets. Kara and Sönmez (1995) show that the set of stable matchings, but no subset of the core, is Nash implementable. Alcalde and Romero-Medina (2000) provide simple mechanisms, which implement the core in SPNE, for many-to-one matching markets. Sotomayor (2004) deals with the implementation problem in many-to-many matching markets. This paper is the first paper which works on implementation problem for many-to-one matching markets with preferences over colleagues. Finally, for general cooperative games in coalitional form Perez-Castrillo (1994), Perry and Reny (1994), Serrano (1995), and Serrano and Vohra (1997) address the question of implementation of the core.

The rest of the paper is organized as follows. In Section 3.2 we give a brief introduction to college admissions problems when colleagues matter and define our notion of subgame perfect implementation. Section 3.3 presents the sequential matching mechanism and we show that the matching mechanism implements in subgame perfect equilibrium the core correspondence in our framework. Section 3.4 introduces a two-stage simple mechanism which can implement the core when the core is non-empty. In Section 3.5 we show that if we exchange the roles of colleges and students, the two-stage simple mechanism cannot implement the core. The concluding remarks follows in Section 3.6.

3.2 Preliminaries

3.2.1 The Model

We consider an extended version of a college admissions problem with n students and m colleges. Let $S = \{s_1, ..., s_n\}$ be a set of students and $C = \{c_1, ..., c_m\}$ be a set of colleges. Each college c_i has a preference $P(c_i)$ defined over groups of students 2^S. Each s_i has a preference $P(s_i)$ defined over $(C \times S_{s_i}) \cup \{(\emptyset, \emptyset)\}$, where $S_{s_i} = \{A | A \in 2^S, s_i \in A\}$ is the set of subsets of S which contains s_i. Preferences are linear orders. A preference profile is a list $P = (P(x))_{x \in C \cup S} \in \mathbb{P}$. We assume that being unmatched or $\emptyset$ is not the last choice for each college.

A **matching** is a mapping μ from $C \cup S$ into $2^S \cup (C \times S_{s_i}) \cup \{(\emptyset, \emptyset)\}$ which satisfies for all $c_i \in C$ and $s_i \in S$:

(1) $\mu(s_i) \in (C \times S_{s_i}) \cup \{(\emptyset, \emptyset)\}$,

(2) $\mu(c_i) \in 2^S$,

(3) If $s_i \in \mu(c_i)$ then $\mu(s_i) = (c_i, \mu(c_i))$,

(4) If $\mu(s_i) = (c_j, S')$ then $\mu(c_j) = S'$.

$\mu(s_i) = (\emptyset, \emptyset)$ means that the student s_i is not matched to any college. Similarly, if $\mu(c_i) = \emptyset$ then there are no students matched to college c_i. Given a set of students

$A \subseteq S$, we denote by $Ch(A, P(c))$ the maximal element on 2^A under the linear order $P(c)$.

A matching μ is **individually rational** if no student prefers to be unmatched and no college prefers to be matched with a subset of the current set of students. Formally, for all $s_i \in S$ and all $c_i \in C$, $\mu(s_i)P(s_i)(\emptyset, \emptyset)$ and $\mu(c_i) = Ch(\mu(c_i), P(c_i))$. Let $\mathcal{M}$ be the set of all matchings and $\mathcal{N}$ be the set of individually rational matchings.

Given any profile of preferences P, a matching μ is in the **core**, denoted $C(P)$, if there is no $C' \subseteq C$, $S' \subseteq S$, and a matching μ' such that:

1. $C' \cup S' \neq \emptyset$;

2. For all $c \in C'$ and $s \in S'$, $\mu'(c) \in 2^{S'}$ and $\mu'(s) \in C' \times S'_s \cup \{(\emptyset, \emptyset)\}$;

3. For all $a \in C' \cup S'$, $\mu'(a)R(a)\mu(a)$;

4. There exists $a \in C' \cup S'$ such that $\mu'(a)P(a)\mu(a)$.

If such an C', S', and μ' exist, then we say that μ is blocked by $C' \cup S'$. It is also easy to see that if a matching is not in the core, then either it is not individually rational or it is blocked by a single college and some students. Hence the following definition is equivalent to the above core definition.[1] A matching is in the **core** if it is individually rational and there does not exist a student-group–college pair $(B, c) \in 2^S \times C$ where $B \cap \mu(c) = \emptyset$ and $A \subseteq \mu(c)$ such that for all $s' \in A \cup B$, $(c, A \cup B)P(s')\mu(s')$ and $A \cup BP(c)\mu(c)$.

Given a set of students $A \subseteq S$, we denote by $Ch(A, P(c))$ the maximal element on 2^A under the linear order $P(c)$. A matching μ is **pairwise-stable** if there is no

[1] See Echenique and Yenmez (2007) for a proof.

pair $(s, c) \in S \times C$ such that $s \notin \mu(c)$ implies $s \in Ch(\{s\} \cup \mu(c), P(c))$ and for all $s' \in Ch(\{s\} \cup \mu(c), P(c))$ implies $(c, Ch(\{s\} \cup \mu(c), P(c)))P(s')\mu(s')$. If there exists such a pair, we will say that μ is pairwise-blocked by (s, c). Let $S(P)$ be the set of pairwise-stable matchings.

The following result states the relationship between the core and pairwise-stable matchings. The core is a subset of the set of pairwise-stable matchings. A proof is given by Echenique and Yenmez (2007).

Proposition 4 *$C(P) \subseteq S(P)$.*

In the classical college admission problems where students are indifferent about their colleagues, the set of pairwise-stable matchings coincides with the core when colleges' preferences satisfy substitutability.[2] That is, for any student $s \neq s'$, if s belongs to $Ch(A, P(c))$, then she will also belong to $Ch(A \setminus \{s'\}, P(c))$. Moreover, the substitutability assumption guarantees non-emptiness of the core and implies that the core has a lattice structure.

In our setting the core can be empty. Dutta and Massó (1997) show that the set of matchings in the core remains non-empty when students' preferences are college-lexicographic and colleges' preferences satisfy substitutability. That is student s_i's preferences are college-lexicographic if there is a strict ordering $\overline{P}_i$ over colleges such that for all $(C, S), (C', S'), (C \neq C')$, $(C, S)P(s_i)(C', S') \Leftrightarrow C\overline{P}_iC'$, and $(C, S)P(s_i)s_i \Leftrightarrow C\overline{P}_is_i$. Pycia (2007) also presents a sufficient and, in certain sense, necessary condition for non-emptiness of the core.

[2]See Roth and Sotomayor (1990) for more on this.

3.2.2 Subgame Perfect Implementation

We now introduce our definition of subgame perfect implementation. By a sequential game form, we mean a finite-horizon extensive game form with perfect information. A sequential game form Γ is said to (fully) implement in subgame perfect equilibrium a solution $\Im$ if for every preference profile $P \in \mathbb{P}$, $\Im(P)$ is the unique **pure strategy** subgame perfect equilibrium outcome of the game (Γ, P). A sequential game form Γ is said to partially implement in subgame perfect equilibrium a solution $\Im$ if for every preference profile P, the pure-strategy subgame perfect equilibrium outcomes of the game (Γ, P) is a subset of $\Im(P)$. By a sequential game form $\Gamma = (M, h) = (\prod_{i \in N} M_i, h)$, we mean a finite horizon extensive game form with perfect information where M_i is agent i's strategy space and $h : M \rightarrow \mathcal{M}$ is an outcome function.

Let $SPNE(\Gamma, P)$ denote the set of pure strategy subgame perfect equilibria for the game (Γ, P) and let $SP(\Gamma, P) = h(SPNE(\Gamma, P))$ denote the set of all outcomes corresponding to the subgame perfect equilibria of (Γ, P). Then Γ is said to (fully) implement the core in subgame perfect equilibrium if $SP(\Gamma, P) = C(P)$ for all $P \in \mathbb{P}$.

3.3 The Matching Mechanism

Before defining our mechanism (or game form), we need some additional notation. A college-permutation is a bijection $\pi_i \in \Pi$ from C to C. Given a profile of permutations $\pi = (\pi_i)_{i \in C}$, let $f(\pi) = \pi_{c_1} \circ ... \circ \pi_{c_m}$ be the composition of the permutations and

$f_i(\pi)$ be the i-th college of $f(\pi)$.[3] This type of permutation is useful in endogenously defining the order of moves in our sequential game form. Note that any $i \in C$ can make a unilateral change in π_i to make itself the first player in the order of f.

We can now define the matching mechanism which is a sequential game form that implements the core.

In stage 1, every college c_i simultaneously announces an individually rational matching $\mu_{c_i} \in \mathcal{N}$ and a permutation $\pi_{c_i} \in \Pi$. If for any c_i and c_j, $\mu_{c_i} \neq \mu_{c_j}$, then the game ends with $\mu(f_i(\pi)) = \emptyset$ for all $i < m$ and $\mu(f_m(\pi)) \in 2^S$ such that $\emptyset P(f_m(\pi))\mu(f_m(\pi))$. If $\mu_{c_i} = \mu_{c_j} = \mu^1$ for all $c_i, c_j \in C$, then proceed to stage 2. In stage 2, a college $f_1(\pi)$ announces the set of students $S_{f_1(\pi)} \subseteq S$ that it wants to admit. Students who received an offer from the college respond sequentially to this proposal (the order of students is fixed and the order is not important to the results). If all students in $S_{f_1(\pi)}$ accept the offer, then we proceed to stage 3. If any student in $S_{f_1(\pi)}$ rejects the proposal, the game ends with $\mu^2 = \mu^1$ as the outcome. In stage 3, a college $f_2(\pi)$ announces the set of students $S_{f_2(\pi)} \subseteq S \setminus S_{f_1(\pi)}$ that it wants to admit. Students who received an offer from the college respond sequentially to this proposal. If all students in $S_{f_2(\pi)}$ accept the offer, then we proceed to stage 4. If any student in $S_{f_2(\pi)}$ rejects the proposal, then the game ends with μ^3 as the outcome. Outcomes of the game form when a student rejects a proposal is formally defined in the next paragraphs.

$\vdots \quad \vdots$

[3]This type of permutation device is proposed by Thomson (2005) in implementation of the "no-envy" correspondence and its variants in abstract economic environments.

In stage $m+1$, a college $f_m(\pi)$ announces the set of students $S_{f_m(\pi)} \subseteq S \setminus \cup_{i<m} S_{f_i(\pi)}$ that it wants to admit. Students who received an offer from the college respond sequentially to this proposal. If all students in $S_{f_m(\pi)}$ accept the offer, then $\mu(f_i(\pi)) = S_{f_i(\pi)}$ for all $i \leq m$, $\mu(s_i) = (f_j(\pi), S_{f_j(\pi)})$ for all students such that $s_i \in S_{f_j(\pi)}$, and any student who did not receive any offer is unmatched. If any student in $S_{f_m(\pi)}$ rejects the proposal, the game ends with μ^{m+1} as the outcome.

To formally summarize, the matching mechanism Γ is defined as follows:

Stage 1. Every college c_i simultaneously announces $\eta^1_{c_i} \in \mathcal{N} \times \Pi$.

If for any c_i and c_j, $\mu_{c_i} \neq \mu_{c_j}$, then the game ends with $\mu(f_i(\pi)) = \emptyset$ for all $i < m$ and $\mu(f_m(\pi)) \in 2^S$ such that $\emptyset P(f_m(\pi))\mu(f_m(\pi))$. If $\mu_{c_i} = \mu_{c_j} = \mu^1$ for all $c_i, c_j \in C$, then go to the next stage.

Stage 2. The college $f_1(\pi)$ announces a set of students $\eta^2_{f_1(\pi)} = S_{f_1(\pi)} \subseteq S$. Then, each $s_i \in S_{f_1(\pi)}$ sequentially announces $\lambda^2_{s_i} \in \{$"*accept*", "*reject*"$\}$.[4]

If there is $s_i \in S_{f_1(\pi)}$ such that $\lambda^2_{s_i} =$ "*reject*", the game ends with $\mu^2 = \mu^1$ as the outcome. Otherwise, go to the next stage.

Stage 3. The college $f_2(\pi)$ announces a set of students $\eta^3_{f_2(\pi)} = S_{f_2(\pi)} \subseteq S \setminus S_{f_1(\pi)}$. Next, each $s_i \in S_{f_2(\pi)}$ sequentially announces $\lambda^3_{s_i} \in \{$"*accept*", "*reject*"$\}$.

If $\lambda^3_{s_i} =$ "*accept*" for all $s_i \in S_{f_2(\pi)}$, then go to the next stage. If there is $s_i \in S_{f_2(\pi)}$ such that $\lambda^3_{s_i} =$ "*reject*", the game ends with μ^3 as the outcome where;

$$\mu^3(f_1(\pi)) = S_{f_1(\pi)}, \quad \mu^3(s_i) = (f_1(\pi), S_{f_1(\pi)}) \ for \ all \ s_i \in S_{f_1(\pi)},$$

[4]We assume that each student announces "accept" if she is indifferent between announcing "accept" and "reject". The order of students who respond to the offer of the college is fixed in every stage and the order is not important to the results.

$$\forall \;\; 1 < i \leq m, \;\; \mu^3(f_i(\pi)) = \begin{cases} \mu^1(f_i(\pi)) & \text{if } \mu^1(f_i(\pi)) \in S \setminus S_{f_1(\pi)} \\ \emptyset & \text{otherwise.} \end{cases}$$

and

$$\forall \;\; s_i \in S \setminus S_{f_1(\pi)}, \;\; \mu^3(s_i) = \begin{cases} (f_i(\pi), \mu^1(f_i(\pi)) & \text{if } \mu^1(f_i(\pi)) \in S \setminus S_{f_1(\pi)} \\ (\emptyset, \emptyset) & \text{otherwise.} \end{cases}$$

$$\vdots \quad \vdots$$

Stage m+1. The college $f_m(\pi)$ announces a set of students $\eta^{m+1}_{f_m(\pi)} = S_{f_m(\pi)} \subseteq S \setminus \cup_{i<m} S_{f_i(\pi)}$. Then, each $s_i \in S_{f_m(\pi)}$ sequentially announces $\lambda^{m+1}_{s_i} \in \{$"*accept*", "*reject*"$\}$.

If $\lambda^{m+1}_{s_i} =$ "*accept*" for all $s_i \in S_{f_m(\pi)}$, then the game ends with $\mu(f_i(\pi)) = S_{f_i(\pi)}$ for all $i \leq m$, $\mu(s_i) = (f_j(\pi), S_{f_j(\pi)})$ for all students such that $s_i \in S_{f_j(\pi)}$ and any student who did not get any offer is unmatched, as the outcome. If there is $s_i \in S_{f_m(\pi)}$ such that $\lambda^{m+1}_{s_i} =$ "*reject*", the game ends with $\mu^{m+1}(f_i(\pi)) = S_{f_i(\pi)}$ for all $i < m$, $\mu^{m+1}(f_m(\pi)) = \emptyset$, $\mu^{m+1}(s_k) = (f_i(\pi), S_{f_i(\pi)})$ for all $s_k \in S_{f_i(\pi)}$ and $\mu^{m+1}(s_j) = (\emptyset, \emptyset)$ for all $s_j \in S \setminus \cup_{i<m} S_{f_i(\pi)}$ as the outcome.

Note that in any stage if a college announces the empty set, then we directly move to the next stage and the college is unmatched as an outcome of the game induced by our game form.

3.3.1 Implementing the Core

In our setting, the core might be empty even if we assume that colleges' preferences satisfy substitutability. We need very strong assumptions on the preferences to guarantee the non-emptiness of the core. In our main result, we do not put these restrictions on the preferences and we characterize the core as the subgame perfect equilibrium outcomes of the game induced by the matching mechanism. This also implies in a matching market with an empty core, the subgame perfect equilibrium outcomes of the game form are also empty. The following is our main result.

Theorem 4 *The matching mechanism Γ implements in subgame perfect Nash equilibrium the core of college admission problems with preferences over colleagues.*

The main structure of the proof is the following. We first assume the existence of the core and pure-strategy subgame perfect equilibrium. Then, we characterize pure strategy subgame perfect equilibrium outcomes as the core. We will also show that the core is empty if and only if there does not exist a subgame perfect equilibrium whose proof directly follows from the proof of our main result.

Proof. We fix $P \in \mathbb{P}$. First, we prove that $C(P) \subseteq SP(\Gamma, P)$. Let $\mu \in C(P)$ and consider the following strategy profile: (i) $\eta^1_{c_i} = (\mu, \pi^I_{c_i}) = (\mu, \pi^I)$ and $\eta^1_{s_i} = \pi^I_{s_i} = \pi^I$ for all $c_i \in C$ and all $s_i \in S$, where π^I is the identity permutation. (ii) $S_{f_i(\pi)} = \mu(f_i(\pi))$ for all $i \leq m$, where $f_i(\pi) = c_i$ for all $c_i \in C$, since every agent announces the identity permutation. (iii) For all $s_i \in S_{f_{t-1}(\pi)}$, all $2 \leq t \leq m+1$ and all $s_i \in S$, $\lambda^t_{s_i} =$ "*accept*" if and only if $(f_{t-1}(\pi), S_{f_{t-1}(\pi)})R(s_i)\mu^t(s_i)$.

We only need to check that this strategy profile is an equilibrium. Since $\mu \in C(P)$, there is no blocking student-group–college pair. This implies no college can gain by announcing $S_{f_i(\pi)} \neq \mu(f_i(\pi))$ since it will be rejected by at least one student who got offer from the college $f_i(\pi)$. Therefore, it is a dominant strategy for all

$f_i(\pi) \in C$ to announce $S_{f_i(\pi)} = \mu(f_i(\pi))$. This also implies no college can gain by announcing a different permutation and hence changing the order of colleges. It is also easy to see that announcing "accept" is an optimal strategy if the condition in (iii) holds for all students. Note that we assume that if a student is indifferent between announcing "accept" and "reject", she announces without loss of generality "accept". An equilibrium path corresponding to the above strategy profile is one where every college makes an offer to the students with whom she is matched in the core matching and all students announce "accept". The equilibrium outcome corresponding to the strategy profile is μ. Thus, $\mu \in SP(\Gamma, P)$, which implies $C(P) \subseteq SP(\Gamma, P)$.

Now, we show that $SP(\Gamma, P) \subseteq C(P)$. Suppose by the way of contradiction that there is a matching μ such that $\mu \in SP(\Gamma, P)$ and $\mu \notin C(P)$. Then μ is not individually rational or there exists a student-group–college pair $(B, c) \in 2^S \times C$ which blocks μ. The first case is not possible because each student by rejecting a proposal can be single or guarantee an individually rational match since $\mu^1 \in \mathcal{N}$. Moreover, it is easy to see that in every SPNE, $\mu_{c_i} = \mu_{c_j} = \mu^1$ for all $c_i, c_j \in C$. If this is not the case, college $f_m(\pi)$ can gain by announcing different permutation since $\emptyset P(f_m(\pi))\mu(f_m(\pi))$. Then we move to stage 2 in every SPNE. This implies each college cannot be matched with unacceptable set of students since each college has the option of remaining single by announcing the empty set. Now, consider the second case. We know that $B \cap \mu(c) = \emptyset$ and there is $A \subseteq \mu(c)$ such that for all $s' \in A \cup B$, $(c, A \cup B)P(s')\mu(s')$ and $A \cup BP(c)\mu(c)$. We claim that $c \neq f_1(\pi)$. Suppose $c = f_1(\pi)$. Then $\eta^2_{f_1(\pi)} = B \subseteq S$ and each student s_i in B accepts this offer.

This implies $\mu(c) = B$, a contradiction. Suppose $c = f_k(\pi)$ where $1 < k \leq m$. Then $f_k(\pi)$ can make a unilateral change in $\pi_{f_k(\pi)}$ to make itself the first announcing college and announce $\eta_c^2 = B$. Each student in B accepts this offer. This implies college c was not playing its best response, a contradiction to $\mu \in SP(\Gamma, P)$. ■

Corollary 1 *For all* $P \in \mathbb{P}$, $C(P) = \emptyset \Leftrightarrow SPE(\Gamma, P) = \emptyset$.

Proof. Immediately follows from our main result. ■

Corollary 2 *The matching mechanism* Γ *partially implements in subgame perfect Nash equilibrium the pairwise stable correspondence of college admission problems with preferences over colleagues.*

Proof. Follows from our main result and Proposition 4. ■

3.4 The Colleges Propose and Students Choose Mechanism

This mechanism is widely studied in subgame perfect implementation of the core correspondence of many-to-one and many-to-many matching models. See, for example, Alcalde and Romero-Medina (2000) for college admissions problems and Sotomayor (2004) for many-to-many matching markets. In this mechanism, each college selects the set of potential students. Then, once each student receives the admission letters, she accepts her most preferred college by also considering the classmates. In this section we show that extension of such a two-stage game to our environment cannot implement the core in subgame perfect Nash equilibrium without strong restrictions to the preferences of colleges and students.

Consider the following mechanism, called Γ^{CS}. This is a two-stage mechanism. In this mechanism, offers are made by colleges and students select the best college by also considering the students who were also admitted to the same college. In the first stage, colleges have to decide simultaneously. Each college message space is 2^S. In the second stage, students knowing colleges' announcements select simultaneously the college in which they want to study and their classmates.

Given this strategy space, outcome function h^{CS} selects a matching μ such that:

$$\mu(s) = \begin{cases} (m_1(s), m_2(s)) & \text{if } m(m_1(s)) = m_2(s) \text{ for all } s \in m_2(s) \\ (\emptyset, \emptyset) & \text{otherwise,} \end{cases}$$

and, for each college $c \in C$,

$$\mu(c) = \begin{cases} m(c) & \text{if } m_1(s) = c \text{ and } m_2(s) \supseteq m(c) \text{ for all } s \in m(c) \\ \emptyset & \text{otherwise.} \end{cases}$$

This mechanism Γ^{CS} also cannot implement in SPNE the core correspondence of college admissions problems with preferences over colleagues without strong restrictions on the preferences. The following example shows that there is a preference profile P such that $C(P) \subset SP(\Gamma^{CS}, P)$ where $C(P) = \emptyset$.

Example 1 *Consider two colleges c_1, c_2 and three students s_1, s_2, s_3 with the following preference profile P:*

$P(c_1) : \{s_1, s_2\}, \{s_2, s_3\}$

$P(c_2) : \{s_1, s_3\}$

$P(s_1) : (c_2, \{s_1, s_3\}), (c_1, \{s_1, s_2\})$

$P(s_2) : (c_1, \{s_1, s_2\}), (c_1, \{s_2, s_3\})$

$P(s_3) : (c_1, \{s_2, s_3\}), (c_2, \{s_1, s_3\})$.

This notation means that c_1 prefers $\{s_1, s_2\}$ to $\{s_2, s_3\}$. Only acceptable sets of students for c_1 are listed. That is, potential groups of students not listed are worse for c_1 than being single.

For this example Echenique and Yenmez (2007) show that $C(P) = \emptyset$ using their algorithm. We show that $SP(\Gamma^{CS}, P) \neq \emptyset$ for this example. Consider the following set of strategies:

$$m(c_1) = \{s_2, s_3\} \ ; \ m(c_2) = \{s_1, s_3\} \text{ or } \emptyset$$

$$m(s_1) = (c_2, \{s_1, s_3\}) \ ; \ m(s_2) = (c_1, \{s_2, s_3\}) \ ; \ m(s_3) = (c_1, \{s_2, s_3\}).$$

This constitutes a SPNE for the game Γ^{CS} whose outcome coincides with μ where

$$\mu = \begin{matrix} c_1 & c_2 \\ s_2 s_3 & \emptyset \end{matrix}$$

$\mu \in SP(\Gamma^{CS}, P)$ since college one and other students cannot reach higher utility by deviating unilaterally. Suppose college one c_1 deviates and announces $m'(c_1) =$

$\{s_1, s_2\}$. This leads to the following matching μ':

$$\mu' = \begin{matrix} c_1 & c_2 \\ \emptyset & s_1 s_3 \end{matrix}.$$

However, $\{s_2, s_3\}P(c_1)\emptyset$. As a result $\mu \in SP(\Gamma^{CS}, P)$ but $C(P) = \emptyset$. This implies the extension of the "colleges propose and students choose" mechanism cannot implement the core correspondence with an unrestricted domain of preferences. That is: $C(P) \subset SP(\Gamma^{CS}, P)$.

In the next result, we will assume that workers' preferences are firm-lexicographic and colleges' preferences over sets of students satisfy substitutability. We show that with these strong restrictions on preferences, the core correspondence can be implemented by the Γ^{CS} mechanism.

Theorem 5 *If students' preferences are college-lexicographic and colleges' preferences over sets of students satisfy substitutability then the mechanism Γ^{CS} implements in SPNE the core correspondence.*

Proof. First, we show that every SPNE outcome is in the core. $SP(\Gamma^{CS}, P) \subseteq C(P) \neq \emptyset$. Suppose by way of contradiction that there is a matching μ such that $\mu \in SP(\Gamma^{CS}, P)$ and $\mu \notin\subseteq C(P)$. Then μ is not individually rational or there exists a student-group–college pair $(B, c) \in 2^S \times C$ which blocks μ. The first case is not possible because each student s_i can remain unmatched by reporting $m'(s_i) = (\emptyset, \emptyset)$ and for each college c_j, $Ch(\mu(c_j), P(c_j)) = \mu(c_j)$. Suppose there is a college c such that $D = Ch(\mu(c), P(c)) \subset \mu(c)$. Then college c can be matched with the set of students

D by reporting $m'(c) = D$ by the construction of the mechanism, a contradiction to $\mu \in SP(\Gamma^{CS}, P)$. Now consider the second case. We know that $B \cap \mu(c) = \emptyset$ and there is $A \subseteq \mu(c)$ such that for all $s' \in A \cup B$, $(c, A \cup B) P(s') \mu(s')$ and $A \cup B P(c) \mu(c)$. This implies college c can be better off by announcing $m'(c) = A \cup B \neq \mu(c)$. Notice that in the second stage for all $s_i \in A \cup B$, $m'_1(s_i)$ has to be equal to college c since students' preferences are college-lexicographic. That is, for all $s \in B$, $cP_s\mu_1(s)$. Moreover, $m'_2(s_i) = A \cup B$ for all $s_i \in A \cup B$ since the students play their best response after knowing the announcements of the colleges. This cannot be the case if $\mu \in SP(\Gamma^{CS}, P)$, a contradiction.

Now, we show that $C(P) \subseteq SP(\Gamma^{CS}, P)$. Let $\mu \in C(P) \neq \emptyset$. Consider the following strategies for the agents. For each college $c_j \in C$, its message is $m(c_j) = \mu(c_j)$. In the second stage, for each student s_i, her message is $m(s_i) = \mu(s_i)$. Clearly, $\mu \in SP(\Gamma^{CS}, P)$. We only need to show that this strategy profile is an equilibrium. Suppose by way of contradiction there is a college c_j who can reach higher utility by deviating to $m'(c_j) = \mu'(c_j) = A$ and hence $AP(c_j)\mu(c_j)$. Since μ is individually rational, $\mu(c_j) = Ch(\mu(c_j), P(c_j))$. This implies $A \not\subseteq \mu(c_j)$ but there is a set of students $B \subset A$ such that $B \subset \mu(c_j)$ and $A \setminus B \notin \mu(c_j)$. This implies for all $s \in A \setminus B$, $c_j P_s \mu_1(s)$ since there exists a strict order P_s over colleges. However, we know that $\mu \in C(P)$ and hence there is a student $s' \in B \neq \emptyset$ such that $(c_j, \mu(c_j)) P(s') (c_j, A)$. This implies $m'(s') = m(s') = \mu(s') = (c_j, \mu(c_j))$ and college c_j cannot be matched to the set of students A by construction of the mechanism ($\mu'(c_j) \neq h_{c_j}(m')$), a contradiction to our supposition. Moreover, a student cannot deviate since her dominant strategy is to choose the college that wants her. Therefore μ is a SPNE. This completes the proof. ■

3.5 The Students Propose and Colleges Choose Mechanism

In this mechanism, each student selects the college at which she wants to study and a potential set of classmates. Then, once each college has received all of the application forms, it accepts its most preferred set of students. In this section we show that if we exchange the role of students and colleges, the core correspondence cannot be implemented with or without strong restrictions to the preferences of colleges and students

Let Γ^{SC} be the "students propose and college choose mechanism". In the first stage, each student selects the college at which she wants to study and set of students with whom she wants to go to the college. That is $m(s) = (m_1(s), m_2(s)) \in (C \times S_s) \cup \{(\emptyset, \emptyset)\}$ for all students. At the second stage each college, knowing the announcements of the students, accepts its most preferred set of students. Each college message space is equivalent to 2^S.

Given this strategy space, outcome function h^{SC} selects a matching μ such that: for any student $s \in S$,

$$\mu(s) = \begin{cases} (m_1(s), m_2(s)) & \text{if } m(m_1(s)) = m_2(s) \text{ for all } s \in m_2(s) \\ (\emptyset, \emptyset) & \text{otherwise,} \end{cases}$$

and, for each college $c \in C$,

$$\mu(c) = \begin{cases} m(c) & \text{if } m_1(s) = c \text{ and } m_2(s) \supseteq m(c) \text{ for all } s \in m(c) \\ \emptyset & \text{otherwise.} \end{cases}$$

The following example shows that this mechanism Γ^{SC} cannot implement in SPNE the core correspondence of college admissions problems with preferences over colleagues without strong restrictions on the preferences. That is there is a preference profile P such that $C(P) \neq SP(\Gamma^{SC}, P)$.

Example 2 *Consider two colleges* c_1, c_2 *and four students* s_1, s_2, s_3, s_4 *with the following preference profile* P*:*

$P(c_1) : \{s_1, s_2, s_4\}, \{s_1, s_3, s_4\}, \{s_1\}, \{s_2\}, \{s_3\}$

$P(c_2) : \{s_2, s_3, s_4\}, \{s_3\}, \{s_2\}$

$P(s_1) : (c_1, \{s_1, s_2, s_4\}), (c_1, \{s_1, s_2\}), (c_1, \{s_1, s_3\}), (c_1, \{s_1\})$

$P(s_2) : (c_2, \{s_2, s_3, s_4\}), (c_1, \{s_1, s_2, s_4\})$

$P(s_3) : (c_2, \{s_2, s_3, s_4\}), (c_2, \{s_3\})$

$P(s_4) : (c_1, \{s_3, s_4\}), (c_2, \{s_2, s_3, s_4\}), (c_1, \{s_1, s_2, s_4\})$.

This notation means that c_1 prefers $\{s_1, s_2, s_4\}$ to $\{s_1, s_3, s_4\}$, $\{s_1, s_3, s_4\}$ to $\{s_1\}$, and so on. Only acceptable students for c_1 are listed. That is, potential groups of students not listed are worse for c_1 than being single.

Now, lets look at the following two matchings,

$$\mu_1 = \begin{matrix} c_1 & c_2 \\ s_1 & s_2 s_3 s_4 \end{matrix} \qquad \mu_2 = \begin{matrix} c_1 & c_2 \\ s_1 s_2 s_4 & s_3 \end{matrix}.$$

It is easy to see that $\mu_1 \in C(P)$ and μ_2 can be supported in SPNE by strategies $m(c_1) = \{s_1, s_2, s_4\}$, $m(c_2) = \{s_3\}$, $m(s_1) = (c_1, \{s_1, s_2, s_4\})$, $m(s_2) = (c_1, \{s_1, s_2, s_4\})$, $m(s_3) = (c_2, \{s_3\})$, and $m(s_4) = (c_1, \{s_1, s_2, s_4\})$. That is $\mu_2 \in SP(\Gamma^{SC}, P)$. However, it can be checked that μ_2 is blocked by $\langle \{c_2\}, \{s_2, s_3, s_4\}, \mu_1 \rangle$. This implies $C(P) \neq SP(\Gamma^{SC}, P)$.

In the next example, we assume that students' preferences are college-lexicographic and colleges' preferences over sets of students satisfy substitutability. We will see that even with these strong restrictions on preferences, the core correspondence cannot be implemented by the Γ^{SC} mechanism.

Example 3 *Consider two colleges $C = \{c_1, c_2\}$ and four students $S = \{s_1, s_2, s_3, s_4\}$ with the following preference profile P:*

$P(c_1) : \{s_1, s_2, s_4\}, \{s_1, s_2\}, \{s_1, s_4\}, \{s_2, s_4\}, \{s_1\}, \{s_2\}, \{s_3\}, \{s_4\}$

$P(c_2) : \{s_2, s_3\}, \{s_2, s_4\}, \{s_3\}, \{s_4\}, \{s_3\}, \{s_2\}, \{s_1\}, \{s_4\}$

$P_{s_1} : c_1$

$P_{s_2} : c_2, c_1$

$P_{s_3} : c_2$ *and* $argmax P(s_3) = (c_2, \{s_3\})$

$P_{s_4} : c_2, c_1.$

This notation means that $(c_2, \{s_3\}) P(s_3) (c_2, S')$ for all $S' \in 2^S \backslash \{s_3\}$, $(c_2, S'') P(s_3) (c_1, S''')$, and $(c_2, S'')P(s_3)(\emptyset, \emptyset)$ for all $S'', S''' \in 2^S$.

Consider the following strategy profile:

$$m(c_1) = \{s_1, s_2, s_4\}, m(c_2) = \{s_3\}$$

$$m(s_1) = m(s_2) = m(s_4) = (c_1, \{s_1, s_2, s_4\}), \text{ and } m(s_3) = (c_2, \{s_3\}).$$

It is easy to check that,

$$\mu = \begin{matrix} c_1 & c_2 \\ \\ s_1 s_2 s_4 & s_3 \end{matrix} \in SP(\Gamma^{SC}, P).$$

However, $\langle \{c_2\}, \{s_2, s_4\}, \mu'(c_2) = \{s_2, s_4\} \rangle$ blocks μ. Therefore, $\mu \notin C(P)$. This implies that a symmetrical result for Theorem 5 cannot be obtained by exchanging the role of students and colleges.

3.6 Final Remarks

This chapter presents sequential mechanisms to implement the core correspondence of college admissions problems when colleagues matter. Without any restrictions on the preferences, we propose a multi-stage mechanism that implements the core. This implies that there is no SPNE if the core is empty for this multi-stage mechanism. Under strong restrictions, we show that the extension of the "colleges propose and students choose" mechanism can be used as a noncooperative game to reach the core. Moreover, a symmetric version of the mechanism where colleges and students change their roles cannot implement the core with or without the restrictions on the preferences. This shows that we should take into consideration whether students care about their classmates or not when designing institutions for the real-life matching problem.

Chapter 4

A Theory of Stability in Assignment Games

4.1 Introduction

An assignment game is a two-sided matching market with monetary transfers. In this market, there are two exogenously specified disjoint sets of agents, say firms and workers. The agents engage in bilateral transactions (if worker i works for firm j then firm j employs worker i) and make monetary transfers. Each firm can employ no more than one worker and each worker can not work for more than one firm. A natural solution concept for such markets is the core. The core outcomes specify which partnership we can expect to observe and how the agents will divide their gains. In this chapter, we provide an alternative way to characterize the core of assignment games. We also construct the first algorithm to reach the all core outcomes for such markets.

Shapley and Shubik (1972) show that every assignment game has non-empty core and core payoffs have a nice structure. It is a non-empty complete lattice and there is a polarization of interests in the core. This means that there is a stable outcome which is the most preferred by every agent on one side of the market and at the same time it is the least preferred by every agent on the other side of the market. Geometrically, the core is a closed, convex polyhedron whose dimension is equal to at most the minimum of the number of members in one group or in the other.

We construct a map T on a set of feasible payoffs such that the set of fixed points of T is the core. Then we present a way to reach the core outcomes from fixed points of the map. By recognizing that the mapping is monotone increasing, the lattice property of the set of stable payoffs, as well as its non-emptiness, is proved as an immediate implication of Tarski's fixed point theorem. Furthermore, we show that there is a polarization of interests in the core by using our formulation.

Our characterization is useful because:

1) It allows us to construct an algorithm to find all core outcomes by iterating T.

2) It provides a very simple proof for the lattice structure of the core and the polarization of interests in the core.

This type of fixed point argument has been used in assignment problems with side payments before, but they only characterized certain points in the interior of the core (a subset of the core—*symmetrically bargained allocations*) as stationary points of a rebargaining process between players, see Rochford (1984). By observing a certain monotonicity in the rebargaining process, Roth and Sotomayor (1988) show that these interior points have a lattice property. Moreover, fixed point methods have been used in assignment problems without side transfers (NTU games), see for example Adachi (2000), Echenique and Oviedo (2004), or Echenique and Yenmez (2007) for applications of a fixed point approach for different environments.

The organization of the rest of the chapter is as follows: In the next section, we give a brief introduction to the Shapley and Shubik assignment game and provide some of the well-known results using linear programming formulation. In Section 4.3, we provide some preliminary definitions and state Tarski's fixed point theorem. Section 4.4 formalizes our fixed point approach to the core and shows that our formulation fully characterizes the set of stable outcomes which coincides with the core in this setup. In Section 4.5, we study the lattice structure of stable payoffs. We show that the stable payoffs form a non-empty complete lattice using our formulation. In Section 4.6, we introduce the algorithm. The discussion and future research agenda follows in Section 4.7.

4.2 Assignment Games with Money

This section gives a brief description of the assignment games and provides some well-known results via linear programming proofs. We refer the reader to Shapley and Shubik (1972) or Roth and Sotomayor (1990) for more discussion and justification of the setup.

The game in coalitional function form with side payments is defined by three-tuple $\Gamma = \langle F, W, \alpha \rangle$ where

1. $F = \{f_1, ..., f_m\}$ is a set of firms,

2. $W = \{w_1, ..., w_n\}$ is a set of workers,

3. α is a $m \times n$ matrix of nonnegative numbers $\{\alpha_{fw} \in \mathbb{R}_+ : (f, w) \in F \times W\}$ where α_{fw} is the value of pairwise partnership. Note that $\alpha_{kk} = 0$ for all $k \in F \cup W$.

An **assignment** $\mu : F \cup W \to F \cup W$ is a one-to-one mapping of order two (that is $\mu^2(k) = k$) such that if $\mu(f) \neq f$ then $\mu(f) \in W$ and if $\mu(w) \neq w$ then $\mu(w) \in F$. Let $\mathcal{M}$ be the set of all assignments. An assignment μ can also be represented as a vector $x \in \{0, 1\}^{F \times W}$, such that $x_{fw} = 1$ if $\mu(f) = w$ and $x_{fw} = 0$, otherwise. Hence, $\sum_{w \in W} x_{fw} \leq 1$ for all $f \in F$ and $\sum_{f \in F} x_{fw} \leq 1$ for all $w \in W$.

An assignment x is **optimal** if for all $x' \in \mathcal{M}$,

$$\sum_{(f,w)\in F\times W} \alpha_{fw} x_{fw} \geq \sum_{(f,w)\in F\times W} \alpha_{fw} x'_{fw}.$$

Let $\mathcal{X}$ be the set of optimal assignments. The optimal assignment is usually unique. If there is more than one optimal assignment, a slight perturbation of the values of the pairwise partnerships will result in a unique optimal assignment.

Any agent is free to remain single and receive zero, and the worth of an arbitrary coalition is equal to the sums of the pairwise coalitions it can form with pairs consisting of one agent from F and one from W. That is for all coalitions S,

$$V(S) = \begin{cases} 0 & \text{if } |S| = 0 \text{ or } 1 \\ 0 & \text{if } S \subseteq F \text{ or } S \subseteq W \\ \max_{\mu: F\cap S \to W\cap S} \sum_{f\in F\cap S} \alpha_{f\mu(f)} & \text{if } |F\cap S| \leq |W \cap S| \\ \max_{\mu': W\cap S \to F\cap S} \sum_{w\in W\cap S} \alpha_{\mu(w)w} & \text{if } |F\cap S| \geq |W \cap S|. \end{cases}$$

Definition 2 *The pair of vectors (u, v), with $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$, is called a feasible payoff for $\Gamma = \langle F, W, \alpha\rangle$ if there is an assignment x such that*

$$\sum_{f\in F} u_f + \sum_{w\in W} v_w = \sum_{(f,w)\in F\times W} \alpha_{fw} x_{fw}.$$

In this case we say (u, v) and x are **compatible** with each other, and we call $((u, v); x)$ a feasible outcome.

Definition 3 *A feasible outcome $((u, v); x)$ is **stable** (or the payoff (u, v) with an* *assignment x is stable) if*

(i) $u_f \geq 0, \quad v_w \geq 0$ *(individual rationality)*

(ii) $u_f + v_w \geq \alpha_{fw}$ *for all* $(f, w) \in F \times W$.

Note that condition (ii) only eliminates deviations by pair of agents since the set of pairwise stable outcomes coincides with the set of group stable outcomes in this framework. Let $S(\Gamma)$ be the set of stable payoffs.

Consider just the assignment problem for the coalition of all players:

(AP)

$$\max \qquad z = \sum_{(f,w) \in F \times W} \alpha_{fw} x_{fw}$$

$$s.t. \qquad \sum_{w \in W} x_{fw} \leq 1 \quad \forall \quad f \in F,$$

$$\sum_{f \in F} x_{fw} \leq 1 \quad \forall \quad w \in W,$$

$$x_{fw} \geq 0 \quad \forall \quad (f, w) \in F \times W.$$

This optimization problem is associated with dual linear program having the form:

(DAP)

$$\min \qquad d = \sum_{f \in F} u_f + \sum_{w \in W} v_w$$

$$s.t. \qquad u_f + v_w \geq \alpha_{fw} \quad \forall \quad (f, w) \in F \times W,$$

$$u_f, v_w \geq 0.$$

Therefore, (DAP) formulates the problem of finding payoff vectors in the core of the assignment game. The existence of optimal solutions of (AP) and duality theorem show that the set of stable payoff vectors is non-empty. Moreover, in the game the set of stable outcomes and the core are the same.

Theorem 6 (Shapley and Shubik 1972) *The core of an assignment game* ($C(\Gamma)$ $= S(\Gamma) \times \mathcal{X}$) *is non-empty and is precisely equal to the set of solutions of the (DAP).*

4.3 Preliminary Definitions

Tarski's fixed point theorem is crucial in our formulation. Before stating the theorem, the following definitions will be useful. A **partial order** is a binary relation which is reflexive, transitive, and antisymmetric. A set X endowed with a partial order $\leq$ is denoted $\langle X, \leq \rangle$. $\langle X, \leq \rangle$ is a **complete lattice** if, for all non-empty $B \subseteq X$, the greatest lower bound $\bigwedge_X B$ and the least upper bound $\bigvee_X B$ exist in X. Let $\langle X, \leq_X \rangle$ be a (complete) lattice and $\langle Y, \leq_Y \rangle$ be a partially ordered set. A function $F : X \to Y$ is **monotone increasing** if $x \leq_X y$ implies $F(x) \leq_Y F(y)$. Let $\mathcal{E}(F) = \{x \in X : x = F(x)\}$ be the set of fixed points of F. See Topkis (1998) for more on these concepts.

Theorem 7 *Let* $\langle X, \preceq \rangle$, *be a complete lattice. If* $F : X \to X$ *is monotone increasing, then* $\langle \mathcal{E}(F), \preceq \rangle$, *is a non-empty complete lattice.*

Proof. See Tarski (1955) or Echenique (2005) for a short and constructive proof. ■

4.4 The Core as a Set of Fixed Points

In this section, we present a formulation that allows us to fully characterize the core as the set of fixed points of a certain function. We assume that $|F| = |W| = n$ to simplify the formulation.[1] We shall also assume that for all $f \in F$ $u_f \in \{0, 1, ..., \max_{w\in W} \alpha_{fw}\}$ and for all $w \in W$ $v_w \in \{0, 1, ..., \max_{f\in F} \alpha_{fw}\}$ to make the payoff space discrete. These assumptions simplify the notation, but all results hold without these assumptions.

We can now proceed to define our formulation. Let Y be the set of possible payoffs such that:

$$Y = \{((u_f)_{f\in F}, (v_w)_{w\in W}) \mid \forall f \in F,\ 0 \leq u_f \leq \max_{w\in W} \alpha_{fw}\ ;\ \forall w \in W,\ 0 \leq v_w \leq \max_{f\in F} \alpha_{fw}\}.$$

Given (u, v), let

$$U_{f_1}(u, v) = \max_{w\in W}(\alpha_{f_1 w} - v_w),$$

$$U_{f_2}(u, v) = \max_{w\in W^2 = W\setminus\{argmax_{w\in W}(\alpha_{f_1 w} - v_w)\}}(\alpha_{f_2 w} - v_w),$$

$$\vdots$$

$$U_{f_n}(u, v) = \max_{w\in W^n}(\alpha_{f_n w} - v_w),$$

$$V_{w_1}(u, v) = \max_{f\in F}(\alpha_{f w_1} - u_f),$$

$$V_{w_2}(u, v) = \max_{f\in F^2 = F\setminus\{argmax_{f\in F}(\alpha_{f w_1} - u_f)\}}(\alpha_{f w_2} - u_f),$$

[1] It is easy to extend the formulation to a more general model that allows for different number of firms and workers. This can be done by adding null firm (or null worker) which represent being unemployed for workers (not having any worker for firms).

$$\vdots$$

$$V_{w_n}(u,v) = \max_{f \in F^n}(\alpha_{fw_n} - u_f).$$

Note that $argmax_{w \in W'}(\alpha_{f_i w} - v_w)$ and $argmax_{f \in F'}(\alpha_{fw_j} - u_f)$ may not be a singleton. Hence, $U(.)$ and $V(.)$ may not be well defined. We use a tie breaking rule to avoid the cases when these sets are not singleton. A tie breaking rule guarantees that $|argmax_{w \in W'}(\alpha_{f_i w} - v_w)| = |argmax_{f \in F'}(\alpha_{fw_j} - u_f)| = 1$ for all $(f,w) \in F' \times W' \subseteq F \times W$. We also want our mapping to be consistent in the sense that it is independent of the sequence of agents in which the final payoffs are realized.

Let $\mathcal{F} \subseteq Y$ be the set of firms consistent payoffs, $\mathcal{W} \subseteq Y$ be the set of workers consistent payoffs, $\mathcal{B} \subseteq Y$ be the set of firms and workers consistent payoffs, and $\mathcal{Z} \subseteq Y$ be the set of inconsistent payoffs. $(u,v) \in \mathcal{F}$ if and only if either there exist a firm-consistent tie breaking rule or for all $f \in F$ $|argmax_{w \in W}(\alpha_{fw} - v_w)| = 1$, such that for all $n \geq 2$, and all $f_n \in F$,

$$argmax_{w \in W}(\alpha_{f_n w} - v_w) \in W^n = W^{n-1} \setminus \{argmax_{w \in W^{n-1}}(\alpha_{f_{n-1} w} - v_w)\}, \qquad (4.1)$$

where $W^1 = W$. This implies $U(.)$ is independent of the order of firms. $(u,v) \in \mathcal{W}$ if and only if either there exist a worker-consistent tie breaking rule or for all $w \in W$ $|argmax_{f \in F}(\alpha_{fw} - u_f)| = 1$, such that for all $n \geq 2$, and all $w_n \in W$,

$$argmax_{f \in F}(\alpha_{fw_n} - u_f) \in F^n = F^{n-1} \setminus \{argmax_{f \in F^{n-1}}(\alpha_{fw_{n-1}} - u_f)\}, \qquad (4.2)$$

where $F^1 = F$. This implies $V(.)$ is independent of the order of workers. $(u,v) \in \mathcal{B}$ if and only if $(u,v) \in \mathcal{F}$ and $(u,v) \in \mathcal{W}$. $(u,v) \in \mathcal{Z}$ if and only if there does not exist any type of consistent tie breaking rule or there is no need for a tie breaking rule, such that neither (4.1) nor (4.2) does hold. Notice that $\mathcal{Z} = Y \setminus \mathcal{F} \cup \mathcal{W}$, $\mathcal{B} \subseteq \mathcal{F}$, and $\mathcal{B} \subseteq \mathcal{W}$.

Lemma 5 *$(u,v) \in \mathcal{B}$ if and only if there exists a tie breaking rule such that $w' = argmax_{w \in W}(\alpha_{f'w} - v_w)$ if and only if $f' = argmax_{f \in F}(\alpha_{fw'} - u_f)$.*

Proof. It is enough to show that the tie breaking rule defined is consistent. Let $(u,v) \in Y$ and suppose that there exists a tie breaking rule such that

$$w' = argmax_{w \in W}(\alpha_{f'w} - v_w) \Leftrightarrow f' = argmax_{f \in F}(\alpha_{fw'} - u_f).$$

This implies there is no $f, f' \in F$, and $f \neq f'$ such that $argmax_{w \in W}(\alpha_{fw} - v_w) = argmax_{w \in W}(\alpha_{f'w} - v_w)$. Moreover, there is no $w, w' \in W$, and $w \neq w'$ such that $argmax_{f \in F}(\alpha_{fw} - u_f) = argmax_{f \in F}(\alpha_{fw'} - u_f)$. Then for all $f_n \in F$, $argmax_{w \in W}(\alpha_{f_n w} - v_w) \in W^n = W^{n-1} \setminus \{argmax_{w \in W^{n-1}}(\alpha_{f_{n-1}w} - v_w)\}$ and for all $w_n \in W$, $argmax_{f \in F}(\alpha_{fw_n} - u_f) \in F^n = F^{n-1} \setminus \{argmax_{f \in F^{n-1}}(\alpha_{fw_{n-1}} - u_f)\}$. Therefore, the tie breaking rule is consistent. Note that we do not need a tie breaking rule if all argmax$\{.\}$ are singleton.

■

Remark 1 *There may be more than one consistent tie breaking rule.*

Now, define a map $T : Y \to Y$ such that if $(u,v) \in \mathcal{F} \setminus \mathcal{B}$ then $T_f(u,v) = U_f(u,v) \vee 0 = (\max_{w \in W}(\alpha_{fw} - v_w)) \vee 0$ for all $f \in F$ and $T_w(u,v) = V'_w(u,v) \vee 0$

for all $w \in W$; if $(u,v) \in \mathcal{W} \setminus \mathcal{B}$ then $T_f(u,v) = U'_f(u,v) \vee 0$ for all $f \in F$ and $T_w(u,v) = V_w(u,v) \vee 0 = (\max_{f\in F}(\alpha_{fw} - u_f)) \vee 0$ for all $w \in W$; if $(u,v) \in \mathcal{B}$ then $T_f(u,v) = U_f(u,v) \vee 0 = (\max_{w\in W}(\alpha_{fw} - v_w)) \vee 0$ for all $f \in F$ and $T_w(u,v) = V_w(u,v) \vee 0 = (\max_{f\in F}(\alpha_{fw} - u_f)) \vee 0$ for all $w \in W$; if $(u,v) \in \mathcal{Z}$ then $U(u,v)$ and $V(u,v)$ are not order independent and hence we use an arbitrary fixed tie breaking rule in which if an agent is indifferent between two agents from other side of the market, the agent prefers the one with a lower index. That is

$$U'_{f_1}(u,v) = \max_{w\in W}(\alpha_{f_1 w} - v_w),$$

$$U'_{f_2}(u,v) = \max_{w\in W^2 = W\setminus\{w_i | i\leq j \ \forall \ w_i, w_j \in argmax_{w\in W}(\alpha_{f_1 w} - v_w)\}}(\alpha_{f_2 w} - v_w),$$

$$\vdots$$

$$U'_{f_n}(u,v) = \max_{w\in W^n}(\alpha_{f_n w} - v_w),$$

$$V'_{w_1}(u,v) = \max_{f\in F}(\alpha_{f w_1} - u_f),$$

$$V'_{w_2}(u,v) = \max_{f\in F^2 = F\setminus\{f_i | i\leq j \ \forall \ f_i, f_j \in argmax_{f\in F}(\alpha_{f w_1} - u_f)\}}(\alpha_{f w_2} - u_f),$$

$$\vdots$$

$$V'_{w_n}(u,v) = \max_{f\in F^n}(\alpha_{f w_n} - u_f).$$

Let $\mathcal{E}(T) = \{(u,v) \in Y : (u,v) = T(u,v)\}$ be the set of fixed points of T. We suppose that $\mathcal{E}(T) \neq \emptyset$ for this section. In Section 4.5, we will prove that the set of fixed points of T is non-empty.

Lemma 6 *If* $(u, v) \in \mathcal{E}(T)$ *then* $(u, v) \in \mathcal{B}$*.*

Proof. Suppose $(u, v) \in \mathcal{F} \setminus \mathcal{B}$. This implies there exists a tie breaking rule which is firm consistent but not worker consistent. Then there is $w_j \in W$ such that $T_{w_j}(u, v) \neq max_{f \in F}(\alpha_{fw_j} - v_f)$. Moreover, there is $f_i, f_k \in F$ such that $f_i = argmax_{f \in F}(\alpha_{fw_j} - u_f) \notin F^j$ and $f_k = argmax_{f \in F^j}(\alpha_{fw_j} - u_f)$. This implies $T_{w_j}(u, v) = \alpha_{f_k w_j} - u_{f_k} < \alpha_{f_i w_j} - u_{f_i} = v_{w_j}$. Therefore, $(u, v) \notin \mathcal{E}(T)$. Now, suppose $(u, v) \in \mathcal{W} \setminus \mathcal{B}$. By reversing the roles of workers with firms in the above argument, we get $(u, v) \notin \mathcal{E}(T)$. The argument for $(u, v) \in \mathcal{Z}$ implies $(u, v) \notin \mathcal{E}(T)$ is also identical to the first step. As a result, $\mathcal{E}(T) \subseteq Y \setminus \{\{\mathcal{F} \cup \mathcal{W} \cup \mathcal{Z}\} \setminus \mathcal{B}\} = \mathcal{B}$. ∎

Remark 2 *The other direction of the above result is not true. The following example shows that there is* $(u, v) \in \mathcal{B}$ *but* $(u, v) \notin \mathcal{E}(T)$*.*

Remark 3 *It is also easy to see that* $S(\Gamma) \subseteq \mathcal{B}$ *using Lemma* 5 *and the definition of core payoffs.*

The following example illustrates the structure of the mapping T.

Example 4 [Shapley-Shubik (1972)]. *Let* $\Gamma = \langle \{f_1, f_2, f_3\}, \{w_1, w_2, w_3\}, \alpha \rangle$ *be an assignment game where* α *is*

	w_1	w_2	w_3
f_1	*5*	*8*	*2*
f_2	*7*	*9*	*6*
f_3	*2*	*3*	*0*

.

In this example $(0,0,0,7,9,6) \in \mathcal{F}\backslash\mathcal{B}$, $(8,9,3,0,0,0) \in \mathcal{W}\backslash\mathcal{B}$, $(3,5,0,2,5,1) \in \mathcal{B}$, and $(0,0,0,7,9,1) \in \mathcal{Z}$. Then, $T(0,0,0,7,9,6) = (0,0,0,7,8,0)$, $T(8,9,3,0,0,0) = (8,7,0,0,0,0)$, $T(3,5,0,2,5,1) = (3,5,0,2,5,1)$, and $T(0,0,0,7,9,1) = (1,0,0,7,8,0)$. Moreover, $(4,5,0,2,4,0) \in \mathcal{B}$ but $(4,5,0,2,4,0) \notin \mathcal{E}(T)$.

Two of our main results can now be stated. The first one (Proposition 5) shows that the core (or stable) payoffs of the assignment game are equal to the set of fixed points of the aforementioned mapping. Note that core outcomes are the Cartesian product of the core payoffs and the set of optimal assignments. The second one (Corollary 3) which directly follows from the first main result, states that the core is equivalent to the Cartesian product of the set of fixed points of T and the set of optimal assignments.

Proposition 5 *$\mathcal{E}(T) = S(\Gamma)$.*

Proof. First we show that $S(\Gamma) \subseteq \mathcal{E}(T)$. Let $(u,v) \in S(\Gamma)$. This implies that there is at least one optimal assignment x (or μ) such that $\sum_{f\in F} u_f + \sum_{w\in W} v_w = \sum_{(f,w)\in F\times W} \alpha_{fw}x_{fw}$ and for all $(f,w) \in F\times W$, $u_f + v_w = \alpha_{fw}$ if f and w are matched and $u_f + v_w \geq \alpha_{fw}$ if f and w are not matched. Moreover, $u_f \geq 0$ for all $f \in F$ and $v_w \geq 0$ for all $w \in W$. Then for a given $f \in F$, $u_f \geq \alpha_{fj} - v_j$ for all $j \in W$ and $\mu(f) \in argmax_{j\in W}(\alpha_{fj} - v_j)$. This implies

$$u_{f_1} = \alpha_{f_1\mu(f_1)} - v_{\mu(f_1)} = \left(\max_{w\in W}(\alpha_{f_1w} - v_w)\right) \vee 0 = U_{f_1}(u,v),$$

$$u_{f_2} = \alpha_{f_2\mu(f_2)} - v_{\mu(f_2)} = \left(\max_{w\in W\backslash\{\mu(f_1)\}}(\alpha_{f_2w} - v_w)\right) \vee 0 = U_{f_2}(u,v),$$

$$\vdots$$

$$u_{f_n} = \alpha_{f_n\mu(f_n)} - v_{\mu(f_n)} = \left(\max_{w\in W\setminus\{\cup_{i\leq(n-1)}\mu(f_i)\}}(\alpha_{f_nw} - v_w)\right) \vee 0 = U_{f_n}(u,v).$$

This implies $u_f = (U_f) \vee 0$ for all $f \in F$. Similar arguments also reveal that, $v_w = (V_w) \vee 0$ for all $w \in W$. Therefore, $T(u,v) = (u,v)$, which implies $(u,v) \in \mathcal{E}(T)$.

Next, we prove that $\mathcal{E}(T) \subseteq S(\Gamma)$. Let $(u,v) \in \mathcal{E}(T)$. This implies $(u,v) \in \mathcal{B}$ by Lemma 6. We first need to show that there exists an assignment x such that it is compatible with (u,v) implying $((u,v);x)$ is a feasible outcome. Let $x_{f_nw'} = 1$ and $x_{f_nw''} = 0$ for all $w'' \neq w'$, where $w' = argmax_{w\in W^n}(\alpha_{f_nw} - v_w) = argmax_{w\in W}(\alpha_{f_nw} - v_w)$. By Lemma 5, $f_n = argmax_{f\in F}(\alpha_{fw'} - u_f)$. Then $u_{f_n} \geq \alpha_{f_nw'} - v_{w'} \geq \alpha_{f_nw''} - v_{w''}$ for all $w'' \neq w'$. Suppose that $u_{f_n} > \alpha_{f_nw'} - v_{w'}$. This is only possible if $u_{f_n} = 0$ and $v_{w'} > \alpha_{f_nw'}$ since $u_{f_n} = \max_{w\in W}(\alpha_{f_nw'} - v_{w'}) \vee 0$. But then $v_{w'} \neq max_{f\in F}(\alpha_{fw'} - u_f) \vee 0 = \alpha_{f_nw'}$, contradicting the assumption that (u,v) is a fixed point. This implies $u_{f_n} + v_{w'} = \alpha_{f_nw'}$ and $u_{f_n} + v_{w''} \geq \alpha_{f_nw''}$ for all $w'' \neq w'$ when $x_{f_nw'} = 1$. Define x by repeating the process for all $f \in F \setminus \{f_n\}$. Observe that x is an assignment by construction. Then, $\sum_{f\in F} u_f + \sum_{w\in W} v_w = \sum_{(f,w)\in F\times W} \alpha_{fw}x_{fw}$, which implies that $((u,v);x)$ is a feasible outcome.

Now, we want to show that the payoff (u,v) with the assignment x is stable (i.e., x is an optimal assignment). First note that by the definition of fixed point $u_f \geq 0$ for all $f \in F$ and $v_w \geq 0$ for all $w \in W$. Hence individual rationality is satisfied. We already showed that if $x_{fw} = 1$, $u_f + v_w = \alpha_{fw}$ and for all $f' \neq f$, $u_{f'} + v_w \geq \alpha_{f'w}$ and all $w' \neq w$, $u_f + v_{w'} \geq \alpha_{fw'}$. This implies that $u_f + v_w \geq \alpha_{fw}$ for all $(f,w) \in F \times W$. Therefore $(u,v) \in S(\Gamma)$. ■

Corollary 3 *$\mathcal{E}(T) \times \mathcal{X} = S(\Gamma) \times \mathcal{X} = C(\Gamma)$.*

Proof. The result directly follows from Proposition 5 and the definition of core. ■

Remark 4 *The set of optimal assignments is constructed using the following process. Let $(u, v) \in S(\Gamma) \subseteq B$. Then there is a consistent tie breaking rule such that $w' = argmax_{w \in W}(\alpha_{f'w} - v_w)$ if and only if $f' = argmax_{f \in F}(\alpha_{fw'} - u_f)$. Let $\mu(w') = f'$ and $\mu(f') = w'$. Then, $\mu \in \mathcal{X}$. Note that there may be more than one consistent tie breaking rule such that $w'' = argmax_{w \in W}(\alpha_{f''w} - v_w)$ if and only if $f'' = argmax_{f \in F}(\alpha_{fw''} - u_f)$. We need to repeat the process for all such tie breaking rules to reach the set of optimal assignments.*

4.5 The Lattice Structure of Stable Payoffs

In this section we show that $\mathcal{E}(T)$ is non-empty and forms a complete lattice. We introduce a partial order on Y such that T is a monotone increasing map. Tarski's fixed point theorem then leads to a lattice structure on $\mathcal{E}(T)$, and thus on $S(\Gamma)$. First we define the following binary relations on Y.

Definition 4 *Let $(u, v) \in Y$.*

(i) Define a binary relation $\geq_F$ by $(u, v) \geq_F (u', v') \Leftrightarrow u \geq u'$.

(ii) Define a binary relation $\geq_W$ by $(u, v) \geq_W (u', v') \Leftrightarrow v \geq v'$.

(iii) Define a partial ordering $\succeq_F$ by

$$(u, v) \succeq_F (u', v') \Leftrightarrow (u, v) \geq_F (u', v') \text{ and } (u', v') \geq_W (u, v).$$

(iv) Define a partial ordering $\succeq_W$ by

$$(u, v) \succeq_W (u', v') \Leftrightarrow (u, v) \geq_W (u', v') \text{ and } (u', v') \geq_F (u, v).$$

The following lemma shows that T is a monotone increasing function under both partial orders, $\succeq_F$ and $\succeq_W$.

Lemma 7 *For all $(u, v), (u', v') \in Y$, $(u, v) \succeq_F (u', v')$ implies $T(u, v) \succeq_F T(u', v')$ and $(u, v) \succeq_W (u', v')$ implies $T(u, v) \succeq_W T(u', v')$.*

Proof. Let $(u, v) \succeq_F (u', v')$. We need to show that $T(u, v) \succeq_F T(u', v')$. Since $(u, v) \succeq_F (u', v')$, $u_f \geq u'_f$ for all $f \in F$ and $v'_w \geq v_w$ for all $w \in W$. Then for all $f_i \in F$,

$$\begin{aligned} T_{f_i}(u, v) &= \max_{w \in W^i} (\alpha_{f_i w} - v_w) \vee 0 \\ &\geq \max_{w \in W^i} (\alpha_{f_i w} - v'_w) \vee 0 = T_{f_i}(u', v'). \end{aligned}$$

Therefore, $T(u, v) \geq_F T(u', v')$. Similarly, $T(u', v') \geq_W T(u, v)$. Hence $T(u, v) \succeq_F T(u', v')$. The proof of $(u, v) \succeq_W (u', v')$ implies $T(u, v) \succeq_W T(u', v')$ is similar. ■

The following results show that the set of core payoffs forms a complete lattice under the partial orders $\succeq_F$ and $\succeq_W$. We can prove them by applying Tarski's fixed point theorem.

Proposition 6 *$\langle \mathcal{E}(T), \succeq_F \rangle$ is a non-empty complete lattice.*

Proof. We showed in Lemma 5 that T is monotone increasing with respect to $\succeq_F$. Moreover, $\langle Y, \geq_F \rangle$ and $\langle Y, \geq_W \rangle$ are complete lattices since Y is closed and bounded. Then $\langle Y, \succeq_F \rangle$ is also a complete lattice since it is a product set endowed with a product order. Tarski's fixed point theorem (Theorem 7) implies that $\mathcal{E}(T) \neq \emptyset$ and $\langle \mathcal{E}(T), \succeq_F \rangle$ is a complete lattice. ■

Proposition 7 *$\langle \mathcal{E}(T), \succeq_W \rangle$ is a non-empty complete lattice.*

Proof. Omitted. The proof is similar to the proof of Proposition 6. ■

The above propositions (Proposition 6 and 7) also imply that there is a polarization of interest in the core. That is, there is a F-optimal stable payoff $(\bar{u}, \underline{v})$ that is simultaneously the best for all firms and the worst for all workers, and opposite is true for W-optimal stable payoff $(\underline{u}, \overline{v})$. This type of polarization seems to be a general property of two-sided matching markets.

The following lemma about the structure of the core is useful for the next section.

Lemma 8 *Let $(u, v) \in \mathcal{E}(T)$,$(u', v') \in \mathcal{E}(T)$, and $(u, v) \succeq_F (u', v')$. If $u_f - u'_f = t$ then there is $w \in W$ such that $v'_w - v_w = t$.*

Proof. By Lemma 6, $(u, v) \in B$ and $(u', v') \in B$. Then T is order independent. $T_f(u, v) = \max_{w \in W}(\alpha_{fw} - v_w) \vee 0 = u_f$ and $T_f(u', v') = \max_{w \in W}(\alpha_{fw} - v'_w) \vee 0 = u'_f = u_f - t$. This implies $u_f = \max_{w \in W}(\alpha_{fw} - v'_w + t)$. Therefore, there is $w \in W$ such that $v_w = v'_w - t$. ■

4.6 The Algorithm

The T-algorithm is very simple and uses our formulation. It starts at some $(u, v) \in Y$ and iterate $T(u, v)$ until two iterations are identical. The algorithm stops when two iterations are identical. We prove that when the algorithm stops, it must be at a stable payoff. Moreover, we show that all stable payoffs can be reached through the algorithm.

T-algorithm:

1. Set $(u^0, v^0) = (u, v)$. Set $(u^1, v^1) = T(u^0, v^0)$ and $k = 1$.

2. While $(u^k, v^k) \neq (u^{k-1}, v^{k-1})$, do:

(a) set $k = k + 1$

(b) set $(u^k, v^k) = T(u^{k-1}, v^{k-1})$.

3. Set $\tau = (u^k, v^k)$. Stop.

Proposition 8 *If the T-algorithm stops at $\tau \in Y$, then τ is a stable payoff and there is an optimal assignment $x \in \mathcal{X}$ such that $(\tau; x)$ is in the core. If (u^k, v^k) is in the set of stable payoffs, for some iteration k of the T-algorithm, then the algorithm stops at $\tau = (u^k, v^k)$.*

Proof. If the algorithm stops at $\tau \in Y$, then $(u^k, v^k) = (u^{k-1}, v^{k-1}) = \tau$. Then, $\tau = T(u^{k-1}, v^{k-1}) = T(\tau)$, so $\tau \in \mathcal{E}(T)$. By Proposition 1, $\tau \in S(\Gamma)$. Moreover, by Corollary 3, there is an optimal assignment x such that $(\tau; x) \in C(\Gamma)$. To prove the second part, observe that if (u^k, v^k) is a stable payoff, then (u^k, v^k) is a fixed point of T by Proposition 5. Then the algorithm stops at $\tau = (u^k, v^k)$. ■

We now provide the second algorithm to find all core payoffs. Let

$$(\bar{u}_Y, \underline{v}_Y) = (\max_{w \in W} \alpha_{f_1 w}, ..., \max_{w \in W} \alpha_{f_n w}, 0, ..., 0),$$

$$(\underline{u}_Y, \bar{v}_Y) = (0, ..., 0, \max_{f \in F} \alpha_{f w_1}, ..., \max_{f \in F} \alpha_{f w_n}).$$

Moreover, let e_l^n be the l-th unit vector in $\mathbb{R}^n$, i.e. $e_l^n = (0, ...1, 0, ..., 0) \in \mathbb{R}^n$, where 1 is the l-th element of e_l^n.

Algorithm 2:

1. Set $(u^0, v^0) = (\bar{u}_Y, \underline{v}_Y)$. Set $(u^1, v^1) = T(u^0, v^0)$ and $k = 1$.

2. While $(u^k, v^k) \neq (u^{k-1}, v^{k-1})$, do:

(a) set $k = k + 1$

(b) set $(u^k, v^k) = T(u^{k-1}, v^{k-1})$.

3. Set $\tau = (u^k, v^k)$.

4. Let $\hat{\mathcal{E}} = \tau$. The possible states of the algorithm is Y. Start at state Ω^0 where

$$\Omega^0 = \{(\bar{u}_Y \wedge u^k + e_l^n, 0 \vee v^k - e_m^n), (0 \vee u^k - e_l^n, \bar{v}_Y \wedge v^k + e_m^n)\} \subset Y$$

for all $1 \leq l, m \leq n$. Let the state of the algorithm be Ω. While $\Omega' \neq \emptyset$ do the following subroutine to get a new state Ω'. Then set $\Omega = \Omega'$.

SUBROUTINE: Let $\Omega' = \emptyset$. For each $(u, v) \in \Omega$, run $T(u, v)$. If $T(u, v) = (u, v)$ add (u, v) to $\hat{\mathcal{E}}$ and add $\{(\bar{u}_Y \wedge u + e_l^n, 0 \vee v - e_m^n), (0 \vee u - e_l^n, \bar{v}_Y \vee v + e_m^n)\} \setminus \hat{\mathcal{E}}$ for all $1 \leq l, m \leq n$ to Ω'.

Theorem 8 *The set $\hat{\mathcal{E}}$ produced by Algorithm 2 coincides with the core payoffs $S(\Gamma)$ of the assignment games.*

Proof. First I prove that the algorithm reaches a fixed point after a finite k number of iterations. Then, we know that $\tau = (u^k, v^k) \in S(\Gamma)$ by Proposition 8. Then I show that $\hat{\mathcal{E}} \subseteq S(\Gamma)$, and then $S(\Gamma) \subseteq \hat{\mathcal{E}}$.

We want to show that the first part of Algorithm 2, T-algorithm, reaches a fixed point. That is for some finite k, $\tau = (u^k, v^k) = (u^{k-1}, v^{k-1})$. Assume this does not hold for any k. Then, $\{(u^k, v^k)\}$ is an infinite sequence of distinct payoffs in Y. However, there exists a finite number of payoffs that is for all $f \in F$ $u_f \in \{0, 1, ..., \max_{w \in W} \alpha_{fw}\}$ and for all $w \in W$ $v_w \in \{0, 1, ..., \max_{f \in F} \alpha_{fw}\}$, contradicting to the initial assumption. This implies there is $k < \infty$ such that T-algorithm reaches a fixed point.

Now we show that the rest of Algorithm 2 stops after a finite number of steps. Let $M \subseteq Y$ be the collection of states visited by the algorithm. Let $d^1(\Omega)$, where $\Omega \subseteq M$, be the minimum of the Euclidean distance between payoffs in Ω and $(\bar{u}_Y, \underline{v}_Y)$ and $d^2(\Omega)$ be the minimum of the Euclidean distance between payoffs in Ω and $(\underline{u}_Y, \bar{v}_Y)$. If $\Omega = \emptyset$, let $d^1(\Omega) = d^2(\Omega) = 0$. We consider $d^1(\Omega)$ and $d^2(\Omega)$ because if the state is $\{(\bar{u}_Y, \underline{v}_Y), (\underline{u}_Y, \bar{v}_Y)\}$, $\{(\bar{u}_Y, \underline{v}_Y)\}$, or $\{(\underline{u}_Y, \bar{v}_Y)\}$ the next state is $\emptyset$ by the definition of the subroutine. Let Ω' and Ω'' be successive states in the algorithm. It is clear from the definition that $d^1(\Omega') > d^1(\Omega'')$ and $d^2(\Omega') > d^2(\Omega'')$. Since M is a finite set, $d^1(.)$ and $d^2(.)$ takes only a finite number of values. Thus after a finite number of steps the algorithm stops, i.e., $\Omega = \emptyset$.

$\hat{\mathcal{E}} \subseteq S(\Gamma)$. Let $(u, v) \in \hat{\mathcal{E}}$. This implies $(u, v) = T(u, v)$ by the definition of the algorithm and hence $(u, v) \in \mathcal{E}(T)$. By Proposition 5, $\mathcal{E}(T) = S(\Gamma)$. Therefore $(u, v) \in S(\Gamma)$ which proves $\hat{\mathcal{E}} \subseteq S(\Gamma)$.

$S(\Gamma) \subseteq \hat{\mathcal{E}}$. Let $(u, v) \in S(\Gamma) = \mathcal{E}(T)$. Suppose, by way of contradiction, that $(u, v) \notin \hat{\mathcal{E}}$. This implies $\tau = (u^k, v^k) \neq (u, v)$ and $(u, v) \notin M$ so that the algorithm's states does not contain (u, v). Then either $\tau \succeq_F (u, v)$ or $\tau \preceq_F (u, v)$. Suppose, without loss of generality, $\tau \succeq_F (u, v)$ and $\max_{f \in F}(u^k_f - u_f) = t$. By Lemma 8, there is $w \in W$ such that $v_w - v^k_w = t$. Now we show that $\{(\bar{u}_Y \wedge u + e^n_f, 0 \vee v - e^n_g)\} \notin M$ for all $1 \leq f, g \leq n$. Suppose this is not the case. Then there is a state Ω^c of the algorithm and $a, b \in [1, n]$ such that $(\bar{u}_Y \wedge u + e^n_a, 0 \vee v - e^n_b) \in \Omega^c \subseteq M$. This is only possible if (u, v) is in the previous state $\Omega^{c-1} \subseteq M$ by the definition of the subroutine; a contradiction since we assumed that $(u, v) \notin M$. Using the same argument, we can also conclude that for all $1 \leq h, k \leq n$ $\{(\bar{u}_Y \wedge u + e^n_f + e^n_h, 0 \vee v - e^n_g - e^n_k)\} \notin M$. Repeating the same argument $t-1$ times implies $(\bar{u}_Y \wedge u^k - e^n_l, 0 \vee v^k + e^n_g) \notin M$, which is a contradiction since we have shown that there is $\tau = (u^k, v^k) \in \hat{\mathcal{E}}$ and $(\bar{u}_Y \wedge u^k - e^n_l, 0 \vee v^k + e^n_g) \in \Omega^0 \subseteq M$. This implies $(u, v) \in M$ and hence $(u, v) \in \hat{\mathcal{E}}$. The case where $\tau \preceq_F (u, v)$ is also similar. ■

Now we use Example 4 to show the details of the algorithm. Algorithm 2 starts at $(u^0, v^0) = (8, 9, 3, 0, 0, 0)$ and does $T(8, 9, 3, 0, 0, 0) = (8, 7, 0, 0, 0, 0)$, $T(8, 7, 0, 0, 0, 0) = (8, 7, 0, 2, 2, 0)$, $T(8, 7, 0, 2, 2, 0) = (6, 6, 0, 2, 2, 0)$, $T(6, 6, 0, 2, 2, 0) = (5, 6, 0, 2, 3, 0)$, $T(5, 6, 0, 2, 3, 0) = (5, 6, 0, 2, 3, 0)$. This implies $\tau = (5, 6, 0, 2, 3, 0)$. Now,

$$\Omega^0 = \{(6, 6, 0, 1, 3, 0), (6, 6, 0, 2, 2, 0), (6, 6, 0, 2, 3, 0), (5, 7, 0, 1, 3, 0), (5, 7, 0, 2, 2, 0),$$

$$(5, 7, 0, 2, 3, 0), (5, 6, 1, 1, 3, 0), (5, 6, 1, 2, 2, 0), (5, 6, 1, 2, 3, 0), (4, 6, 0, 3, 3, 0),$$

$$(4, 6, 0, 2, 4, 0), (4, 6, 0, 2, 3, 1), (5, 5, 0, 3, 3, 0), (5, 5, 0, 2, 4, 0), (5, 5, 0, 2, 3, 1),$$

$$(5, 6, 0, 3, 3, 0), (5, 6, 0, 2, 4, 0), (5, 6, 0, 2, 3, 1)\}.$$

Note that for all $(u, v) \in \{(5,6,1,1,3,0),(4,6,0,2,4,0)\} \subset \Omega^0$, $T(u, v) = (u, v)$. Then add $\{(5,6,1,1,3,0),(4,6,0,2,4,0)\}$ to $\hat{\mathcal{E}}$. The new state is

$$\Omega = \{(5,6,1,1,3,0) + (e_l^3, -e_m^3), (5,6,1,1,3,0) + (-e_l^3, +e_m^3),$$
$$(4,6,0,2,4,0) + (e_l^3, -e_m^3), (4,6,0,2,4,0) + (-e_l^3, e_m^3)\} \setminus \{(5,6,0,2,3,0)\}.$$

For all $(u, v) \in \{(4,6,1,1,4,0),(4,5,0,2,4,1),(3,6,0,2,5,0)\} \subset \Omega$, $T(u, v) = (u, v)$. Then add $\{(4,6,1,1,4,0),(4,5,0,2,4,1),(3,6,0,2,5,0)\}$ to $\hat{\mathcal{E}}$. The new state is

$$\Omega' = \{(4,6,1,1,4,0) + (e_l^3, -e_m^3), (4,6,1,1,4,0) + (-e_l^3, +e_m^3), (4,5,0,2,4,1) + (e_l^3, -e_m^3),$$
$$(4,5,0,2,4,1) + (-e_l^3, +e_m^3), (3,6,0,2,5,0) + (e_l^3, -e_m^3), (3,6,0,2,5,0) + (-e_l^3, +e_m^3)\} \setminus \hat{\mathcal{E}}.$$

It is only the case that for $(3,5,0,2,5,1) \in \Omega'$, $T(3,5,0,2,5,1) = (3,5,0,2,5,1)$. Then add $(3,5,0,2,5,1)$ to $\hat{\mathcal{E}}$. The new state is

$$\Omega'' = \{(3,5,0,2,5,1) + (e_l^3, -e_m^3), (3,5,0,2,5,1) + (-e_l^3, +e_m^3)\} \setminus \hat{\mathcal{E}}.$$

Note that there is not any $(u, v) \in \Omega''$ such that $T(u, v) = (u, v)$. Then the new state is $\emptyset$. This implies the algorithm stops and the core of the assignment game is

$$\hat{\mathcal{E}} = \{(5,6,0,2,3,0),(5,6,1,1,3,0),(4,6,0,2,4,0),(4,6,1,1,4,0),(4,5,0,2,4,1),$$
$$(3,6,0,2,5,0),(3,5,0,2,5,1)\}.$$

4.7 Formulation with Core Outcomes Is Not Possible

It would be nice to find a construction such that fixed points will directly provide the core outcomes. However unlike assignment literature without money, it is not possible to work with core outcomes in this setup. In the rest of this section, we define a reasonable construction which can work with outcomes. Then we provide examples to show that this type of formulation is not possible.

Let π be a **pre-assignment** if $\pi : F \cup W \to F \cup W$ such that $\pi(f) \in W \cup \{f\}$ for all $f \in F$, and $\pi(w) \in F \cup \{w\}$ for all $w \in W$. Let Π be the set of all pre-assignment vectors. Define a map $T' : Y \times \Pi \to Y \times \Pi$ such that

$$T'_f((u,v);\pi(f)) = ((\max U_f(u,v)) \vee 0; w) \textit{ where } w \in argmax\ (\alpha_{fw} - v_w) \ \ \forall f \in F,$$

and

$$T'_w((u,v);\pi(w)) = ((\max V_w(u,v); f)) \vee 0 \textit{ where } f \in argmax\ (\alpha_{fw} - u_f) \ \ \forall w \in W.$$

Then we could show that the fixed points of T' are equivalent to the core. However, this type of formulation is not possible in this framework since there might be more than one optimal assignment and different (pre)assignments might correspond to same payoffs. Then, fixed point of T' may fail to induce an assignment. Moreover, proving the existence of a fixed point is problematic. On the other hand, by using our formulation core payoffs can always be found, and core outcomes will be equal to the Cartesian product of the fixed points and the set of optimal assignments which is constructed.

Example 5 *[Shapley-Shubik (1972)] Let* $\Gamma = \langle \{f_1, f_2, f_3\}, \{w_1, w_2, w_3\}, \alpha \rangle$ *be an assignment game where* α *is*

	w_1	w_2	w_3
f_1	*0*	*2*	*0*
f_2	*2*	*0*	*2*
f_3	*0*	*2*	*0*

.

There are four optimal assignments given by

$$\mathcal{X} = \{(0,1,0;1,0,0;0,0,1), (0,1,0;0,0,1;1,0,0), (0,0,1;1,0,0;0,1,0),$$

$$(1,0,0;0,0,1;0,1,0)\}$$

with value $\sum_{(f,w) \in F \times W} \alpha_{fw} x_{fw} = 4$. The core of the game is given by

$$C(\Gamma) = (0,2,0,0,2,0) \times \mathcal{X}.$$

Moreover, $((0,2,0,0,2,0);\pi)$ where $\pi(f_1) = w_1, \pi(f_2) = w_3, \pi(f_3) = w_2, \pi(w_1) = f_3, \pi(w_2) = f_1, \pi(w_3) = f_3$ is a fixed point of T' with appropriate tie breaking rule but π is not an assignment. Hence $((0,2,0,0,2,0);\pi) \notin C(\Gamma)$.

The following example shows that a construction like T' will not work even though there is a unique optimal assignment.

Example 6 (The same as Example 4). *Let* $\Gamma' = \langle \{f_1, f_2, f_3\}, \{w_1, w_2, w_3\}, \alpha' \rangle$ *be an assignment game where* α' *is*

	w_1	w_2	w_3
f_1	*5*	*8*	*2*
f_2	*7*	*9*	*6*
f_3	*2*	*3*	*0*

.

There is one optimal assignment given by

$$\mathcal{X} = \{(0, 1, 0; 0, 0, 1; 1, 0, 0)\}$$

with value $\sum_{(f,w)\in F\times W} \alpha_{fw} x_{fw} = 16$. It is easy to see that $(3, 5, 0, 2, 5, 1) \in S(\Gamma')$. Moreover, $((3, 5, 0, 2, 5, 1); \pi)$ where $\pi(f_1) = w_1, \pi(f_2) = w_3, \pi(f_3) = w_2, \pi(w_1) = f_3, \pi(w_2) = f_1, \pi(w_3) = f_2$ is a fixed point of T' with appropriate tie breaking rule but π is not an assignment. Hence $((3, 5, 0, 2, 5, 1); \pi) \notin C(\Gamma')$.

Note that we only work on the lattice structure of core payoffs since different optimal assignments may correspond to same payoffs (See Example 5). Hence, it is not possible to construct a binary relation which is antisymmetric on $\mathcal{E}(T) \times \mathcal{X}$ ($\succeq_F$ *or* $\succeq_W$ is not a partial order on the set of core outcomes.)

4.8 Final Remarks

The paper presents an alternative way to formulate the core of assignment games as the fixed points of a certain mapping, via a powerful algebraic fixed point theorem of Tarski.

In our formulation, we work with payoffs and construct optimal assignments rather than directly working with outcomes. The main reason for that is different (pre)assignments might lead to a same payoff structure and the mapping defined on feasible outcomes may fail to induce an assignment. Moreover, defining a partial order on the Cartesian product of the payoffs and (pre)assignments is a problem. Such a formulation (if it is not impossible) which works also with outcomes, seems to be an important follow-up to our work.

The extension of this new formulation to many-to-one and many-to-many assignment games will be a subject of our future work. Sotomayor (1999) showed that the results in one-to-one assignment game can be extended to these more complex assignment models. Therefore, it might be possible to extend the techniques developed in this paper to any such markets.

Bibliography

[1] Adachi, H. (2000), "On a Characterization of Stable Matchings," *Economic Letters* 68, 43–49.

[2] Alcalde, J. (1996), "Implementation of Stable Solutions to Marriage Problems," *Journal of Economic Theory* 69, 240–254.

[3] Alcalde, J., and A. Romero-Medina (2000), "Simple Mechanisms to Implement the Core of College Admissions Problems," *Games and Economic Behavior* 31, 294–302.

[4] Dasgupta, P., and E. Maskin (2000), "Efficient Auctions," *Quarterly Journal of Economics*, 115, 341–388.

[5] Dutta, B., and J. Massó (1997), "Stability of Matchings When Individuals Have Preferences over Colleagues," *Journal of Economic Theory* 75, 464–475.

[6] Echenique, F. (2005), "A Short and Constructive Proof of Tarski's Fixed-Point Theorem," *International Journal of Game Theory* 33, 215–218.

[7] Echenique, F., and J. Oviedo (2004), "Core many-to-one matchings by fixed point methods," *Journal Economic Theory* 115, 358–376.

[8] Echenique, F., and M. Yenmez (2007), "A Solution to Matching with Preferences Over Colleagues," *Games and Economic Behavior*, 59, 46–71.

[9] Fehr, E., and K. Schmidt (1999), "A Theory of Fairness, Competition and Cooperation," *Quarterly Journal of Economics* 114, 817–868.

[10] Fehr, E., and K. Schmidt (2006), "The Economics of Fairness, Reciprocity and Altruism: Experimental Evidence and New Theories," in S. Kolm, J.M. Ythier, eds., *Handbook of the Economics of Giving, Altruism and Reciprocity*: *Volume 1*, *Foundations*, 615–691.

[11] Figueroa, N., and V. Skreta (2006), "The Role of Outside Options in Auction Design," *Discussion Paper*, University of California, Los Angeles.

[12] Gale, D., and M. Sotomayor (1985), "Ms. Machiavelli and The Stable Matching Problem," *American Mathematical Monthly* 92, 261–268.

[13] Gresik, T. (1991), "Ex Ante Incentive Efficient Trading Mechanisms without the Private Valuation Restriction," *Journal of Economic Theory*, 55, 41–63.

[14] Holmstrom, B., and R.B. Myerson (1983), "Efficient and Durable Decision Rules with Incomplete Information," *Econometrica* 51, 1799–1819.

[15] Jackson, M (2003), "Mechanism Theory," *The Encyclopedia of Life Support Systems*, EOLSS Publishers Co. Ltd, United Kingdom.

[16] Jehiel, P., and B. Moldovanu (2001), "Efficient Design with Interdependent Values," *Econometrica* 69, 1237–1259.

[17] Jehiel, P., B. Moldovanu, and E. Stacchetti (1999), "Multidimensional Mechanism Design for Auctions with Externalities," *Journal of Economic Theory* 85, 258–293.

[18] Kara, T., and T. Sönmez (1995), "Implementation of College Admission Rules," *Economic Theory* 9, 197–218.

[19] Krishna, V., and M. Perry (1998), "Efficient Mechanism Design," *Discussion Paper*, Pennsylvania State University.

[20] Laussel, D., and T.R. Palfrey (2003), "Efficient Equilibria in the Voluntary Contributions Mechanism with Private Information," *Journal of Public Economic Theory* 5, 449–478.

[21] Ledyard, J.O. (1995), "Public Goods: A Survey of Experimental Research," in John Kagel and Alvin Roth, eds., *Handbook of Experimental Economics*, Princeton University Press, Princeton, New Jersey, 111–194.

[22] Ledyard, J.O., and T.R. Palfrey (2007), "A General Characterization of Interim Efficient Mechanisms for Independent Linear Environments," *Journal of Economic Theory* 133, 441–466.

[23] Mezzetti, C. (2004), "Mechanism Design with Interdependent Valuations: Efficiency," *Econometrica* 72, 1617–1626.

[24] Milgrom, P., and R. Weber (1982), "A Theory of Auctions and Competitive Bidding," *Econometrica* 50, 1089–1122.

[25] Myerson R., (1981): “Optimal Auction Design,” *Mathematics of Operations Research* 6, 58–73.

[26] Myerson R., and M. Satterthwaite (1983), “Efficient Mechanisms for Bilateral Trading,” *Journal of Economic Theory* 28, 265–281.

[27] Perez-Castrillo, D. (1994), “Cooperative Outcomes Through Non-Cooperative Games,” *Games and Economic Behavior* 7, 428–440.

[28] Perry, M., and P. Renny (1994), “A Non-Cooperative View of Coalition Formation and The Core,” *Econometrica* 62, 795–817.

[29] Perry, M., and P.J. Reny (2002), “An Efficient Auction,” *Econometrica* 70, 1199–1212.

[30] Pycia, M. (2007), “Many-to-One Matching with Complementarities and Peer Effects,” *Working Paper*, Pennsylvania State University.

[31] Revilla, P. (2004), “Many-to-One Matching When Colleagues Matter,” *Working Paper*, Pennsylvania State University.

[32] Rochet, J.C. (1987), “A Necessary and Sufficient Condition for Rationalizability in a Quasi-Linear Context,” *Journal of Mathematical Economics* 16, 191–200.

[33] Rochford, S.C. (1984), “Symmetrically Pairwise-Bargained Allocations in an Assignment Market,” *Journal of Economic Theory* 34, 262–281.

[34] Roth, A.E. (1984), "Misrepresentation and Stability in The Marriage Problem," *Journal of Economic Theory* 34, 383–387.

[35] Roth, A.E., and M. Sotomayor (1990), "Two sided matching: A Study in Game-Theoretic Modelling and Analysis," Econometric Society Monograph Series, Cambridge University Press, Cambridge.

[36] Roth, A.E., and M. Sotomayor (1988), "Interior Points in the Core of Two-Sided Matching Markets," *Journal of Economic Theory* 45, 85–101.

[37] Serrano, R. (1995), "A Market to Implement The Core," *Journal of Economic Theory* 67, 285–294.

[38] Serrano, R., and R. Vohra (1997), "Non-cooperative Implementation of the Core," *Social Choice and Welfare* 14, 513–525.

[39] Shapley, L.S., and J. Shubik (1972), "The Assignment Game I: The Core," *International Journal of Game Theory* 1, 111–130.

[40] Sotomayor, M. (1999), "The Lattice Structure of the Set of Stable Outcomes of the Multiple Partners Assignment Game," *International Journal of Game Theory* 28, 567–583.

[41] Sotomayor, M. (2003), "Reaching The Core of The Marriage Market Through a Non-Revelation Matching Mechanism," *International Journal of Game Theory* 32, 241–251.

[42] Sotomayor, M. (2004), "Implementation in the many-to-many matching market," *Games and Economic Behavior* 46, 199–212.

[43] Tarski, A. (1955), "A Lattice Theoretical Fixed Point Theorem and Its Applications," *Pacific Journal of Mathematics* 5, 285–309.

[44] Thomson, W. (2005), "Divide and permute," *Games and Economic Behavior* 52, 186–200.

[45] Topkis, D.M. (1998), "*Supermodularity and Complementarity*," Princeton University Press, Princeton, New Jersey.

[46] Wilson, R. (1985), "Incentive Efficiency of Double Auctions," *Econometrica* 53, 1101–1116.