Essays on Microeconomic Theory

by

Qinggong Wu

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2016

Doctoral Committee:

Professor Tilman Börgers, Co-Chair
Professor Stephan Lauermann, University of Bonn, Co-Chair
Professor Kfir Eliaz, Tel Aviv University
Assistant Professor David Miller
Professor Scott Page

ACKNOWLEDGEMENTS

My ancient Chinese ancestors believed that a great work was due more to divine inspiration than to its earthly deliverer. The work presented here, albeit laying no claim to greatness, has likewise been nourished by many an inspiring mind. In bringing it to fruition I am greatly indebted to my Ph.D. advisors Tilman Börgers and Stephan Lauermann for their tireless mentoring and encouragement, and also to the other members on my Ph.D. committee, namely Kfir Eliaz, David Miller and Scott Page, whose guidances have rendered the work much more wholesome than it could have become. The rest of the long list of intellectual creditors are separately acknowledged in each chapter. Special thanks with love to my wife Haosi, grace of my life, whose experience of the travail entailed by the work is no less profound than mine.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF APPENDICES

Appendix

ABSTRACT

Essays on Microeconomic Theory
by
Qinggong Wu

Chairs: Tilman Börgers and Stephan Lauermann

The present work collects three essays on microeconomic theory.

In the first essay, I study a model in which a finite number of men and women look for future spouses via random meetings. I ask whether equilibrium marriage outcomes are stable matchings when search frictions are small. The answer is they can but need not be. For any stable matching there is an equilibrium leading to it almost surely. However unstable—even Pareto-dominated—matchings may still arise with positive probability. In addition, inefficiency due to delay may remain significant despite vanishing search frictions. Finally, a condition is identified under which all equilibria are outcome equivalent, stable, and efficient.

In the second essay, a joint work Kfir Eliaz, we model a competition between two teams as an all-pay auction with incomplete information. The teams may differ in size and individuals exert effort to increase the performance of one's own team via an additively separable aggregation function. The team with a higher performance wins, and its members enjoy the prize as a public good. The value of the prize is identical to members of the same team but is unknown to the other team. We show that there exists a unique monotone equilibrium in which everyone actively participates, and in this equilibrium a bigger team is more likely to win if the aggregation function is concave, less likely if convex, or equally likely if linear.

In the third essay, I study a situation in which a group of people working on a common objective want to share information. Oftentimes information sharing via precise communication is impossible and instead information is aggregated by institutions within which communication is coarse. The paper proposes a unified framework for modeling a general class of such information-aggregating institutions. Within this class, it is shown that institution A outperforms institution B for any common objective if and only if the underlying communication infrastructure of A can be obtained from that of B by a sequence of elementary operations. Each operation either removes redundant communication instruments from B or introduces effective ones to it.

CHAPTER I

A Finite Decentralized Marriage Market with Bilateral Search

1.1 Introduction

The stable matching is the main solution concept for cooperative two-sided matching problems under nontransferable utility. Many centralized mechanisms are designed to implement stable matchings.[1] However, whether outcomes of decentralized two-sided matching markets correspond to stable matchings remains unclear. The present paper addresses this question by considering a decentralized two-sided matching market modeled as a search and matching game. Following Gale and Shapley (1962) I inherit the interpretation that the game represents the situation in which unmarried men and women gather in a marketplace to look for future spouses. The game starts with an initial market à la Gale-Shapley, henceforth referred to as a marriage market, consisting of finitely many men and women with heterogeneous preferences. In every period a meeting between a randomly selected pair of a man and a woman takes place, during which they sequentially decide whether to marry each other. Mutual agreement leads to marriage. Married couples leave the game. Disagreement leads to separation. Separated people continue searching. The game ends when no mutually acceptable pairs of a man and a woman are left. Search is costly due to frictions parametrized as a common discount factor that diminishes the value of a future marriage. A game outcome, reflecting who has married whom and who stays single, corresponds to a matching for the initial market. The central question addressed in the paper is whether matchings that obtain in equilibria are stable matchings for the initial market when search frictions are small. The analysis focuses on a near-frictionless setting in order to test the general conjecture that if in a decentralized market the participants have easy access to each other with low costs then equilibrium outcomes would be in the core of the underlying market.[2] The paper shows that the answer to the central question is indeterminate at best and No in general, in contrast to what has been

[1]See Roth and Sotomayor (1990) for a survey of the theory of the stable matching. Roth (2008) surveys applications in designing centralized mechanisms for two-sided matching markets.

[2]It is well known that the core of the marriage market is the set of all stable matchings.

conjectured on this matter.[3] First, for any stable matching there is an equilibrium leading to that matching almost surely (Proposition 1.5.2), that is, every player expects to marry according to the pairing scheme implied by the matching. This result establishes that the set of all stable matchings is contained in the set of all matchings that may arise in equilibria. Then it is shown that the latter set may contain unstable matchings as well: Under certain preference structures there are equilibria leading to an unstable matching almost surely (Example 1). The paper proceeds to propose two conditions, each of which rules out such equilibria: 1. The players do not condition their behavior on the actions during any past failed meeting (Proposition 1.5.4). 2. The players' preferences satisfy the Sequential Preference Condition, a condition that implies a certain degree of preference alignment (Proposition 1.5.6). However the two conditions, separately or combined, are not sufficient to rule out equilibria in which unstable matchings arise with positive probability; some of the probable matchings may even be Pareto-dominated (Example 3). Another source of inefficiency is delay: Significant loss of efficiency due to delay may be present in an equilibrium even if search frictions are arbitrarily small (Examples 4 and 5). The paper ends with a uniqueness result that is pro-stability and efficiency: If the players' preferences satisfy a strengthening of the Sequential Preference Condition which implies a stronger degree of alignment, then all equilibria are outcome equivalent, stable, and efficient (Proposition 1.5.8).

1.2 Literature

The present paper contributes to the literature on search and matching games in which a marriage market is embedded. The central question of the literature agrees with that of the present paper: Do equilibrium outcomes correspond to stable matchings? An early paper in this literature, Roth and Vande Vate (1990) studies the steady state of a search and matching game with short-sighted players and concludes that a stable matching obtains almost surely. Later papers consider sophisticated players. McNamara and Collins (1990), Burdett and Coles (1997), Eeckhout (1999), Bloch and Ryder (2000) and Smith (2006) assume that the underlying marriage market admits a unique stable matching that is positively assortative. Their results confirm that equilibrium outcomes retain some extent of assorting. Adachi (2003) and Lauermann and Nöldeke (2014) consider a market with a general preference structure. Adachi (2003) studies a model in which the steady state stock of active players is exogenously maintained and confirms that equilibrium outcomes converge to stable matchings as search frictions vanish. Lauermann and Nöldeke (2014) considers endogenous steady states and finds that all limit outcomes are stable if and only if the underlying market has a unique stable matching.

The model considered in this paper also embeds a marriage market in a search and

[3] Roth and Sotomayor (1990), page 245.

matching game. In contrast to the previously cited papers, all of which study steady state equilibria in a stationary setting, the present model features a nonstationary search situation. Indeed, the market shrinks as players marry and leave. Moreover, all but Roth and Vande Vate (1990) consider a market with a continuum of nameless players, whereas in the present model the market is finite and the players are identifiable. Nonstationarity and finiteness make the present model qualitatively different from most models considered in the literature. It follows that the set of matchings that may obtain in equilibria of the present model is in general different from that of a stationary and continuum model.

Another related literature investigates models embedding a marriage market in a sequential bargaining game reminiscent of the deferred acceptance protocol in Gale and Shapley (1962). This literature includes Alcalde (1996), Diamantoudi et al. (2015), Pais (2008), Suh and Wen (2008), Niederle and Yariv (2009), Bloch and Diamantoudi (2011), and Haeringer and Wooders (2011).[4] Like the present paper, these papers consider a finite marriage market that shrinks as players marry and leave. The difference between models in this literature and those in the search and matching literature, including the present model, is the search technology. A sequential bargaining game models a market with directed search: When it is his or her turn to move, a player can reach and deal with any player of the opposite sex without delay or uncertainty. In contrast, a search and matching game models a market with undirected search: Bilateral meetings are stochastic; one needs patience and luck to encounter a particular person. One common finding among papers with a sequential bargaining model is that some or all stable matchings can be supported in equilibria. Such equilibria bear resemblance to equilibria that lead to a particular stable matching almost surely in the present model, see Proposition 1.5.2. On the other hand, unstable matchings may also obtain in equilibria of a sequential bargaining game, which is the case in Diamantoudi et al. (2015), Suh and Wen (2008) and Haeringer and Wooders (2011). This common finding is also in accordance with results in the present paper. However, because of random search, the model in the present paper may have equilibria that have no counterpart in a sequential bargaining model. For instance, in a typical sequential bargaining model, an equilibrium in pure strategies leads to one matching deterministically, whereas in this model an equilibrium in pure strategies may lead to several possible matchings, because the players' strategies may depend on which of the multiple probable paths the history has taken. In this respect, nonstationarity has little influence in a sequential bargaining model because a player's expected payoff remains unchanged as the game unfolds, whereas in the present model exogenous uncertainty may drastically change a player's continuation prospect.

[4] Marriage problems belong to a class of coalitional games under nontransferable utility called "hedonic games". Alcalde and Romero-Medina (2000) and Alcalde and Romero-Medina (2005) investigate decentralized implementation of stable outcomes of many-to-one matching problems, which are also hedonic games. Bloch and Diamantoudi (2011) study implementability of the core of a general hedonic game.

A third related literature[5] studies whether the Walrasian price can be supported in equilibria of a search and bargaining game in which an exchange economy, instead of a marriage market, is embedded. Papers from this literature and the present paper are united under the theory of non-cooperative foundation of cooperative solution concepts. Indeed, the present model can be seen as the nontransferable utility version of the models considered in Rubinstein and Wolinsky (1990) and Gale and Sabourian (2006).

The layout of the paper is as follows: Section 2 introduces the game. Section 3 sets up an analytic framework. Section 4 provides the analysis. Section 5 concludes. Lengthy proofs and additional examples are found in the Appendices.

1.3 The Game

The Marriage Market

There are two disjoint sets of players: the set of men M and the set of women W. A generic man is denoted as m, a woman as w, and a pair of a man and a woman as (m, w). A man might end up marrying some $w \in W$ or remaining single. All men's preferences over $W \cup \{s\}$, where s stands for being single, are represented by $u : M \times (W \cup \{s\}) \mapsto \mathbb{R}$ where $u(m, \cdot)$ is m's Bernoulli utility function over $W \cup \{s\}$. Likewise all women's preferences are represented by $v : (M \cup \{s\}) \times W \mapsto \mathbb{R}$ where $v(\cdot, w)$ is w's Bernoulli utility function over $M \cup \{s\}$. The **marriage market** (or simply **market**) is summarized by the tuple (M, W, u, v).

Let $\succeq_m$ denote man m's preference relation over $W \cup \{s\}$ induced by $u(m, \cdot)$: $w \succeq_m w'$ if and only if $u(m, w) \geq u(m, w')$. Likewise let $\succeq_w$ denote woman w's preference relation over $M \cup \{s\}$. Player y is **acceptable** to player x if $y \succ_x s$. A market is **trivial** if it does not have a mutually acceptable pair. A game starts with a market satisfying the following:

A1 Preferences are strict: $u(m, \cdot)$ is one-to-one for any $m \in M$; $v(\cdot, w)$ is one-to-one for any $w \in W$.

A2 Normalization: $u(m, s) = 0$ for any $m \in M$; $v(s, w) = 0$ for any $w \in W$.

A3 The market is finite: $|M| < \infty$ and $|W| < \infty$.

A4 The market is nontrivial.

[5]Surveyed in Osborne and Rubinstein (1990) and Gale (2000).

The Game Rules

The game starts on day one ($t = 1$) with an **initial market** (M, W, u, v) and unfolds indefinitely into the future ($t = 2, 3, ...$). On each day a randomly selected pair $(m, w) \in M \times W$ meet. The random meeting process will be described in detail later. As they meet, m moves first to either accept or reject w. If m rejects w then the pair separate and return to the market. If m accepts w then it is w's turn to either accept or reject m. If w accepts m then (m, w) marry and leave the game for good; otherwise the pair separate and return to the market. Either the separation or the marriage concludes the current day. m and w receive one-time payoffs of $u(m, w)$ and $v(m, w)$, respectively, upon marrying each other. The value of a marriage delayed by τ days is discounted by δ^τ where the common discount factor $\delta \in (0, 1)$ is meant to capture the overall search frictions. The game ends when there is no longer a mutually acceptable pair left in the market. A player receives a payoff of 0 when the game ends if he or she stays unmarried at that time.[6] Information is complete and past actions are perfectly observable.

Notations and Terminology

Let H denote the set of all histories. Let $\hat{H}$ denote the set of all nonterminal histories after which a new day starts but the pair to meet on that day has not been determined. For $h \in H$ let $\Gamma(h)$ denote the subsequent subgame given h is reached. Note that $\Gamma(h)$ per se is a proper game if and only if $h \in \hat{H}$.

Let Z denote the set of all terminal histories. A terminal history may be infinite. The **outcome matching** of $h \in Z$ is a mapping $\mu_h : M \cup W \mapsto M \cup W \cup \{s\}$ such that $\mu_h(x)$ is player x's corresponding spouse if x managed to marry at some point along h, or otherwise $\mu_h(x) = s$. In the latter case x is said to be **single** under h. If h is finite then x is single if he or she stays unmarried until the game ends at h. If h is infinite then x is single if he or she is unmarried after any finite subhistory of h.

The market (M', W', u', v') is a **submarket** of the initial market (M, W, u, v) if $M' \subset M$, $W' \subset W$, u' is u restricted to $M' \times (W' \cup \{s\})$, and v' is v restricted to $(M' \cup \{s\}) \times W'$. Abuse notation to write (M', W', u, v) for simplicity. Let $\mathscr{S}$ denote the set of all nontrivial submarkets of the initial market. Given $S := (M', W', u, v)$, respectively use the notations $x \in S$ to denote $x \in M' \cup W'$, $(m, w) \in S$ to denote $(m, w) \in M' \times W'$, and $S \backslash (m, w)$ to denote the submarket $(M' \backslash \{m\}, W' \backslash \{w\}, u, v)$.

For $S \in \mathscr{S}$ and $x \in S$, let $A^S(x)$ denote the set $\{y \in S : y \succ_x s \text{ and } x \succ_y s\}$. $A^S(x)$

[6] It might be more natural to let the game end until no man or woman is left. Lemma 1.5.1 to appear later, which still holds under this alternative game-ending rule, implies under the alternative rule no one will marry and everyone's expected payoff is 0 in any subgame perfect equilibrium when the remaining market is trivial. Thus the default game-ending rule neither creates nor destroys equilibria in effect, yet it simplifies the exposition.

is thus the set of all players in S with whom x forms a mutually acceptable pair. Let $\alpha^S(x)$ denote the greatest element in $A^S(x) \cup \{s\}$ according to $\succeq_x$.

For $h \in H$ the **remaining market** after h, denoted as $S(h)$, consists of the men and women who are unmarried after h. Obviously $S(h) \in \mathscr{S}$ for any nonterminal $h \in H$.

The Contact Function

Recall that on each day a pair of a man and a woman are randomly selected to meet each other. The random meeting process is modeled by the **contact function** $C : M \times W \times \mathscr{S} \mapsto [0, 1]$, where $C(m, w, S)$ is the probability that (m, w) meet on a day at the beginning of which the remaining market is S. The game rules thus require that for any $S \in \mathscr{S}$,

B1 Only unmarried people meet: $C(m, w, S) = 0$ if $m \notin S$ or $w \notin S$.

B2 A meeting takes place on each day: $\sum_{(m,w)\in S} C(m, w, S) = 1$.

In addition, assume the meeting probability of any remaining pair is considerably large:

B3 There exists $\epsilon > 0$ such that $C(m, w, S) > \epsilon$ if $(m, w) \in S$.

Note that by the definition of the contact function, the meeting probabilities on a given day are determined by the remaining market at the beginning of that day. This implies that $\Gamma(h)$ and $\Gamma(h')$ are isomorphic for any histories h and h' such that $S(h) = S(h')$.

The game is summarized by the tuple (M, W, u, v, C, δ).

1.4 From Equilibria to Matchings

Equilibria

For most of the analysis the solution concept that will be applied is the subgame perfect equilibrium. In addition I consider two equilibrium selection criteria to accommodate more restrictive information settings.

For history h let $g(h)$ denote the sequence $(m_t, w_t, R_t)_{t=1:\tau(h)}$ where m_t and w_t are the man and woman who met on date t under h, $R_t \in \{marriage, separation\}$ is the result of that meeting, and $\tau(h)$ is the date of the last concluded meeting under h. A strategy profile σ satisfies the **private-dinner condition** if $g(h) = g(h')$

implies σ restricted to $\Gamma(h)$ is the same as σ restricted to $\Gamma(h')$. The private-dinner condition accommodates the information setting in which players are aware of who met whom in the past and the results of those meetings but not what happened during those meetings, presumably because the meetings took place over private dinners. In particular, if a meeting ended in separation there is no telling whether it was the man or the woman who said no. Note that the private-dinner condition implies a player's strategy cannot depend on actions taken during a failed meeting even if himself or herself participated in it.

A strategy profile satisfies the **Markov condition** if for any history h the player who moves after h conditions his or her behavior only on $S(h)$. The Markov condition is stronger than the private-dinner condition because $g(h) = g(h')$ implies $S(h) = S(h')$ but not vice versa. The Markov condition is compatible with the more restrictive information setting in which players are only aware of the current market.

When describing strategies, I will simply say "player x accepts/rejects player y under condition K" to represent the statement that x accepts/rejects y at every decision point satisfying condition K where it is x's turn to make the pertinent decision. I say (m, w) **marry upon first meeting** under strategy profile σ if on the equilibrium path the first meeting between (m, w) results in marriage. (m, w) marry upon first meeting if and only if on the equilibrium path (m, w) always accept each other.

Matchings

A **matching** for (M, W, u, v) is a scheme that pairs some players into married couples and leaves others single. A matching is formalized as a function $\mu : M \cup W \mapsto M \cup W \cup \{s\}$ such that $\mu(x) \in W \cup \{s\}$ if $x \in M$, $\mu(x) \in M \cup \{s\}$ if $x \in W$, and $\mu(\mu(x)) = x$ if $\mu(x) \neq s$. μ is **unstable** if there is a player x such that $s \succ_x \mu(x)$, in which case μ is **individually blocked** by x, or if there is a pair (m, w) such that $w \succ_m \mu(m)$ and $m \succ_w \mu(w)$, in which case μ is **pairwise blocked** by (m, w). μ is **stable** if it is not unstable. Gale and Shapley (1962) shows that at least one stable matching exists for any marriage market, and moreover there is a **men-optimal matching** commonly agreed by all men as the best stable matching and likewise there is a **women-optimal matching**. Given a matching μ for the market (M, W, u, v) let $\mathscr{S}_\mu$ denote the set of all nontrivial submarkets (M', W', u, v) such that $W \backslash W' = \mu(M \backslash M')$ where $\mu(M \backslash M')$ denotes the μ-image of $M \backslash M'$. Observe that if $S \in \mathscr{S}_\mu$ then μ restricted to S is a matching for S; moreover if μ is a stable matching for (M, W, u, v) then μ restricted to S is a stable matching for S.

Obviously the outcome matching of any $h \in Z$ is a matching for the initial market. A strategy profile σ and the contact function C jointly induce a probability measure on 2^Z and hence also induce a probability mass function on the set of all matchings for the initial market. We say that a matching **obtains** if it arises as an outcome matching. A strategy profile σ **enforces** a matching μ if μ obtains almost surely under σ. μ being enforced implies the players will almost surely be coupled together or left single according to μ.

Near-Frictionless Analysis

This paper focuses on analyzing game outcomes when search frictions are small. With respect to this approach we introduce the following terminology: An **environment** $(M, W, u, v, C) := \{(M, W, u, v, C, \delta) : \delta \in (0, 1)\}$ is the set of all games that share the same initial market and contact function. A strategy profile σ is a **limit equilibrium** of the environment (M, W, u, v, C) if there exists some $d < 1$ such that σ is a subgame perfect equilibrium of the game (M, W, u, v, C, δ) for any $\delta > d$.

1.5 Analysis

Preliminary Results

The following lemma collects some useful results for future reference.

Lemma 1.5.1. *For a subgame perfect equilibrium σ let $\pi(x)$ denote the expected payoff for player x under σ. The following are true for σ:*

(a) $\pi(x) \geq 0$ *for any* $x \in M \cup W$.

(b) (m, w) *marry with positive probability only if* m *is acceptable to* w. [7]

(c) (m, w) *marry with positive probability only if* $\alpha^{S_I}(m) \succeq_m w$ *and* $\alpha^{S_I}(w) \succeq_w m$ *where* S_I *denotes the initial market.*

(d) $\pi(m) \leq u(m, \alpha^{S_I}(m))$ *for any* $m \in M$. $\pi(w) \leq v(\alpha^{S_I}(w), w)$ *for any* $w \in W$.

Proof: (a) follows from the observation that a player secures an expected payoff of 0 by rejecting everyone forever. The same observation implies a woman's equilibrium continuation payoff from rejecting a man is nonnegative, thus (b) follows. To show (c), first observe that if $w \succ_m \alpha^{S_I}(m)$ then m is unacceptable to w, thus (m, w) will not marry by (b). Suppose $m \succ_w \alpha^{S_I}(w)$ yet (m, w) marry with positive probability. w is unacceptable to m. That (m, w) marry with positive probability implies m accepts w with positive probability after some history h. Let h' denote the immediate history following h as m has accepted w. w rejects m with positive probability after h' because otherwise m's expected payoff from accepting w after h is $u(m, w) < 0$, less

[7] However, a pair (m, w), where w is unacceptable to m, might marry in equilibrium with positive probability. For an example of this possibility see Appendix A.3. Before reading the example the reader is recommended to go through Proposition 1.5.2 and Example 1 in the main text.

than the payoff of 0 from rejecting everyone forever. Let V denote w's expected payoff in the subsequent subgame Γ as she has rejected m after h'. That w rejects m with positive probability after h' implies $\delta V \geq v(m,w)$, then in turn implies there exists some $m' \succ_w m$ such that (m',w) marry with positive probability in Γ. Since $m' \succ_w \alpha^{S_I}(w)$, we can apply the same argument for m' and conclude there exists some m'' such that $m'' \succ_w m'$ and (m'',w) marry with positive probability in some subgame. Iteratively applying the same argument leads to the necessary contradiction because M is finite. (d) follows from (c). □

Enforcing Stable Matchings

The foremost question of whether stable matchings may be enforced in equilibria is addressed in this subsection. Proposition 1.5.2 below gives a positive answer by showing that any stable matching can be enforced in a limit equilibrium. The proof is based on a construction per se worth highlighting: For a matching μ of the initial market, the **μ-strategy profile** σ_μ is described by the following table that specifies what man m and woman w do if they meet each other when the remaining market is S. In the table μ^S denotes the women-optimal matching for S.

	m	w
If $S \in \mathscr{S}_\mu$	Accept w if $w \succeq_m \mu(m)$	Accept m if $m \succeq_w \mu(w)$
If $S \notin \mathscr{S}_\mu$	Accept w if $w \succeq_m \mu^S(m)$	Accept m if $m \succeq_w \mu^S(w)$

Let $V_\mu(x,\delta)$ be player x's expected payoff under σ_μ if the discount factor is δ.

Proposition 1.5.2. *If μ is a stable matching for the initial market of the environment (M,W,u,v,C) then:*

(a) σ_μ *enforces* μ.

(b) (m,w) *marry upon first meeting under* σ_μ *if* $w = \mu(m)$.

(c) $\lim_{\delta\to1} V_\mu(m,\delta) = u(m,\mu(m))$ *and* $\lim_{\delta\to1} V_\mu(w,\delta) = v(\mu(w),w)$.

(d) σ_μ *satisfies the Markov condition.*

(e) σ_μ *is a limit equilibrium of* (M,W,u,v,C).

Proof: The initial market is in $\mathscr{S}_\mu$. Observe that if (m,w) meet when the remaining market is in $\mathscr{S}_\mu$ then the meeting results in marriage if and only if $w=\mu(m)$ because μ being stable implies $w \succeq_m \mu(m)$ and $m \succeq_w \mu(w)$ hold simultaneously if and only if $w = \mu(m)$. Thus the remaining market after any history on the equilibrium path is in $\mathscr{S}_\mu$, which combined with the previous observation implies (b). It also follows that

(m, w) will not marry on the equilibrium path if $w \neq \mu(m)$. Consequently if $\mu(m) \neq s$ then $(m, \mu(m))$ remain in the market until the first meeting between them takes place. Thus the probability that $(m, \mu(m))$ marry is equal to the probability that they meet eventually, the latter being bounded from below by $\sum_{n=0}^{\infty} \epsilon(1-\epsilon)^n = 1$, implying (a). Following (a) and (b) we have $u(m, \mu(m)) \geq V_\mu(m, \delta) \geq \sum_{n=0}^{\infty} \epsilon[\delta(1-\epsilon)]^n u(m, \mu(m))$. (c) follows from $\lim_{\delta \to 1} \sum_{n=0}^{\infty} \epsilon[\delta(1-\epsilon)]^n u(m, \mu(m)) = u(m, \mu(m))$ and the analogous equality for any w. (d) follows from the observation that a player's behavior depends on only the current remaining market.

Now show (e). Suppose (m, w) meet on a day when the remaining market is S. Let $\overline{\mu}$ denote μ restricted to S if $S \in \mathscr{S}_\mu$ or μ^S otherwise. Thus $\overline{\mu}$ is a stable matching for S. Let Γ denote the subsequent subgame resulting from (m, w)'s separation. By construction σ_μ restricted to Γ is equal to the $\overline{\mu}$-strategy profile $\sigma_{\overline{\mu}}$ of Γ. Let $V_\mu(x, \delta|\Gamma)$ denote the expected payoff for player $x \in S$ under σ_μ restricted to Γ. Then $\lim_{\delta \to 1} \delta V_\mu(x, \delta|\Gamma) = u(m, \overline{\mu}(m))$ by (c) and $\delta V_\mu(x, \delta|\Gamma) < u(m, \overline{\mu}(m))$ by (a); the analogous equality and inequality respectively hold for w. Apply one-deviation analysis for m and w. For δ sufficiently close to 1, $\delta V_\mu(w, \delta|\Gamma) < v(m, w)$ if and only if $m \succeq_w \overline{\mu}(w)$, where the left side of the inequality is w's expected payoff from rejecting m and the right side that from accepting m. Thus accepting m if and only if $m \succeq_w \overline{\mu}(w)$ is optimal for w. m's expected payoff from rejecting w is $\delta V_\mu(m, \delta|\Gamma)$ whereas that from accepting w is $u(m, \overline{\mu}(m))$ if $w = \overline{\mu}(m)$, $\delta V_\mu(m, \delta|\Gamma)$ if $w \succ_m \overline{\mu}(m)$, or $pu(m, w) + (1-p)\delta V_\mu(m, \delta|\Gamma)$ if $\overline{\mu}(m) \succ_m w$ where p is either 0 or 1. Thus accepting w if and only if $w \succeq_m \overline{\mu}(m)$ is optimal for m if δ is sufficiently close to 1. $\square$

Proposition 1.5.2 agrees with the common finding that the core of a coalitional game can be supported in equilibria of a non-cooperative counterpart. For results along this line in a similar search and matching context see Rubinstein and Wolinsky (1990), Adachi (2003), and Lauermann and Nöldeke (2014). The limit equilibria constructed above will also be used as important building blocks for more complicated equilibria.

Enforcing Unstable Matchings

In this subsection, the question of whether unstable matchings are enforceable in limit equilibria is addressed with an affirmative example.

Example 1. Reward and punishment

To describe the initial market, player x's preferences are represented by a list $P(x)$ such that $P(x) = a, ..., b$ if and only if $a \succ_x ... \succ_x b \succ_x s$. Note that players unacceptable to x are omitted from $P(x)$. The initial market is represented as

$$P(m_1) = w_2, w_1, \qquad P(w_1) = m_1, m_2, m_3$$
$$P(m_2) = w_1, w_2, \qquad P(w_2) = m_2, m_3, m_1,$$
$$P(m_3) = w_2.$$

A limit equilibrium σ is constructed to enforce μ such that $\mu(m_1) = w_2$, $\mu(m_2) = w_1$ and $\mu(m_3) = s$. μ is unstable because the pair (m_3, w_2) blocks it. σ is specified by an automaton with the following states:

q_0 : The initial state. In q_0, m_1 accepts w_2; m_2 accepts no one; m_3 accepts no one; w_1 accepts m_1; w_2 accepts m_1 and m_2. The transition rules are:

$$q_0 \longrightarrow \begin{cases} q_1 & \text{if } (m_1, w_2) \text{ marry,} \\ q_2 & \text{if for some } (m, w) \neq (m_1, w_2)\text{: } w \text{ rejects } m \text{ or } (m, w) \text{ marry,} \\ q_0 & \text{otherwise.} \end{cases}$$

q_1 : An absorbing state. As the state has just become q_1, the remaining market is $S_1 := S_0 \backslash (m_1, w_2) \in \mathscr{S}_\mu$ where S_0 denotes the initial market. In q_1 the players follow the μ^{S_1}-strategy profile where μ^{S_1} is μ restricted to S_1.

q_2 : An absorbing state. Let S_2 denote the (history-dependent) remaining market as the state has just become q_2. In q_2 the players follow the μ^{S_2}-strategy profile where μ^{S_2} denotes the women-optimal matching for S_2.

On the equilibrium path (m_1, w_2) marry first. Then the state becomes q_1 in which μ^{S_1} (μ restricted to S_1) is enforced, as implied by Proposition 1.5.2(a), because μ^{S_1} is a stable matching for S_1. Thus μ obtains under any finite terminal history on the equilibrium path. It is straightforward to verify that the game ends almost surely, implying μ obtains almost surely. Thus σ enforces μ.

Now verify that σ is indeed a limit equilibrium. By Proposition 1.5.2(e), σ restricted to subgames in q_1 and q_2 is a limit equilibrium of the respective subgames. Thus it suffices to check q_0. Consider the situation that the blocking pair (m_3, w_2) meet on a day in q_0. Apply one-deviation analysis. Suppose m_3 has accepted w_2. w_2's action of rejecting m_3 will switch the state to q_2 in which the μ^W-strategy profile will be implemented where μ^W is the women-optimal matching for the initial market. By Proposition 1.5.2(c), w_2's expected payoff from rejecting m_3 is approximately $v(\mu^W(w_2), w_2) = v(m_2, w_2)$ for δ sufficiently close to 1, strictly greater than $v(m_3, w_2)$. Thus rejecting m_3 is optimal when near-frictionless. Now consider m_3. If he rejects w_2 then μ is enforced; otherwise if he accepts w_2 then w_2 will reject him, switching the state to q_2 in which μ^W is enforced. $\mu(m_3) = \mu^W(m_3) = s$ implies rejecting w_2 is (weakly) optimal for m_3.

Consider the situation that (m_2, w_1) meet in q_0. As in w_2's case above, it is optimal for w_1 to reject m_2 for δ sufficiently close to 1. m_2's case is slightly different from m_3's case above. If m_2 accepts w_1, the μ^W-strategy profile will be implemented under which m_2's expected payoff is approximately $u(m, \mu^W(m)) = u(m_2, w_2)$ for δ sufficiently close to 1. If m_2 rejects w_1, m_2 will marry w_1 eventually but only after (m_1, w_2) marry. A lower bound for m_2's expected payoff from rejecting w_1 can thus be computed as $\left[\frac{\delta\epsilon}{1-\delta(1-\epsilon)}\right]^2 u(m_2, w_1)$, strictly greater than $u(m_2, w_2)$ for δ sufficiently close to 1. Rejecting w_1 in q_0 is optimal when near-frictionless. The optimality of σ in other cases is either similar to those discussed above or can be verified by routine inspection. □

A blocking pair would profit from marrying each other to circumvent an unstable matching. To enforce an unstable matching such circumvention must be discouraged. In Example 1, a reward-punishment scheme, implemented in q_2, is employed to prevent the blocking attempt from (m_3, w_2). To see the point, note that if m_3 initiated a blocking attempt by accepting w_2, w_2 would not oblige because she would receive a reward, which is the promise of marrying the more preferable man m_2, from rejecting m_3. In contrast m_2 would be (weakly) punished [8] for initiating the blocking attempt by being forced to stay single. Meanwhile, to ensure the reward for w_2 is credible, m_2 needs to be available until either m_3 or w_2 has married. In Example 1, m_2 may marry only after (m_1, w_2) have married. m_2's potential attempt to marry w_1 early is discouraged by a similar reward-punishment scheme. Should m_2 accept w_1 when (m_1, w_2) have not married, he would be strictly punished (by marrying w_2 eventually) for deviating and w_1 would be rewarded (by marrying m_1 eventually) for not obliging. The reward-punishment schemes resemble those used in Proposition 1 in Rubinstein and Wolinsky (1990) supporting non-core outcomes. In their model, a reward-punishment scheme targeted at the blocking attempt between a buyer and a seller entails reaction from at most three players (those whose welfare would be affected should the attempt succeed), because all sellers are identical and so are all buyers. In contrast, in the present model, because of a more complicated preference structure, a reward-punishment scheme may require the entire market to re-coordinate, which would make its implementability more difficult. Indeed, Proposition 1.5.6 to appear later will show for certain markets no unstable matching can be enforced.

Sufficient Conditions for Enforced Matchings to be Stable

In this subsection two points are made regarding the enforceability of unstable matchings. First, disabling reward-punishment schemes excludes unstable matchings from matchings enforceable in equilibria. Second, enforceability of unstable matchings de-

[8] Example 1 relies on the knife-edge case that m_3 rejects w_2 when indifferent. Such fragility need not be present in enforcing an unstable matching. In an earlier version of this paper I provided a more complicated example in which all circumventing attempts are strictly punished.

pends on the preference structure. Each point is made by a condition under which all matchings enforceable in equilibria are stable.

The following lemma will be useful for the proofs of further results.

Lemma 1.5.3. *If an unstable matching μ is enforced in a subgame perfect equilibrium σ then for any pair (m, w) blocking μ, σ prescribes the following for any on-equilibrium-path meeting between (m, w): m rejects w, moreover if m deviated to accepting w then w would reject m with positive probability.*

Proof: Consider any on-equilibrium-path meeting between a blocking pair (m, w). Suppose w would accept m with probability 1 after she has been accepted by m. m must reject w in the first place, because otherwise (m, w) would marry with positive probability in equilibrium, a contradiction. It follows that m's continuation payoff from rejecting w is weakly higher than $u(m, w)$, the latter strictly higher than $u(m, \mu(m))$ since the pair (m, w) blocks μ, implying m would marry someone other than $\mu(m)$ with positive probability in equilibrium, a contradiction. Thus w's equilibrium strategy is to reject m with positive probability in the current meeting. This implies w's expected payoff from rejecting m is weakly greater than $v(m, w)$, the latter strictly greater than $v(\mu(w), w)$. Hence in the subsequent subgame Γ resulting from w having rejected m, w will marry with someone other than $\mu(w)$ with positive probability. Suppose m accepts w with positive probability in the current meeting, then Γ is reached with positive probability, implying w would marry someone other than $\mu(w)$ with positive probability in equilibrium, a contradiction. □

Lemma 1.5.3 implies that in order to deter a blocking pair (m, w) from circumventing the enforcement of an unstable matching, m must be punished for initiating a blocking attempt and w rewarded for not obliging. Obviously such a scheme cannot be implemented if the other players cannot tell whether (m, w)'s separation resulted from a failed blocking attempt initiated by m but turned down by w, or from m having rejected w as prescribed. As the upcoming proposition shows, the private-dinner condition, which disallows the players from conditioning their behavior on the actions during any failed meeting, indeed rules out unstable matchings from matchings enforceable in equilibria.

Proposition 1.5.4. *No unstable matching can be enforced in a private-dinner equilibrium.*

Proof: Prove by contradiction. Suppose an unstable matching μ for the initial market is enforced in a private-dinner equilibrium σ. μ is individually rational by Lemma 1.5.1(a), thus some pair (m, w) blocks μ. Consider the situation that (m, w) meet on the first day of the game. Let $h \in \hat{H}$ denote the history corresponding to m having rejected w on the current day. Let $h' \in \hat{H}$ denote the history corresponding to m having accepted w and w having rejected m on the current day. Thus $g(h)$ and $g(h')$ are both the one-entry sequence $(m, w, \textit{separation})$. Let V and V' respectively denote w's expected payoffs in $\Gamma(h)$ and $\Gamma(h')$. The private-dinner condition implies $V = V'$.

By Lemma 1.5.3, w would reject m with positive probability on the current day if she was accepted, implying $\delta V' \geq v(m, w)$. h is on the equilibrium path by Lemma 1.5.3 because m rejects w on the current day in equilibrium. Thus μ is enforced by σ restricted to $\Gamma(h)$, implying $v(\mu(w), w) \geq V$. Since the pair (m, w) blocks μ, we have $v(m, w) > v(\mu(w), w)$, implying $\delta V = \delta V' \geq v(m, w) > v(\mu(w), w) \geq V$, a contradiction. □

Proposition 1.5.4 identifies a condition on the equilibrium that negates enforceability of unstable matchings *for any initial market*. In contrast the upcoming proposition identifies a condition on the initial market that negates enforceability of unstable matchings *in any equilibrium*.

Call (m, w) a **top pair** for submarket S if $m = \alpha^S(w)$ and $w = \alpha^S(m)$. (m, w) is a top pair for S if for m, w is the best woman among those in S who find m acceptable, and vice versa. A marriage market satisfies the **Sequential Preference Condition** if there is an ordering of the men $m_1, ..., m_{|M|}$, an ordering of the women $w_1, ..., w_{|W|}$, and a positive integer k such that:

1. For any $i \leq k$, (m_i, w_i) is a top pair for $S_i := (\{m_j : j \geq i\}, \{w_j : j \geq i\}, u, v)$.
2. $S_{k+1} := (\{m_j : j \geq k+1\}, \{w_j : j \geq k+1\}, u, v)$ is trivial.

A stronger condition, introduced in Eeckhout (2000), implies that the market has a unique stable matching.[9] The present Sequential Preference Condition, albeit weaker, is still sufficient for uniqueness. The unique stable matching pairs m_i to w_i for any $i \leq k$ and leaves m_i and w_i single for any $i > k$. The Sequential Preference Condition implies that the players' preferences are aligned in a certain way, see Eeckhout (2000) for a different formalization that highlights the alignment.

Lemma 1.5.5. *If (m, w) is a top pair for the initial market then (m, w) marry with positive probability, and upon first meeting, in any subgame perfect equilibrium.*

Proof: If (m, w) is a top pair for the initial market then clearly they are a top pair for any submarket in which both are present. w always accepts m in any subgame perfect equilibrium because the equilibrium payoff from rejecting m is strictly less than $v(m, w)$ by Lemma 1.5.1(d). It follows that in any subgame perfect equilibrium m always accepts w because his payoff from accepting w is $u(m, w)$ whereas that from rejecting w is strictly less than $u(m, w)$ by Lemma 1.5.1(d). Thus (m, w) marry upon first meeting in any subgame perfect equilibrium. That they marry with positive probability follows from the assumption that they meet on the first day of the game with positive probability. □

[9] The condition in Eeckhout (2000), sometimes also called the Sequential Preference Condition in the literature, is equivalent to the present condition if $|M| = |W|$, every man is acceptable to every woman and vice versa.

Proposition 1.5.6. *If the initial market satisfies the Sequential Preference Condition then no unstable matching can be enforced in any subgame perfect equilibrium.*

Proof: Let the men and women be ordered as in the definition of the Sequential Preference Condition and k be the corresponding index threshold. Fix a subgame perfect equilibrium σ that enforces some matching μ. (m_1, w_1) marry with positive probability under σ by Lemma 1.5.5. It follows that $\mu(m_1) = w_1$. It also follows that a subgame with remaining market S_2 is reached with positive probability, because if (m_1, w_1) meet on the first day, which occurs with positive probability, then they marry and the remaining market becomes S_2. Suppose for any $i < n \leq k$ for some n, $\mu(m_i) = w_i$ and a subgame with remaining market S_{i+1} is reached with positive probability. By the inductive hypothesis a subgame Γ with remaining market S_n is reached with positive probability. By Lemma 1.5.5, (m_n, w_n) marry with positive probability in Γ under σ, implying that (m_n, w_n) marry with positive probability in the initial game under σ. Hence $\mu(m_n) = w_n$. It also follows that a subgame with remaining market S_{n+1} is reached with positive probability. By induction, for any $i \leq k$, $\mu(m_i) = w_i$ and a subgame with remaining market S_{i+1} is reached with positive probability, implying the subgame with remaining market S_{k+1} is reached with positive probability. As the remaining market becomes S_{k+1} the game ends and any player $x \in S_{k+1}$ remains single. Thus $\mu(m_i) = \mu(w_i) = s$ for any $i > k$. The proposition follows from the observation that μ is the unique stable matching. □

For the sequential bargaining model considered in Suh and Wen (2008), if the underlying marriage market satisfies the Sequential Preference Condition then there is a unique equilibrium implementing the unique stable matching (Theorem 1).[10] However, the Sequential Preference Condition does *not* guarantee a unique subgame perfect equilibrium for the present model, as will be shown in a later example (Example 3). In some of the additional equilibria, a lottery of matchings, instead of a single matching, is induced due to the uncertainty in the search process, which is absent from the model in Suh and Wen (2008).

Equilibria Inducing a Lottery of Matchings

Instead of enforcing a single matching, there may exist equilibria inducing a lottery of multiple matchings. In this subsection examples are given showing that unstable matchings, even Pareto-dominated ones, may obtain with positive probability in such equilibria.

Example 2. Regret

[10] Theorem 1 in Suh and Wen (2008) uses the stronger condition from Eeckhout (2000), but remains true under the present Sequential Preference Condition.

The initial market is described by the following lists:

$$\begin{aligned}
P(m_1) &= w_2, w_1, w_3, & P(w_1) &= m_1, m_2, m_3,\\
P(m_2) &= w_3, w_2, w_1, & P(w_2) &= m_2, m_1, m_3,\\
P(m_3) &= w_2, w_3, w_1, & P(w_3) &= m_3, m_2, m_1,\\
P(m_1') &= w_2', w_1', w_3', & P(w_1') &= m_1', m_2', m_3',\\
P(m_2') &= w_3', w_2', w_1', & P(w_2') &= m_2', m_1', m_3',\\
P(m_3') &= w_2', w_3', w_1', & P(w_3') &= m_3', m_2', m_1'.
\end{aligned}$$

Moreover, $\frac{1}{2}v(m_2, w_2) + \frac{1}{2}v(m_3, w_2) > v(m_1, w_2)$ and $\frac{1}{2}v(m_2', w_2') + \frac{1}{2}v(m_3', w_2') > v(m_1', w_2')$. The contact function satisfies $C(m, w, S) = C(\hat{m}, \hat{w}, S)$ for any pairs $(m, w) \in S$ and $(\hat{m}, \hat{w}) \in S$ for any $S \in \mathscr{S}$. The initial market has a unique stable matching μ such that $\mu(m_i) = w_i$ and $\mu(m_i') = w_i'$ for $i = 1, 2, 3$.

In the limit equilibrium σ to be described shortly, two matchings, μ_a and μ_b given below, obtain, each with probability 0.5:

$$\mu_a(m_1) = w_1, \quad \mu_a(m_2) = w_3, \quad \mu_a(m_3) = w_2, \quad \mu_a(m_i') = w_i', i = 1, 2, 3,$$

$$\mu_b(m_1') = w_1', \quad \mu_b(m_2') = w_3', \quad \mu_b(m_3') = w_2', \quad \mu_b(m_i) = w_i, i = 1, 2, 3.$$

Neither μ_a nor μ_b is stable: the pair (m_1, w_2) blocks μ_a and the pair (m_1', w_2') blocks μ_b. σ is specified by an automaton with the following states:

q_0: The initial state. m_1 accepts w_1 and w_2. m_1' accepts w_1' and w_2'. Every other man accepts no one. Every woman accepts her first choice. The transition rules are

$$q_0 \longrightarrow \begin{cases} q_a & \text{if } (m_1, w_1) \text{ marry,}\\ q_b & \text{if } (m_1', w_1') \text{ marry,}\\ q_3 & \text{if some couple other than } (m_1, w_1) \text{ or } (m_1', w_1') \text{ marry,}\\ q_0 & \text{otherwise.} \end{cases}$$

q_a: An absorbing state. As the state has just become q_a the remaining market is $S_a := S_0 \backslash (m_1, w_1) \in \mathscr{S}_{\mu_a}$ where S_0 denotes the initial market. In q_a the players follow the $\hat{\mu}_a$-strategy profile where $\hat{\mu}_a$ is μ_a restricted to S_a.

q_b: An absorbing state. As the state has just become q_b the remaining market is $S_b := S_0 \backslash (m_1', w_1') \in \mathscr{S}_{\mu_b}$. In q_b the players follow the $\hat{\mu}_b$-strategy profile where $\hat{\mu}_b$ is μ_b restricted to S_b.

q_3 : An absorbing state. Let S_3 denote the (history-dependent) remaining market as the state has just become q_3. In q_3 the players follow the μ^{S_3}-strategy profile where μ^{S_3} denotes the women-optimal matching for S_3.

Note that σ satisfies the private-dinner condition. μ_a obtains almost surely conditional on q_a being reached. μ_b obtains almost surely conditional on q_b being reached. q_a is reached if the first meeting between (m_1, w_1) takes place before the first meeting between (m'_1, w'_1); q_b is reached if the first meeting between (m'_1, w'_1) takes place before the first meeting between (m_1, w_1). It is easy to verify that μ_a and μ_b both obtain with probability 0.5.

To see that σ is indeed a limit equilibrium it suffices to check the absence of profitable one-deviations in q_0 because μ_a restricted to S_a, μ_b restricted to S_b, and μ^{S_3} restricted to S_3 are stable matchings for the respective submarkets and hence Proposition 1.5.2 is applicable to these cases. Let $V(x, \delta)$ denote the expected payoff for player x under σ if the discount factor is δ. For any $h \in \hat{H}$ such that the state of h is q_0, player x's expected payoff in $\Gamma(h)$ is also $V(x, \delta)$. We have $\overline{V}(m) := \lim_{\delta \to 1} \delta V(m, \delta) = \frac{1}{2}u(m, \mu_a(m)) + \frac{1}{2}u(m, \mu_b(m))$ for any $m \in M$ and $\overline{V}(w) := \lim_{\delta \to 1} \delta V(w, \delta) = \frac{1}{2}v(\mu_a(w), w) + \frac{1}{2}v(\mu_b(w), w)$ for any $w \in W$. Observe that $v(m, w) > \overline{V}(w)$ if and only if m is w's first choice. Thus accepting only her first choice is optimal for every woman in state q_0 given δ sufficiently close to 1. In particular it is optimal for w_2 to reject m_1 despite they form a blocking pair against μ_a and for w'_2 to reject m'_1 despite they form a blocking pair against μ_b. The men's incentives can be verified by routine inspection. □

Example 2 shows that, despite Proposition 1.5.4, unstable matchings may obtain in private-dinner equilibria. Suppose (m_1, w_2), a pair that blocks one of the probable outcome matchings μ_a, meet when the remaining market is the initial market. At this point (m_1, w_2) do not both profit from circumventing the equilibrium by marrying each other, because the game can still go either way, leading to μ_a or μ_b. Indeed, m_1 would accept w_2 yet w_2 would reject m_1, because w_2 still has a chance to marry a more preferred man, m_2, with probability 0.5. If (m'_1, w'_1) is the first couple to marry then w_2 will marry m_2 eventually. However, if instead (m_1, w_1) is the first couple to marry, then for w_2, m_2 is impossible and m_1 is no more, forcing her to marry the last choice m_3. w_2 would regret that she had rejected m_1, as an old English saying goes: *He that will not when he may; when he will, he shall have Nay.*

Regret occurs in a nonstationary setting because the search prospect may change as the game unfolds. Regret does not occur in a model with a completely stationary search setting, such as the one in Adachi (2003). The model in Lauermann and Nöldeke (2014) also assumes a stationary search setting but has a trace of nonstationarity: Players may be forced to leave the market as singles, at the point of which one's continuation payoff drops to zero. It is this trace of nonstationarity that makes possible the emergence of regret and unstable matchings in some equilibria in which players are forced to leave as singles after having rejected acceptable options in the past.

Example 3. Coordination failure

The initial market is described by the following lists:

$$\begin{aligned} P(m_1) &= w_1, w_2, & P(w_1) &= m_1, m_3, \\ P(m_2) &= w_2, w_3, & P(w_2) &= m_2, m_1, \\ P(m_3) &= w_3, w_1, & P(w_3) &= m_3, m_2. \end{aligned}$$

Moreover, for any $m \in M$, if w and w' are m's first and second choices respectively then $u(m, w') > \frac{2}{3}u(m, w) + \frac{1}{6}u(m, w')$. Similarly for any $w \in W$, if m and m' are w's first and second choices respectively then $v(m', w) > \frac{2}{3}v(m, w) + \frac{1}{6}v(m', w)$. The contact function satisfies $C(m, w, S) = C(\hat{m}, \hat{w}, S)$ for any pairs $(m, w) \in S$ and $(\hat{m}, \hat{w}) \in S$ for any $S \in \mathscr{S}$. Observe that the initial market satisfies the Sequential Preference Condition and so do all of its submarkets.

In the limit equilibrium σ, when the remaining market is the initial market, each man accepts every acceptable woman and each woman accepts every acceptable man. After the first marriage is realized the players follow the μ^S-strategy profile where μ^S is the unique stable matching for the remaining market S right after the first marriage. σ satisfies the Markov condition.

The following four matchings obtain with positive probability:

$$\begin{aligned} \mu_0 :\quad & m_1 \mapsto w_1, & \mu_1 :\quad & m_1 \mapsto w_2, & \mu_2 :\quad & m_1 \mapsto w_1, & \mu_3 :\quad & m_1 \mapsto s, \\ & m_2 \mapsto w_2, & & m_2 \mapsto s, & & m_2 \mapsto w_3, & & m_2 \mapsto w_2, \\ & m_3 \mapsto w_3, & & m_3 \mapsto w_3, & & m_3 \mapsto s, & & m_3 \mapsto w_1, \\ & & & w_1 \mapsto s, & & w_2 \mapsto s, & & w_3 \mapsto s. \end{aligned}$$

μ_0 is the unique stable matching for the initial market. Each of $\mu_i, i = 1, 2, 3$ is Pareto-dominated by μ_0. It is easy to verify the following: μ_0 obtains almost surely conditional on the event E_0 that a man meets his first choice on the first day of the game. $\mu_i, i = 1, 2, 3$, obtains almost surely conditional on the event E_i that m_i meets his second choice on the first day of the game. Let $p(\mu)$ denote the unconditional probability that μ obtains under σ. For any i,

$$p(\mu_i) = \Pr(E_i) \times 1 + \sum_{j \neq i} \Pr(E_j) \times 0 + \Big(1 - \sum_{j=0}^{3} \Pr(E_j)\Big) p(\mu_i).$$

To see that the equality holds, note that conditional on none of E_i, $i = 0, 1, 2, 3$, having occurred on the first day, the probability that μ_i obtains remains $p(\mu_i)$ because σ satisfies the Markov condition. By the specification of the contact function we have $\Pr(E_0) = 1/3$ and $\Pr(E_i) = 1/9$ for $i = 1, 2, 3$. Thus $p(\mu_0) = 1/2$ and $p(\mu_i) = 1/6$ for $i = 1, 2, 3$. Observe that under σ each player marries his or her first choice with probability 2/3, second choice with probability 1/3, and stays single with probability 1/6.

To verify that σ is indeed a limit equilibrium, it suffices to check the absence of profitable one-deviations when the remaining market is the initial market, since Proposition 1.5.2 covers the other cases. Let $V(x, \delta)$ denote player x's expected payoff under σ if the discount factor is δ. $V(x, \delta)$ is also x's expected payoff in $\Gamma(h)$ for any $h \in \hat{H}$ such that $S(h)$ is the initial market since σ satisfies the Markov condition. It is easy to verify that $\overline{V}(w) := \lim_{\delta \to 1} V(w, \delta) = \frac{2}{3}v(m, w) + \frac{1}{6}v(m', w)$ where m and m' are w's first and second choices respectively. We have $\delta V(w, \delta) < \overline{V}(w) < v(m', w)$ where the second inequality is by assumption, implying it is optimal for w to accept both m and m'. Similarly it is optimal for each man to accept every acceptable woman. □

In Example 3, σ is strictly Pareto-dominated by the μ_0-strategy profile, which by Proposition 1.5.2 is also a limit equilibrium. Inefficiency arises because of a coordination failure due to self-confirmation of mutual doubts. Despite being each other's first choice, m_1 does not commit to waiting for w_1 because w_1 does not commit to waiting for m_1 because m_1 does not commit to waiting for w_1 and so on *ad infinitum*. The example shows that the Markov condition, which is stronger than the private-dinner condition, and the Sequential Preference Condition combined are not sufficient to rule out limit equilibria in which unstable matchings obtain with positive probability.

Delay

This subsection studies whether equilibrium delay may cause significant efficiency loss even if search frictions are small. Two types of delay come to mind. The first type refers to the situation that a game never ends. Recall a game ends when the remaining market becomes trivial. A never-ending game implies at least one mutually beneficial marriage is not realized while the pertained players stay unmarried into the infinite future. Such a situation is not unlike a never-ending negotiation over how to split the money on the table. The upcoming proposition essentially rules out this type of delay in equilibrium.

Proposition 1.5.7. *A game ends almost surely in any subgame perfect equilibrium.*

The proof, found in Appendix A.1, hinges on the observation that in any subgame perfect equilibrium, after any nonterminal history, the probability that no marriage occurs during the next T days is less than some constant $\overline{p} < 1$ if T is sufficiently large. It follows that the probability that no marriage occurs for a duration of kT days is less than $\overline{p}^k$, thus the probability that a marriage will occur during the next kT days becomes arbitrarily close to 1 as k tends to ∞, implying the initial market will eventually shrink to a trivial market.

In contrast to a never-ending game, the second type of delay is in its literal sense: Some marriages are realized too late. We define what it means to be "too late" as follows: Let $\mathscr{M}$ denote the set of all matchings for the initial market. For any strategy profile σ and man m define the efficiency loss due to delay under σ as

$$L_\sigma(m,\delta) := \sum_{\mu\in\mathscr{M}} p_\sigma(\mu)u(m,\mu(m)) - V_\sigma(m,\delta)$$

where $p_\sigma(\mu)$ denotes the probability that μ obtains under σ and $V_\sigma(m,\delta)$ denotes m's expected payoff under σ if the discount factor is δ. Define $L_\sigma(w,\delta)$ for each woman w analogously. Efficiency loss is thus measured as the difference between a player's expected payoff from the immediate resolution of the lottery induced by σ and his or her expected payoff under σ. $\delta < 1$ implies $L_\sigma(x,\delta) \geq 0$. We ask whether equilibrium efficiency loss due to delay vanishes as search frictions vanish, that is, whether the equality $\lim_{\delta\to 1}\sup_{\sigma\in\Sigma(\delta)} L_\sigma(x,\delta) = 0$ holds for any player x where $\Sigma(\delta)$ denotes the set of all subgame perfect equilibria of the game with discount factor δ. Examples 4 and 5 show that equilibrium efficiency loss due to delay may remain as search frictions vanish.

Example 4. Wait and see

The initial market is described by the following lists:

$$\begin{aligned} P(m_1) &= w_1, w_2, & P(w_1) &= m_2, m_1,\\ P(m_2) &= w_2, w_1, & P(w_2) &= m_1, m_2. \end{aligned}$$

Each player receives a payoff of 3 from marrying the first choice and 1 from marrying the second choice. The contact function satisfies $C(m,w,S) = C(\hat{m},\hat{w},S)$ for any pairs $(m,w)\in S$ and $(\hat{m},\hat{w})\in S$ for any $S\in\mathscr{S}$.

Fix $\eta\in(0.5,1)$. Given δ sufficiently close to 1 there exists $\tau(\delta)\in\mathbb{N}$ such that $0.5 < \delta^{\tau(\delta)}$ and $\delta^{\tau(\delta)-1} < \eta$. Consider the strategy profile $\sigma(\delta)$ specified by an automaton with the following states:

q_0 : The initial state. In q_0, each man accepts no one; each woman accepts her first choice. The transition rules are

$$q_0 \longrightarrow \begin{cases} q_M & \text{if } (m_1,w_1) \text{ or } (m_2,w_2) \text{ meet on the } \tau(\delta)\text{th day,}\\ q_W & \text{if } (m_1,w_2) \text{ or } (m_2,w_1) \text{ meet on the } \tau(\delta)\text{th day,}\\ q_3 & \text{if some pair marry before the } \tau(\delta)\text{th day,}\\ q_0 & \text{otherwise.} \end{cases}$$

q_M: An absorbing state. When the state has just become q_M the remaining market is the initial one. In q_M the players follow the μ^M-strategy profile where μ^M is the men-optimal matching for the initial market.

q_W: Symmetric to q_M in which the μ^W-strategy profile is followed where μ^W is the women-optimal matching for the initial market.

q_3: An absorbing state. The remaining pair (m, w) accept each other.

Note that $\sigma(\delta)$ satisfies the private-dinner condition. Under $\sigma(\delta)$, no player marries during the first $\tau(\delta) - 1$ days. The pair that meet on the $\tau(\delta)$th day marry on that day and the other pair marry on the next day. It is easily verified that μ^M and μ^W both obtain with probability 0.5, and that on the $\tau(\delta) - n$th day player x's expected payoff $K(x, n, \delta)$ is greater than $\delta^{n+1}(0.5 \times 3 + 0.5 \times 1) = 2\delta^{n+1}$ for any n such that $0 < \tau(\delta) - n < \tau(\delta)$. By the choice of $\tau(\delta)$, $2\delta^{n+1} \geq 2\delta^{\tau(\delta)} > 1$. To verify that $\sigma(\delta)$ is a subgame perfect equilibrium given δ sufficiently close to 1, it suffices to check the absence of profitable one-deviations in q_0 because Proposition 1.5.2 and Lemma 1.5.1 cover the other states. Suppose (m_1, w_1) meet in q_0 on the $\tau(\delta) - n$th day. w_1's expected payoff from accepting m_1 is 1 whereas that from rejecting him is $K(w_1, n, \delta) > 1$, thus rejecting m_1 is optimal. m_1's expected payoffs from accepting and rejecting w_1 are the same as he will be rejected anyway, thus rejecting w_1 is optimal. The other cases are similar or can be verified by routine inspection.

Note that $V_{\sigma(\delta)}(x, \delta) < \delta^{\tau(\delta)-1}(0.5 \times 3 + 0.5 \times 1) = 2\delta^{\tau(\delta)-1}$ for any player x since x cannot marry before the $\tau(\delta)$th day. Thus

$$L_{\sigma(\delta)}(x, \delta) = 0.5 \times 3 + 0.5 \times 1 - V_{\sigma(\delta)}(x, \delta) > 2(1 - \delta^{\tau(\delta)-1}) > 2(1 - \eta) > 0.$$

That $\sigma(\delta)$ is a subgame perfect equilibrium of the game with discount factor δ sufficiently close to 1 implies $\lim_{\delta \to 1} \sup_{\sigma \in \Sigma(\delta)} L_\sigma(x, \delta) \geq 2(1 - \eta)$ for any x. □

In Example 4, each player waits to see if oneself will be lucky to marry his or her first choice, the revelation of which is on the $\tau(\delta)$th day. As search frictions vanish, players become increasingly willing to wait longer. Efficiency loss lingers as the length of waiting grows in pace with the vanishing search frictions.

Example 5. War of attrition

Take the game in Example 4. Consider strategy profile $\sigma(\delta)$ under which each player accepts his or her first choice in the remaining market with certainty and second choice (if there is one) with probability $q(\delta) \in (0, 1)$. Note that $\sigma(\delta)$ satisfies the Markov condition. We want to choose $q(\delta)$ so that $\sigma(\delta)$ is a subgame perfect equilibrium. Note that when the remaining market is the initial one, w_1 randomizes between accepting and rejecting m_1. Conditional on w_1 having rejected m_1 on a given day, all of the following four events occur with probability $q(\delta)/4$: (1) (m_1, w_1) marry tomorrow, (2) (m_2, w_1) marry tomorrow, (3) (m_2, w_2) marry tomorrow, and then (m_1, w_1) marry on the day after tomorrow, (4) (m_1, w_2) marry tomorrow, and then (m_2, w_1) marry on the day after tomorrow. With the remaining probability $1 - q(\delta)$ no one marries tomorrow. To make w_1 indifferent between accepting and rejecting m_1, we must have

$$v(m_1, w_1) = \delta \Bigg[\frac{q(\delta)}{4} \Big(v(m_1, w_1) + v(m_2, w_1) + \delta v(m_1, w_1) + \delta v(m_2, w_1) \Big) + (1 - q(\delta)) v(m_1, w_1) \Bigg].$$

Substituting in $v(m_1, w_1) = 1$ and $v(m_2, w_1) = 3$ we conclude that $q(\delta) = (1-\delta)/\delta^2$. The same indifference argument applies to each of the other players. $(1-\delta)/\delta^2 \in (0, 1)$ for any $\delta > (\sqrt{5} - 1)/2 \approx 0.618$. It is easily verified that for δ sufficiently close to 1, $\sigma(\delta)$ is a subgame perfect equilibrium if and only if $q(\delta) = (1-\delta)/\delta^2$.

Now evaluate the efficiency loss due to delay. $\sigma(\delta)$ induces a lottery in which each player marries either player on the other side with probability 0.5. For any player x and $\delta > (\sqrt{5} - 1)/2$ obviously $V_{\sigma(\delta)}(x, \delta) = 1/\delta$. Thus

$$L_{\sigma(\delta)}(x, \delta) = 0.5 \times 3 + 0.5 \times 1 - 1/\delta.$$

Note that $L_{\sigma(\delta)}(x, \delta)$ is positive for any $\delta > (\sqrt{5} - 1)/2$ and $\lim_{\delta \to 1} L_{\sigma(\delta)}(x, \delta) = 1$. Thus $\lim_{\delta \to 1} \sup_{\sigma \in \Sigma(\delta)} L_\sigma(x, \delta) \geq 1$. □

Example 5 resembles a war of attrition. A player "chickens out" to accept his or her second choice with probability $(1 - \delta)/\delta^2$, which tends to 0 as δ tends to 1. It is notable that $L_{\sigma(\delta)}(x, \delta)$ increases with δ: Ironically, the expected total cost of search goes up as the average cost goes down. In both Examples 4 and 5, the efficiency gain from vanishing search frictions is undone in equilibrium by increasing delay in such a way that the incentives of the players are preserved, leaving the net efficiency loss significant.

A Sufficient Condition for Equilibrium Uniqueness

In this subsection a condition on the preference structure is identified under which the game has an essentially unique subgame perfect equilibrium if search frictions are small.

For $h \in Z$ define $o(h) := \big(\mu_h(m), t_h(m)\big)_{m \in M}$ where μ_h is the outcome matching of h and $t_h(m) \in \{1, 2, ...\}$ is the date of $(m, \mu_h(m))$'s marriage under h if $\mu_h(m) \in W$ or the date of the last day of h if $\mu_h(m) = s$. $t_h(m) = \infty$ if $\mu_h(m) = s$ and h is infinite. $o(h)$ records each realized marriage and its date under h. If $o(h) = o(h')$ then under h and h' every player is married to the same person on the same day or else stays single. Two strategy profiles are **outcome equivalent** if they induce the same probability measure on $\{o(h) : h \in Z\}$. A game has an **essentially unique subgame perfect equilibrium** if all of its subgame perfect equilibria are outcome equivalent.

A pair (m, w) is **woman-acceptable** if m is acceptable to w. Lemma 1.5.1(b) implies equilibrium marriages may occur only between woman-acceptable pairs. For a submarket S, player $x \in S$ is a **top player** for S if x and some $y \in S$ form a top pair for S and moreover one of the following is true: (1) $|A^S(x)| = 1$ or (2) $|A^S(x)| > 1$ and x is a top player for $S \backslash (m, w)$ for any woman-acceptable pair $(m, w) \in S$ such that $x \notin \{m, w\}$.

An immediate observation is that if x is a top player for S, and S' is derived from S as a result of a sequence of other woman-acceptable pairs having left, then either $|A^S(x)| = 0$ or x is still a top player for S'. The observation also implies that, despite the recursive definition, whether x is a top player for S can be verified mechanically in a finite number of steps. Since x (assumedly male) being a top player for S implies he is part of a top pair for S, it follows that he continues to be part of a top pair for the remaining market as other woman-acceptable pairs leave S (although his partner in the current top pair might change), until x can no longer find an acceptable woman who also finds him acceptable. In the case that every man is acceptable to every woman and vice versa, x being a top player for the initial market implies he is the favorite man of his favorite woman in the remaining market at any moment when he is still unmarried.

Proposition 1.5.8. *Suppose there is an ordering of the men $m_1, ..., m_{|M|}$, an ordering of the women $w_1, ..., w_{|W|}$ and a positive integer k such that:*

1. *For any $i \leq k$, (m_i, w_i) is a top pair for $S_i := (\{m_j : j \geq i\}, \{w_j : j \geq i\}, u, v)$, and moreover m_i or w_i is a top player for S_i.*

2. *$S_{k+1} := (\{m_j : j \geq k+1\}, \{w_j : j \geq k+1\}, u, v)$ is trivial.*

There exists $d < 1$ such that for any $\delta > d$, (M, W, u, v, C, δ) has an essentially unique subgame perfect equilibrium.

The proof is provided in Appendix A.2. Given the premise of Proposition 1.5.8, the initial market has a unique stable matching μ such that $\mu(m_i) = w_i$ for $i \leq k$, and $\mu(m_i) = \mu(w_i) = s$ for $i > k$. To sketch out the proof, for simplicity suppose every man is acceptable to every woman and vice versa. The first step is proving by induction on the market size that a top player w_1 (assumedly female) for the initial market marries her favorite man almost surely and upon first meeting in any subgame perfect equilibrium for δ sufficiently close to 1. The inductive step goes as follows. Suppose m_1 is w_1's favorite man in the initial market. By Lemma 1.5.5, (m_1, w_1) accept each other when they meet. Fix an equilibrium. Suppose w_1 deviates to an alternative strategy under which she accepts only m_1 when the remaining market is the initial market, and switches back to following the equilibrium strategy after someone has married. For δ close to 1, w_1's expected payoff from the deviation is strictly larger than $v(m', w_1)$ where m' is w_1's second favorite man, because if someone has married before (m_1, w_1)'s first meeting then in the continuation subgame, as the market has become smaller to which the inductive hypothesis is applicable, w_1 marries her favorite man in the remaining market, which can be m_1 or m', almost surely and upon first meeting, whereas if (m_1, w_1) meet before anyone has married then they marry each other. Thus w_1's equilibrium payoff must as well be strictly larger than $v(m', w_1)$, implying if no one has married, w_1 rejects any man worse than m_1. Given w_1's behavior pinned down, it is shown that m_1's best response is accepting only w_1 when no one has married. Combined with the inductive hypothesis, it follows that (m_1, w_1) marry almost surely and upon first meeting. With (m_1, w_1)'s equilibrium behavior determined this way, we may treat them as nonstrategic "dummy players" who only accept each other. Then since one of m_2 or w_2 is a top player for the submarket without (m_1, w_1), applying the analogous logic it is shown that (m_2, w_2) also marry almost surely and upon first meeting, and so on and so forth until the market unravels to the trivial S_{k+1}. From Lemma 1.5.1 it follows that every player in S_{k+1} must stay single.

The condition proposed in Proposition 1.5.8, for convenience referred to as Condition 1.5.8, is stronger than the Sequential Preference Condition. The initial market from Example 3, for instance, satisfies the Sequential Preference Condition but fails Condition 1.5.8. Indeed, in Example 3, although (m_1, w_1) is a top pair for the initial market, if w_1 goes "astray" and marries m_3 then in the remaining market m_1 is no longer part of a top pair, and consequently his search prospect drops drastically because he has no chance of marrying his second choice w_2. It is this possibility of a large drop in the search prospect that makes m_1 willing to accept his second choice, w_2, even when w_1 is still in the market. For a market that satisfies Condition 1.5.8, in contrast, if m (assumedly male) is a top player for the initial market and forms a top pair with w, then even if w marries someone else, in the remaining market m still forms a top pair with w' whom he likes just next to w. Loosely speaking, the search prospect of m would only experience a gradual drop if w marries someone else. This ensures that m will wait for w if she is still in the market, preventing self-confirmation of mutual doubt present in Example 3 from arising.

An acyclicity condition under which a marriage market has a unique stable matching is proposed in Romero-Medina and Triossi (2013),[11] which requires that the preferences of the players on one side over acceptable players on the other side do not display cycles. It can be shown that the acyclicity condition implies Condition 1.5.8. Condition 1.5.8 is weaker, most importantly because it allows preference cycles. For instance, the market with preference lists

$$P(m_1) = w_1, w_2, \qquad P(w_1) = m_1, m_2,$$
$$P(m_2) = w_2, w_1, \qquad P(w_2) = m_2, m_1$$

displays preference cycles on both sides but satisfies Condition 1.5.8.[12] It is worth-

[11] I am indebted to a referee for referring me to this paper.

[12] A weaker condition, the absence of simultaneous cycles, is proposed in Romero-Medina and

while noting that Condition 1.5.8 is an implication of the commonly seen preference structure such that players on at least one side of the market share the same preference ordering over those on the other side.

1.6 Conclusion

This paper studies a search and matching game with a marriage market embedded, and analyzes whether matchings that arise in equilibria are stable when search frictions are small. It is found that this is not the case in general. Unstable matchings may arise for many reasons and under restrictive conditions. Moreover, significant loss of efficiency due to delay may be incurred in equilibria even if search frictions are small. A condition that implies preference alignment in a strong sense ensures equilibrium uniqueness, restoring stability and efficiency.

Acknowledgements

I wish to thank the editor and two referees of the Journal of Economic Theory, V. Bhaskar, Tilman Börgers, In-Koo Cho, Scott Kominers, John Quah, Bartley Tablante, and the participants at the 2013 Asian Meeting of the Econometric Society and the 2014 Stony Brook International Conference on Game Theory for enlightening discussions and helpful comments. I am especially grateful to Stephan Lauermann for his encouragement and guidance which made possible the fruition of this paper.

Triossi (2013) that also guarantees the stable matching is unique. Condition 1.5.8 is neither necessary nor sufficient for that weaker condition. The initial market from Example 3 satisfies the absence of simultaneous cycles but violates Condition 1.5.8. An example that violates the absence of simultaneous cycles but satisfies Condition 1.5.8 is given in Appendix A.4.

CHAPTER II

A Simple Model of Competition Between Teams, *jointly with Kfir Eliaz*

2.1 Introduction

Many economic, political and social activities are performed by groups or organizations rather than individuals. When firms compete, the strategic interaction is really between *collectives* of individuals that make up the firms. Electoral competition between candidates involves strategic interaction between *teams* consisting of the candidates themselves, their consultants and the activists that support them. Lobbying efforts are carried out by interest *groups* who need to coordinate the actions of their members in response to the actions of other interest groups. Likewise, ethnic conflicts involve different *peoples* who are united by a common background such as religion, origin or economic status.

Despite the ubiquity of strategic interactions between groups, the majority of economic analysis treats players in game theoretic models as individual entities. While this may be a helpful simplification it ignores the interplay between *intra*-group strategizing (how each member of a group reasons about the actions of other members of the *same* group) and *inter*-group strategizing (how each member of a group reasons about the actions of the members of the *opposing* groups), which may have important implications for the outcome of the interaction. Furthermore, by explicitly modeling each participating unit as a collective of decision-makers, one may be able to gain insights on how different group attributes (such as size, for instance) can affect the outcome.

In light of this we propose and analyze a model of competition between groups. We focus on the case of two competing groups, or teams, of possibly different sizes. Following the literature on group contests (see below) we model the interaction between teams as a generalized all-pay auction with incomplete information. Individual members of each team exert effort to increase the performance of one's own team via an additively separable aggregation function. The team with a higher performance wins, and its members enjoy the prize as a public good. The value of the prize is identical to members of the same team but is unknown to the other team. Each player individually bears the cost of his own effort regardless of winning or not.

We consider three cases: That the aggregation function is concave, convex, or linear. We show that if the aggregation function is strictly concave or convex then there is a unique monotone equilibrium in which every player actively participates, that is, everyone makes positive effort in expectation. If the aggregation function is linear then all monotone equilibria are equivalent at the team level. Subsequent analysis focuses on equilibria with active participation. Our main interest is in understanding the implication of team size on the probability of winning and on the members' payoffs. The main finding is that these implications are determined by the curvature of the aggregation function. We first show that the bigger team is more likely to win if the aggregation function is concave, less likely if convex, or equally likely if linear. The underlying intuition is that the curvature of the aggregation function determines whether *in equilibrium* additional members augment or reduce the productivity of existent members. Second, we show that when the aggregation function is concave or linear, then the expected payoff for a player in the bigger team is higher than that in the smaller team. Moreover, there also exist convex aggregation functions under which the same result holds, despite that the bigger team is less likely to win, because a member in the bigger team is also responsible for less work.

Since group size can have important implications, we investigate how teams might form. We consider a two stage game such that in the first stage players split into teams, and in the second stage the teams compete. Our main result is that team formation depends on how each member's payoff changes with the sizes of the two teams. If a member's payoff increases with his own team size and decreases with the size of the opponent team, then there exists a unique nontrivial equilibrium in the first stage in which the teams are stochastically formed. While this is true for a linear aggregation rule, it remains an open question how more general aggregation functions affect team formation.

2.2 Literature

Our analysis is closely related to the literature on contest theory. Most papers in this literature can be classifed according to how their models fit the following binary categorizations:

1. Who are competing: individuals or teams?

2. How is the winner chosen: stochastically or deterministically?

3. Information structure: complete or incomplete?

The literature can thus be organized into a $2 \times 2 \times 2$ design:

Individuals	Complete	Incomplete
Stochastic		
Deterministic		

Teams	Complete	Incomplete
Stochastic		
Deterministic		

By now there is a vast literature that fills the cells in the "Individuals" table. Some of the prominent works in the complete information column include Hillman and Riley (1989), Michael R. Baye (1996) and Siegel (2009) for the "deterministic" case and Siegel (2009) and Cornes and Hartley (2005) for the "stochastic" case . The incomplete information column includes Amann and Leininger (1996), Lizzeri and Persico (2000), Kirkegaard (2013) and Siegel (2014) for the "deterministic" case, and Ryvkin (2010), Ewerhart and Quartieri (2013) and Ewerhart (2014) for the "stochastic" case.

There is also an extensive literature on team contests with complete information. This literature includes Skaperdas (1998), Nitzan (1991), Esteban and Ray (2001, 2008), Nitzan and Ueda (2009, 2011), Münster (2007, 2009), Konrad and Leininger (2007), and Konrad and Kovenock (2009), among many others. In particular, our modeling approach of assuming that the value of winning is a public good among team members follows that of Baik, Kim, and Na (2001), Topolyan (2013) Chowdhury and Topolyan (2015) and Chowdhury, Lee, and Topolyan (2016). These studies assume that the group bid is either the minimum or the maximum of the individual bids, whereas we assume that a possible non-linear function aggregates the individual bids into a total group bid.

Our work falls into the "incomplete information column, which has only been filled recently by Fu, Lu, and Pan (2015) and Barbieri and Malueg (2015). The first paper analyzes a general model that accomodates each of the cells in the "Teams" table. However, their work differs from ours in that they study a multi-battle contest: Players from two equal-size teams form pairwise matches to compete in distinct two-player all-pay auctions, and a team wins if and only if its players win a majority of the auctions. In contrast, we analyze a contest in which the members of both teams participate simultaneously in one big all-pay auction.

The second paper by Barbieri and Malueg (2015) is more closely related to our work since it also analyzes a static incomplete information (static) all-pay auction between teams that may differ in size. In contrast to us, they assume that the value of winning is an independent private value of each team member, and that the team's bid is equal to the maximal bid among its members. Under this specification they show that in the case of two teams with different cdfs, a team's probability of winning increases (decreases) with size if its cdf is inelastic (elastic). We assume that all members of a team have the same commonly known value of winning, but this value is unknown to the opponent team. As stated above, we link the size advantage/disadvantage to the curvature of the bid-aggregation function. In addition, we also analyze the implication of group size on individual welfare, and show that this also depends on the curvature of the aggregation function.

The implication of group size on the probability of winning address the well-known "group size paradox", which argues that free-riding makes a bigger team weaker. The paradox fails in many of the teams-stochastic-complete models (e.g., Esteban and Ray (2001) and Nitzan and Ueda (2011)), but is satisfied in the teams-deterministic-complete model of Barbieri, Malueg, and Topolyan (2013).[1]

2.3 The Model

Two teams, B and S, compete for a prize. Team B has n_B players and team S has n_S players where $n_B \geq n_S$. Thus B stands for "big" and S "small". Sometimes we use X to denote a generic team and Y the other team. Given any player i let $i \in X$ denote that i is a member of team X.

Competition takes the following form: All players simultaneously choose some action from $\mathbb{R}^+$. Player i's chosen action e_i is interpreted as the amount of of effort that player i exerts in the contest. Team X's overall performance, measured by the *score*, is given by the *aggregation function* $H\Big((e_i)_{i\in X}\Big) = \sum_{i\in X} h(e_i)$ where h is some real valued function. Assume that h is strictly increasing, twice differentiable, and normalize $h(0) = 0$. Clearly H is convex, concave or linear if and only if h is respectively convex, concave or linear.

The higher scoring team wins the prize. A tie is broken by a fair coin. Every member in team X receives a payoff of $v_X \in [0, 1]$ from winning the prize. v_X is known to members in team X before the contest starts, but is unknown to members of the other team. It is common knowledge that the ex ante distribution of v_B and v_S are both F, where F admits a strictly positive density function f. Regardless of winning the prize or not, each player pays a cost equal to the amount of effort he has exerted. Thus the net payoff to player i in team X who has exerted effort e_i is $\mathbf{1}_X v_X - e_i$ where $\mathbf{1}_X$ is equal to 1 if team X has won or 0 otherwise.

2.4 Analysis

In the paper we focus on pure strategy Bayesian Nash equilibria (BNE).[2] A BNE can be characterized by a vector of *effort functions* (e_i) such that $e_i(v)$ is the amount of

[1] The papers demonstrating the failure of the group-size paradox do so by assuming diminishing marginal team performance to the cost born by each individual. However, since these papers analyze a very different framework than ours, our result does not follow from theirs.

[2] One of the reasons we study the incomplete information model, instead of the complete information counterpart, is that pure strategy equilibrium typically does not exist in the latter. On the other hand, since we impose no additional assumption on the valuation distribution F, all results in the paper hold for a model with "almost complete" information in which F is arbitrarily close to a degenerate distribution.

effort player i exerts if the value of the prize is v. Given any BNE there are associated equilibrium *score functions* P_B and P_S such that $P_X(v) := \sum_{i \in X} h(e_i(v))$ is the score of team X if the value of the prize is v. A BNE is *monotone* if every player's effort is weakly increasing in his valuation of the prize. For the rest of the paper we focus on monotone BNE. It is straightforward that in a monotone BNE P_B and P_S are weakly increasing as well. Let G_X denote the ex ante distribution of P_X.

The following lemma extends Lemmas 1-3 and 5 of Amann and Leininger (1996) to the present setting.

Lemma 2.4.1. *For any monotone BNE:*

1. *G_B and G_S have common support.*
2. *G_X is continuous on $[0, P_X(1)]$.*
3. *The support of G_X is $[0, P_X(1)]$.*
4. *$P_B(1) = P_S(1)$.*
5. *$\min\{G_B(0), G_S(0)\} = 0$.*

Proof: The proofs of Properties 1, 2, 3, 5 are respectively similar to those of Lemmas 1, 2, 4, 5 of Amann and Leininger (1996). That in the current setting the contest is between teams instead of individuals, and that a team's aggregation function is not linear, do not fundamentally change the original proofs apart from some superficial technical difference. Property 4 is an implication of Properties 1 and 3. □

Given a monotone BNE, for any team X, player $i \in X$ and value $v \in [0, 1]$, $e_i(v)$ is the solution to the following maximization problem:

$$\max_{e \geq 0} G_Y\Big(\sum_{j \in X, j \neq i} h(e_j(v)) + h(e)\Big)v - e.$$

It follows from Lemma 2.4.1 that the first order condition

$$G'_Y\Big(P_X(v)\Big)h'(e_i(v)) = 1 \tag{1}$$

holds if G_Y is differentiable at $P_X(v)$ and $e_i(v) > 0$.

We say that a BNE satisfies the *active-participation* property if there does not exist a player who always exerts zero effort. The following proposition characterizes the set of all monotone BNE for each of the three cases: (1) h is strictly concave, (2) h is strictly convex, (3) h is linear. Moreover it establishes the uniqueness of active-participation BNE in the first two cases, and shows that this BNE is in-team symmetric.

Proposition 2.4.2. *1. If h is strictly concave then there is a unique monotone BNE; this BNE satisfies active-participation and is in-team symmetric.*

2. If h is strictly convex then there is a unique active-participation monotone BNE; this BNE is in-team symmetric.

3. If h is linear then there is a continuum of monotone BNE, among which there is a unique in-team symmetric monotone BNE. Moreover, every monotone BNE of every contest (characterized by the team size configurations (n_B, n_S)) has the same equilibrium team score functions (P_B, P_S).

Proof:

Suppose h is strictly concave or convex. Pick any two players i and j in team X. If $e_i(v) > 0$ and $e_j(v) > 0$ then equation (1) implies $h'(e_i(v)) = h'(e_j(v))$, which in turn implies $e_i(v) = e_j(v)$ because h' is strictly monotone. We have established the fact that in any monotone BNE, if h is strictly concave or convex then given any v if two players in the same team are not shirking then they exert the same effort.

Now we show Part 1. Suppose h is strictly concave. For team X pick any v such that $e_i(v) > 0$ for some $i \in X$. Thus first order condition (1) holds for i. If there is some $j \in X$ such that $e_j(v) = 0$ then since $h'(0) > h'(e_i(v))$ we have

$$G'_Y\Big(P_X(v)\Big)h'(0) > 1,$$

implying that player j can profit by increasing his effort, a contradiction. Thus there does not exist such v such that some players work and some players shirk. Hence if h is strictly concave then any monotone BNE is in-team symmetric.

Clearly for any $t > 0$, $G_Y(t) = \Pr(P_Y(v) \leq t) = \Pr(v \leq P^{-1}(t)) = F(P_Y^{-1}(t))$. Thus $G'_Y(P_X(v)) = f\Big(P_Y^{-1}(P_X(v))\Big)(P_Y^{-1})'(P_X(v))$. In-team symmetry implies equation (1) can be rewritten as

$$f\Big(P_Y^{-1}(P_X(v))\Big)(P_Y^{-1})'(P_X(v))h'\Big(h^{-1}(P_X(v)/n_X)\Big)v = 1. \tag{2}$$

For any v such that $P_X(v) > 0$, equation (2) can be simplied by a change of variable $t = P_X(v)$ and becomes

$$f\Big(P_Y^{-1}(t)\Big)(P_Y^{-1})'(t)P_X^{-1}(t)h'\Big(h^{-1}(t/n_X)\Big) = 1.$$

Given Lemma 2.4.1 it is straightforward to verify that (P_B, P_S) determines a monotone BNE if and only if $P_B = \max(\beta, 0)$ and $P_S = \max(\sigma, 0)$ where (β, σ) solves the following boundary value problem:

$$f\Big(\sigma^{-1}(t)\Big)(\sigma^{-1})'(t)\beta^{-1}(t)h'\Big(h^{-1}(t/n_B)\Big) = 1 \tag{3}$$
$$f\Big(\beta^{-1}(t)\Big)(\beta^{-1})'(t)\sigma^{-1}(t)h'\Big(h^{-1}(t/n_S)\Big) = 1 \tag{4}$$

with boundary conditions:

$$\beta(1) = \sigma(1) \tag{5}$$
$$\max\{\beta(0), \sigma(0)\} = 0. \tag{6}$$

The boundary value problem is exactly the same boundary value problem that characterizes the monotone BNE of an all-pay contest between two players B and S whose valuations are independently distributed according to F, such that the score of player X is the same as his chosen amount of effort, and the cost function of exerting effort e is equal to $c_X(e) := \int_0^e \frac{1}{h'\left(h^{-1}(t/n_X)\right)}$. By Proposition 1 of Kirkegaard (2013) the auxiliary two-player game has a unique monotone BNE. [3] Part 1 immediately follows.

Now show Part 2. Suppose h is strictly convex. That an in-team symmetric monotone BNE is unique is established exactly as above. Clearly that BNE satisfies active-participation. Now we show that an active-participation monotone BNE must be in-team symmetric. Pick any active-participation monotone BNE. Suppose in team X there are players i and j whose effort functions are different. Thus there exists some v such that $e_i(v) \neq e_j(v)$. Without loss of generality assume $e_i(v) > e_j(v)$. It follows that $e_j(v) = 0$. Since e_j is weakly increasing and is not constant, there is some $\overline{v} \geq v$ such that $e_j(\overline{v} - \epsilon) = 0$ and $e_j(\overline{v} + \epsilon) > 0$ for any $\epsilon > 0$ if $\overline{v} < 1$, or $e_j(\overline{v} - \epsilon) = 0$ and $e_j(\overline{v}) = 1$ for any $\epsilon > 0$ if $\overline{v} = 1$. Suppose $\overline{v} < 1$. We have $e_j(\overline{v} + \epsilon) = e_i(\overline{v} + \epsilon) \geq e_i(v) > 0$. It follows that

$$\begin{aligned}\lim_{\epsilon \to 0}\Big(P_X(\overline{v} + \epsilon) - P_X(\overline{v} - \epsilon)\Big) &\geq \lim_{\epsilon \to 0}\Big(h(e_j(\overline{v} + \epsilon)) - h(e_j(\overline{v} - \epsilon))\Big) \\ &\geq h(e_i(v)) - h(0) \\ &= h(e_i(v)) \\ &> 0.\end{aligned}$$

However this is a contradiction because G_X is continuous at $\overline{v}$ by Lemma 2.4.1. Similarly $\overline{v} = 1$ also leads to a contradiction. Thus $e_i(v) = e_j(v)$ for any v. This establishes Part 2.

Now show Part 3. If h is linear then h' is some constant $\gamma > 0$. It is easy to verify

[3] Kirkegaard (2013) imposes additional constraints on c_X, but those constraints are irrelevant for the existence and uniqueness of the solution to the boundary value problem.

that (P_B, P_S) are monotone BNE team score functions if and only if $P_B = \max(\beta, 0)$ and $P_S = \max(\sigma, 0)$ where (β, σ) solves boundary value problem given by (3)-(6), where $h'(h^{-1}(t/n_X)) = \gamma$. Since the BVP has a unique solution, any two monotone BNE have the same team score functions. Moreover since the BVP does not depend on (n_B, n_S), neither does its solution. □

In the rest of the paper we focus on active-participation BNE because we are interested in the implication of team size on performance. If a player is passive, then he has no impact on the contest and is essentially absent from his team. Hence, the "effective" team size should not take him into account. Proposition 2.4.2 implies that restricting attention to active-partitipation BNE is without loss of generality if h is concave because in this case it is the unique monotone BNE, and is without loss of "much" generality if h is linear because all monotone BNE are equivalent at the team level. However, it is worth noting that if h is convex, then there are additional monotone BNE. It is easy to verify each of those non-active-participation BNE looks exactly like the unique active-participation BNE of the smaller contest with all the passive members removed.

2.4.1 Team Size and Performance

Since a team in our model is characterized by its size, a natural question that arises is whether a bigger team is also more likely to win. Recall that if members of both teams exert the same amount of effort, then the bigger team has a higher score and wins. Therefore, a bigger team has a size advantage. However, because the prize is a pure public good, free-riding may be more serious in the bigger team and hence may undermine its performance. It is therefore not apriori clear which of the two forces is stronger, the size advantage or the free-riding problem.

Our next result establishes that a bigger team is more (less) likely to win if there are diminishing (increasing) returns to effort. Formally, the effect of team size on the probability of winning is determined by the curvature of h. This means that a bigger team has an advantage in situations where greater expertise (which increases the rate of return) requires higher effort. On the other hand, in tasks where the biggest contribution to effort occurs early on, a smaller team will have an advantage.

Proposition 2.4.3. *In an active-participation monotone BNE:*

1. *If h is strictly concave then $P_B(v) \geq P_S(v)$ and team B is more likely to win.*

2. *If h is strictly convex then $P_B(v) \leq P_S(v)$ and team S is more likely to win.*

3. *If h is linear then $P_B(v) = P_S(v)$ and both teams win with the same probability.*

Proof: Fix an active-participation monotone BNE. Suppose h is strictly concave and

$P_B(v) = P_S(v) = t > 0$ for some v. Thus by equations (3) and (4) we have

$$(P_S^{-1})'(t)h'\Big(h^{-1}(v/n_B)\Big) = \frac{1}{f(v)v}$$
$$(P_B^{-1})'(t)h'\Big(h^{-1}(v/n_S)\Big) = \frac{1}{f(v)v}.$$

Since h is strictly increasing, $n_B \geq n_S$ implies $h^{-1}(v/n_B) < h^{-1}(v/n_S)$, which in turn implies $h'(h^{-1}(v/n_B)) > h'(h^{-1}(v/n_S))$. Thus $(P_S^{-1})'(t) < (P_B^{-1})'(t)$, which implies that $P_S'(v) > P_B'(v)$. Since $P_B(1) = P_S(1)$ by Lemma 2.4.1(4), it follows that $v = 1$. Given Lemma 2.4.1(5), that $P_S'(1) > P_B'(1)$ then implies $P_B(v) > P_S(v)$ for any $v \in (0,1)$. Part 1 of the present proposition immediately follows. Part 2 is established with the symmetric argument.

To show Part 3, notice that if h is linear then the boundary value problem given by equations (3)-(6) are symmetric in (β, σ). Thus uniqueness of the solution implies that $\beta = \sigma$, in turn implying Part 3. □

Proposition 2.4.3 implies that the existence of the "group-size paradox" depends on whether a player's mariginal return to effort is diminishing or increasing. To give some intuition for this result consider the case of a strictly concave h.

1. If a team with n members incurs a *total cost* of C, then the team score will be $nh(C/n)$. Taking the team as a whole, the marginal productivity is thus $\frac{d}{dC}nh(C/n) = h'(C/n)$. Since h is strictly concave, it follows that for a given C, the marginal productivity of a team increases with team size. Thus given the same total cost C we have $n_B h(C/n_B) > n_S h(C/n_S)$, implying that the score of the bigger team is higher. In other words, the bigger team is more efficient in generating score than a smaller team if h is concave.

2. For an n member team to achieve a total score of T, each member's effort must be $h^{-1}(T/n)$, and therefore each member's individual mariginal productivity is $h'^{-1}(T/n))$. This individual marginal productivity is also increasing in n. In other words, additional members make existent members more "productive": To achieve the same team score everyone on the bigger team can now decrease his effort, which simultaneously increases his marginal productivity. This effect is to some extent similar to that of strategic complementarity.

2.4.2 Team Size and Individual Welfare

Would a player would prefer to be in the bigger team or in the smaller team? We cannot answer this question by merely comparing the winning probabilities because a higher winning probability may require a higher level of individual effort, which may offset the gain from a higher winning probability. Suppose h is either concave, convex or linear, and that the players coordinate on the unique in-team symmetric[4] monotone BNE. It is clear that the ex ante equilibrium expected payoff for a player in team X depends only on the number of players in each team. Let $u(n_X, n_Y)$ denote this expected payoff for a player in team X. The following proposition shows that if h is weakly concave, then a player in the bigger team is better off than a player in the smaller team. Thus, diminishing marginal returns to effort imply that in equilibrium, the members of the bigger team are better off than the member of the smaller team.

Proposition 2.4.4. *If h is weakly concave, then $u(n_B, n_S) > u(n_S, n_B)$ if $n_B > n_S$.*

Proof: Let e_X denote the effort function of a player in team X in the symmetric monotone BNE. Define $P_X^{-1}(0) = 0$. Thus

$$u(n_X, n_Y) = \int_0^1 F\Big(P_Y^{-1}(P_X(v))\Big) v dF(v) - \int_0^1 e_X(v) dF(v).$$

First suppose h is strictly concave. Thus by Proposition 2.4.3(1), $P_B(v) \geq P_S(v)$ with the inequality being strict for a set of values with positive measure under F. Suppose a player in team B unilaterally deviates to using the effort function

$$\hat{e}(v) = \max\Big\{0, h^{-1}\Big(P_S(v) - (n_B - 1)h(e_B(v))\Big)\Big\},$$

that is, the player shirks as much as he can to ensure that for any v the resulting team score $\hat{P}_B(v)$ is as high as $P_S(v)$. That $P_B(v) \geq P_S(v)$ implies $\hat{e}(v) \leq e_B(v)$. It follows from $P_B(v) \geq \hat{P}_B(v) \geq P_S(v)$ that

$$P_S^{-1}(\hat{P}_B(v)) \geq \hat{P}_B^{-1}(P_S(v)) \geq P_B^{-1}(P_S(v))$$

with at least one of the two above inequalities being strict for any v such that $P_B(v) \geq P_S(v)$.

Now we show that $\hat{e}(v) \leq e_S(v)$. This is clearly true for any v such that $e_B(v) \leq e_S(v)$. For any v such that $e_B(v) \geq e_S(v)$ we have $(n_B - 1)h(e_B(v)) \geq n_S h(e_B(v)) \geq n_S h(e_S(v)) = P_S(v)$, implying $h^{-1}\Big(P_S(v) - (n_B - 1)h(e_B(v)) < 0\Big) < 0$, which in turn implies $\hat{e}(v) = 0 \leq e_S(v)$.

[4] We get in-team symmetry for free by Proposition 2.4.2 if h is strictly convex or concave.

Thus we have

$$\begin{aligned}
u(n_B, n_S) &\geq \int_0^1 F\Big(P_S^{-1}(\hat{P}_B(v))\Big) v dF(v) - \int_0^1 \hat{e}(v) dF(v) \\
&\geq \int_0^1 F\Big(\hat{P}_B^{-1}(P_S(v))\Big) v dF(v) - \int_0^1 \hat{e}(v) dF(v) \\
&> \int_0^1 F\Big(P_B^{-1}(P_S(v))\Big) v dF(v) - \int_0^1 \hat{e}(v) dF(v) \\
&\geq \int_0^1 F\Big(P_B^{-1}(P_S(v))\Big) v dF(v) - \int_0^1 e_S(v) dF(v) \\
&= u(n_S, n_B).
\end{aligned}$$

If h is linear then $P_B = P_S$ by Proposition 2.4.3. Thus $v = P_S^{-1}(P_B(v)) = P_B^{-1}(P_S(v))$ and $e_B(v) \leq e_S(v)$ with the inequality being strict if $e_B(v) > 0$. Thus

$$\begin{aligned}
u(n_B, n_S) &= \int_0^1 F(v) v dF(v) - \int_0^1 e_B(v) dF(v) \\
&> \int_0^1 F(v) v dF(v) - \int_0^1 e_S(v) dF(v) \\
&= u(n_S, n_B).
\end{aligned}$$

□

If h is convex, then the argument used in the proof of Proposition 2.4.4 does not work, and the welfare comparison in general is unclear. However, we can establish that for *some* convex h, it is still the case that each member of the bigger team receives a higher expected payoff than each member of the smaller team, even though the smaller team is more likely to win.

Proposition 2.4.5. *For any $n_B > n_S$ there exists some convex h such that $u(n_B, n_S) > u(n_S, n_B)$.*

Proof: Pick any n_B and n_S where $n_B > n_S$. Define $h_\alpha(x) = x^\alpha$. Let $v(\alpha | n_X, n_Y)$ be equal to $u(n_X, n_Y)$ for the game with $h = h^\alpha$. Clearly $v(\alpha | n_X, n_Y)$ is continuous in α. Since $v(1 | n_B, n_S) > v(1 | n_S, n_B)$ by Proposition 2.4.4, there exists some $\alpha > 1$ such that $v(\alpha | n_B, n_S) > v(\alpha | n_S, n_B)$. Clearly h^α is convex if $\alpha > 1$. □

2.4.3 Endogenous Team Formation

So far we assumed that the competing teams are exogenously given. In this section we investigate how teams come into being. To do this, we consider the following two-stage game.

1. In the first stage, an even number of $N > 2$ players simultaneously choose a letter R or L, interpreted as choosing to join team R or L.

2. In the second stage, players who chose the same letter form a team that competes with the players who chose the other letter, where the competition takes the contest format described in Section 2.

Assume that the value of the prize is realized only after the teams are formed. Also, if all players choose the same team, then there is no contest and the prize is awarded to that single team.

Obviously, there is always a trivial equilibrium in which everyone chooses the same team. However, this equilibrium is uninteresting, and also demands a lot of coordination from the players. We therefore explore other equilibria. To do this, assume that h is concave, convex or linear, and that in the second stage the teams coordinate on the unique symmetric monotone BNE. It turns out that team formation depends on how a player's ex-ante equilibrium expected payoff u changes with the teams' sizes.

Proposition 2.4.6. *If $u(n_X, n_Y)$ is weakly increasing in n_X and weakly decreasing in n_Y (but is not constant), then in every nontrivial equilibrium each player chooses either team with equal probability.*

Proof:

Pick any nontrivial equilibrium. We first show that every player is indifferent between writing down either letter. Suppose there is some player i who strictly prefers writing down R. Pick another player j and denote the probability that he writes down R as p. Let $r(n)$ denote the probability that n players other than i and j write down R, and $l(n)$ the analogous probability for L. Player i's expected payoff from writing down R is

$$\begin{aligned} v^R(p) &= \sum_{n=0:N-2} r(n)\Big[pu\Big(n+2, N-n-2\Big) + (1-p)u\Big(n+1, N-n-1\Big)\Big] \\ &= p\sum_{n=0:N-2} r(n)\Big[u\Big(n+2, N-n-2\Big) - u\Big(n+1, N-n-1\Big)\Big] \\ &\quad + \sum_{n=0:N-2} r(n)u\Big(n+1, N-n-1\Big) \end{aligned}$$

Likewise his expected payoff from writing down L is

$$\begin{aligned} v^L(p) &= \sum_{n=0:N-2} l(n)\Big[(1-p)u\Big(n+2, N-n-2\Big) + pu\Big(n+1, N-n-1\Big)\Big] \\ &= (1-p) \sum_{n=0:N-2} l(n)\Big[u\Big(n+2, N-n-2\Big) - u\Big(n+1, N-n-1\Big)\Big] \\ &\quad + \sum_{n=0:N-2} l(n)u\Big(n+1, N-n-1\Big) \end{aligned}$$

By assumption $u\Big(n+2, N-n-2\Big) - u(n+1, N-n-1) > 0$ for each $n = 0 : N-2$. Thus $v^R(p)$ is increasing in p and $v^L(p)$ is decreasing in p.

It is easy to verify that player j's expected payoff from writing down R is $v^R(1)$ and that from writing down L is $v^L(1)$, because player i writes down R with certainty. That player i strictly prefers R to L implies $v^R(p) > v^L(p)$, which in turn implies that $v^R(1) > v^L(1)$. Therefore player j also writes down R with certainty. It follows that every player writes down R with certainty, a contradiction because we have arrived at a trivial equilibrium.

Suppose player j chooses R with probability p and player i with probability q. Inheriting the notation from the previous paragraph, we have $v^R(p) = v^L(p)$ and $v^R(q) = v^L(q)$ because by the previous paragraph both i and j are indifferent between writing down either letter. Thus $p = q$ since v^R is increasing and v^L is decreasing. It follows that in the nontrivial equilibrium, for any player the following indifference condition holds:

$$\begin{aligned} &\sum_{n=0:N-1} \binom{N-1}{n} p^n (1-p)^{N-1-n} u(n+1, N-1-n) \\ = &\sum_{n=0:N-1} \binom{N-1}{n} (1-p)^n p^{N-1-n} u(n+1, N-1-n). \end{aligned}$$

That u is increasing in its first argument and decreasing in its second argument implies the left hand side is increasing in p and the right hand side is decreasing in p. Thus the only solution to this equation is $p = 0.5$. □

Propositions 2.4.1 and 2.4.3 imply that if h is linear then regardless of the number of players in each team, the unique symmetric monotone BNE has the same team score functions (P_B, P_S) and moreover $P_B = P_S =: P$. Thus

$$u(n_X, n_Y) = \int_0^1 \Big(F(v) - h^{-1}(P(v)/n_X)\Big) dF(v).$$

$u(n_X, n_Y)$ is strictly in increasing in n_X because h is increasing, and is invariant with respect to n_Y. Thus we have the following corollary of Proposition 2.4.6:

Corollary 2.4.7. *If h is linear, then in any non-trivial equilibrium each player choses either team with the same probability.*

We were not able to establish how u changes with the teams' sizes for more general h functions. Numerical simulations suggest that u is increasing in n_X and decreasing in n_Y also when h is strictly convex and strictly concave.[5]

2.5 Conclusion

We proposed to model competition between teams as a contest between two groups of players, where each single player incurs the cost of his own effort and the team's overall effort is some aggregation of the individual efforts of its members. What makes a collection of individuals a "team" is the fact that the award from winning is a pure public good among them, and the value of this public good is common knowledge among the members. In contrast, the value of the award to the opposing team is not observed and is treated as a random variable.

This model allowed us to analyze whether a bigger team has an advantage, and whether the members of a bigger team are better off. Our results shed new light on the "group-size paradox" by showing that the strategic effect of size depends on whether the marginal effect of individual effort is diminishing or not. We interpret this to mean that size advantage may depend on the particular task at hand, which determines how the marginal contribution of effort changes with the level of effort.

Future work should try and explore how other characteristics of teams - such as the composition of heterogenous teams, or the communication protocols among their members - affect the outcome of competition. The ultimate goal is to try and incorporate into standard models of strategic interaction the idea that the players are actually groups of individuals who have to consider the actions of their peers as well as those of the competing group.

Acknowledgements

We thank Qiang Fu, Johannes Hörner, Dan Kovenock, Jingfeng Lu, Ron Siegel and Iryna Topolyan for helpful comments.

[5] An alternative model of group formation i sone where team members sequentially decide which team to join. Although this model can be solved via backwards induction, it does require to put more structure on the function u. For the linear case, sequential participation would necessarily lead to the trivial equilibrium in which everyone joins the same team.

CHAPTER III

Coarse Communication and Institution Design

3.1 Introduction

This paper studies information-aggregating institutions in which precise internal communication is not possible.

It is often very difficult for an informed person to convey his information precisely to an uninformed person. In the first place, the communication instrument available to the informed person, such as words, gestures, or some other medium, may not be sufficient to express the complexity of his information. In addition, upon receiving the message that carries the information, the uninformed person may fail to digest the message to its utmost precision. For example, imagine an expert who wishes to report an important number to a manager. The decimal expansion of this number is very long. To save time, the expert reports this number only to its fourth decimal place. To save memory, the manager memorizes the report only to its second decimal place. In this example, time and memory impose constraints on communication and render it coarse.

Many institutions are established to aggregate information coarsely from a group of people working on a common objective when precise communication within the group is impossible or costly due to restrictions on time, memory, or other resources. For example, a voting system is established to aggregate information coarsely from the public via votes, because gathering each individual's precise opinion about the candidates is costly. Similar institutions include surveys, polls and ratings. As another example, a firm's hierarchical briefing system is established to aggregate information coarsely from different divisions of the firm via briefs, because collecting and processing detailed reports in a centralized fashion is costly. As a third example, when it is difficult for a person to meticulously remember everything he has in mind, he often uses the "mnemonic institution" of taking short notes or forming crude impressions to remind his future self of what he knows in the present.

The infeasibility of precise communication within institutions gives rise to interesting design problems that would otherwise be trivial. Indeed, given common interest induced by the common objective, first-best information aggregation can be easily achieved in equilibrium by an institution permitting precise communication. It is plain to see that, regardless of institutional details, every such institution is reducible to the direct mechanism that implements the efficient outcome. On the other hand, an institution that does not permit precise communication cannot be reduced to a direct mechanism, and consequently institutional details become important to determine its performance.

To elaborate on the design problem regarding institutions not permitting precise communication, this paper first analyzes the optimal binary voting system in a common values environment. In this problem, which is known as the Condorcet Jury Problem, a group of privately informed jurors have to arrive at a verdict to find a defendant guilty or innocent. The jury must use a binary voting system to solve the problem. A binary voting system requires that each juror is given one opportunity to cast a vote of either "guilty" or "innocent". Every binary voting system is characterized by two institutional components: the procedure and the rule. The procedure specifies the number of stages the voting has and who votes in each stage. The procedure may imply simultaneous voting, sequential voting, or a mixture of both. The rule specifies the verdict for each vote profile. The rule may correspond to a simple majority rule, a unanimous rule, or even some non-standard rule.

Three design problems are considered. In the first problem, given that there is no restriction on the number of stages of voting, among the optimal binary voting systems there is a system that has a sequential procedure and a rule under which the last voter serves as a dictator. In the second problem, given that voting has to finish in one stage, among the optimal one-stage systems there is a system that has a weighted majority rule. In the third problem, given that the procedure is fixed and has two stages, it turns out that the optimal system may have a counterintuitive rule: If a voter unilaterally changes his vote from "guilty" to "innocent", the verdict may change from "innocent" to "guilty".

The paper then proposes a framework for modeling a general class of information-aggregating institutions including typical voting systems, hierarchical organizations, and institutions of other sorts. An institution within this class is built on a communication infrastructure which provides each participant with (1) a set of messages for him to convey information to other participants, and (2) a set of units called perceptions for him to receive and process messages from other participants. Communication constraints on the sending side are captured by the fact that the set of messages available to a participant may be smaller than the set of all pieces of information he may wish to convey. Communication constraints on the receiving side are captured by the fact that a participant may be unable to distinguish between distinct message profiles if they correspond to the same perception.

A robust Pareto order is introduced to compare institutions within this class: Institution A is said to dominate institution B if, for *any* common objective, the best equilibrium under A generates a weakly higher common expected payoff than the best equilibrium under B. The purpose of focusing on the dominance order is three-fold. First, for a more specific design problem, the designer can use the dominance order to eliminate dominated institutions without having to know the environment parameters that determine the common objective. Second, understanding why an institution dominates another provides insight to understanding the advantages or disadvantages of specific institutional details. Third, from a theoretical perspective, the comparison of institutions in terms of dominance is parallel to the comparison of experiments analyzed in Blackwell (1951). This parallelism is discussed in more detail in Section 3.2.

The paper provides two characterizations of the dominance order. First, institution A dominates institution B if and only if A can induce weakly more social choice functions in pure strategies than B. Therefore, "better" and "more versatile" turn out to be equivalent regarding institutions. Second, if the set of all pieces of information each player may have is sufficiently rich, then institution A dominates institution B if and only if the underlying communication infrastructure of A can be obtained from that of B by a sequence of operations, each operation either removes redundant messages or perceptions from B, or introduces effective messages or perceptions to B. The sufficiency part is relatively straightforward, although identifying redundant messages and perceptions calls for care, in particular if the institution is complicated. Necessity is more difficult to establish. The argument is based on the observation that there is a social choice function inducible by institution B in pure strategies such that any other institution that can induce the same social choice function in pure strategies must embed the effective part of B, that is, the underlying communication infrastructure of B with redundant messages and perceptions removed. Thus it is possible to construct a sequence of redundancy-reducing operations that transforms B to its effective part, and then there is a sequence of complementing operations that transforms the effective part of B to A

The general analysis is applied to two specific problems. The first application investigates the design problem regarding generalized voting systems. It is shown that among the optimal generalized voting systems there is one that has a sequential procedure, a full disclosure policy, and a rule under which the last voter is always pivotal. The second application analyzes the marginal benefit of perceptions and messages within a hierarchical organization. It is shown that under mild conditions the marginal benefit of either perceptions or messages is always strictly positive. In particular, the result implies that even if an organization has very limited message-processing capacity, it will still strictly benefit from having more available messages.

In the rest of the paper, Section 3.2 discusses the relevant literature. Section 3.3 analyzes the Condorcet Jury Problem. Section 3.4 introduces the general model. Section 3.5 presents analysis of the dominance order. Section 3.6 applies the general analysis to two problems. Section 3.7 concludes. The Appendices include the proofs and some technical details.

3.2 Literature

Two literatures within the field of mechanism design and implementation have explicitly considered communication constraints. One of them investigates the minimal amount of communication necessary to implement a given social choice function. Within this literature, Nisan and Segal (2006) and Segal (2007) show that for a class of social choice functions the minimal communication mechanisms are generalized price mechanisms, assuming sincere players. Fadel and Segal (2009) and Segal (2010) consider similar implementation questions with strategic players. In spirit this literature may be seen as solving the dual problem to the problem studied in the present paper. The present paper asks for the efficiency maximizing mechanism subject to communication constraints, whereas papers in the above literature ask for the communication minimizing mechanism that achieves certain efficiency level.

The other literature considers mechanism design problems subject to given communication constraints. Recent papers in this literature include Blumrosen et al. (2007), Van Zandt (2007), Kos (2012), Blumrosen and Feldman (2013), Kos (2014), and Mookherjee and Tsumagari (2014). All of the papers assume the presence of conflicts of interest among the players, and a main emphasis of the literature is the interplay between communication constraints and incentive constraints. The present paper, by considering an environment with common interest, removes the tension between information aggregation and preference aggregation, and thus allows a sharper focus on implications of communication constraints alone.

The quest for the optimal voting system for the Condorcet Jury Problem may be viewed as a specific mechanism design exercise regarding institutions subject to communication constraints. The formulation of the Condorcet Jury Problem is historically attributed to Condorcet (1785). Papers discussing information aggregation efficiency of various voting systems include Feddersen and Pesendorfer (1998), McLennan (1998), and Duggan and Martinelli (2001), all of which emphasize asymptotic efficiency as the jury size grows large. For a fixed jury, Costinot and Kartik (2007) show that the optimal voting system is invariant to the possibility of boundedly rational voters. Dekel and Piccione (2000) show that in a symmetric environment, any equilibrium under simultaneous voting stays an equilibrium under sequential voting. They conclude with a note that sequentiality does not bring improvement in terms of information aggregation. The present paper shows that, if a voting system is evaluated by its best equilibrium, instead of by its worst one as is implicitly assumed in Dekel and Piccione (2000), then sequentiality may bring improvement.

As mentioned in the Introduction, the comparison of institutions studied in the present paper is analogous to the comparison of experiments studied in Blackwell (1951). Indeed, similar to an experiment *à la* Blackwell, an institution may be viewed as a device that generates signals (messages) to facilitate decision-making based on the true state (the dispersed information). On the other hand, unlike an experiment, an institution has a multi-player dynamic structure and the signals (messages) are generated endogenously by strategic players. In Blackwell (1951), an experiment dominates another if and only the former can be obtained from the latter by a well-defined transformation. This corresponds in spirit to the finding of the present paper that an institution dominates another if and only if the former can be obtained from the latter by a sequence of well-defined operations.

Chapter 8 of Marschak and Radner (1972) extends Blackwell's single-player model to multiple players. In their model, players move sequentially, and later players make decisions based on their own information and noisy realization of messages from earlier players. The model in the present paper has a similar sequential procedure. In two important aspects the two models are different. First, in the present model players are strategic, whereas they are non-strategic in the model of Marschak and Radner (1972). Second, imperfection of communication in the present model is due to coarseness, whereas it is due to noise in the model of Marschak and Radner (1972). The analysis of Marschak and Radner (1972) is restricted to specific examples and has a different focus.

3.3 The Condorcet Jury Problem

In this section we view the classical Condorcet Jury Problem from an institution design perspective.

A jury $J = \{1, ..., n\}$ has to reach a verdict on a defendant. The verdict is either "guilty" (G) or "innocent" (I). Let $\omega = G$ denote the event that the defendant is in fact guilty, and $\omega = I$ that the defendant is in fact innocent. $\omega = G$ with probability π where $0 < \pi < 1$, and $\omega = I$ with probability $1 - \pi$. Every juror receives a payoff of 1 if the verdict matches the fact ω, or 0 otherwise.

Each juror i has a private signal x_i that carries some information about ω. Specifically, x_i is independently drawn from a finite, but possibly very large, subset X_i of $\mathbb{R}$ with probability $f^i_\omega(x_i) > 0$ conditional on ω. [1] Assume that for any $i \in J$,

$$\text{(MLRP)} \quad \frac{f^i_G(x_i)}{f^i_I(x_i)} > \frac{f^i_G(x'_i)}{f^i_I(x'_i)} \text{ if } x_i > x'_i.$$

Assumption MLRP implies that a higher signal carries a stronger evidence for $\omega = G$.

Suppose that, due to a shortage of resources necessary for precise communication among the jurors to fully reveal their information, the jury has to use a **binary voting system** to reach the verdict. In a binary voting system, each juror can send a message once, and the set of messages available to him is $\{G, I\}$. The messages

[1] The finiteness assumption is to simplify analysis. See also Footnote 3.

may be interpreted as votes, and the action of sending a message may be interpreted as voting.

Each binary voting system is characterized by two institutional components. The first is the **procedure** that specifies the order in which the jurors vote. The procedure is formulated as a function $r : J \to \mathbb{N}$, with the interpretation that $r(i) > r(j)$ means i votes after j, and $r(i) = r(j)$ means i and j vote simultaneously. Assume that a juror can see all previously casted votes. [2] The second institutional component is the **rule** that specifies a verdict for each vote profile. The rule is formulated as a function $d : \{G, I\}^n \to \{G, I\}$ such that $d(z_1, ..., z_n)$ is the verdict given vote profile $(z_1, ..., z_n)$. A binary voting system with procedure r and rule d is denoted as (r, d).

Let $\Sigma_V(r)$ denote the set of all strategy profiles of the game induced by voting system (r, d). (Note that $\Sigma_V(r)$ does not depend on d.) Given the common payoff function and the common prior, all voters' *ex ante* preferences over $\Sigma_V(r)$ can be represented by the same expected payoff function $u(\cdot|r, d)$. The **value** of (r, d) is defined as the highest common expected payoff achieved by any perfect Bayesian equilibrium of the game induced by (r, d). Let $U(r, d)$ denote the value of (r, d). We can Pareto-rank voting systems by their values, and say that the optimal system is the one with the highest value.

Calculating the value of a voting system from definition can be computationally heavy, because it is often tedious to determine the set of equilibria of a game. Fortunately, as asserted by the following lemma, the common interest environment allows us to simplify the calculation by circumventing that step. The lemma, which follows immediately from the proof of Proposition 3.5.1 to appear later, generalizes Theorem 1 in McLennan (1998).

Lemma 3.3.1. $U(r, d) = \max_{\sigma \in \Sigma_V(r)} u(\sigma|r, d)$.

Voting takes place in multiple stages under a sequential or partially sequential procedure, and may require a long time to finish if there are many stages. It is thus reasonable to consider situations in which the jury can only use a voting system whose procedure has a limited number of stages. Below we study the optimal voting systems in three cases. In the first case, there is no restriction on the number of stages of voting. In the second case, voting has to take place in one stage. In the third case, a particular procedure with two stages is used.

3.3.1 Unlimited Number of Stages

When there is no restriction on the number of stages, the following proposition asserts that, to find an optimal voting system, it suffices to focus on ones that have a

[2] Here, to keep the example simple, we do not consider partial disclosure of previously casted votes. Partial disclosure policies are discussed in Section 3.6.

sequential procedure and a rule that depends only on the last voter's vote.

Proposition 3.3.2. *Among the optimal binary voting systems there is* (r^*, d^*) *such that:*

1. r^* *is a sequential procedure.*

2. d^* *depends only on the last voter's vote.*

The proof is based on the observation that any pure strategy profile under (r, d) can be "replicated" by an outcome-equivalent pure strategy profile under (r^*, d^*). It follows that the best strategy profile under (r^*, d^*) must be no worse than the best strategy profile under (r, d), which then implies that the value of (r^*, d^*) must be no less than that of (r, d) by Lemma 3.3.1.

Section B.1 in the Appendices gives a direct proof of this proposition. It will become clear that the proposition can also be derived as a corollary of Proposition 3.6.3, which is proved using the machinery to be introduced in Section 3.5. Moreover, the proof of Proposition 3.6.3 gives an alternative explanation for the superiority of (r^*, d^*): The system with a sequential procedure and a dictatorial rule makes full use of the communication instruments allowed by binary voting, whereas any other binary voting system does not.

3.3.2 One Stage

Suppose that voting must take place in a single stage, that is, the procedure must be simultaneous. Finding an optimal voting system in this case is equivalent to finding an optimal rule for simultaneous voting. The following proposition asserts that, for this quest, it suffices to focus on weighted majority rules. Rule d is a weighted majority rule if there is a vector of nonnegative weights $(w_1, ..., w_n)$ and a threshold k such that $d(z_1, ..., z_n) = G$ if and only if $\sum_{i \in J} w_i \mathbf{1}(z_i) \geq k$, where $\mathbf{1}(z_i) = 1$ if $z_i = G$ or $\mathbf{1}(z_i) = 0$ otherwise.

Proposition 3.3.3. *Among the optimal rules for simultaneous voting there is a weighted majority rule.*

The proof is built on the observation that, given any vote profile $(z_1, ..., z_n)$, an optimal rule d^* produces verdict G if and only if $\omega = G$ is more likely than $\omega = I$ conditional on players following the best strategy profile s^* under d^* and the realized vote profile being $(z_1, ..., z_n)$. In other words, if the rule is replaced with an uninformed decision maker whose interest is aligned with those of the players, the decision maker would also choose verdict G given $(z_1, ..., z_n)$ based on Bayesian updating. Proposition 3.3.3 then follows from the fact that the probability of $\omega = G$ conditional on s^* and $(z_1, ..., z_n)$ is log-linear in the number of G-votes contained in $(z_1, ..., z_n)$.

3.3.3 Two Stages: Non-Monotonicity

In this subsection we present an example showing that, if the procedure is partially sequential, the optimal voting systems may display counterintuitive properties.

There are three jurors. Suppose the procedure is fixed: Juror 1 votes in the first stage, Jurors 2 and 3 vote simultaneously in the second stage. Again we look for the optimal rule for this particular procedure. The environment parameters are as follows. $\pi = 1 - \pi = 0.5$. Tables 3.1 and 3.2 respectively show the conditional probabilities of x_1 and x_2. The conditional probabilities of x_3 are the same as x_2.

Table 3.1: $f^1_\omega(x_1)$

	$\omega = G$	$\omega = I$
$x_1 = 1$	0.6	0.4
$x_1 = 0$	0.4	0.6

Table 3.2: $f^2_\omega(x_2)$

	$\omega = G$	$\omega = I$
$x_1 = 1$	0.4	0.1
$x_1 = 1/2$	0.5	0.5
$x_1 = 0$	0.1	0.4

Using Lemma 3.3.1 we find multiple optimal rules. The unique best equilibrium under each optimal rule is in cutoff strategies. There is only one optimal rule d^* under which the best equilibrium is "truthful", in the sense that each juror votes G if and only if his signal is above the cutoff. Recall that Assumption MLRP implies higher signals are more indicative of $\omega = G$. Thus in a truthful equilibrium jurors always use a G-vote to express a stronger evidence supporting $\omega = G$, whereas in an equilibrium that is not truthful, a juror sometimes uses an I-vote to express a stronger evidence supporting $\omega = G$. It is reasonable to consider d^* as superior to the other optimal rules, because presumably the jurors can more easily coordinate on a truthful equilibrium.

However, d^* displays a counterintuitive property: It is not monotone in the vote profile. In particular, it is the case that

$$d^*(G, G, I) = I, \qquad d^*(I, G, I) = G,$$

that is, if Juror 1 unilaterally changes his vote from G to I, the verdict changes in the opposite direction from I to G. This property is unexpected because a G vote from Juror 1 carries stronger evidence supporting $\omega = G$ than an I vote.

To explain the phenomenon, notice that in the best equilibrium under d^*, the aggregate evidence contained in the (G, I) vote combination from Jurors 2 and 3 depends on the vote from Juror 1. If Juror 1's vote is G then Juror 2 votes G if $x_2 \in \{1/2, 1\}$, whereas if Juror 1's vote is I then Juror 2 votes G if $x_2 = 1$. Juror 3's strategy is the same as Juror 2's. It is easy to verify that the aggregate evidence contained in the

(G, I) vote combination from Jurors 2 and 3 is against, and overrules, the evidence contained in the vote from Juror 1.

3.4 The Model

In this section we introduce a framework for modeling a general class of institutions that is broad enough to capture many specific real life institutions, including binary voting systems analyzed in the previous section.

3.4.1 The Common Objective

A group $\mathscr{N} = \{1, ..., N\}$ of players work on a common objective. Each player $i \in \mathscr{N}$ contributes by choosing an action $a_i \in A_i$ where A_i is the set of actions available to i. A_i can be a singleton. The vector of actions $\mathbf{a} = (a_1, ..., a_N)$ is called an **outcome**. Let $A = A_1 \times ... \times A_N$ denote the set of all outcomes.

The value of an outcome depends on the **state**, which is an $N-$tuple $\mathbf{x} = (x_1, ..., x_N)$. Every player receives the same payoff of $\phi(\mathbf{a}, \mathbf{x})$ if the outcome is $\mathbf{a}$ and the state is $\mathbf{x}$. $F(\mathbf{x})$ denotes the *ex ante* probability that the state is $\mathbf{x}$.

Each player i privately observes the ith component x_i of the state. x_i may take any value from a set X_i. Let $X = X_1 \times ... \times X_N$ denote the set of all possible states. Assume $|A| < \infty$ and $|X| < \infty$. [3]

Assume that the set of states X and the set of outcomes A are the same for every common objective the group may face. A common objective is denoted by the objective-specific parameters (ϕ, F).

3.4.2 The Institution

Within the model, an **institution** refers to the indirect mechanism described as follows. Players move sequentially according to their indices: Player 1 moves first, Player 2 moves second, etc. Every player moves only once. When it is player i's turn to move, he chooses action a_i from A_i, and he also chooses a message m_i from some set M_i to send to players who have not moved yet. M_i is the set of messages provided to player i by the institution. Player N does not send any message. For Player N to have a non-trivial role in the institution, we assume that $|A_N| > 1$.

[3] This assumption is made for simplicity, with which we may avoid some minor technical complications. The analysis should not change significantly if $|A|$ and $|X|$ are infinite.

Players do not observe actions chosen by the other players. A player imperfectly observes the messages he has received. Let $T_i = M_1 \times ... \times M_{i-1}$ denote the set of message profiles that player i may receive. i's observation of received messages is determined by a partition P_i of T_i. Each element p of P_i is called a **perception** of i. i can distinguish between two message profiles if and only if they are in different perceptions of his. Therefore it can be said that player i perceives all message profiles in the same perception $p \in P_i$ as if they are the same.

Define $T = \cup_{i \in \mathscr{N}} T_i$, $P = \cup_{i \in \mathscr{N}} P_i$, and $M = \cup_{i<N} M_i$. The institution is denoted by the tuple (T, P, M).

Below we show various real life institutions that the general model captures.

Example 1. Voting

Consider a voting system that generalizes the binary voting system analyzed in Section 3.3. A group J of voters have to collectively choose from a set Y of candidates. Each voter $i \in J$ receives a private signal x_i. The value of each candidate is the same to all voters and depends on the vector of private signals $(x_i)_{i \in J}$.

Each voter has to cast one vote from the set of votes Z. [4] To avoid trivial cases assume $|Y| > 1$ and $|Z| > 1$. The voting system consists of three institutional components: the procedure $r : J \to \mathbb{N}$ that assigns each voter i to the $r(i)$th stage of the voting, the rule $d : Z^{|J|} \to Y$ that elects candidate $d(\mathbf{z})$ given vote profile $\mathbf{z}$, and a disclosure policy t which specifies how past votes are disclosed to those who have not voted yet. Examples of disclosure policies include full disclosure, disclosing the votes but not the voters' indices (anonymous voting), etc.

To capture voting system (r, d, t) using the general institution model introduced above, index the voters as $1, ..., |J|$ such that $i > j$ if $r(i) > r(j)$. Let $\mathscr{N} = \{1, ..., |J|, |J| + 1\}$ where player $|J| + 1$ represents the rule d. The action set A_i is the singleton $\{\overline{a}\}$ for every $i \leq |J|$, whereas $A_{|J|+1} = Y$. Note that only player $|J| + 1$ has a non-singleton action set, because he represents d and thus him alone determines the real outcome, that is, the chosen candidate from Y.

For each player $i \leq |J|$, X_i is the set of private signals that voter i may observe. $X_{|J|+1} = \{\overline{x}\}$, implying that the voter representing the rule always receives the uninformative signal $\overline{x}$.

For any outcome $\mathbf{a} = (\overline{a}, ..., \overline{a}, y)$ and vector of private signals $\mathbf{x} = (x_1, ..., x_{|J|}, \overline{x})$, $\phi(\mathbf{a}, \mathbf{x})$ is the expected value of candidate y conditional on each voter i observing private signal x_i.

Voting system (r, d, t) is represented by institution (T, P, M) as follows:

[4] Typically $Z = Y$. The present formulation allows other voting protocols, for example the inclusion of blank votes or abstention.

1. For each $i \leq |J|$, M_i is equal to Z, that is, a vote is interpreted as a message provided by the institution.

2. For each $i \leq |J|$, P_i is consistent with voter i's observation of votes from voters $1, ..., i-1$ under procedure r and disclosure policy t. For example, if two vote profiles $(z_1, ..., z_{i-1})$ and $(z'_1, ..., z'_{i-1})$ differ only at votes from those who vote simultaneously with i, that is, if $z_j = z'_j$ for every j where $r(j) < r(i)$, then the two vote profiles are in the same perception of i. As another example, if according to the disclosure policy voter i may only observe the number of votes already casted for one particular candidate y, then $(z_1, ..., z_{i-1})$ and $(z'_1, ..., z'_{i-1})$ are in the same perception of i if the number of votes for y from those who voted before i are the same in both vote profiles.

3. For player $|J| + 1$, vote profiles $(z_1, ..., z_{|J|})$ and $(z'_1, ..., z'_{|J|})$ are in the same perception if and only if $d(z_1, ..., z_{|J|}) = d(z'_1, ..., z'_{|j|})$.

It is plain to see that the game induced by voting system (r, d, t) is essentially the same game as that induced by institution (T, P, M) in which player $|J| + 1$ is committed to choose $d(z_1, ..., z_{|J|})$ given his perception that contains vote profile $(z_1, ..., z_{|J|})$. [5]

Example 2. Reporting

If we replace the rule in the voting model with an uninformative decision maker, the consequent model captures the situation in which a group of consultants advise an uninformed boss on choosing a project from the set of alternatives Y. The voters are reinterpreted as consultants, Player $|J| + 1$ as the boss, the votes as internal reports, the procedure and the disclosure policy as a protocol that organizes reporting.

Example 3. Organization

An organization has N levels. The official of level i gathers intelligence $x_i \in X_i$, and has to take immediate action $a_i \in A_i$ in response. Moreover, he also sends a message $m_i \in M_i$ to inform the officials of levels above him. The organization is modeled by institution (T, P, M). M_i is interpreted as internal codes available to the official of level i. P_i can either be used to describe how the official of level i interprets the codes he has received, or the code-processing protocol of the organization.

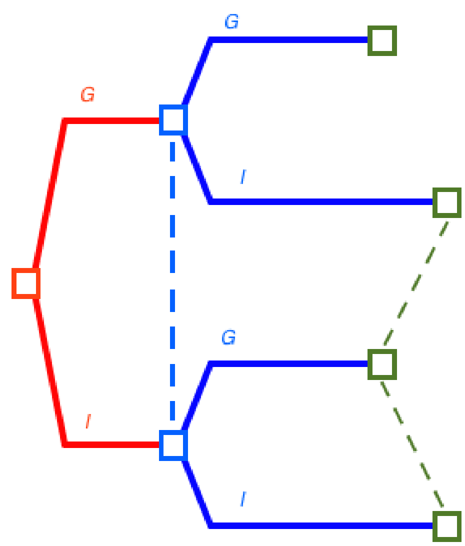
Example 4. Memory

Institution (T, P, M) can also model the dynamic optimization problem of a single person with imperfect recall. There is only one man, who on each day i learns something $x_i \in X_i$, and does something $a_i \in A_i$. Moreover, to remind himself of what he has learnt or done, he takes down a note, or forms a crude impression, in the form of $m_i \in M_i$. P_i represents how the person digests past notes or recalls past

[5] Apart from the modeling artifact that, in the game induced by institution (T, P, M), the voters take the dummy action $\overline{a}$ along the way, and player $|J| + 1$ receives an uninformative signal $\overline{x}$.

impressions. For example, if distinct $(m_1, ..., m_{i-1})$ and $(m'_1, ..., m'_{i-1})$ belong to the same $p \in P_i$, it may reflect the fact that on the ith day impressions $(m_1, ..., m_{i-1})$ and $(m'_1, ..., m'_{i-1})$ strike the person as the same, or the fact that the person does not read his notes very carefully. The decision maker with limited memory studied in Wilson (2014), for example, can be reformulated, with small modifications, as an infinite-horizon extension of the present model in which the decision maker can only recall the note he took down in the previous period. □

Institution (T, P, M) can be schematically depicted as a rooted tree equipped with a partition. Each node of the tree represents a message profile, and each edge represents a message. The root of the tree is the empty message profile. Two nodes are linked if one is extended from the other by one message, and the message is the edge that links the two nodes. The partitioning of the nodes at level i, which are $i-1$ degrees from the root, agrees with the partitioning of T_i by P_i. Figure **??** shows the graph representing the voting system with two voters, the simultaneous procedure, and the rule that chooses verdict G only if both votes are G.



The graph of an institution visually resembles that of an extensive form game in which histories correspond to message profiles and information sets correspond to perceptions. It should be noted that the graph of an institution does not fully depict the game induced by the institution, because for each player i the choice of action a_i and the acquisition of information x_i are not reflected in the graph.

3.5 Comparing Institutions

The common interest environment provides us with a natural measure to evaluate an institution.

Definition. The **value** of institution (T, P, M) for common objective (ϕ, F) is the highest common expected payoff achievable by any perfect Bayesian equilibrium of the game induced by (T, P, M) and (ϕ, F). [6]

For a given common objective, all institutions are totally Pareto-ordered by their values. Finding an optimal institution among a set of alternatives becomes a standard optimization problem. However, it is often the case that when an institution is to be established, the objective-specific parameters are unknown to the designer. Moreover, the same institution may be used repeatedly for variable common objectives. Due to these concerns, we would like to Pareto-order institutions without knowledge of the objective-specific parameters. It is thus natural to consider the parameter-free **dominance order** defined as follows.

Definition. Institution (T', P', M') **dominates** another institution (T, P, M) if the value of (T', P', M') is weakly higher than the value of (T, P, M) for *any* (ϕ, F) where ϕ is a real valued function on $A \times X$ and F is a probability function on X.

Clearly the dominance order is reflexive and transitive. It may not be complete, though.

Remark: When an institution is defined, it is assumed that players move sequentially according to their indices. Hence when comparing (T', P', M') and (T, P, M) we implicitly assume that players move in the same order under both institutions. A natural question that follows is whether it is possible to compare two mechanisms, both describable by the general model if players are indexed appropriately, but under which the players *de facto* move in different orders. This indeed is possible, because, after all, if the common objective is the same, then any two mechanisms can be compared by their values. Thus, the dominance order can naturally be extended to the set of all mechanisms without restriction on the order of moves. In this paper, as a first step of the research agenda, we focus only on institutions with the same order of moves.

A pure strategy profile s under institution (T, P, M) induces a **social choice function** $\boldsymbol{\alpha}(\cdot|s) : X \to A$, such that if the players follow s, then in state $\mathbf{x}$ the outcome is $\boldsymbol{\alpha}(\mathbf{x}|s)$. Let $C(T, P, M)$ denote the set of all social choice functions inducible by a pure strategy profile under (T, P, M). The following proposition gives a simple characterization of the dominance order.

Proposition 3.5.1. *Institution (T', P', M') dominates another institution (T, P, M) if and only if $C(T, P, M) \subset C(T', P', M')$.*

The proposition is a corollary of a more general result, Proposition B.4.2 in Section B.4, that asserts an analogous statement for any two finite mechanisms. Within the present context, the proof is based on the observation that among the strategy profiles that maximize the common expected payoff there is a pure strategy equilibrium.

[6] The value exists because the induced game is a finite.

Therefore (T', P', M') must have a (weakly) higher value than (T, P, M) for any common objective because it can induce more social choice functions in pure strategies. On the other hand, for any pure strategy profile s under (T, P, M) there is a common objective for which $\boldsymbol{\alpha}(\cdot|s)$ is efficient, therefore if (T', P', M') has a weakly higher value for this common objective, (T', P', M') must also be able to induce $\boldsymbol{\alpha}(\cdot|s)$ in pure strategies.

The reader is reminded of the main message of Blackwell (1951) that an experiment is more valuable if and only if it is more informative. The more informative experiment *à la* Blackwell allows more state-to-action mappings for the concerned single-player decision problem. Proposition 3.5.1 thus strikes a similar note. Indeed, because of common interest, the institution can be interpreted as a dynamic decision situation that a single player faces, as Example 4 in Section 3.4 suggests. It should be noted, however, that the information structure induced by an institution is endogenously generated via messages, whereas that induced by an experiment is exogenously generated via noisy signals.

In Blackwell (1951) the "more informative" order of experiments has a very simple structural characterization, which can be verified by examining the distribution functions representing the experiments. The rest of the section is dedicated towards a goal in the same spirit, and gives us a method to compare institutions by examining their structures.

To preview, the structural characterization of the dominance order takes the following form: Institution (T', P', M') dominates institution (T, P, M) if and only if (T', P', M') can be obtained from (T, P, M) by a sequence of operations, each being of one of the following five types: (1) expanding, (2) refining, (3) trimming, (4) relabeling, and (5) merging. Each type of operation involves adding or removing messages, perceptions, or both. Before elaborating on the operations, we first introduce a useful auxiliary concept, the improper institution, which is the "intermediate product" produced in the process of obtaining one institution from another using the operations.

Improper Institutions

An **improper institutions** has a similar message-perception backbone as a (proper) institution introduced in Section 3.4. The only difference is that the set of messages available to a player in an improper institution may depend on which message profile he has received. Note that the improper institution is a generalization of the (proper) institution. We extend the system of notation denoting components of a (proper) institution to denote the components of an improper institution: For each $i \in \mathscr{N}$, T_i is the set of message profiles player i may receive. P_i is a partition of T_i that represents i's perception. If $i < N$ then for each $h \in T_i$, $M(h)$ is the set of messages available to i if the message profile he receives is h. Denote $T = \cup_{i \in \mathscr{N}} T_i$ and $P = \cup_{i \in \mathscr{N}} P_i$.

The tuple (T, P, M), where M now denotes the correspondence that determines the set of available messages for each received message profile, is now used to represent an improper institution. Clearly, if (T, P, M) is a proper institution then $M(h) = M_i$ for each $h \in T_i$.

For a generic message profile h, let h_j denote the jth component of h, let $h(j)$ denote the first j components of h, and let $|h|$ denote the length of h. Thus for any $h \in T_i$ we have $h = (h_1, ..., h_{i-1})$, $h(j) = (h_1, ..., h_j)$, and $|h| = i - 1$.

Given $h \in T_i$, message profile $g \in T$ is said to be a **descendant** of h, and h is said to be an **ancestor** of g, if g is an extension of h. Clearly, if g is a descendant of h then $h = g(|h|)$. Moreover, if g is a one-component extension of h, then h is said to be the **parent** of g, and g is said to be a **child** of h. Clearly, if h is the parent of g then $h = g(|g| - 1)$. If h is the parent of g and the last component of g is m, we sometimes denote g as $h \times m$.

If distinct g and g' are descendants of h such that (1) $|g| = |g'|$, and (2) $g_j = g'_j$ for any $j \neq |h| + 1$, then g and g' are said to be $h-$**cousins**. Note that two $h-$cousins only differ at the $|h|$th component. [7]

Let $P(h)$ denote the perception $p \in P$ that contains message profile h.

We require that an improper institution (T, P, M) must satisfy the following regularity conditions:

C1 T_i is nonempty for any $i \in \mathscr{N}$.

C2 If $h \in T$ then every ancestor of h is in T.

C3 If $h \in T$ then $h \times m \in T$ if and only if $m \in M(h)$.

C4 If $P(h) = P(h')$ then $M(h) = M(h')$.

C1 and C2 imply that the graph representing (T, P, M) is a rooted tree with N levels. C3 implies that the tree is generated by M. C4 implies that the sets of available messages given different message profiles in the same perception are the same.

Note that given C4, conditional on perception $p \in P_i$, player i cannot acquire additional information on which particular message profile in p is the one he has received by examining the set of currently available messages. Therefore it causes no ambiguity to define $M(p)$ as the set of messages available to player i conditional on p.

Figure 3.1 shows the graph of an improper institution.

Given C1-C4, an improper institution induces a well-defined dynamic game in which

[7] Two message profiles with the same parent are also cousins by this definition.

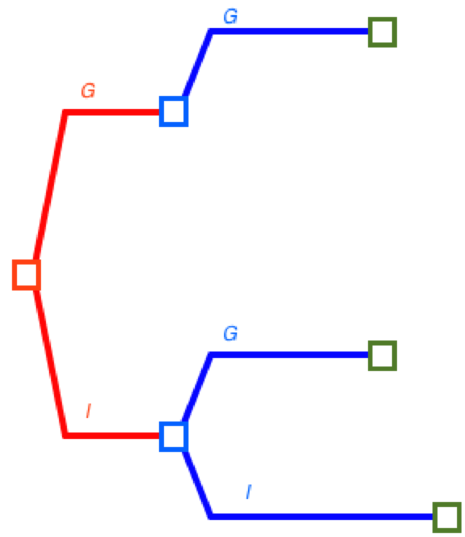


Figure 3.1

players sequentially take actions and send messages where the set of available messages might be perception-dependent. The concepts of value and dominance are naturally extended to improper institutions. Proposition 3.5.1 extends to improper institutions as well.

Proposition 3.5.2. *Improper institution* (T', P', M') *dominates another improper institution* (T, P, M) *if and only if* $C(T, P, M) \subset C(T', P', M')$.

For the rest of the section we will stop distinguishing between improper institutions and (proper) institutions, as all results apply to improper institutions and thus to proper institutions as special cases. The word "institution" will be used to refer to either.

Expanding

Expanding is the operation of creating a new institution (T', P', M') by adding more messages to an existing institution (T, P, M), while maintaining the perceptibility of the existing message profiles, that is, for any two message profiles h and g in T, they are in the same perception under P' if and only if they are in the same perception under P. The operation is defined formally as follows.

Definition. (T', P', M') is obtained from (T, P, M) by **expanding** if:

E1 $T \subset T'$.

E2 For any $h, g \in T$, $P(h) = P(g)$ if and only if $P'(h) = P'(g)$.

We say that (T, P, M) is a **sub-institution** of (T', P', M') if the latter is obtained from the former by expanding. For example, the institution depicted in Figure 3.3 is obtained from that depicted in Figure 3.2 by expanding, in particular, by making the additional message I available to Player 2.

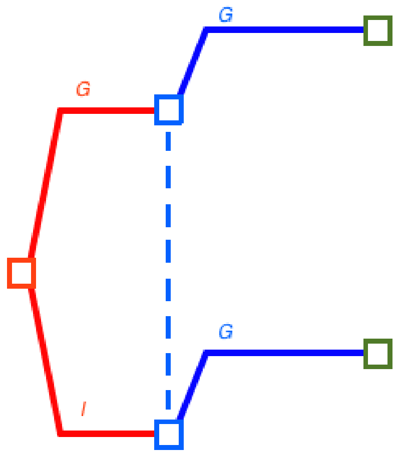


Figure 3.2

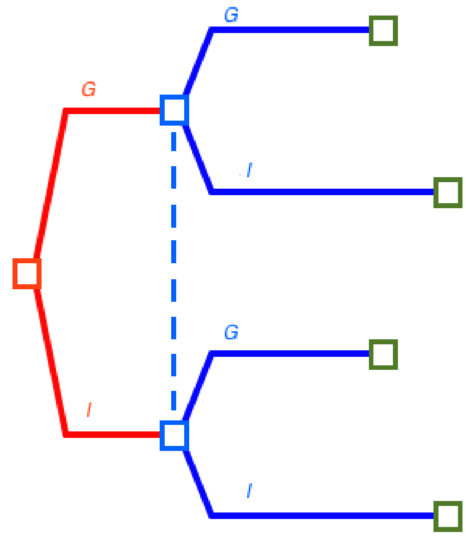


Figure 3.3

The following lemma states that expansion creates a better institution.

Lemma 3.5.3. *If* (T', P', M') *is obtained from* (T, P, M) *by expanding then* (T', P', M') *dominates* (T, P, M).

Refining

Refining is the operation that improves the players' observation of received messages. Refining is formally defined as follows.

Definition. (T', P', M') is obtained from (T, P, M) by **refining** if $T' = T$, $M' = M$, and P_i' is a weak refinement of P_i for each $i \in \mathcal{N}$.

Thus under (T', P', M') a player perceives received messages (weakly) more accurately than he does under (T, P, M). For example, the institution depicted in Figure 3.5 is obtained from that depicted in Figure 3.4 by refining, in particular Player 2's observation of received messages is strictly improved.

It is plain to see that any strategy profile of (T, P, M) can be "replicated" in (T', P', M') to produce the same outcome. Therefore (T', P', M') dominates (T, P, M) by Proposition 3.5.2. This observation is formally asserted in the following lemma. The proof is omitted.

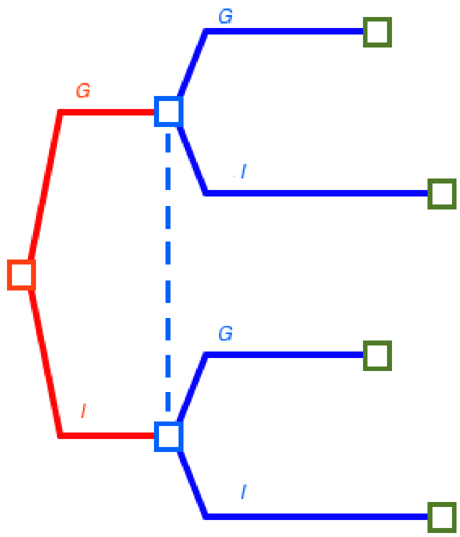


Figure 3.4

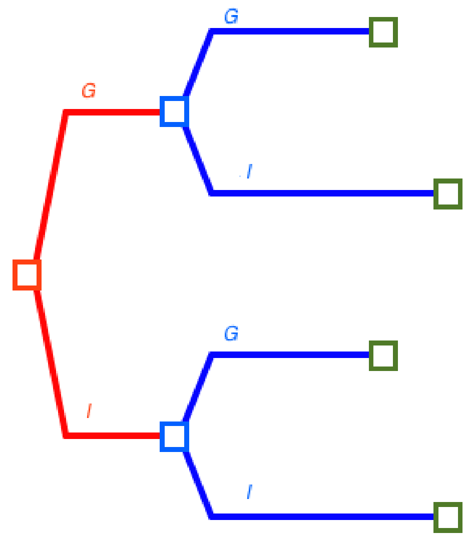


Figure 3.5

Lemma 3.5.4. *If* (T', P', M') *is obtained from* (T, P, M) *by refining then* (T', P', M') *dominates* (T, P, M).

Trimming

In real life, two words have the same communicative function if they are synonymous. Therefore one of the synonymous words may be viewed as functionally redundant. Excluding the redundant word from the vocabulary does not compromise communication. Trimming is the analogous operation of excluding a redundant message from an institution.

Before elaborating on trimming, it is helpful to first understand synonymity and redundancy within an institution. Whether two words in real life are synonymous or not often depends on the context. In one context they are synonymous; in another they have different meanings. Within an institution (T, P, M), whether two messages available to player i are synonymous or not also depends on the context, and the context is the message profile $h \in T_i$ that player i has received. Given $h \in T_i$, messages $m_i \in M(h)$ and $m'_i \in M(h)$ are considered to be synonymous if:

- Player $i + 1$ cannot distinguish between $h \times m_i$ and $h \times m'_i$.
- Regardless of what message m_{i+1} that player $i + 1$ sends, player $i + 2$ cannot distinguish between $h \times m_i \times m_{i+1}$ and $h \times m'_i \times m_{i+1}$.
- Regardless of what message m_{i+2} that player $i + 2$ sends, player $i + 3$ cannot distinguish between $h \times m_i \times m_{i+1} \times m_{i+2}$ and $h \times m'_i \times m_{i+1} \times m_{i+2}$.

- And so on for every player who moves after.

Therefore, keeping the messages from everyone else fixed, if player i unilaterally deviates from sending m_i to sending m_i', no other player would perceive the difference. The formal definition is given as follows.

Definition. Fix institution (T, P, M). For any $h \in T_i$ where $i < N$, messages $m_i \in M(h)$ and $m_i' \in M(h)$ are **synonymous given** h **within** (T, P, M) if $P(g) = P(g')$ for any $h-$cousins g and g' where $g, g' \in T$, $g_i = m_i$ and $g_i' = m_i'$.

Within the institution depicted in Figure 3.6, Player 2's messages G and I are synonymous given message profile (G), because Player 3 cannot distinguish between (G, G) and (G, I). Similarly G and I are synonymous given message profile (L) as well.

If m_i and m_i' are synonymous given every $h \in p$ within (T, P, M) for some $p \in P_i$ then we say m_i and m_i' are **synonymous given** p **within** (T, P, M). Message m_i may be considered as redundant given p. Trimming is the operation that excludes the redundant message m_i from the set of available messages given p. The formal definition is as follows.

Definition. (T', P', M') is obtained from (T, P, M) by **trimming** if there exist $i < N$, $p \in P_i$ and $m_i \in M(p)$ such that:

T1 (T', P', M') is a sub-institution of (T, P, M). Moreover $h \in T \backslash T'$ implies $h(i-1) \in p$ and $h_i = m_i$.

T2 m_i is synonymous to some $m_i' \in M(p)$ given p within (T, P, M).

By T1, (T', P', M') is the institution corresponding to player i not provided with message m_i given perception p. T2 emphasizes that m_i is indeed redundant given p.

The institution depicted in Figure 3.7 is obtained from that depicted in Figure 3.6 by trimming off message I, which is synonymous to G and is therefore redundant given Player 2's only perception.

The following lemma asserts that trimming does not change an institution functionally.

Lemma 3.5.5. *If (T', P', M') is obtained from (T, P, M) by trimming then (T', P', M') and (T, P, M) dominate each other.*

The proof is based on the observation that given perception p, player i cannot use synonymous messages m_i and m_i' to effectively communicate different pieces of information, because if i unilaterally changes from sending m_i to sending m_i' other players cannot perceive the change and hence will not react differently. Therefore, excluding the redundant message m_i from $M(p)$ does not compromise communication.

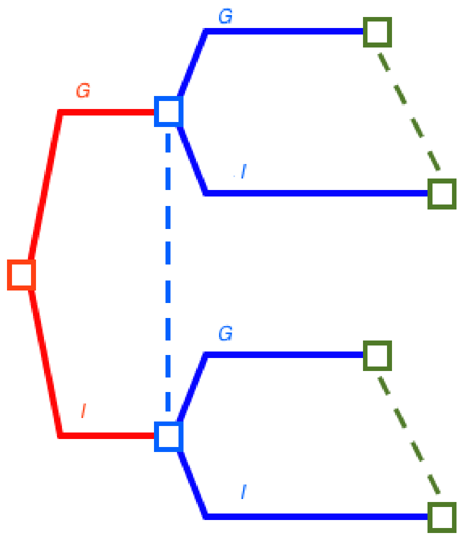


Figure 3.6

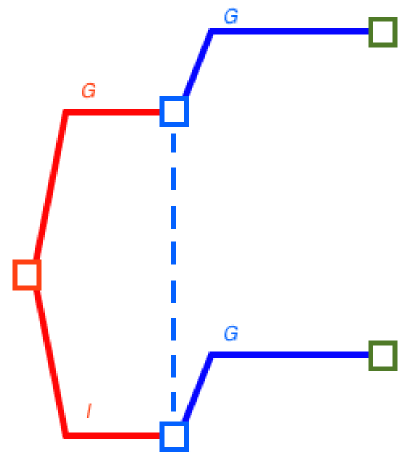


Figure 3.7

Relabeling

Relabeling is the operation of changing the message labels of an institution without changing its essential structure. Intuitively, as along as distinct messages have distinct labels, what those labels are should not matter to the use of the messages. In the jury voting case, for example, changing the vote labels from *Guilty* and *Innocent* to G and I only changes the institution superficially.

For an institution, the operation of relabeling changes message labels on a perception-by-perception basis. The operation can be thought of as the following process: First relabel the messages available to Player 1 given his only perception. Then relabel the messages available to Player 2 given each of his perceptions. The process continues until Player $N-1$'s messages are relabeled. The following is a formal definition.

Definition. (T', P', M') is obtained from (T, P, M) by **relabeling** if there is a **relabeling function** $\gamma : T \to T'$ such that:

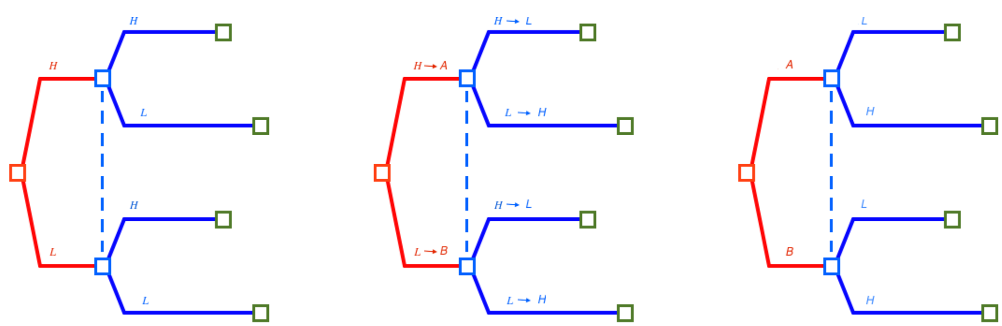
- R1 γ is a bijection.
- R2 γ preserves parent-child relation.
- R3 $P(h) = P(g)$ if and only if $P'(\gamma(h)) = P'(\gamma(g))$.
- R4 For any $i < N$, $p \in P_i$ and $m_i \in M(p)$ there is a message $\kappa(m_i, p)$ such that $\gamma(h \times m_i) = \gamma(h) \times \kappa(m_i, p)$ for any $h \in p$.

The definition is given in terms of the final product instead of the construction. To link the definition to the construction, note that R1, R2 and R3 imply that the graphs of (T, P, M) and (T', P', M') are isomorphic if the edges are label-less. $h \in T$ and $\gamma(h) \in T'$ are "essentially the same" message profile except that the labels of the messages they contain are different. R2 implies that if after relabeling the message profile $(m_1, ..., m_i)$ becomes $(m'_1, ..., m'_i)$, then for any message profile h whose first i components are $(m_1, ..., m_i)$, the first i components of the relabeled counterpart $\gamma(h)$ are $(m'_1, ..., m'_i)$. R4 is related to the perception-by-perception basis on which relabeling is conducted: Every edge with label m_i issued from perception p is given the same new label $\kappa(m_i, p)$.

Intuitively, relabeling should be invertible, that is, we should be able to retrieve (T, P, M) from (T', P', M') by "labelling back", where "labelling back" itself an operation of relabeling. The following lemma confirms this intuition.

Lemma 3.5.6. *If (T', P', M') is obtained from (T, P, M) by relabeling with relabeling function γ then (T, P, M) is obtained from (T', P', M') by relabeling with relabeling function γ^{-1}.*

Figure 3.8 shows an example of relabeling. The institution depicted in Panel (c) is obtained from that depicted in Panel (a) by relabeling, where Panel (b) shows the implied κ described in R4. Note that both edges with the label H issuing from the only perception of Player 2 are relabeled to L, and both edges with label L issuing from the same perception are relabeled to H, as R4 requires that edges with the same label issuing from the same perception are relabeled identically.



(a) (b) (c)

Figure 3.8

The following lemma asserts that relabeling does not change an institution functionally.

Lemma 3.5.7. *If (T', P', M') is obtained from (T, P, M) by relabeling then (T', P', M') and (T, P, M) dominate each other.*

Merging

Given institution (T, P, M), the purpose of merging is to produce a new institution in which two perceptions $p \in P_i$ and $q \in P_i$ of some player $i < N$ where $|A_i| = 1$ are combined into one perception $p \cup q$, where the new institution is functionally equivalent to the old institution.

It is useful to first motivate the idea of merging. Suppose a speaker wishes to convey his private information x to a listener using one of the two messages L and R. There are two possible contexts, A and B, which are relevant to the listener's decision problem. If the speaker knows the context, he can use a communication strategy that depends on the context and x. Suppose instead the speaker cannot observe the context but has four available messages $\{m_1, m_2, m_3, m_4\}$. In this situation, by using the following communication strategy the speaker can convey x as precisely as if he knew the context but only had two available messages:

1. Send m_1 if given x he would send L in context A and L in context B.
2. Send m_2 if given x he would send R in context A and R in context B.
3. Send m_3 if given x he would send L in context A and R in context B.
4. Send m_4 if given x he would send R in context A and L in context B.

Merging is based on the same idea. On one hand we make the observation of player i less accurate by combining his perceptions p and q into one perception $p \cup q$. On the other hand we compensate the possible loss of communicative capacity by providing player i with more messages given the combined perception $p \cup q$.

In practice, merging involves the following four steps:

Step 1 If $|M(p)| < |M(q)|$ then the institution is expanded by making more messages available to player i given p, so that the new sets of available messages $\hat{M}(p)$ and $\hat{M}(q)$ are equal in size. Each of the new messages is set to be synonymous to some existing message given p.

Step 2 Relabel the messages in $\hat{M}(p)$ and $\hat{M}(q)$ so that the new sets of messages $\tilde{M}(p)$ and $\tilde{M}(q)$ are the same, not only in terms of size but also in terms of labels.

Step 3 For each pair of distinct messages $(m_i, m_i') \in \tilde{M}(p) \times \tilde{M}(p)$, expand the institution by making a new message $n(m_i, m_i')$ available to player i given both p and q. $n(m_i, m_i')$ is set to be synonymous to m_i given p, and to m_i' given q.

Step 4 Combine p and q into one perception $p \cup q$.

Steps 1 and 2 make the technical preparation so that after combining p and q the resulting structure is an institution. Step 3 introduces redundant messages. However, some of the redundant messages will no longer be redundant after the merge of p and q. Given the less accurate perception $p \cup q$, player i can use message $n(m_i, m_i')$ to convey the private information that he would use m_i to convey if he knew the perception was p and would use m_i' to convey if he knew the perception was q.

The formal definition of merging is complicated, and is thus relegated to Section B.9 in the Appendices.

Figure 3.9 shows merging in those steps. Panel (a) depicts the original institution. The two perceptions $\{H\}$ and $\{L\}$ of Player 2 are to be merged. Panel (b) depicts the end product of Step 1, that is, an additional message A, set to be synonymous to B given perception $\{L\}$, is introduced so that the number of messages available to Player 2 given both perceptions are the same. Panel (c) depicts the end product of Step 2, that is, after relabeling, messages available to Player 2 given both perceptions have the same labels. Panel (d) depicts the end product of Step 3, that is, introducing additional redundant messages to both perceptions, where, for example, $n(H, L)$ is set to be synonymous to H given perception $\{H\}$, and to L given perception $\{L\}$. Panel (d) depicts the end product of Step 4, that is, combining the two perceptions.

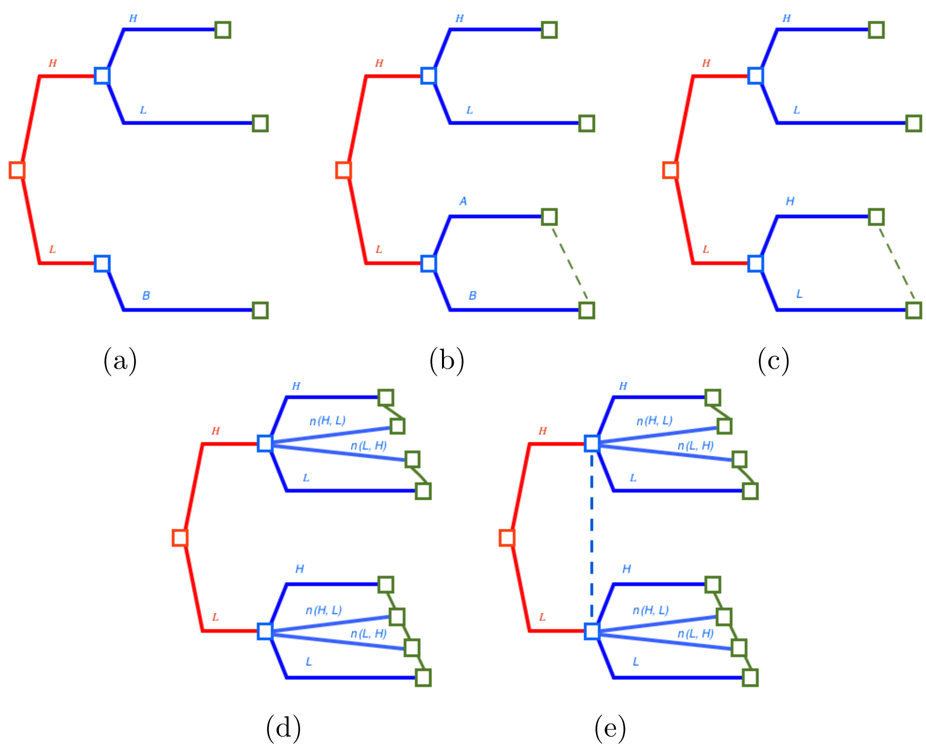


Figure 3.9

Steps 1, 2 and 3 involve either relabeling or introducing redundant messages (the inverse operation of trimming), and therefore by Lemmas 3.5.7 and 3.5.5 do not change the institution functionally. Following the intuition of the earlier example, going from Step 3 to Step 4 should not change the institution functionally either. The following lemma confirms the intuition.

Lemma 3.5.8. *If (T', P', M') is obtained from (T, P, M) by merging then (T', P', M') and (T, P, M) dominate each other.*

Now we are ready to state the structural characterization of dominance.

Theorem 3.5.9. *Given two institutions (T, P, M) and (T', P', M'):*

1. *(T', P', M') dominates (T, P, M) if (T', P', M') is obtained from (T, P, M) by a sequence of operations of expanding, refining, trimming, relabeling or merging.*

2. *For any $i \in \mathscr{N}$ there is $t_i \in \mathbb{N}$ such that if $|X_i| \geq t_i$ then (T', P', M') dominates (T, P, M) only if (T', P', M') can be obtained from (T, P, M) by a sequence of operations of expanding, refining, trimming, relabeling or merging.*

Part 1, the "if" direction, of the theorem is an immediate consequence of Lemmas 3.5.3, 3.5.4, 3.5.5, 3.5.7 and 3.5.8.

The proof of Part 2, the "only if" direction, can be broken into the following steps. First, it is clear that after applying merging operations to (T, P, M) for finitely many times we can obtain some $(\hat{T}, \hat{P}, \hat{M})$ such that $|\hat{T}_i| = 1$ for any $i < N$ where $|A_i| = 1$. If $|X_i|$ is sufficiently large for each i, there is a strategy profile $\hat{s} \in S(\hat{T}, \hat{P}, \hat{M})$ such that every non-redundant message in $(\hat{T}, \hat{P}, \hat{M})$ is utilized. Moreover, any institution (T', P', M') that can induce $\boldsymbol{\alpha}(\cdot|\hat{s})$ in pure strategies must embed the relabeled version of $(\hat{T}, \hat{P}, \hat{M})$'s non-redundant backbone (T^*, P^*, M^*) as its sub-institution. (T^*, P^*, M^*) is shown to be obtained from $(\hat{T}, \hat{P}, \hat{M})$ by refining and trimming. Since the relabeled version of (T^*, P^*, M^*) is embedded in (T', P', M'), (T', P', M') can be obtained from (T^*, P^*, M^*) by relabeling and expanding.

If the state space is not rich enough, that is, if $|X_i| < t_i$ for some $i \in \mathscr{N}$, then Part 2 of Theorem 3.5.9 need not be true. Indeed, if $|X_i|$ is sufficiently small for every $i \in \mathscr{N}$ then (T, P, M) and (T', P', M') may both accommodate *precise* communication, despite that one may not be obtainable from the other by a sequence of the five types of operations. It is worth noting that the value of t_i depends on the dominated institution (T, P, M) only.

3.6 Applications

3.6.1 Voting Revisited

In Example 1 of Section 3.4 we have described how to model voting system (r, d, t) in terms of an institution. Because voting systems are finite mechanisms for a common objective, the definition of value extends to them, and the dominance order can also be extended to compare them. Let $C_V(r,d,t)$ denote the set of all mappings from $X_1 \times ... \times X_{|J|}$ to Y inducible in pure strategies under voting system (r,d,t). A result analogous to Proposition 3.5.1 holds for voting systems.

Lemma 3.6.1. *Voting system (r',d',t') dominates another voting system (r,d,t) if and only if $C_V(r,d,t) \subset C_V(r',d',t')$.*

Like Proposition 3.5.1, Lemma 3.6.1 is also a corollary of Proposition B.4.2, because Proposition B.4.2 applies to any pair of finite mechanisms.

Let (T,P,M) be the institution representing (r,d,t). Recall that player $|J|+1$ represents the rule and $P(\mathbf{z}) = P(\mathbf{z}')$ if and only if $d(\mathbf{z}) = d(\mathbf{z}')$ for any $\mathbf{z}, \mathbf{z}' \in T_{|J|+1}$, where $T_{|J|+1}$ is the set of all complete vote profiles. Let $S(T,P,M|d)$ denote the set of pure strategy profiles of (T,P,M) satisfying the following condition:

V1 For any $\mathbf{z} \in T_{|J|+1}$, player $|J|+1$ chooses the action $d(\mathbf{z})$ given perception $P(\mathbf{z})$.

V1 requires player $|J|+1$ to exactly follow the choice rule determined by d. It is straightforward to see that for any pure strategy profile s under voting system (r,d,t) there is a corresponding strategy profile $s' \in S(T,P,M|d)$ under institution (T,P,M) such that s and s' lead to the same choice of $y \in Y$ given any vector of private signals $(x_1,...,x_{|J|})$, and vice versa, because, as discussed in Example 1 in Section 3.4, the game induced by (r,d,t) and the game induced by (T,P,M) are essentially the same if player $|J|+1$ has to choose according to d.

Let $C(T,P,M|d)$ denote the set of all social choice functions inducible by any $s \in S(T,P,M|d)$. The following lemma results immediately from the observation in the previous paragraph. The proof is omitted.

Lemma 3.6.2. *Voting system (r',d',t') dominates another voting system (r,d,t) if and only if $C(T,P,M|d) \subset C(T',P',M'|d')$ where (T,P,M) and (T',P',M') are institutions respectively representing (r,d,t) and (r',d',t').*

It is possible to compare voting systems by analyzing the institutions representing them. However, it should be noted that one institution dominating another institution is usually not sufficient for the voting system that one represents to dominate the voting system that the other represents, due to the additional constraint imposed on the strategy of player $|J|+1$. Despite this caveat, the machinery developed in Section 3.5 still provides tools for us to conclude the following results.

Proposition 3.6.3. *For any voting system (r,d,t):*

1. *(r,d,t) is dominated by voting system (r,d,t') where t' is the full disclosure policy.*

2. *If t is the full disclosure policy, then (r,d,t) is dominated by voting system*

(r', d, t) where r' is a sequential procedure.

3. *If t is the full disclosure policy, then (r, d, t) is dominated by voting system (r, d', t) where d' is a rule under which the collective choice is not determined before voting in the last stage (according to r) has taken place.*

Parts 1 and 2 are based on the observation that allowing full disclosure or making the procedure sequential corresponds to refining the institution that represents the voting system. Moreover, since these modifications do not change player $|J| + 1$'s perception, player $|J| + 1$ can still choose according to the rule d.

Part 3 is based on the observation that, if the collective choice is determined before voting in the last stage has taken place, then votes in the last stage become redundant, because they cannot effectively carry private information from those who vote in the last stage to affect the collective choice. Changing the rule to one that allows consideration of those last stage votes renders them useful and thus leads to improvement.

An immediate implication of Proposition 3.6.3 is that, to find the optimal voting system for any collective choice problem, it is sufficient to focus on ones that have a sequential procedure, full disclosure policy, and a rule under which the last voter is always pivotal.

It should be noted, however, that the last voter being always pivotal does not mean that the rule depends entirely on his vote. For example, earlier votes may effectively determine the set of candidates that the last voter can choose from.

3.6.2 The Benefit of Complexity

An institution offers two kinds of instruments that facilitate communication: messages and perceptions. A more complex institution has more messages and perceptions. Lemmas 3.5.3 and 3.5.4 imply that complex institutions weakly outperform less complex ones. In this section we investigate whether the benefit of additional complexity is always strictly positive.

The complexity of institution (T, P, M) has two dimensions, one that concerns the messages and the other the perceptions. The message-complexity of (T, P, M) is measured by the vector $(|M_1|, ..., |M_{N-1}|)$. The perception-complexity of (T, P, M) is measured by the vector $(|P_1|, ..., |P_N|)$.

We ask two questions:

1. Whether increasing the perception-complexity of an institution while keeping the message-complexity fixed leads to a strictly better institution in terms of dominance.

2. Whether increasing the message-complexity of an institution while keeping the perception-complexity fixed leads to a strictly better institution in terms of dominance.

The first question can be thought of as concerning the situation in which messages are costly to provide, whereas perceptions are relatively cheap, so that it is worthwhile to increase the number of perceptions as long as it strictly improves the institution. The second question can be thought of as concerning the situation in which perceptions are costly but messages are cheap.

For the rest of the subsection we assume that $|X_i|$ is very large for each $i \in \mathscr{N}$, that is, it is arbitrarily close to ∞, so that the lower bound requirement on $|X_i|$ in Part 2 of Theorem 3.5.9 is satisfied for any institution we are going to consider. Moreover, assume that every player has a non-singleton action set. The assumptions are important to the results of the present subsection.

The answer to the first question is affirmative.

Proposition 3.6.4. *Any institution (T, P, M) where $|P_i| < |T_i|$ for some $i \in \mathscr{N}$ is strictly dominated by some institution (T', P', M') where*

1. $|M'_j| = |M_j|$ *for any* $j < N$.
2. $|P'_j| \geq |P_j|$ *for any* $j \in \mathscr{N}$.

The proof is based on the observation that if $|P'_i| > |P_i|$ for some i then (T, P', M) is not dominated by (T, P, M), because no operation of expanding, refining, trimming, relabeling or merging can decrease the number of perceptions of i (merging does not, because it only merges perceptions of players with singleton action sets). If player i's observation of received message profiles is not perfect under P_i (implied by $|P_i| < |T_i|$), then refining (T, P, M) by strictly refining P_i strictly improves the institution.

The answer to the second question is also affirmative if the institution is "mildly" complex.

Proposition 3.6.5. *Suppose $N \geq 3$. Any institution (T, P, M) where $|P_{i+1}| \geq 2$ for some $i \geq 2$ is strictly dominated by some institution (T', P', M') where*

1. $|P'_j| = |P_j|$ *for any* $j \in \mathscr{N}$.
2. $|M'_j| \geq |M_j|$ *for any* $j < N$.

The result may not seem surprising at first sight, but let us illustrate a concern which would suggest that additional messages might be of no additional value at all. Suppose there are only two players: the speaker and the listener. The listener has two perceptions. Clearly, if there are already two messages available to the speaker, any additional message is going to be redundant because it will be synonymous to one of the existing messages. This example, which shows that the decision maker (the listener) is not able to make use of more data (messages) because of the constraints on his data-processing capacity (the number of perceptions), reflects a prominent phenomenon, termed as data overload, in many real life situations.

Since an institution may face stringent message-processing constraints due to limited perceptions, it is natural to expect that data overload will eventually occur, in particular when the existing message-complexity is already very high. However, Proposition 3.6.5 implies that even if there is only one player that has multiple perceptions, data overload can still be avoided regardless of the message-complexity of the existing institution. That there are more than two players within the institution is crucial for this result. Indeed, if there are only two players then data overload will eventually occur, as in the speaker-listener example. However, if there are more than two players, it is possible to simultaneously enlarge the message sets for multiple players and carefully arrange how other players perceive message profiles containing these newly introduced messages, so that no additional redundancy is created by the modification.

3.7 Conclusion

This paper proposes a framework for modeling a general class of information-aggregating institutions, introduces a robust Pareto order on institutions thus modeled, and derives two characterizations of this order.

It is not difficult to extend the model to capture more complex institutions, for example, those in which the players engage in conversation-like interactive communication, those in which actions are observable to certain degree, or those in which the players take actions after the communication phase is over. In fact, any mechanism that tackles a common objective by information aggregation can be captured by a straightforward extension of the present model, because the essential part of the model is no more than a partial structure of the extensive form game induced by the corresponding mechanism. As a generalization of Proposition 3.5.1, Proposition B.4.2 in the Appendices provides a characterization of the dominance order on all finite mechanisms. It is natural to ask if Theorem 3.5.9 can also be generalized so that other mechanisms can be compared structurally in a similar way. As a first step in this direction, it is worthwhile investigating whether institutions that only differ in the order in which players move may be compared structurally.

To extend the analysis in a different direction, we can consider mildly relaxing the common interest assumption to the extent that institutions can still be Pareto-ordered in a non-trivial way. One possibility, for example, is that every player's payoff only depends on his own action. Along with some extension to the model, we may compare pure information sharing systems, for example social networks, in which there is no need to coordinate actions.

Acknowledgements

I am indebted to Tilman Börgers, Kfir Eliaz, Stephan Lauermann and David Miller for constant support and encouragement throughout this project. I am also grateful to Miguel Ballester, Pierpaolo Battigalli, Jimmy Chan, Mehmet Ekmekci, Shaowei Ke, Daniel Krähmer, Anqi Li, Bart Lipman, Heng Liu, Taejun Oh, Efe Ok, Scott Page, John Quah, Anne-Katrin Roesler, Ariel Rubinstein, Todd Sarver, Ran Shorrer, Philipp Strack, Alexander White, Chenyu Yang and Fudong Zhang for great conversations and feedback.

APPENDICES

APPENDIX A

Additional Proofs and Examples in Chapter 1

A.1 Proof of Proposition 1.5.7

Proposition 1.5.7 is proved with the assistance of Lemma A.1.1.

Lemma A.1.1. *For any game* (M, W, u, v, C, δ) *there exist* $\tau \in \mathbb{N}^+$, $\overline{q} \in (0,1)$ *and* $(m, w) \in M \times W$ *such that for any subgame perfect equilibrium the following is true: If a meeting between* (m, w) *ends in separation with positive probability, then the probability that no player marries during the first* τ *days of the subgame resulting from that meeting having ended in separation is less than* $\overline{q}$.

Proof: Let S_I denote the initial market. There exists a pair (m, w) such that $m = \alpha^{S_I}(w)$ because S_I is nontrivial. Let $\overline{w}$ and $\underline{w}$ respectively denote m's first and last choices in $A^{S_I}(m)$ and define $\overline{\pi} := u(m, \overline{w})$ and $\underline{\pi} := u(m, \underline{w})$. Thus $\overline{\pi} \geq u(m, w) \geq \underline{\pi} > 0$ and therefore we can pick $\tau \in \mathbb{N}^+$ and $\overline{q} \in (0,1)$ such that $\delta[(1-\overline{q})\overline{\pi}+\overline{q}\delta^\tau\overline{\pi}] < \underline{\pi}$.

Fix a subgame perfect equilibrium σ. Suppose (m, w) meet on a day. Lemma 1.5.1(d) implies w accepts m if she has been accepted since $m = \alpha^{S_I}(w)$. Thus if the meeting ends in separation with positive probability then it must be that m rejects w with positive probability. Let q denote the equilibrium probability that no player marries during the first τ days in the subgame resulting from m having rejected w. By Lemma 1.5.1(d), m's expected payoff from rejecting w is no more than $\delta[(1-q)\overline{\pi}+q\delta^\tau\overline{\pi}]$. If m rejects w with positive probability in equilibrium then $\delta[(1-q)\overline{\pi}+q\delta^\tau\overline{\pi}] \geq u(m, w) \geq \underline{\pi}$ because m's expected payoff from accepting w is $u(m, w)$ as he will be accepted. Thus $q < \overline{q}$ by the choice of $\overline{q}$. □

Proof of Proposition 1.5.7: Fix a game (M, W, u, v, C, δ) and denote the initial market as S_I. Prove by induction. Suppose $|M| = 1$. Let m be the only man and $w := \alpha^{S_I}(m)$. Thus (m, w) is a top pair for S_I. In any subgame perfect equilibrium the game ends with certainty after (m, w) meet because they marry each other upon first meeting by Lemma 1.5.5. The proposition then follows.

Suppose the proposition is true if $|M| < n$ for some n. Consider $|M| = n$. Fix a subgame perfect equilibrium σ. By Lemma A.1.1 there exist $\tau \in \mathbb{N}^+$, $\overline{q} \in (0, 1)$ and $(m, w) \in S_I$ such that if (m, w) meet and in equilibrium the meeting ends in separation with positive probability, then the probability that no one marries during the first τ days of the subgame resulting from that meeting having ended in separation is less than $\overline{q}$. For any $h \in \hat{H}$ such that $S(h) = S_I$ let $\beta(h)$ denote the equilibrium probability that no one marries in the first $\tau + 1$ days of $\Gamma(h)$; let $\phi(h)$ denote the equilibrium probability that (m, w) marry on the first day of $\Gamma(h)$ conditional on them meeting on that day. Then by Lemma A.1.1, for any $h \in \hat{H}$

$$\beta(h) \leq \epsilon\Big(\phi(h) \times 0 + (1 - \phi(h)) \times \overline{q}\Big) + (1 - \epsilon) \leq \epsilon\overline{q} + (1 - \epsilon). \tag{I1}$$

Let h_0 denote the initial (empty) history. For any $k \in \mathbb{N}^+$ let $\hat{H}_k$ denote the set of all histories h in $\hat{H}$ such that the first day of $\Gamma(h)$ is the $k(\tau + 1) + 1$st day of the initial game and $S(h) = S_I$. Thus $h \in \hat{H}_k$ being reached implies no player has married during the first $k(\tau + 1)$ days of the initial game. Let E_k denote the event that any $h \in \hat{H}_k$ is reached. Inequality I1 implies $\Pr_\sigma(E_1) \leq \epsilon\overline{q} + (1 - \epsilon)$ because $h_0 \in \hat{H}$. I1 also implies $\Pr_\sigma(E_k|E_{k-1}) \leq \epsilon\overline{q} + (1 - \epsilon)$ for any $k > 1$. By the inductive hypothesis the probability Q that the game does not end at all is equal to the probability that every player remains in the market into the infinite future. Hence for any $k > 1$ we have $Q \leq \Pr \sigma(E_k) = \Pr_\sigma(E_1) \prod_{j=2}^{k} \Pr_\sigma(E_j|E_{j-1})$. The present proposition follows as $Q \leq \lim_{k\to\infty}[\epsilon\overline{q} + (1 - \epsilon)]^k = 0$. □

A.2 Proof of Proposition 1.5.8

Introduce additional terminology for the proof. Let $\Sigma(\delta)$ denote the set of all subgame perfect equilibria of the game with discount factor δ in the environment $E := (M, W, u, v, C)$. A (possibly empty) set P of disjoint pairs of a man and a woman from the initial market $S_I := (M, W, u, v)$ is a **settled set** for E if there exists some $d < 1$ such for any $\sigma \in \Sigma(\delta)$ where $\delta > d$, σ implies any $(m, w) \in P$ marry almost surely and upon first meeting in $\Gamma(h)$ if $h \in \hat{H}$, $(m, w) \in S(h)$, and moreover $S(h)$ can be derived from S_I as a result of a (possibly empty) sequence of pairs having left, among which each pair (m', w') is either in P or is a woman-acceptable pair from $S_I \backslash P$, where $S_I \backslash P$ denotes the submarket that complements P in S_I. d is called a **settling discount factor** for P. Note that any $d' > d$ is also a settling discount factor for P.

Proposition 1.5.8 is proved with the assistance of Lemma A.2.1.

Lemma A.2.1. *For any settled set P for the environment E, if m is a top player for*

$S_I \backslash P$ *and* $w := \alpha^{S_I \backslash P}(m)$, *or* w *is a top player for* $S_I \backslash P$ *and* $m := \alpha^{S_I \backslash P}(w)$, *then* $P \cup \{(m, w)\}$ *is also a settled set for* E.

Proof: Prove by induction on $|M|$. Suppose $|M| = 1$. Let m be the only man and $w := \alpha^{S_I}(m)$. For δ sufficiently close to 1, it is easy to see that in any subgame perfect equilibrium, (m, w) marry almost surely and upon first meeting. Thus there are two settled sets, the empty set P_1 and $P_2 := \{(m, w)\}$. The lemma is true for P_1 because m is a top player in $S_I = S_I \backslash P_1$ and as is concluded $\{(m, w)\}$ is a settled set. The lemma is true for P_2 vacuously.

Suppose the lemma is true if $|M| < n$ for some n. Consider $|M| = n$. Pick a settled set P for E and let $d < 1$ be a settling discount factor for P. Suppose $\overline{w}$ is a top player for $\overline{S} := S_I \backslash P$. Let $\overline{m} := \alpha^{\overline{S}}(\overline{w})$ and denote $\underline{m}$ as $\overline{w}$'s second choice in $A^{\overline{S}}(\overline{w}) \cup \{s\}$. By definition of the top player we have $\overline{w} = \alpha^{\overline{S}}(\overline{m})$. Notice that $\underline{m}$ may be s. Define $\hat{H}_I := \{h \in \hat{H} : S(h) = S_I\}$. Thus $\Gamma(h)$ is isomorphic to the initial game for $h \in \hat{H}_I$. Consider $\sigma \in \Sigma(\delta)$ for $\delta > d$. It is helpful to bear in mind that the choice of σ depends on δ. In σ, by assumption $\overline{w}$ may only marry a man from $\overline{S}$ in the subgame $\Gamma(h)$ where $h \in \hat{H}_I$, thus her expected payoff in $\Gamma(h)$ is no higher than $v(\overline{m}, \overline{w})$ by an argument similar to the proof for Lemma 1.5.1(d). It follows that $\overline{w}$ always accepts $\overline{m}$ when the remaining market is S_I. Given that, $\overline{m}$ always accepts $\overline{w}$ when the remaining market is S_I by a similar argument. Thus in σ, (m, w) marry upon first meeting in $\Gamma(h)$ for any $h \in \hat{H}_I$.

Suppose $\overline{w}$ deviates to the strategy under which she only accepts $\overline{m}$ when the remaining market is S_I, and switches back to following σ if the remaining market is no longer S_I. Let σ' denote the strategy profile due to $\overline{w}$'s unilateral deviation. For each $h \in \hat{H}_I$ let $K(h, \delta, \sigma)$ denote $\overline{w}$'s expected payoff in $\Gamma(h)$ under σ'. Define $\underline{K}(\delta, \sigma) := \inf_{h \in \hat{H}_I} K(h, \delta, \sigma)$. Let $p(h, \sigma, m, w)$ denote the probability that $(m, w) \in S_I$ meet and marry on the first day of $\Gamma(h)$ under σ'. Let $h'(h, m, w) \in \hat{H}$ denote the immediate history resulting from (m, w) having married on the first day of $\Gamma(h)$. For $w \neq \overline{w}$ let $U(h, \delta, \sigma, m, w)$ denote $\overline{w}$'s expected payoff in $\Gamma(h'(h, m, w))$ under σ'. σ instead of σ' appears as an argument for K, p and U because σ' is determined by σ. By the definition of $\underline{K}$, for any $\eta_1 > 0$ there exists $h_{\eta_1} \in \hat{H}_I$ such that

$$
\begin{aligned}
\underline{K}(\delta, \sigma) + \eta_1 \quad &> \quad K(h_{\eta_1}, \delta, \sigma) \\
&\geq \quad p(h_{\eta_1}, \sigma, \overline{m}, \overline{w}) v(\overline{m}, \overline{w}) \\
&\quad + \sum_{m \in M} \sum_{w \neq \overline{w}} p(h_{\eta_1}, \sigma, m, w) \delta U(h_{\eta_1}, \delta, \sigma, m, w) \\
&\quad + \Big(1 - p(h_{\eta_1}, \sigma, \overline{m}, \overline{w}) - \sum_{m \in M} \sum_{w \neq \overline{w}} p(h_{\eta_1}, \sigma, m, w)\Big) \delta \underline{K}(\delta, \sigma).
\end{aligned} \tag{I2}
$$

$p(h_{\eta_1}, \sigma, \overline{m}, \overline{w}) > \epsilon$ because $(\overline{m}, \overline{w})$ accept each other in $\Gamma(h_{\eta_1})$. Fix (m, w) such that $w \neq \overline{w}$ and $p(h_{\eta_1}, \sigma, m, w) > 0$. Because players other than $\overline{w}$ follow σ,

$(m,w) \in P$ or (m,w) is a woman-acceptable pair from $\overline{S}$. Let $P' := P \backslash \{(m,w)\}$ if $(m,w) \in P$ or $P' := P$ otherwise. Let $\overline{S}'$ denote the submarket complementing P' in $S(h'(h,m,w)) = (M \backslash \{m\}, W \backslash \{w\}, u, v)$. If $(m,w) \in P$ then $\overline{S}' = \overline{S}$ and thus $\overline{w}$ is a top player for $\overline{S}'$ and $\alpha^{\overline{S}'}(\overline{w}) = \overline{m}$. If instead (m,w) is a woman-acceptable pair from $\overline{S}$ then $\overline{S}' = \overline{S} \backslash (m,w)$ and one of the following is true:

1. $m \neq \overline{m}$: $\overline{w}$ is a top player for $\overline{S}'$ and $\alpha^{\overline{S}'}(\overline{w}) = \overline{m}$.

2. $m = \overline{m}$, $\underline{m} \neq s$: $\overline{w}$ is a top player for $\overline{S}'$ and $\alpha^{\overline{S}'}(\overline{w}) = \underline{m}$.

3. $m = \overline{m}$, $\underline{m} = s$: $\alpha^{\overline{S}'}(\overline{w}) = s$.

It is easy to verify from definition that P being a settled set for E implies P' is a settled set for the sub-environment $E' := (M \backslash \{m\}, W \backslash \{w\},\ u, v, C)$. Note that σ' restricted to $\Gamma(h'(h_{\eta_1}, m, w))$ is a subgame perfect equilibrium of it because $\overline{w}$ has switched back to following σ. If $(m,w) \in P$, or if case 1 above is true, then since the number of men in $S(h'(h,m,w))$ is less than n, the inductive hypothesis applies, implying $P' \cup \{(\overline{m}, \overline{w})\}$ is a settled set for E', thus for δ sufficiently close to 1, $\sigma \in \Sigma(\delta)$ implies $(\overline{m}, \overline{w})$ marry almost surely and upon first meeting in $\Gamma(h'(h_{\eta_1}, m, w))$, and thus $\delta U(h_{\eta_1}, \delta, \sigma, m, w) \to v(\overline{m}, \overline{w})$ as $\delta \to 1$. Likewise if case 2 is true then $\delta U(h_{\eta_1}, \delta, \sigma, m, w) \to v(\underline{m}, \overline{w})$ as $\delta \to 1$. If case 3 is true then by the assumption that P' is a settled set for E' and Lemma 1.5.1(a), $\delta U(h_{\eta_1}, \delta, \sigma, m, w) = 0 = v(\underline{m}, \overline{w})$. In general, for any $\eta_2 > 0$ there exists $d_{\eta_2} < 1$ such that if $\delta > d_{\eta_2}$ then for $w \neq \overline{w}$, $p(h_{\eta_1}, \sigma, m, w) > 0$ implies

$$v(\overline{m}, \overline{w}) > \delta U(h_{\eta_1}, \delta, \sigma, m, w) > v(\underline{m}, \overline{w}) - \eta_2. \tag{I3}$$

Substituting I3 along with $p(h_{\eta_1}, \sigma, \overline{m}, \overline{w}) > \epsilon$, $\underline{K}(\delta, \sigma) \leq v(\overline{m}, \overline{w})$ and $R(h_{\eta_1}, \sigma) := \sum_{m \in M} \sum_{w \neq \overline{w}} p(h_{\eta_1}, \sigma, m, w)$ back to I2, we have

$$\underline{K}(\delta, \sigma) > \frac{\epsilon v(\overline{m}, \overline{w}) + R(h_{\eta_1}, \sigma)(v(\underline{m}, \overline{w}) - \eta_2) - \eta_1}{1 - \delta(1 - \epsilon - R(h_{\eta_1}, \sigma))} \tag{I4}$$

for η_2 sufficiently small and $\delta > d_{\eta_2}$. The right hand side of I4 is decreasing in $R(h_{\eta_1}, \sigma)$ for δ close 1 and η_1, η_2 close to 0. For such extreme values, $R(h_{\eta_1}, \sigma) < 1 - \epsilon$ implies for $\delta > d_{\eta_2}$,

$$\underline{K}(\delta, \sigma) > \epsilon v(\overline{m}, \overline{w}) + (1 - \epsilon) v(\underline{m}, \overline{w}) - (1 - \epsilon)\eta_2 - \eta_1.$$

Taking limits $\eta_1 \to 0$, $\eta_2 \to 0$ and correspondingly $\delta \to 1$ we have

$$\lim_{\delta \to 1} \inf_{\sigma \in \Sigma(\delta)} \underline{K}(\delta, \sigma) \geq \epsilon v(\overline{m}, \overline{w}) + (1 - \epsilon) v(\underline{m}, \overline{w}) > v(\underline{m}, \overline{w}). \tag{I5}$$

Inequality I5 implies $\overline{w}$'s expected payoff under σ in $\Gamma(h)$ where $h \in \hat{H}_I$ is strictly greater than $v(\underline{m}, \overline{w})$ for δ sufficiently close to 1. For such δ, $\sigma \in \Sigma(\delta)$ implies $\overline{w}$

rejects any m such that $\overline{m} \succ_{\overline{w}} m$ when the remaining market is S_I.

Now consider $\overline{m}$. Suppose other players (including $\overline{w}$) follow σ and $\overline{m}$ deviates to the strategy under which he only accepts $\overline{w}$ when the remaining market is S_I, and switches back to following σ if the remaining market is no longer S_I. Let σ'' denote the strategy profile due to $\overline{m}$'s unilateral deviation.

Claim A.2.2. *For δ sufficiently close to 1, $\sigma \in \Sigma(\delta)$ implies under σ'', $(\overline{m}, \overline{w})$ marry almost surely and upon first meeting in $\Gamma(h)$ for any $h \in \hat{H}_I$.*

Proof: Observe that for δ sufficiently close to 1, almost surely one of the following scenarios occur under σ'':

1. $(\overline{m}, \overline{w})$ meet when the remaining market is S_I.

2. The first marriage is between (m, w) where $m \neq \overline{m}$, $w \neq \overline{w}$, $(m, w) \in P$ or (m, w) is a woman-acceptable pair from $\overline{S}$.

That $\overline{w}$ will not marry with $m \neq \overline{m}$ as the first married couple is due to the assumption that she may not marry a man from a pair in P, plus the conclusion from the above that she rejects every man worse than $\overline{m}$ when the remaining market is S_I. If scenario 1 occurs then $(\overline{m}, \overline{w})$ marry immediately. If scenario 2 occurs, then the inductive hypothesis[1] implies $(\overline{m}, \overline{w})$ marry almost surely and upon first meeting in the subsequent subgame. □

By Claim A.2.2, $\overline{m}$'s expected payoff under σ'' in $\Gamma(h)$ where $h \in \hat{H}_I$ tends to $u(\overline{m}, \overline{w})$ as $\delta \to 1$, implying $\overline{m}$'s expected payoff under σ in $\Gamma(h)$ also tends to $u(\overline{m}, \overline{w})$. Thus for δ sufficiently close to 1, $\sigma \in \Sigma(\delta)$ implies when the remaining market is S_I, $\overline{m}$ rejects any w such that $\overline{w} \succ_{\overline{m}} w$ if w would accept $\overline{m}$ with positive probability. Then an analogous claim to Claim A.2.2 is true for σ with an analogous proof.

Now consider any $h \in \hat{H}$ such that $S(h) \neq S_I$, $(\overline{m}, \overline{w}) \in S(h)$, and moreover $S(h)$ can be derived from S_I as a result of a sequence of pairs having left, among which each pair (m, w) is either in P or is a woman-acceptable pair from $\overline{S}$. Let $P(h)$ denote the set of pairs in P which are present in $S(h)$, and let $E(h)$ denote the sub-environment with initial market $S(h)$. It is easy to verify from definition that $P(h)$ is a settled set for $E(h)$, $\overline{w}$ is a top player for $\overline{S}(h) := S(h) \backslash P(h)$, and $\overline{m} = \alpha^{\overline{S}(h)}(\overline{w})$. Thus the inductive hypothesis implies $P(h) \cup \{(\overline{m}, \overline{w})\}$ is a settled set for $E(h)$. Therefore for δ sufficiently close to 1, $\sigma \in \Sigma(\delta)$ implies $(\overline{m}, \overline{w})$ marry almost surely and upon first meeting in $\Gamma(h)$. It follows that $P \cup \{(\overline{m}, \overline{w})\}$ is a settled set for E.

The proof for the case that a male player is a top player for $\overline{S}$ is similar. □

Proof of Proposition 1.5.8: Denote $E := (M, W, u, v, C)$. To prove the present propo-

[1] The argument for the applicability of the inductive hypothesis, and its actual application, are the same as those used above for $\overline{w}$'s case.

sition it suffices to show there exists some $d < 1$ such that for any $\sigma \in \Sigma(\delta)$ where $\delta > d$, (m_i, w_i) marry almost surely and upon first meeting if $i \leq k$, and m_i or w_i stays single if $i > k$.

Since the empty set is a settled set for E and m_1 or w_1 is a top player for the initial market S_1, Lemma A.2.1 implies $P_1 := (\{m_1, w_1\})$ is a settled set for E. Let $P_i := \{(m_j, w_j) : j \leq i\}$. Suppose P_i is a settled set for E if $i < n$ for some $n \leq k$. Since m_n or w_n is a top player for $S_n = S_1 \backslash P_{n-1}$, Lemma A.2.1 implies that P_n is a settled set for E. Thus P_k is a settled set for E. Let d be a settling discount factor for P_k. Pick $\sigma \in \Sigma(\delta)$ where $\delta > d$. P_k being a settled set for E implies (m_i, w_i) marry almost surely and upon first meeting under σ if $i \leq k$. Pick a man $m \in S_{k+1}$. (m, w) marry with positive probability in σ only if $w \in S_{k+1}$ and m is acceptable to w. Since S_{k+1} is trivial, if m is acceptable to $w \in S_{k+1}$ then w is unacceptable to m. Thus if m marries at all with positive probability, his expected payoff would be negative. Lemma 1.5.1(a) then implies m stays single in σ. It follows that a woman in S_{k+1} stays single in σ as well. □

A.3 Example: A man might marry an unacceptable woman

The following example demonstrates that a man may marry an unacceptable woman with positive probability in equilibrium.

The initial market is described by the following lists:

$$P(m_1) = w_1, w_2, \qquad P(w_1) = m_2, m_1,$$
$$P(m_2) = w_2, w_1, \qquad P(w_2) = m_1, m_2,$$
$$P(m_3) = w_3, \qquad P(w_3) = m_3, m_1.$$

Each of the players m_1, m_2, w_1, w_2 receives a payoff of 16 from marrying the first choice, 1 the second choice, and -1 the unacceptable choice; m_3 receives a payoff of 16 from marrying w_3; w_3 receives a payoff of 16 from marrying m_3 and 4 from marrying m_1. Set $\delta = 0.5$. The contact function satisfies $C(m, w, S) = C(\hat{m}, \hat{w}, S)$ for any pairs $(m, w) \in S$ and $(\hat{m}, \hat{w}) \in S$ for any $S \in \mathscr{S}$.

Consider strategy profile σ given as follows:

- If (m_1, w_3) do not meet on the first day, the players follow the μ^m strategy profile where μ^m is the men-optimal matching for the initial market.
- If (m_1, w_3) meet on the first day, m_1 accepts w_3 with certainty and w_3 accepts m_1 with probability 0.5.
  - If m_1 accepts w_3 on the first day, then regardless of whether m_1 is accepted,

after the first day the players follow the $\hat{\mu}^m$-strategy profile where $\hat{\mu}^m$ is the men-optimal matching for the remaining market.

– If m_1 rejects w_3 on the first day, then after the first day the players follow the μ^w strategy profile where μ^w is the women-optimal matching for the initial market.

It is straightforward to verify that σ is a subgame perfect equilibrium of the game. Notably, if (m_1, w_3) meet on the first day, m_1 accepts w_3 with positive probability despite the latter being unacceptable, because by doing so there is a probability of 0.5 that w_3 will reject m_1 and in the subsequent subgame m_1 will marry his first choice w_1 eventually, whereas if he rejects w_3 he will have to marry his second choice w_2 eventually. Meanwhile, because of discounting, w_3 is indifferent between marrying m_1 today and marrying her first choice m_3 in the future, justifying her randomizing between accepting and rejecting m_1. Under σ there is a positive probability of $1/9 \times 0.5 = 1/18$ that m_1 marries the unacceptable w_3.

A.4 Example: A market that satisfies Condition 1.5.8 but violates the absence of simultaneous cycles

It is easy to verify that the marriage market described by the following lists of preferences satisfies Condition 1.5.8 but violates the absence of simultaneous cycles given in Romero-Medina and Triossi (2013).

$$
\begin{aligned}
P(m_1) &= w_1, w_3, w_4, & P(w_1) &= m_1, \\
P(m_2) &= w_2, w_4, w_3, & P(w_2) &= m_2, \\
P(m_3) &= w_3, & P(w_3) &= m_3, m_2, m_1, \\
P(m_4) &= w_4, & P(w_4) &= m_4, m_1, m_2.
\end{aligned}
$$

Observe that there are two preference cycles: $w_3 \succ_{m_1} w_4 \succ_{m_2} w_3$ and $m_2 \succ_{w_3} m_1 \succ_{w_4} m_2$, implying the simultaneous cycle $(w_4, m_1, w_3, m_2, w_4)$ in the notation of Romero-Medina and Triossi (2013).

APPENDIX B

Additional Proofs and Notation in Chapter 3

B.1 Proof of Proposition 3.3.2

Proof. Choose any voting system (r, d) and $s \in \Sigma_V(r)$. There is a sequential procedure r^* such that $r^*(i) > r^*(j)$ for any $i, j \in J$ such that $r(i) \geq r(j)$. Thus if j's vote is observable to i under r then it is also observable to i under r^*. Let $s_i(x_i, (z_j)_{r(j)<r(i)})$ denote i's strategy given past votes $(z_j)_{r(j)<r(i)}$ and signal x_i. Construct $s' \in \Sigma_V(r^*)$ such that for every $i \in J$, $s'_i(x_i, (z'_j)_{r^*(j)<r^*(i)}) = s_i(x_i, (z_j)_{r(j)<r(i)})$ if $z'_j = z_j$ for any j such that $r(j) < r(i)$. Obviously s and s' are outcome equivalent, implying $u(s|r, d) = u(s'|r^*, d)$. Thus $\max_{\sigma \in \Sigma_V(r)} u(\sigma|r, d) \leq \max_{\sigma \in \Sigma_V(r^*)} u(\sigma|r^*, d)$.

Let k be the last voter according to r^*. Construct d^* such that $d^*(z_1, ..., z_n) = z_k$. Choose any $\hat{s} \in \Sigma_V(r^*)$. Construct $\tilde{s} \in \Sigma_V(r^*)$ such that $\tilde{s}$ and $\hat{s}$ agree for every player $i \neq k$, and for any signal x_k and past votes $\mathbf{z}_{-k}$,

$$\tilde{s}_k(x_k, \mathbf{z}_{-k}) = \begin{cases} d(\mathbf{z}_{-k}, G) & \text{if } d(\mathbf{z}_{-k}, G) = d(\mathbf{z}_{-k}, I) \\ \hat{s}_k(x_i, \mathbf{z}_{-k}) & \text{if } d(\mathbf{z}_{-k}, G) \neq d(\mathbf{z}_{-k}, I) \end{cases} .$$

It is straightforward to verify that $\tilde{s}$ and $\hat{s}$ are outcome equivalent, implying $u(\hat{s}|r^*, d) = u(\tilde{s}|r^*, d^*)$. Thus $\max_{\sigma \in \Sigma_V(r^*)} u(\sigma|r^*, d) \leq \max_{\sigma \in \Sigma_V(r^*)} u(\sigma|r^*, d^*)$. It follows that $\max_{\sigma \in \Sigma_V(r)} u(\sigma|r, d) \leq \max_{\sigma \in \Sigma_V(r^*)} u(\sigma|r^*, d^*)$, implying $U(r, d) \leq U(r^*, d^*)$ by Lemma 3.3.1. □

B.2 Proof of Proposition 3.3.3

For any procedure r and $s \in \Sigma_V(r)$, let $\Pr(G|\mathbf{z}, s)$ denote the probability that $\omega = G$ conditional on the jurors following s and the realized votes are $\mathbf{z}$. The proof is assisted by the following lemma.

Lemma B.2.1. *For any procedure r, if $d^* \in argmax_d U(r,d)$ then for any $s^* \in \Sigma_V(r)$ such that $u(s^*|r,d^*) = U(r,d^*)$ and vote profile $\mathbf{z} \in \{G,I\}^n$,*

$$d^*(\mathbf{z}) = \begin{cases} G & \textit{if} \quad \Pr(G|\mathbf{z}, s^*) > 0.5, \\ I & \textit{if} \quad \Pr(G|\mathbf{z}, s^*) < 0.5. \end{cases}$$

Proof. Fix procedure r. Suppose there is $d^* \in \text{argmax}_d U(r,d)$ and $s^* \in \Sigma_V(r)$ where $u(s^*|r,d^*) = U(r,d^*)$ such that d^* does not satisfy the condition in the lemma. Define $K = \Big\{\mathbf{z} \in \{G,I\}^n : d^*(\mathbf{z}) = I \text{ and } \Pr(G|\mathbf{z},s^*) > 0.5\Big\}$ and $L = \Big\{\mathbf{z} \in \{G,I\}^n : d^*(\mathbf{z}) = G \text{ and } \Pr(G|\mathbf{z},s^*) < 0.5\Big\}$. By assumption $K \cup L \neq \emptyset$. Let $q_\omega(\mathbf{z})$ be the probability that the realized vote profile is $\mathbf{z}$ conditional on ω and the jurors following s^*. Construct rule d' such that for any $\mathbf{z} \in \{G,I\}^n$,

$$d'(\mathbf{z}) = \begin{cases} G & \text{if} \quad \Pr(G|\mathbf{z}, s^*) > 0.5, \\ I & \text{if} \quad \Pr(G|\mathbf{z}, s^*) < 0.5, \\ d^*(\mathbf{z}) & \text{if} \quad \Pr(G|\mathbf{z}, s^*) = 0.5. \end{cases}$$

It follows that $d'(\mathbf{z}) = d^*(\mathbf{z})$ if $\mathbf{z} \notin K \cup L$. We have

$$u(s^*|r,d') - u(s^*|r,d^*) = \sum_{\mathbf{z} \in K} \Big(\pi q_G(\mathbf{z}) - (1-\pi) q_I(\mathbf{z})\Big) + \sum_{\mathbf{z} \in L} \Big((1-\pi) q_I(\mathbf{z}) - \pi q_G(\mathbf{z})\Big).$$

If $\mathbf{z} \in K$ then $\Pr(G|\mathbf{z},s^*) = \frac{\pi q_G(\mathbf{z})}{\pi q_G(\mathbf{z}) + (1-\pi) q_I(\mathbf{z})} > 0.5$, implying $\pi q_G(\mathbf{z}) - (1-\pi) q_I(\mathbf{z}) > 0$. Similarly if $\mathbf{z} \in L$ then $(1-\pi) q_I(\mathbf{z}) - \pi q_G(\mathbf{z}) > 0$. It follows that $u(s^*|r,d') - u(s^*|r,d^*) > 0$ because $K \cup L \neq \emptyset$. Therefore $U(r,d') > U(r,d^*)$, contradicting the assumption that $d^* \in \text{argmax}_d U(r,d)$. □

Proof of Proposition 3.3.3

Fix the simultaneous procedure r. Choose any $d' \in \text{argmax}_d U(r,d)$ and $s' \in \Sigma_V(r)$ such that $u(s'|r,d') = U(r,d')$. Lemma B.2.1 implies that for any vote profile $\mathbf{z} \in \{G,I\}^n$,

$$d'(\mathbf{z}) = \begin{cases} G & \text{if} \quad \Pr(G|\mathbf{z}, s') > 0.5, \\ I & \text{if} \quad \Pr(G|\mathbf{z}, s') < 0.5. \end{cases}$$

For each $i \in J$ let p^i_ω denote the probability that juror i votes G conditional on ω and s'. Note that since $f^i_\omega(x_i) > 0$ for any $i \in J$, $x_i \in X_i$ and $\omega \in \{G,I\}$, $p^i_I = 0$ if and only if $p^i_G = 0$. Construct s^* that satisfies the following for any $i \in J$:

- If $p_G^i \notin \{0,1\}$:
  - If $p_G^i/p_I^i \geq 1$ then s^* prescribes the same strategy for i as s'.
  - If $p_G^i/p_I^i < 1$ then given any signal x_i, i votes G with the same probability that he votes I given x_i under s'.
- If $p_G^i \in \{0,1\}$ then i votes G with probability 0.5 regardless of x_i.

Construct d^* such that $d^*(\mathbf{z}) = G$ if and only if $\Pr(G|\mathbf{z}, s^*) \geq 0.5$. It is straightforward to verify that $u(s'|r,d') = u(s^*|r,d^*)$. Let t_ω^i denote the probability that juror i votes G conditional on ω and s^*. Clearly $0 < t_\omega^i < 1$ for $\omega \in \{G, I\}$. Moreover $t_G^i/t_I^i = p_G^i/p_I^i$ if $p_G^i/p_I^i \geq 1$, $t_G^i/t_I^i = (1-p_G^i)/(1-p_I^i)$ if $p_G^i/p_I^i < 1$, and $t_G^i/t_I^i = 1$ if $p_G^i \in \{0,1\}$. Consequently $t_G^i/t_I^i \geq 1$ for any $i \in J$. Recall that for any vote profile $\mathbf{z} = (z_1, ..., z_n)$, $\mathbf{1}(z_i) = 1$ if $z_i = G$ or 0 otherwise. We have

$$\Pr(G|\mathbf{z}, s^*) = 1\Big/\Big[1 + \frac{(1-\pi)}{\pi} \prod_{i\in J, \mathbf{1}(z_i)=1} \frac{t_I^i}{t_G^i} \prod_{i \in J, \mathbf{1}(z_i)=0} \frac{1-t_I^i}{1-t_G^i}\Big].$$

$\Pr(G|\mathbf{z}, s^*) \geq 0.5$ if and only if

$$\frac{(1-\pi)}{\pi} \prod_{i\in J, \mathbf{1}(z_i)=1} \frac{t_I^i}{t_G^i} \prod_{i \in J, \mathbf{1}(z_i)=0} \frac{1-t_I^i}{1-t_G^i} \leq 1,$$

or equivalently

$$\begin{aligned} &\log\frac{1-\pi}{\pi} + \sum_{i\in J, \mathbf{1}(z_i)=1} \log\frac{t_I^i}{t_G^i} + \sum_{i\in J, \mathbf{1}(z_i)=0} \log\frac{1-t_I^i}{1-t_G^i} \leq 0 \\ \Longrightarrow &\log\frac{1-\pi}{\pi} + \sum_{i\in J} \mathbf{1}(z_i) \log\frac{t_I^i}{t_G^i} + \sum_{i\in J} (1-\mathbf{1}(z_i)) \log\frac{1-t_I^i}{1-t_G^i} \leq 0 \\ \Longrightarrow &\sum_{i\in J} \mathbf{1}(z_i) \log\frac{1-t_I^i}{1-t_G^i}\frac{t_G^i}{t_I^i} \geq \log\frac{1-\pi}{\pi} + \sum_{i\in J} \log\frac{1-t_I^i}{1-t_G^i}. \end{aligned}$$

Let $w_i = \log\frac{1-t_I^i}{1-t_G^i}\frac{t_G^i}{t_I^i}$ and $k = \log\frac{1-\pi}{\pi} + \sum_{i\in J}\log\frac{1-t_I^i}{1-t_G^i}$. $w_i \geq 0$ because $\frac{1-t_I^i}{1-t_G^i}\frac{t_G^i}{t_I^i} \geq 1$. Thus $d^*(\mathbf{z}) = G$ if and only if $\sum_{i\in J} w_i \mathbf{1}(z_i) \geq k$. □

B.3 Notation for Proofs of Results in Section 3.5

Introduce the following notation for the game induced by (T, P, M), where (T, P, M) can either be an institution or an improper institution. The notation will be used throughout the Appendices for the proofs of results in Section 3.5.

- $S(T, P, M)$: the set of all *pure* strategy profiles of the game induced by (T, P, M).
- (x_i, p): a typical information set of player i, where x_i is to his private information about the state, and p is his perception that contains the message he has received.
- For $s \in S(T, P, M)$, $i \in \mathscr{N}$ and $\mathbf{x} = (x_1, ..., x_N) \in X$,

  $a_i(x_i, p|s)$: i's choice of action under s given (x_i, p).

  $m_i(x_i, p|s)$: (if $i < N$) i's choice of message under s given (x_i, p).

  $\rho_i(\mathbf{x}|s)$: the message profile i receives conditional on $\mathbf{x}$ and s.

  $\alpha_i(\mathbf{x}|s)$: the action i takes conditional on $\mathbf{x}$ and s.

  $\mu_i(\mathbf{x}|s)$: the message i sends conditional on $\mathbf{x}$ and s.

  $\rho_j(\mathbf{x}|s, i, h, m_i)$: the message profile that player $j > i$ receives conditional on (1) every player after i follows s, (2) player i receives $h \in T_i$, (3) player i sends message $m_i \in M(h)$.

The following equalities hold by definition:

$$\begin{aligned}
\alpha_i(\mathbf{x}|s) &= a_i(x_i, Pt(\rho_i(\mathbf{x}|s))|s), \\
\mu_i(\mathbf{x}|s) &= m_i(x_i, P(\rho_i(\mathbf{x}|s))|s), \\
\rho_{i+1}(\mathbf{x}|s) &= \rho_i(\mathbf{x}|s) \times \mu_i(\mathbf{x}|s).
\end{aligned}$$

B.4 Proof of Proposition 3.5.1

We will prove a more general result which implies Proposition 3.5.1 as a corollary.

It is straightforward to extend the definition of value to any mechanism Γ whose set of outcomes is a subset of A. Then we can also extend the definition of dominance: Mechanism Γ' dominates mechanism Γ if the value of Γ' is weakly higher than the value of Γ for any common objective.

Fix a mechanism Γ. Let $S(\Gamma)$ denote the set of all pure strategy profiles of Γ. Let $C(\Gamma)$ denote the set of all social choice functions inducible by a pure strategy profile of Γ. For $s \in S(\Gamma)$ let $v(s|\Gamma, \phi, F)$ denote the common expected payoff achieved by s in the game induced by Γ and (ϕ, F). Let $V(\Gamma, \phi, F)$ denote the value of Γ for (ϕ, F).

First we show a lemma.

Lemma B.4.1. *$V(\Gamma, \phi, F) = \max_{s \in S(\Gamma)} v(s|\Gamma, \phi, F)$ for any finite mechanism Γ and common objective (ϕ, F).*

Proof: Fix Γ and (ϕ, F). Let $\Sigma(\Gamma)$ denote the set of all strategy profiles of Γ. For $\epsilon \in (0, \bar{\epsilon})$ where $\bar{\epsilon}$ is sufficiently small let $\Gamma(\epsilon)$ denote the perturbed version of Γ such whenever a player chooses a generic action (to be distinguished from the action a_i that player i contributes to the common objective) as a realization of a possibly mixed strategy, his chosen action will realize with probability $1 - (n-1)\epsilon$ where n is the total number of generic actions available at this point, and each of the other $n-1$ generic actions will realize with probability ϵ.

For any $\sigma \in \Sigma(\Gamma)$ let $w(\sigma, \epsilon)$ denote the common expected payoff achieved by σ in $\Gamma(\epsilon)$. Fix $\epsilon \in (0, \bar{\epsilon})$. $\text{argmax}_{\sigma \in \Sigma(\Gamma)} w(\sigma, \epsilon)$ is nonempty because $\Sigma(\Gamma)$ is compact and $w(\cdot, \epsilon)$ is continuous in its first argument. Choose $\hat{\sigma} \in \text{argmax}_{\sigma \in \Sigma} w(\sigma, \epsilon)$. Let $\hat{\beta}$ denote the belief system derived from $\hat{\sigma}$ using Bayes' rule in $\Gamma(\epsilon)$. Suppose there is an information set K of player i such that under $\hat{\sigma}$ player i is not best responding. Thus i would find it profitable to deviate to some strategy σ_i' in K. Let σ' denote the strategy profile under which i unilaterally deviates to σ_i' in K. Clearly i's expected payoff under σ' in $\Gamma(\epsilon)$ is strictly higher than that under $\hat{\sigma}$ because K is reached with strictly positive probability, implying $w(\sigma', \epsilon) > w(\hat{\sigma}, \epsilon)$, contradicting the choice of $\hat{\sigma}$. We have thus established that $(\hat{\sigma}, \hat{\beta})$ is a perfect Bayesian equilibrium of $\Gamma(\epsilon)$.

Construct pure strategy profile $\hat{s}$ such that for each player i, $\hat{s}$ prescribes a pure strategy that is in the support of the (possibly mixed) strategy taken by i under $\hat{\sigma}$. That $\hat{\sigma}$ being a perfect Bayesian equilibrium of $\Gamma(\epsilon)$ implies $w(\hat{s}, \epsilon) = w(\hat{\sigma}, \epsilon)$. Thus $w(\hat{s}, \epsilon) = \max_{\sigma \in \Sigma(\Gamma)} w(\sigma, \epsilon)$, implying $\max_{s \in S(\Gamma)} w(s, \epsilon) = \max_{\sigma \in \Sigma(\Gamma)} w(\sigma, \epsilon)$.

Note that, for any fixed σ, $w(\sigma, \epsilon)$ is a polynomial function of ϵ of finite degrees. Since $S(\Gamma)$ is finite, for some $\eta > 0$ there is $s^* \in S(\Gamma)$ such that $w(s^*, \epsilon) = \max_{s \in S(\Gamma)} w(s, \epsilon)$ for any $\epsilon < \eta$. It follows that $s^* \in \text{argmax}_{\sigma \in \Sigma(\Gamma)} w(\sigma, \epsilon)$ if $\epsilon < \eta$. Let $\beta^*(\epsilon)$ be the belief system derived from s^* using Bayes' rule in $\Gamma(\epsilon)$. $s^* \in \text{argmax}_{\sigma \in \Sigma(\Gamma)} w(\sigma, \epsilon)$ implies $(s^*, \beta^*(\epsilon))$ is a perfect Bayesian equilibrium of $\Gamma(\epsilon)$ by the argument in the second paragraph of the proof. Note that $\beta^*(\epsilon)$ is continuous in ϵ and thus $\beta^* = \lim_{\epsilon \to 0} \beta^*(\epsilon)$ exists. Clearly (s^*, β^*) is a perfect Bayesian equilibrium of the unperturbed game. Suppose there is $\tilde{\sigma} \in \Sigma(\Gamma)$ such that $v(\tilde{\sigma}|\Gamma, \phi, F) > v(s^*|\Gamma, \phi, F)$. Since $\lim_{\epsilon \to 0} w(\sigma, \epsilon) = v(\sigma|\Gamma, \phi, F)$ for any σ, there is some $\tilde{\eta} > 0$ such that $w(\tilde{\sigma}, \epsilon) > w(s^*, \epsilon)$ for any $\epsilon < \tilde{\eta}$, contradicting that $s^* \in \text{argmax}_{\sigma \in \Sigma(\Gamma)} w(\sigma, \epsilon)$ for any $\epsilon < \eta$. Hence $s^* \in \text{argmax}_{s \in \Sigma(\Gamma)} v(s|\Gamma, \phi, F)$. It follows that $V(\Gamma, \phi, F) = v(s^*|\Gamma, \phi, F) = \max_{s \in S(\Gamma)} v(s|\Gamma, \phi, F)$. $\square$

Proposition B.4.2. *If Γ' and Γ are finite mechanisms then Γ' dominates Γ if and only if $C(\Gamma) \subset C(\Gamma')$.*

Proof: (The "if" direction.) Suppose $C(\Gamma) \subset C(\Gamma')$. Fix (ϕ, F) and choose $s \in \text{argmax}_{S(\Gamma)} v(s|\Gamma, \phi, F)$. By assumption there is $s' \in S(\Gamma')$ such that $\boldsymbol{\alpha}(\cdot|s') = \boldsymbol{\alpha}(\cdot|s)$.

It follows that

$$\begin{aligned} v(s'|\Gamma',\phi,F) &= \sum_{\mathbf{x}\in X} F(\mathbf{x})\phi(\boldsymbol{\alpha}(\mathbf{x}|s'),\mathbf{x}) \\ &= \sum_{\mathbf{x}\in X} F(\mathbf{x})\phi(\boldsymbol{\alpha}(\mathbf{x}|s),\mathbf{x}) = v(s|\Gamma,\phi,F). \end{aligned}$$

By Lemma B.4.1, $V(\Gamma',\phi,F) \geq v(s'|\Gamma',\phi,F) = v(s|\Gamma,\phi,F) = V(\Gamma,\phi,F)$, implying that Γ' dominates Γ because (ϕ,F) is arbitrarily chosen.

(The "only if" direction.) Suppose Γ' dominates Γ. Choose any $s \in S(\Gamma)$. Construct ϕ such that $\phi(\mathbf{a},\mathbf{x}) = 1$ if $\mathbf{a} = \boldsymbol{\alpha}(\mathbf{x}|s)$ and $\phi(\mathbf{a},\mathbf{x}) = 0$ otherwise. Choose F that is strictly positive on X. Clearly $v(s|\Gamma,\phi,F) = 1$. By Lemma B.4.1, that Γ' dominates Γ implies

$$\max_{s\in S(\Gamma')} v(s|\Gamma',\phi,F) \geq \max_{s\in S(\Gamma)} v(s|\Gamma,\phi,F) = 1,$$

which in turn implies there is $s' \in S(\Gamma')$ such that $v(s'|\Gamma',\phi,F) \geq 1$. If $\boldsymbol{\alpha}(\mathbf{y}|s') \neq \boldsymbol{\alpha}(\mathbf{y}|s)$ for some $\mathbf{y} \in X$ then

$$v(s'|\Gamma',\phi,F) = \sum_{\mathbf{x}\in X} F(\mathbf{x})\phi(\boldsymbol{\alpha}(\mathbf{x}|s'),\mathbf{x}) \leq 1 - F(\mathbf{y}) < 1,$$

a contradiction. Thus $\boldsymbol{\alpha}(\mathbf{x}|s') = \boldsymbol{\alpha}(\mathbf{x}|s)$ for any $\mathbf{x} \in X$, implying $C(\Gamma) \subset C(\Gamma')$ because s is arbitrarily chosen. □

B.5 Proof of Lemma 3.5.3

Proof. Suppose (T',P',M') is obtained from (T,P,M) by expanding. E2 implies that for any $p \in P'$ where $p \cap T \neq \emptyset$ there is $\tau(p) \in P$ such that $\tau(p) = p \cap T$. E1 implies $M(\tau(p)) \subset M'(p)$. Choose $s \in S(T,P,M)$. Construct $s' \in S(T',P',M')$ such that for any $i \in \mathscr{N}$, $x_i \in X_i$ and $p \in P'_i$ where $p \cap T \neq \emptyset$,

$$\begin{aligned} a_i(x_i,p|s') &= a_i(x_i,\tau(p)|s), \\ \text{(if } i<N) \quad m_i(x_i,p|s') &= m_i(x_i,\tau(p)|s). \end{aligned}$$

To verify that s' is properly defined, observe that for any $p \in P'_i$ where $p \cap T \neq \emptyset$ we have $m_i(x_i,p|s') = m_i(x_i,\tau(p)|s) \in M(\tau(p)) \subset M'(p)$. Choose any $\mathbf{x} = (x_1,...,x_N) \in X$. Clearly $\rho_1(\mathbf{x}|s') = \rho_1(\mathbf{x}|s)$. Suppose $\rho_i(\mathbf{x}|s') = \rho_i(\mathbf{x}|s)$ for any $i \leq k$ for some $k \geq 1$. Thus

$$\begin{aligned} \mu_k(\mathbf{x}|s') &= m_k(x_k, P'(\rho_k(\mathbf{x}|s'))|s') = m_k(x_k, \tau(P'(\rho_k(\mathbf{x}|s')))|s) \\ &= m_k(x_k, P(\rho_k(\mathbf{x}|s'))|s) = m_k(x_k, P(\rho_k(\mathbf{x}|s))|s) = \mu_k(\mathbf{x}|s). \end{aligned}$$

It follows that

$$\rho_{k+1}(\mathbf{x}|s') = \rho_k(\mathbf{x}|s') \times \mu_k(\mathbf{x}|s') = \rho_k(\mathbf{x}|s) \times \mu_k(\mathbf{x}|s) = \rho_{k+1}(\mathbf{x}|s).$$

Thus $\rho_i(\mathbf{x}|s') = \rho_i(\mathbf{x}|s)$ for any $i \in \mathscr{N}$, implying

$$\begin{aligned}\alpha_i(\mathbf{x}|s') &= a_i(x_i, P'(\rho_i(\mathbf{x}|s'))|s') = a_i(x_i, \tau(P'(\rho_i(\mathbf{x}|s')))|s) \\ &= a_i(x_i, P(\rho_i(\mathbf{x}|s'))|s) = a_i(x_i, P(\rho_i(\mathbf{x}|s))|s) = \alpha_i(\mathbf{x}|s).\end{aligned}$$

Thus $\boldsymbol{\alpha}(\cdot|s') = \boldsymbol{\alpha}(\cdot|s)$, implying $C(T,P,M) \subset C(T',P',M')$ since s is arbitrarily chosen. The lemma then follows immediately from Proposition 3.5.2. □

B.6 Proof of Lemma 3.5.5

Two lemmas are used to assist the proof.

Lemma B.6.1. *Fix (T,P,M). For any $i < N$ and $h \in T_i$, if $m_i, m_i' \in M(h)$ are synonymous given h within (T,P,M) then $P(\rho_j(\mathbf{x}|s,i,h,m_i)) = P(\rho_j(\mathbf{x}|s,i,h,m_i'))$ for any $j > i$, $\mathbf{x} \in X$ and $s \in S(T,P,M)$.*

Proof. Choose any $\mathbf{x} = (x_1, ..., x_N) \in X$, $s \in S(T,P,M)$, $i < N$, $h \in T_i$ and $m_i, m_i' \in M(h)$ such that m_i and m_i' are synonymous given h within (T,P,M). For any $j > i$ denote $g^j = \rho_j(\mathbf{x}|s,i,h,m_i)$ and $f^j = \rho_j(\mathbf{x}|s,i,h,m_i')$.

Clearly $g^{i+1} = h \times m_i$ and $f^{i+1} = h \times m_i'$. Thus g^{i+1} and f^{i+1} are h-cousins where $g_i^{i+1} = m_i$ and $f_i^{i+1} = m_i'$. It follows by synonymity that $P(g^{i+1}) = P(f^{i+1})$. Suppose g^j and f^j are h-cousins, where $g_i^j = m_i$ and $f_i^j = m_i'$, for any $j \leq k$ for some $k \geq i+1$.

$P(g^k) = P(f^k)$ by synonymity. Note that $g^{k+1} = g^k \times m_k(x_k, P(g^k)|s)$ and $f^{k+1} = f^k \times m_k(x_k, P(f^k)|s)$. $P(g^k) = P(f^k)$ implies the last components of g^{k+1} and f^{k+1} are identical, and hence g^{k+1} and f^{k+1} are h-cousins where $g_i^{k+1} = m_i$ and $f_i^{k+1} = m_i'$ by the inductive hypothesis. Hence $P(g^{k+1}) = P(f^{k+1})$ by synonymity. The present lemma follows by the principle of induction.□

Lemma B.6.2. *If (T',P',M') is obtained from (T,P,M) by trimming such that there exist $i < N$, $p \in P_i$ and $m_i, m_i' \in M(p)$ satisfying T1 and T2, then for any $j < N$ and $p' \in P_j'$:*

1. *There is $\zeta(p') \in P_j$ such that $p' \subset \zeta(p')$.*
2. *If $\zeta(p') \neq p$ then $M'(p') = M(\zeta(p'))$.*
3. *If $\zeta(p') = p$ then $M'(p') = M(p) \backslash \{m_i\}$.*

Moreover, $|P_j'| = |P_j|$ for any $j \in \mathscr{N}$.

Proof. Part 1 follows from T1 immediately.

To show Parts 2 and 3, choose any $j < N$ and $p' \in P'_j$. Note that $M'(p') \subset M(\zeta(p'))$ by T1. If $\zeta(p') \neq p$ then for any $h \in p'$ and $\hat{m}_j \in M(h)$ we have $h \times \hat{m}_j \in T'$ by T1, implying Part 2. If $\zeta(p') = p$ then $j = i$. Note that for any $h \in p'$ and $\hat{m}_i \in M(h)$ we have $h \times \hat{m}_i \in T'$ if and only if $\hat{m}_i \neq m_i$ by T1, implying Part 3.

Pick any $j \in \mathscr{N}$. If $j \leq i$ then $T_j = T'_j$ and thus $P'_j = P_j$ by T1. If $j > i$ then $|P'_j| \leq |P_j|$ by T1. If $|P'_j| < |P_j|$ then there is some $\hat{p} \in P_j$ such that $h \notin T'$ for any $h \in \hat{p}$. Choose any $h \in \hat{p}$. It follows by T1 that $h(i-1) \in p$ and $h_i = m_i$. It then follows by T2 that there exists some $h' \in \hat{p}$ such that $h'(i-1) \in p$ and $h'_i = m'_i$. Hence $h' \in T'$ by T1, a contradiction. Therefore $|P'_j| = |P_j|$ for any $j \in \mathscr{N}$. ∎

Proof of Lemma 3.5.5. Suppose (T', P', M') is obtained from (T, P, M) by trimming such that there exist $i < N$, $p \in P_i$ and $m_i, m'_i \in M(p)$ satisfying T1 and T2. Since (T', P', M') is a sub-institution of (T, P, M), Lemma 3.5.3 implies (T, P, M) dominates (T', P', M').

Now show that (T', P', M') dominates (T, P, M). By assumption there are $i < N$, $h \in T_i$, $p \in P_i$ and $m_i, m'_i \in M(p)$ such that m_i and m'_i are synonymous given p within (T, P, M). Choose any $s \in S(T, P, M)$. Construct $\hat{s} \in S(T, P, M)$ that agrees with s except in the following: $m_i(x_i, p|\hat{s}) = m'_i$ for any $x_i \in X_i$ such that $m_i(x_i, p|s) = m_i$.

Choose any $\mathbf{x} = (x_1, ..., x_N) \in X$. If $\rho_i(\mathbf{x}|s) \notin p$ or $m_i(x_i, p|s) \neq m_i$ then clearly $\rho_j(\mathbf{x}|\hat{s}) = \rho_j(\mathbf{x}|s)$ for any $j \in \mathscr{N}$. Suppose $\rho_i(\mathbf{x}|s) \in p$ and $m_i(x_i, p|s) = m_i$. Obviously $\rho_j(\mathbf{x}|\hat{s}) = \rho_j(\mathbf{x}|s)$ for any $j \leq i$. Denote $h = \rho_i(\mathbf{x}|s)$. For any $j > i$ we have

$$\begin{aligned} P(\rho_j(\mathbf{x}|\hat{s})) &= P(\rho_j(\mathbf{x}|\hat{s}, i, h, m'_i)) \\ &= P(\rho_j(\mathbf{x}|\hat{s}, i, h, m_i)) \\ &= P(\rho_j(\mathbf{x}|s, i, h, m_i)) \\ &= P(\rho_j(\mathbf{x}|s)) \end{aligned}$$

where the first line is due to $h = \rho_i(\mathbf{x}|\hat{s})$ and $m_i(x_i, p|\hat{s}) = m'_i$, the second line is due to Lemma B.6.1 because m_i and m'_i are synonymous given h within (T, P, M), the third line is due to the fact that $\hat{s}$ and s agree for every player after i, and the fourth line is due to $\rho_i(\mathbf{x}|s) = h$ and $m_i(x_i, p|s) = m_i$. By induction $P(\rho_j(\mathbf{x}|\hat{s})) = P(\rho_j(\mathbf{x}|s))$ for any $j \in \mathscr{N}$. Therefore

$$\alpha_j(\mathbf{x}|\hat{s}) = a_j(x_j, P(\rho_j(\mathbf{x}|\hat{s}))|\hat{s}) = a_j(x_j, P(\rho_j(\mathbf{x}|s))|\hat{s}) = a_j(x_j, P(\rho_j(\mathbf{x}|s))|s)$$

for any $j \in \mathscr{N}$. Hence $\boldsymbol{\alpha}(\cdot|\hat{s}) = \boldsymbol{\alpha}(\cdot|s)$.

By Part 1 of Lemma B.6.2, for each $j \in \mathscr{N}$ and $p' \in P'_j$ there is $\zeta(p') \in P_j$ such that $p' \subset \zeta(p')$. Construct $s' \in S(T', P', M')$ such that for any $j \in \mathscr{N}$, $x_j \in X_j$ and

$p' \in P'_j$,

$$a_j(x_j, p'|s') = a_j(x_j, \zeta(p')|\hat{s}),$$
$$(\text{if } j < N) \quad m_j(x_j, p'|s') = m_j(x_j, \zeta(p')|\hat{s}).$$

s' is properly defined if and only if $m_j(x_j, p'|s') \in M'(p')$ for any $j < N$, $x_j \in X_j$ and $p' \in P'_j$. To verify it, observe that if $\zeta(p') \neq p$ then $m_j(x_j, p'|s') = m_j(x_j, \zeta(p')|\hat{s}) \in M(\zeta(p')) = M'(p')$ by Part 2 of Lemma B.6.2; whereas if $\zeta(p') = p$ then we have $j = i$ and moreover $m_i(x_i, p'|s') = m_i(x_i, p|\hat{s}) \in M(p)\backslash\{m_i\} = M'(p')$ by Part 3 of Lemma B.6.2 because by construction $m_i(x_i, p|\hat{s}) \neq m_i$.

Choose any $\mathbf{x} = (x_1, ..., x_N) \in X$. Clearly $\rho_1(\mathbf{x}|s') = \rho_1(\mathbf{x}|\hat{s})$. Suppose $\rho_j(\mathbf{x}|s') = \rho_j(\mathbf{x}|\hat{s})$ for any $j \leq k$ for some $k \geq 1$. Thus

$$\begin{aligned}\mu_k(\mathbf{x}|s') &= m_k(x_k, P'(\rho_k(\mathbf{x}|s'))|s') = m_k(x_k, \zeta(P'(\rho_k(\mathbf{x}|s')))|\hat{s}) \\ &= m_k(x_k, P(\rho_k(\mathbf{x}|s'))|\hat{s}) = m_k(x_k, P(\rho_k(\mathbf{x}|\hat{s}))|\hat{s}) = \mu_k(\mathbf{x}|\hat{s}).\end{aligned}$$

It follows that

$$\rho_{k+1}(\mathbf{x}|s') = \rho_k(\mathbf{x}|s') \times \mu_k(\mathbf{x}|s') = \rho_k(\mathbf{x}|\hat{s}) \times \mu_k(\mathbf{x}|\hat{s}) = \rho_{k+1}(\mathbf{x}|\hat{s}).$$

Thus $\rho_j(\mathbf{x}|s') = \rho_i(\mathbf{x}|\hat{s})$ for any $j \in \mathscr{N}$, implying

$$\begin{aligned}\alpha_j(\mathbf{x}|s') &= a_j(x_i, P'(\rho_j(\mathbf{x}|s'))|s') = a_j(x_j, \zeta(P'(\rho_j(\mathbf{x}|s')))|\hat{s}) \\ &= a_j(x_j, P(\rho_j(\mathbf{x}|s'))|\hat{s}) = a_j(x_j, P(\rho_j(\mathbf{x}|\hat{s}))|\hat{s}) = \alpha_j(\mathbf{x}|\hat{s}).\end{aligned}$$

It follows that $\boldsymbol{\alpha}(\cdot|s') = \boldsymbol{\alpha}(\cdot|\hat{s})$, implying $\boldsymbol{\alpha}(\cdot|s') = \boldsymbol{\alpha}(\cdot|s)$. Thus $C(T, P, M) \subset C(T', P', M')$ because s is arbitrarily chosen from $S(T, P, M)$. The present lemma then follows from Proposition 3.5.2. □

B.7 Proof of Lemma 3.5.6

Proof. Suppose (T', P', M') is obtained from (T, P, M) by relabeling with relabeling function γ satisfying R1-R4. R3 implies that for any $p \in P'$ there is $\tau(p) \in P$ such that $h \in \tau(p)$ if and only if $\gamma(h) \in p$. Let γ^{-1} be the inverse mapping of γ. Clearly γ^{-1} satisfies R1-R3.

Now show that γ^{-1} satisfies R4. Define $\kappa'(m_i, p)$ where $p \in P'_i$ and $m_i \in M'(p)$ such that $\kappa'(m_i, p) \in M(\tau(p))$ and $\kappa(\kappa'(m_i, p), \tau(p)) = m_i$. First verify that κ' is defined for every $p \in P'_i$ and $m_i \in M'(p)$. Choose any $p \in P'_i$, $m_i \in M'(p)$ and $h \in \tau(p)$. That γ^{-1} satisfies R1 and R3 implies $\gamma^{-1}(\gamma(h) \times m_i) = h \times \hat{m}_i$ for some $\hat{m}_i \in M(\tau(p))$. By definition of κ we have $\kappa(\hat{m}_i, \tau(p)) = m_i$, confirming that indeed $\kappa'(m_i, p)$ exists.

Choose any $p \in P_i'$, $m_i \in M'(p)$ and $h \in p$. $\gamma^{-1}(h) \in \tau(p)$ by R3. By construction $h \times m_i = h \times \kappa(\kappa'(m_i, p), \tau(p))$. Also note that

$$\begin{aligned}\gamma\Big(\gamma^{-1}(h) \times \kappa'(m_i, p)\Big) &= \gamma(\gamma^{-1}(h)) \times \kappa(\kappa'(m_i, p), \tau(p)) \\ &= h \times \kappa(\kappa'(m_i, p), \tau(p)) = h \times m_i.\end{aligned}$$

Applying γ^{-1} to both sides we have $\gamma^{-1}(h) \times \kappa'(m_i, p) = \gamma^{-1}(h \times m_i)$. Thus γ^{-1} satisfies R4. It follows that (T, P, M) is obtained from (T', P', M') by relabeling with relabeling function γ^{-1}. $\square$

B.8 Proof of Lemma 3.5.7

Proof. Suppose (T', P', M') is obtained from (T, P, M) by relabeling with relabeling function γ satisfying R1-R4. R3 implies that for any $p \in P'$ there is $\tau(p) \in P$ such that $h \in \tau(p)$ if and only if $\gamma(h) \in p$.

Choose any $s \in S(T, P, M)$. Construct $s' \in S(T', P', M')$ such that for any $i \in \mathscr{N}$, $x_i \in X_i$ and $p \in P_i'$,

$$\begin{aligned} a_i(x_i, p|s') &= a_i(x_i, \tau(p)|s), \\ (\text{if } i < N) \quad m_i(x_i, p|s') &= \kappa(m_i(x_i, \tau(p)|s), \tau(p)), \end{aligned}$$

where κ is defined in R4. To verify that s' is properly defined, choose any $h \in \tau(p)$ where $p \in P_i'$ for some $i < N$. Thus $\gamma(h) \in p$. R4 and R1 implies

$$\gamma(h \times m_i(x_i, \tau(p)|s)) = \gamma(h) \times \kappa(m_i(x_i, \tau(p)|s), \tau(p)) = \gamma(h) \times m_i(x_i, p|s') \in T'.$$

Since $\gamma(h) \in p$, it follows that $m_i(x_i, p|s') \in M'(\gamma(h))$, implying $m_i(x_i, p|s') \in M'(p)$ because $\gamma(h) \in p$.

Choose any $\mathbf{x} = (x_1, ..., x_N) \in X$. Clearly $\rho_1(\mathbf{x}|s') = \gamma(\rho_1(\mathbf{x}|s))$. Suppose $\rho_i(\mathbf{x}|s') = \gamma(\rho_i(\mathbf{x}|s))$ for any $i \leq k$ for some $k \geq 1$. Denote $h = \rho_k(\mathbf{x}|s)$, $h' = \rho_k(\mathbf{x}|s')$, $p = P(h)$ and $p' = P'(h')$. By the inductive hypothesis $h' = \gamma(h)$ and thus $p = \tau(p')$. We have

$$\begin{aligned}\rho_{k+1}(\mathbf{x}|s') &= h' \times m_k(x_k, p'|s') = \gamma(h) \times \kappa(m_k(x_k, \tau(p')|s), \tau(p')) \\ &= \gamma(h) \times \kappa(m_k(x_k, p|s), p) = \gamma(h \times m_k(x_k, p|s)) = \gamma(\rho_{k+1}(\mathbf{x}|s)).\end{aligned}$$

Thus $\rho_i(\mathbf{x}|s') = \gamma(\rho_i(\mathbf{x}|s))$ for any $i \in \mathscr{N}$, implying $P(\rho_i(\mathbf{x}|s)) = \tau(P'(\rho_i(\mathbf{x}|s')))$. Therefore

$$\begin{aligned}\alpha_i(\mathbf{x}|s') &= a_i(x_i, P'(\rho_i(\mathbf{x}|s'))|s') = a_i(x_i, \tau(P'(\rho_i(\mathbf{x}|s')))|s) \\ &= a_i(x_i, P(\rho_i(\mathbf{x}|s))|s) = \alpha_i(\mathbf{x}|s).\end{aligned}$$

Thus $\boldsymbol{\alpha}(\cdot|s') = \boldsymbol{\alpha}(\cdot|s)$, implying $C(T, P, M) \subset C(T', P', M')$ since s is arbitrarily chosen from $S(T, P, M)$.

Lemma 3.5.6 implies (T, P, M) is obtained from (T', P', M') by relabeling. Thus by an analogous argument as above we have $C(T', P', M') \subset C(T, P, M)$. The present lemma then follows by Proposition 3.5.2.

□

B.9 The Formal Definition of Merging

To introduce merging formally, it is useful to first define a special kind of expanding. Fix institution (T, P, M). Consider the following construction: Pick any $i < N$, $p \in P_i$ and $m_i \in M(p)$. Let T^{m_i} denote the set of all $h \in T$ such that $h(i-1) \in p$ and $h_i = m_i$. Let $T^{m_i'}$ denote the set of all message profiles derived from changing the ith component of some $h \in T^{m_i}$ from m_i to $m_i' \notin M(p)$. Thus for each $h \in T^{m_i'}$ there is $f^h \in T^{m_i}$ such that h and f^h differ only at the ith component. Observe that T and $T^{m_i'}$ are disjoint, for if $h \in T \cap T^{m_i'}$ then $h(i-1) \in p$ and $h_i = m_i'$, implying $m_i' \in M(p)$, a contradiction. Let $T' = T \cup T^{m_i'}$. Let P' be a partition of T' such that

D1 $P'(h) = P'(g)$ if and only if $P(h) = P(g)$ for any $h, g \in T$.

D2 $P'(h) = P'(f^h)$ for any $h \in T^{m_i'}$.

It is plain to see that T' equipped with partition P' is derived from T equipped with P by putting each additional $h \in T^{m_i'}$ in the same partition cell as f^h. Define M' such that

D3 $M'(h) = M(h)$ if $h \in T \backslash p$.

D4 $M'(h) = M(h) \cup \{m_i'\}$ if $h \in p$.

D5 $M'(h) = M(f^h)$ if $h \in T^{m_i'}$.

Observe that (T', P', M') is an improper institution. To verify it, the only non-obvious part is to show that $M'(h) = M'(g)$ for any $h, g \in T'$ such that $P'(h) = P'(g)$. If $h, g \in T$ then $P'(h) = P'(g)$ implies $P(h) = P(g)$ and consequently $M'(h) = M(h) = M(g) = M'(g)$ if $h, g \notin p$, or $M'(h) = M(h) \cup \{m_i'\} = M(g) \cup \{m_i'\} = M'(g)$ if $h, g \in p$. If $h \in T$ and $g \in T^{m_i'}$ then $P'(h) = P'(g)$ implies by D2 that $P'(h) = P'(f^g)$, and it follows from $f^g \in T$ and D5 that $M'(h) = M'(f^g) = M'(g)$. If $h \in T^{m_i'}$ and $g \in T^{m_i'}$ then $P'(h) = P'(g)$ implies $P'(f^h) = P'(f^g)$ and it follows by D5 that $M'(h) = M'(f^h) = M'(f^g) = M'(g)$. Thus (T', P', M') is indeed an improper institution. We say that (T', P', M') is obtained from (T, P, M) by $(i, p, m_i, m_i')-$**duplication**.

The following lemma suggests that $(i, p, m_i, m_i')-$**duplication** does no more than expanding (T, P, M) by duplicating message m_i for p and giving the cloned message the label m_i'.

Lemma B.9.1. *If (T', P', M') is obtained from (T, P, M) by $(i, p, m_i, m_i')-$duplication then:*

1. *$T_j' = T_j$ and $P_j' = P_j$ for any $j \le i$.*
2. *m_i and m_i' are synonymous given p within (T', P', M').*
3. *(T, P, M) can be obtained from (T', P', M') by trimming.*
4. *(T, P, M) and (T', P', M') dominate each other.*

Proof. Inherit the notation $T^{m_i'}$ and f^h from the introduction of duplication.

$T_j = T_j'$ for any $j \le i$ because $T^{m_i'}$ only contains message profiles of length at least i. Thus D1 implies $P_j' = P_j$, establishing Part 1.

Now show Part 2. Choose any $h \in p$. D4 implies $m_i, m_i' \in M'(h)$. Part 1 implies $P'(h) = p$. Thus $m_i, m_i' \in M'(p)$. Let $g, g' \in T'$ be $h-$cousins such that $g_i = m_i$ and $g_i' = m_i'$. Clearly $g = f^{g'}$. It follows from D2 that $P'(g) = P'(g')$, implying m_i and m_i' are synonymous given h within (T', P', M'). Part 2 follows immediately because h is arbitrarily chosen from p.

Now show Part 3. Clearly (T, P, M) is a sub-institution of (T', P', M'). If $h \in T' \backslash T$ then $h \in T^{m_i'}$ and it follows by the construction of $T^{m_i'}$ that $h(i-1) \in p$ and $h_i = m_i'$, establishing T1. T2 follows from Part 2.

Part 4 follows from Part 3 by Lemma 3.5.5. □

Now we introduce merging. The definition is given in terms of the construction. Each step is explained by the accompanying remark.

Definition. (T', P', M') is obtained from (T, P, M) by **merging** if there is $i < N$ where $|A_i| = 1$, $p \in P_i$ and $q \in P_i$ (without loss of generality assume $|M(p)| < |M(q)|$) such that (T', P, M') is the product of the following algorithm:

Step 1 Fix $m_i \in M(p)$. Set $(T^0, P^0, M^0) = (T, P, M)$. Produce (T^k, P^k, M^k) from $(T^{k-1}, P^{k-1}, M^{k-1})$ by $(i, p, m_i, m_i^k)-$duplication [1] where m_i^k is an arbitrary message not in $M^{k-1}(p)$. Stop if $|M^k(p)| = |M^k(q)| = |M(q)|$. [2] Denote the terminal product of this step as $(\tilde{T}, \tilde{P}, \tilde{M})$.

[1]Part 1 of Lemma B.9.1 implies $p \in P^{k-1}$.

[2]Part 1 of Lemma B.9.1 implies $q \in P^{k-1}$. D3 implies $M^k(q) = M(q)$.

Remark: This step equalizes the number of messages available to player i given p and that given q.

Step 2 Choose any bijections $\lambda : \tilde{M}(p) \to \tilde{M}(q)$. Construct mapping γ on $\hat{T}$ such that

(a) If g is not a descendant of some $h \in p$ then $\gamma(h) = h$.

(b) If g is a descendant of some $h \in p$ then $\gamma(g) = e^g$ where e^g is derived from g by replacing the ith component of g from g_i to $\lambda(g_i)$.

Denote the end product of this step as $(\hat{T}, \hat{P}, \hat{M})$.

Remark: It is easy to verify that γ is a relabeling function satisfying R1-R4. This step relabels the messages available to player i given p, so that now $\hat{M}(p) = \hat{M}(q)$.

Step 3 Arbitrarily index messages in $\hat{M}(p)$ as $m^1, ..., m^K$ where $K = |\hat{M}(p)|$. For each $j \neq k$, abritrarily choose a unique message $n^{j,k}$ where $n^{j,k} \notin \hat{M}(p)$. Let $n^{j,j} = m^j$.

Set $(\hat{T}^{1,1}, \hat{P}^{1,1}, \hat{M}^{1,1}) = (\hat{T}, \hat{P}, \hat{M})$. Iterate through $k = 1 : K$ as follows: For each j and $k \neq j-1$ produce $(T^{j,k+1}, P^{j,k+1}, M^{j,k+1})$ from $(T^{j,k}, P^{j,k}, M^{j,k})$ by the following two substeps:

(a) Produce $(T_p^{j,k+1}, P_p^{j,k+1}, M_p^{j,k+1})$ from $(T^{j,k}, P^{j,k}, M^{j,k})$ by $(i, p, m^j, n^{j,k})$-duplication.

(b) Produce $(T^{j,k+1}, P^{j,k+1}, M^{j,k+1})$ from $(T_p^{j,k+1}, P_p^{j,k+1}, M_p^{j,k+1})$ by $(i, q, m^k, n^{j,k+1})$-duplication.

In the case that $k = j-1$, set $(T^{j,j}, P^{j,j}, M^{j,j}) = (T^{j,j-1}, P^{j,j-1}, M^{j,j-1})$.

When $(T^{j,K}, P^{j,K}, M^{j,K})$ is reached, produce $(T^{j+1,1}, P^{j+1,1}, M^{j+1,1})$ from $(T^{j,K}, P^{j,K}, M^{j,K})$ by the following two substeps:

(a) Produce $(T_p^{j+1,1}, P_p^{j+1,1}, M_p^{j+1,1})$ from $(T^{j,K}, P^{j,K}, M^{j,K})$ by $(i, p, m^{j+1}, n^{j+1,1})$-duplication.

(b) Produce $(T^{j+1,1}, P^{j+1,1}, M^{j+1,1})$ from $(T_p^{j+1,1}, P_p^{j+1,1}, M_p^{j+1,1})$ by $(i, q, m^1, n^{j+1,1})$-duplication.

Stop when $(T^{K,K}, P^{K,K}, M^{K,K})$ is reached. Denote $(T^{K,K}, P^{K,K}, M^{K,K})$ as (T^*, P^*, M^*).

Remark: It is straightforward to verify that $M^*(p) = M^*(q) = \{n^{j,k} : (j,k) \in \{1, ..., K\}^2\}$. This step expands $(\hat{T}, \hat{P}, \hat{M})$ in a particular way, so that for each pair j, k where $j \neq k$, the message $n^{j,k}$ is introduced to p and q, where it is synonymous to m^j given p, and to m^k given q.

Step 4 Produce (T', P', M') from (T^*, P^*, M^*) such that $T' = T^*$ and P^* satisfies

(a) $P'(h) = P'(g)$ if and only if $P^*(h) = P^*(g)$ for any $h, g \notin p \cup q$.

(b) $P'(h) = P'(g)$ for any $h, g \in p \cup q$.

Remark: This step combines separate perceptions p and q into one single perception $p \cup q$.

B.10 Proof of Lemma 3.5.8

Proof. Inherit the notation introduced the formal definition of merging in Section B.9. (T^*, P^*, M^*) is obtained from (T, P, M) by a sequence of duplication and relabeling operations. Thus (T^*, P^*, M^*) and (T, P, M) dominate each other by Lemmas 3.5.7 and B.9.1. Clearly (T^*, P^*, M^*) can be obtained from (T', P', M') by refining. Thus (T^*, M^*, P^*) dominates (T', P', M') by Lemma 3.5.4 and hence (T, P, M) dominates (T', P', M').

To establish that (T', P', M') dominates (T, P, M) it suffices to show that (T', P', M') dominates (T^*, M^*, P^*). It is straightforward to verify that $M^*(p) = M^*(q) = \{n^{j,k} : (j, k) \in \{1, ..., K\}^2\}$. Recall that $n^{j,j} = m^j$. Choose any $n^{j,k}$ where $j \neq k$. Since $(T_p^{j,k}, P_p^{j,k}, M_p^{j,k})$ is obtained from $(T^{j,k-1}, P^{j,k-1}, M^{j,k-1})$ (or $(T^{j-1,K}, P^{j-1,K}, M^{j-1,K})$ in the case $k = 1$) by $(i, p, m^j, n^{j,k})$-duplication, it follows by Lemma B.9.1 that m^j and $n^{j,k}$ are synonymous given p within $(T_p^{j,k}, P_p^{j,k}, M_p^{j,k})$. Choose any $h \in p$ and $h-$cousins $g, g' \in T^*$ such that $g_i = m^j$ and $g'_i = n^{j,k}$. Clearly $g, g' \in (T^{j,k}, P^{j,k}, M^{j,k})$. It follows that $P^{j,k}(g) = P^{j,k}(g')$ by synonymity. Since $(T_p^{j,k}, P_p^{j,k}, M_p^{j,k})$ is a sub-institution of (T^*, M^*, P^*), we have $P^*(g) = P^*(g')$. Moreover since for any $l > i$ we have $P'_l = P^*_l$, it follows that $P'(g) = P'(g')$. Therefore m^j and $n^{j,k}$ are synonymous given any $h \in p$ within (T^*, P^*, M^*) and within (T', P', M'). Similarly m^k and $n^{j,k}$ are synonymous given any $h \in q$ within (T^*, P^*, M^*) and within (T', P', M').

Choose any $s \in S(T^*, P^*, M^*)$. Construct $\hat{s}$ such that $\hat{s}$ agrees with s except in the following:

$$m_i(x_i, p|\hat{s}) = m^j \text{ for any } x_i \text{ such that } m_i(x_i, p|s) = n^{j,k},$$

$$m_i(x_i, q|\hat{s}) = m^k \text{ for any } x_i \text{ such that } m_i(x_i, q|s) = n^{j,k}.$$

Clearly $\rho_l(\mathbf{x}|\hat{s}) = \rho_l(\mathbf{x}|s)$ for any $l \in \mathscr{N}$ if $\rho_i(\mathbf{x}|s) \notin p \cup q$. Fix $\mathbf{x} = (x_1, ..., x_N) \in X$ such that $\rho_i(\mathbf{x}|s) \in p$. $P^*(\rho_l(\mathbf{x}|\hat{s})) = P^*(\rho_l(\mathbf{x}|s))$ for any $l \leq i$ because $\rho_l(\mathbf{x}|\hat{s}) = \rho_l(\mathbf{x}|s))$. Denote $h = \rho_i(\mathbf{x}|\hat{s})$. Note that $m_i(x_i, p|\hat{s}) = m^j$ and $m_i(x_i, p|s) = n^{j,k}$ for some $(j,k) \in \{1,...,K\}^2$. For any $l > i$ observe that

$$\begin{aligned} P^*(\rho_l(\mathbf{x}|\hat{s})) &= P^*(\rho_l(\mathbf{x}|\hat{s}, i, h, m^j)) \\ &= P^*(\rho_l(\mathbf{x}|\hat{s}, i, h, n^{j,k})) \\ &= P^*(\rho_l(\mathbf{x}|s, i, h, n^{j,k})) \\ &= P^*(\rho_l(\mathbf{x}|s)). \end{aligned}$$

The first line is due to $h = \rho_i(\mathbf{x}|\hat{s})$ and $m_i(x_i, p|\hat{s}) = m^j$. The second line is due to Lemma B.6.1 because m^j and $n^{j,k}$ are synonymous given h within (T^*, P^*, M^*) if $j \neq k$, or $m^j = n^{j,k}$ if $j = k$. The third line is due to the fact that $\hat{s}$ and s agree for every player after i, and the fourth line is due to $\rho_i(\mathbf{x}|s) = h$ and $m_i(x_i, p|s) = n^{j,k}$. Similarly $P^*(\rho_l(\mathbf{x}|\hat{s})) = P^*(\rho_l(\mathbf{x}|s))$ for any $\mathbf{x}$ such that $\rho_i(\mathbf{x}|s) \in q$. By induction $P^*(\rho_l(\mathbf{x}|\hat{s})) = P^*(\rho_l(\mathbf{x}|s))$ for any $l \in \mathscr{N}$ and $\mathbf{x} \in X$.

Note that $P' = (P^* \backslash \{p, q\}) \cup \{p \cup q\}$. Choose $s' \in S(T', P', M')$ such that

- For any $l \neq i$, $x_l \in X_l$ and $r \in P'_l$:

$$a_l(x_l, r|s') = a_l(x_l, r|\hat{s}).$$

- For any $l < N$, $x_l \in X_l$ and $r \in P'_l$ where $r \neq p \cup q$:

$$m_l(x_l, r|s') = m_l(x_l, r|\hat{s}).$$

- For any $x_i \in X_i$ such that $m_i(x_i, p|\hat{s}) = m^j$ and $m_i(x_i, p|\hat{s}) = m^k$:

$$m_i(x_i, p \cup q|s') = n^{j,k}.$$

We have $\rho_l(\mathbf{x}|s') = \rho_l(\mathbf{x}|\hat{s})$ for any $\mathbf{x}$ such that $\rho_i(\mathbf{x}|s') \notin p \cup q$. Fix $\mathbf{x} = (x_1, ..., x_N)$ such that $\rho_i(\mathbf{x}|s') \in p$. Obviously $\rho_i(\mathbf{x}|s') = \rho_i(\mathbf{x}|\hat{s})$. Denote $h = \rho_i(\mathbf{x}|s')$. By construction of $\hat{s}$ we have $m_i(x_i, p|\hat{s}) = m^j$ and $m_i(x_i, q|\hat{s}) = m^k$ for some $(j,k) \in \{1,...,K\}^2$. Thus for any $l > i$,

$$\begin{aligned} P'(\rho_l(\mathbf{x}|s')) &= P'(\rho_l(\mathbf{x}|s', i, h, n^{j,k})) \\ &= P'(\rho_l(\mathbf{x}|s', i, h, m^j)) \\ &= P^*(\rho_l(\mathbf{x}|\hat{s}, i, h, m^j)) \\ &= P^*(\rho_l(\mathbf{x}|\hat{s})). \end{aligned}$$

Similarly for any $\mathbf{x}$ such that $\rho_i(\mathbf{x}|s') \in q$ we have $P'(\rho_l(\mathbf{x}|s')) = P^*(\rho_l(\mathbf{x}|\hat{s}))$ for any $l > i$. Thus $P'(\rho_l(\mathbf{x}|s')) = P^*(\rho_l(\mathbf{x}|\hat{s}))$ for any $l \neq i$ and $\mathbf{x} \in X$ by induction. For any $l \in \mathscr{N}$ where $|A_l| > 1$ (and hence $l \neq i$ because $|A_i| = 1$) and $\mathbf{x} = (x_1, ..., x_N) \in X$

we have

$$\alpha_l(\mathbf{x}|s') = a_l(x_l, P'(\rho_l(\mathbf{x}|s'))|s') = a_l(x_l, P^*(\rho_l(\mathbf{x}|\hat{s}))|s') = a_l(x_l, P^*(\rho_l(\mathbf{x}|s))|s')$$
$$= a_l(x_l, P^*(\rho_l(\mathbf{x}|s))|\hat{s}) = \alpha_l(x_l, P^*(\rho_l(\mathbf{x}|s))|s) = \alpha_l(\mathbf{x}|s).$$

It follows that $\boldsymbol{\alpha}(\cdot|s') = \boldsymbol{\alpha}(\cdot|s)$, implying $C(T^*, P^*, M^*) \subset C(T', P', M')$ because s is arbitrarily chosen from $S(T^*, P^*, M^*)$. Thus (T', P', M') dominates (T^*, P^*, M^*) by Proposition 3.5.2, establishing the present lemma. □

B.11 Proof of Theorem 3.5.9

Proof. Part 1 is an immediate consequence of Lemmas 3.5.3, 3.5.4, 3.5.5, 3.5.7 and 3.5.8.

Now we prove Part 2, the "only if" direction. Suppose (T', P', M') dominates (T, P, M). The proofs is broken into several steps.

(Step 1. "Merging".) It is plain to see that an institution $(\hat{T}, \hat{P}, \hat{M})$, where $|\hat{P}_i| = 1$ for every player i whose action set is a singleton, can be obtained from (T, P, M) by a sequence of merging operations. It follows from Lemma 3.5.8 that (T, P, M) dominates $(\hat{T}, \hat{P}, \hat{M})$. Thus (T', P', M') dominates $(\hat{T}, \hat{P}, \hat{M})$.

(Step 2. "Refining".) For any $i < N$ define $t_i = \max\{\max_{p \in \hat{P}_i} \hat{M}(p), \frac{\log |\hat{P}_i|}{\log |A_i|}\}$. Suppose $|X_i| \geq t_i$. Since $|X_i| \geq \max_{p \in \hat{P}_i} \hat{M}(p)$, there is a mapping $y_i : \hat{P}_i \times \cup_{p \in \hat{P}_i} \hat{M}(p) \to X_i$ such that $y_i(p, m_i) \neq y_i(p, m_i')$ for any $m_i, m_i' \in \hat{M}(p)$ where $m_i \neq m_i'$. If $|A_i| > 1$ then for each $p \in \hat{P}_i$ there is a mapping $s_p : X_i \to A_i$ such that $s_p \neq s_{p'}$ if $p \neq p'$. To see that, first observe that the total number of mappings from X_i to A_i is $|A_i|^{|X_i|}$. By assumption $|X_i| \geq \frac{\log |\hat{P}_i|}{\log |A_i|}$ or equivalently $|A_i|^{|X_i|} \geq |\hat{P}_i|$. Thus each $p \in \hat{P}_i$ can be assigned with a unique s_p.

Note that for each $p, p' \in \hat{P}_i$ where $p \neq p'$ there is $z_i(p, p') \in X_i$ such that $s_p(z_i(p, p')) \neq s_{p'}(z_i(p, p'))$. Choose any $s \in S(\hat{T}, \hat{P}, \hat{M})$ such that

- For any $i \in \mathscr{N}$ where $|A_i| > 1$, $x_i \in X_i$ and $p \in \hat{P}_i$:

$$a_i(x_i, p|s) = s_p(x_i).$$

- For any $i < N$, $p \in \hat{P}_i$ and $m_i \in \hat{M}(p)$:

$$m_i(y_i(p, m_i), p|s) = m_i.$$

By Proposition 3.5.2 there is $s' \in S(T', P', M')$ such that $\boldsymbol{\alpha}(\cdot|s') = \boldsymbol{\alpha}(\cdot|s)$. For each $h \in \hat{T}$ construct $\mathbf{x}^h = (x_1^h, ..., x_N^h) \in X$ such that $x_i^h = y_i(\hat{P}(h(i-1)), h_i)$ for any $i \le |h|$. It is straightforward to verify that $h = \rho_{|h|+1}(\mathbf{x}^h|s)$. Define the following objects:

- $\gamma(h) = \rho_{|h|+1}(\mathbf{x}^h|s')$ for any $h \in \hat{T}$.
- $\tau(p) = \{h \in \hat{P}_i : \gamma(h) \in p\}$ for any $p \in P'_i$.
- $\tilde{P}_i = \{\tau(p) : p \in P'_i \text{ and } \tau(p) \neq \emptyset\}$ for any $i \in \mathscr{N}$.

$\tilde{P}_i$ is a partition of $\hat{T}_i$ because: (1) for any $h \in \hat{T}_i$, $\gamma(h) \in p$ for some $p \in P'_i$, and (2) if $h \in \tau(p)$ and $h \in \tau(p')$ then $p = P'(\gamma(h)) = p'$.

Note that for any i where $|A_i| = 1$, we have $|\hat{P}_i| = 1$ by construction and hence $\tilde{P}_i$ is a refinement of $\hat{P}_i$. Fix any i where $|A_i| > 1$. Suppose there exist $h, g \in \hat{T}_i$ such that $\hat{P}(h) \neq \hat{P}(g)$ yet $\tilde{P}(h) = \tilde{P}(g)$. Denote $h' = \gamma(h)$ and $g' = \gamma(g)$. $\tilde{P}(h) = \tilde{P}(g)$ implies $P'(h') = P'(g')$. Choose $\mathbf{x} = (x_1, ..., x_N) \in X$ such that $x_j = x_j^h$ for any $j < i$ and $x_i = z_i(\hat{P}(h), \hat{P}(g))$. Choose $\hat{\mathbf{x}} = (\hat{x}_1, ..., \hat{x}_N) \in X$ such that $\hat{x}_j = x_j^g$ for any $j < i$ and $\hat{x}_i = z_i(\hat{P}(h), \hat{P}(g))$. Obviously $\rho_i(\mathbf{x}|s) = \rho_i(\mathbf{x}^h|s) = h$. Similarly $\rho_i(\hat{\mathbf{x}}|s) = g$. Therefore we have

$$\begin{aligned}\alpha_i(\mathbf{x}|s) &= a_i(z_i(\hat{P}(h), \hat{P}(g)), \hat{P}(h)|s) = s_{\hat{P}(h)}(z_i(\hat{P}(h), \hat{P}(g))) \\ &\neq s_{\hat{P}(g)}(z_i(\hat{P}(h), \hat{P}(g))) = a_i(z_i(\hat{P}(h), \hat{P}(g)), \hat{P}(g)|s) = \alpha_i(\hat{\mathbf{x}}|s).\end{aligned}$$

It follows from $h' = \rho_i(\mathbf{x}^h|s')$ that $h' = \rho_i(\mathbf{x}|s')$ because $\mathbf{x}^h$ and $\mathbf{x}$ agree for the first $i-1$ components. Similarly $g' = \rho_i(\hat{\mathbf{x}}|s')$. Thus

$$\alpha_i(\mathbf{x}|s') = a_i(z_i(\hat{P}(h), \hat{P}(g)), P'(h')|s') = a_i(z_i(\hat{P}(h), \hat{P}(g)), P'(g')|s') = \alpha_i(\hat{\mathbf{x}}|s')$$

because $P'(h') = P'(g')$. It follows that

$$\alpha_i(\mathbf{x}|s') = \alpha_i(\mathbf{x}|s) \neq \alpha_i(\hat{\mathbf{x}}|s) = \alpha_i(\hat{\mathbf{x}}|s') = \alpha_i(\mathbf{x}|s'),$$

a contradiction. Thus $\hat{P}(h) \neq \hat{P}(g)$ implies $\tilde{P}(h) \neq \tilde{P}(g)$, in turn implying $\tilde{P}_i$ is a refinement of $\hat{P}_i$. Let $\tilde{P} = \cup_{i \in \mathscr{N}} \tilde{P}_i$. It then follows that $(\hat{T}, \tilde{P}, \hat{M})$ is obtained from $(\hat{T}, \hat{P}, \hat{M})$ by refining.

(Step 3. "Trimming".) Apply a sequence of trimming operations to $(\hat{T}, \tilde{P}, \hat{M})$ to obtain (T^*, P^*, M^*) such that for any $i < N$ and $p \in P^*_i$ there do not exist $m_i, m'_i \in M^*(p)$ that are synonymous given p within (T^*, P^*, M^*).

(Step 4. "Relabeling".) Clearly (T^*, P^*, M^*) is a sub-institution of $(\hat{T}, \tilde{P}, \hat{M})$. Therefore $P^*(h) = P^*(g)$ if and only if $\tilde{P}(h) = \tilde{P}(g)$ for any $h, g \in T^*$. Let T'' denote the range of γ with domain restricted to T^*. Let P'' denote P' restricted to T'', that is, $P''(h) = P''(g)$ if and only if $P'(h) = P'(g)$ for any $h, g \in T''$. For any $i < N$ and $h \in T_i''$ define $M''(h) = \{m_i : h \times m_i \in T''\}$. We want to show that (T'', P'', M'') is obtained from (T^*, P^*, M^*) by relabeling with relabeling function γ. This will be achieved below by establishing that γ with domain restricted to T^* satisfies R1-R4.[3]

Observe that for any $h, g \in T^*$, if h is the parent of g then the first $|h|$ components of $\mathbf{x}^h$ and $\mathbf{x}^g$ are the same. It follows that $\rho_{|h|+1}(\mathbf{x}^g|s') = \rho_{|h|+1}(\mathbf{x}^h|s')$, implying $\gamma(h)$ is the parent of $\gamma(g)$. Thus γ satisfies R2.

Observe that for any $h, g \in T^*$, $P^*(h) = P^*(g)$ if and only if $\tilde{P}(h) = \tilde{P}(g)$ if and only if $P'(h) = P'(g)$ if and only if $P''(h) = P''(g)$. Thus γ satisfies R3.

To establish that γ is a bijection between T^* and T'' it is sufficient to verify that γ restricted to T^* is one-to-one. Suppose γ is not one-to-one, then there are $g, g' \in T^*$ such that $\gamma(g) = \gamma(g')$. Let i denote the largest index such that $g(i-1) = g'(i-1)$. Denote $f = g(i-1)$, $h = g(i)$, $h' = g'(i)$, $m_i = g_i$ and $m_i' = g_i'$. Thus $h = f \times m_i$ and $h' = f \times m_i'$. Since f is the parent of h and h' it follows that $\gamma(f)$ is the parent of $\gamma(h)$ and $\gamma(h')$. That $\gamma(g) = \gamma(g')$ then implies $\gamma(h) = \gamma(h')$. Denote $p = \hat{P}(f)$ and $p' = P'(\gamma(f))$. By construction of γ, $\gamma(h) = \gamma(h')$ implies $\rho_i(\mathbf{x}^h|s') = \rho_i(\mathbf{x}^{h'}|s')$ and $\mu_i(\mathbf{x}^h|s') = \mu_i(\mathbf{x}^{h'}|s')$. Since $\rho_i(\mathbf{x}^h|s') = \rho_i(\mathbf{x}^{h'}|s') = \gamma(f)$, it follows that

$$m_i(x_i^h, p'|s') = \mu_i(\mathbf{x}^h|s') = \mu_i(\mathbf{x}^{h'}|s') = m_i(x_i^{h'}, p'|s').$$

Thus $m_i(y_i(p, m_i), p'|s') = m_i(y_i(p, m_i'), p'|s')$ because by construction we have $x_i^h = y_i(p, m_i)$ and $x_i^{h'} = y_i(p, m_i')$. Choose any $l \in P^*(f)$ and $l-$cousins $u, u' \in T^*$ where $u_i = m_i$ and $u_i' = m_i'$. $P^*(f) = P^*(l)$ implies $\tilde{P}(f) = \tilde{P}(l) = p$. Thus $x_i^u = y_i(p, m_i)$ and $x_i^{u'} = y_i(p, m_i')$. Also $P^*(f) = P^*(l)$ implies $P'(\gamma(f)) = P'(\gamma(l)) = p'$. Since γ satisfies R2, l being the ancestor of u and u' implies $\gamma(l)$ is the ancestor of $\gamma(u)$ and $\gamma(u')$, then implying $\gamma(l) = \rho_i(\mathbf{x}^u|s') = \rho_i(\mathbf{x}^{u'}|s')$. Moreover we have

$$\begin{aligned}
\rho_{i+1}(\mathbf{x}^u|s') &= \rho_i(\mathbf{x}^u|s') \times m_i(x_i^u, P'(\rho_i(\mathbf{x}^u|s'))|s') \\
&= \gamma(l) \times m_i(y_i(p, m_i), p'|s') = \gamma(l) \times m_i(y_i(p, m_i'), p'|s') \\
&= \rho_i(\mathbf{x}^{u'}|s') \times m_i(x_i^{u'}, P'(\rho_i(\mathbf{x}^{u'}|s'))|s') = \rho_{i+1}(\mathbf{x}^{u'}|s')
\end{aligned}$$

where $m_i(y_i(p, m_i), p'|s') = m_i(y_i(p, m_i'), p'|s')$ has been established above. Suppose $\rho_j(\mathbf{x}^u|s') = \rho_j(\mathbf{x}^{u'}|s')$ for any $j \le k$ for some $k \ge i+1$. The inductive hypothesis implies $P'(\rho_k(\mathbf{x}^u|s')) = P'(\rho_k(\mathbf{x}^{u'}|s'))$. Since $\mathbf{x}^{u(k-1)}$ agrees with $\mathbf{x}^u$ for the first $k-1$ components, it follows that $\gamma(u(k-1)) = \rho_k(\mathbf{x}^{u(k-1)}|s') = \rho_k(\mathbf{x}^u|s')$. Similary $\gamma(u'(k-1)) = \rho_k(\mathbf{x}^{u'}|s')$. Thus $P'(\rho_k(\mathbf{x}^u|s')) = P'(\rho_k(\mathbf{x}^{u'}|s'))$ implies $P'(\gamma(u(k-1))) = P'(\gamma(u'(k-1)))$. It follows that $\tilde{P}(u(k-1)) = \tilde{P}(u'(k-1))$ by construction of $\tilde{P}$, which in turn implies $\hat{P}(u(k-1)) = \hat{P}(u'(k-1))$ because $\tilde{P}_k$ is a refinement of $\hat{P}_k$. Since u and u' are cousins, we have $u_k = u_k'$. Thus $y_k(\hat{P}(u(k-1)), u_k) =$

[3] (T'', M'', P'') will be shown to be an improper institution shortly.

$y_k(\hat{P}(u'(k-1)), u'_k)$. Hence

$$\begin{aligned}\rho_{k+1}(\mathbf{x}^u|s') &= \rho_k(\mathbf{x}^u|s') \times m_k\Big(y_k(\hat{P}(u(k-1), u_k)), P'(\rho_k(\mathbf{x}^u))|s'\Big) \\ &= \rho_k(\mathbf{x}^{u'}|s') \times m_k\Big(y_k(\hat{P}(u'(k-1), u'_k)), P'(\rho_k(\mathbf{x}^{u'}))|s'\Big) = \rho_{k+1}(\mathbf{x}^{u'}|s')\end{aligned}$$

where the first line is due to $x^u_k = y_k(\hat{P}(u(k-1)), u_k)$. Therefore $\rho_{|u|+1}(\mathbf{x}^u|s') = \rho_{|u|+1}(\mathbf{x}^{u'}|s')$ by induction, or equivalently $\gamma(u) = \gamma(u')$. It follows from $P'(\gamma(u)) = P'(\gamma(u'))$ that $\tilde{P}(u) = \tilde{P}(u')$, in turn implying $P^*(u) = P^*(u')$. Hence m_i and m'_i are synonymous given l within (T^*, P^*, M^*). Since l is arbitrarily chosen from $P^*(f)$, it follows that m_i and m'_i are synonymous given $P^*(f)$ within (T^*, P^*, M^*), a contradiction, because by construction (T^*, P^*, M^*) admits no synonymous messages given any perception in P^*. Therefore γ restricted to T^* is one-to-one, establishing R1.

Now show that γ satisfies R4. Choose any $i < N$, $p \in P^*_i$ and $m_i \in M^*(p)$. There exists $\hat{p} \in \hat{P}$ such that $p \subset \hat{p}$. There also exists $p' \in P'$ such that $\gamma(h) \in p'$ for any $h \in p$. For any $h \in p$ we have

$$\begin{aligned}\gamma(h \times m_i) &= \rho_{i+1}(\mathbf{x}^{h\times m_i}|s') \\ &= \rho_i(\mathbf{x}^{h\times m_i}|s') \times m_i\Big(y_i(\hat{p}, m_i), P'(\rho_i(\mathbf{x}^{h\times m_i}|s'))|s'\Big) \\ &= \gamma(h) \times m_i(y_i(\hat{p}, m_i), p'|s')\end{aligned}$$

where the second line is due to $\mathbf{x}^{h\times m_i}_i = y_i(\hat{p}, m_i)$ since $\hat{P}(h) = \hat{p}$. Note that $m_i(y_i(\hat{p}, m_i), p'|s')$ does not depend on the choice of h, thus implying R4.

Now we show that (T'', P'', M'') is indeed an improper institution. The only non-obvious part is that $M''(h) = M''(g)$ if $P''(h) = P''(g)$. Suppose $P''(h) = P''(g)$ yet $M''(h) \neq M''(g)$ for some $h, g \in T''$. Without loss of generality suppose for some $i < N$ there is $m_i \in M''(h)$ such that $m_i \notin M''(g)$. R2 implies $\gamma^{-1}(h \times m_i) = \gamma^{-1}(h) \times \hat{m}_i$ for some $\hat{m}_i \in M^*(\gamma^{-1}(h))$. R3 implies $P^*(\gamma^{-1}(h)) = P^*(\gamma^{-1}(g))$. Thus $\hat{m}_i \in M^*(\gamma^{-1}(g))$. R4 then implies $\gamma(\gamma^{-1}(g) \times \hat{m}_i) = \gamma(\gamma^{-1}(g)) \times m_i = g \times m_i$ since $P^*(\gamma^{-1}(h)) = P^*(\gamma^{-1}(g))$, contradicting the supposition that $m_i \notin M''(g)$. Therefore (T'', P'', M'') is indeed an improper institution and is obtained from (T^*, P^*, M^*) by relabeling with relabeling function γ.

(Step 5. Expanding) Since (T'', P'', M'') is a sub-institution of (T', P', M'), (T', P', M') is obtained from (T'', P'', M'') by expanding. □

B.12 Proof of Proposition 3.6.3

The proof is assisted with the following lemma.

Lemma B.12.1. *For institutions* (T, P, M) *and* (T, P', M) *respectively representing voting systems* (r, d, t) *and* (r', d, t')*, if* P'_i *is a refinement of* P_i *for every* $i \le |J|$ *and* $P'_{|J|+1} = P_{|J|+1}$ *then* (r', d, t') *dominates* (r, d, t) *for any rule* d*.*

Proof: Suppose P'_i is a refinement of P_i for every $i \le |J|$, and $P'_{|J|+1} = P_{|J|+1}$. For every $i \le |J|$ and $p \in P'_i$ there is $\tau(p) \in P_i$ such that $p \subset \tau(p)$. Choose any $s \in S(T, P, M|d)$. Construct $s' \in S(T, P', M|d)$ such that $m_i(x_i, p|s') = m_i(x_i, \tau(p)|s)$ for every $i \le |J|$, $x_i \in X_i$ and $p \in P'_i$. Clearly $\rho_1(\mathbf{x}|s') = \rho_1(\mathbf{x}|s)$ for any $\mathbf{x} = (x_1, ..., x_{|J|}, \overline{x}) \in X$. Suppose $\rho_i(\mathbf{x}|s') = \rho_i(\mathbf{x}|s)$ for every $i \le k$ for some $k \ge 1$. Thus $\tau(P'(\rho_k(\mathbf{x}|s'))) = P(\rho_k(\mathbf{x}|s')) = P(\rho_k(\mathbf{x}|s))$. Therefore

$$\begin{aligned}\rho_{k+1}(\mathbf{x}|s') &= \rho_k(\mathbf{x}|s') \times m_k(x_k, P'(\rho_k(\mathbf{x}|s'))|s') \\ &= \rho_k(\mathbf{x}|s) \times m_k(x_k, P(\rho_k(\mathbf{x}|s))|s) = \rho_{k+1}(\mathbf{x}|s).\end{aligned}$$

Hence $\alpha_{|J|+1}(\mathbf{x}|s') = d(\rho_{|J|+1}(\mathbf{x}|s')) = d(\rho_{|J|+1}(\mathbf{x}|s)) = \alpha_{|J|+1}(\mathbf{x}|s)$. It follows that $\boldsymbol{\alpha}(\cdot|s') = \boldsymbol{\alpha}(\cdot|s)$, implying $C(T, P, M|d) \subset C(T, P', M|d)$. The lemma follows from Lemma 3.6.2. □

Proof of Proposition 3.6.3. (Proof of Part 1.) Let (T, P, M) be the institution representing (r, d, t). Let t' be the full disclosure policy. Let (T, P', M) be the institution representing (r, d, t'). Therefore P'_i is a refinement of P_i for every $i \le |J|$, and $P'_{|J|+1} = P_{|J|+1}$. Thus (r, d, t') dominates (r, d, t) by Lemma B.12.1. Part 1 is established.

(Proof of Part 2.) Suppose t is the full disclosure policy, let r' be the sequential procedure such that $r'(i) > r'(j)$ is $r(i) > r(j)$. Using the same argument as in the previous paragraph we conclude that (r', d, t) dominates (r, d, t), establishing Part 2.

(Proof of Part 3.) Suppose t is the full disclosure policy. Let i denote the player with the largest index among those who vote before the last stage (according to r). It follows that each $p \in P_{|J|}$ can be uniquely identified as $p^{(z_1,...,z_i)}$ such that $P(h) = P(h') = p$ if and only if $h(i) = h'(i) = (z_1, ..., z_i)$.

Let $\hat{Z}$ denote the set of vote profiles $(z_1, ..., z_i)$ from voters voting before the last stage such that $d(z_1, ..., z_i, z_{i+1}, ..., z_{|J|}) = d(z_1, ..., z_i, z'_{i+1}, ..., z'_{|J|})$ for any vote profiles $(z_{i+1}, ..., z_{|J|})$ and $(z'_{i+1}, ..., z'_{|J|})$ from voters voting in the last stage. The collective choice is not determined before voting in the last stage takes place if and only if $\hat{Z}$ is empty. It follows that for any $(z_1, ..., z_i) \in \hat{Z}$, every pair of votes (messages) $z, z' \in Z$ are synonymous given $p^{(z_1,...,z_i)}$ within (T, P, M).

Choose any $s \in S(T, P, M|d)$. Construct $\hat{s} \in S(T, P, M|d)$ such that

1. For any $i < |J|$, $x_i \in X_i$ and $p \in P_i$: $m_i(x_i, p|\hat{s}) = m_i(x_i, p|s)$.

2. For any $x_{|J|} \in X_{|J|}$ and $(z_1, ..., z_i) \notin \hat{Z}$:

$$m_{|J|}(x_{|J|}, p^{(z_1,...,z_i)}|\hat{s}) = m_{|J|}(x_{|J|}, p^{(z_1,...,z_i)}|s).$$

3. For any $x_{|J|} \in X_{|J|}$ and $(z_1, ..., z_i) \in \hat{Z}$:

$$m_{|J|}(x_{|J|}, p^{(z_1,...,z_i)}|\hat{s}) = \overline{z}$$

for some fixed $\overline{z} \in Z$.

Fix any $\mathbf{x} = (x_1, ..., x_{|J|}, \overline{x}) \in X$. Clearly $\rho_{|J|}(\mathbf{x}|\hat{s}) = \rho_{|J|}(\mathbf{x}|s)$. Denote $h = \rho_{|J|}(\mathbf{x}|\hat{s})$. If $P(h) = p^{(z_1,...,z_i)}$ where $(z_1, ..., z_i) \notin \hat{Z}$ then $\mu_{|J|}(\mathbf{x}|\hat{s}) = \mu_{|J|}(\mathbf{x}|s)$ and it follows that $\rho_{|J|+1}(\mathbf{x}|\hat{s}) = \rho_{|J|+1}(\mathbf{x}|s)$. If $P(h) = p^{(z_1,...,z_i)}$ where $(z_1, ..., z_i) \in \hat{Z}$ then $\mu_{|J|}(\mathbf{x}|\hat{s})$ is synonymous to $\mu_{|J|}(\mathbf{x}|s)$ given h within (T, P, M) and hence

$$P(\rho_{|J|+1}(\mathbf{x}|\hat{s})) = P(h \times \mu_{|J|}(\mathbf{x}|\hat{s})) = P(h \times \mu_{|J|}(\mathbf{x}|s)) = P(\rho_{|J|+1}(\mathbf{x}|s)).$$

We have established that $P(\rho_{|J|+1}(\mathbf{x}|\hat{s})) = P(\rho_{|J|+1}(\mathbf{x}|s))$ for any $\mathbf{x} \in X$, implying $\alpha_{|J|+1}(\mathbf{x}|\hat{s}) = \alpha_{|J|+1}(\mathbf{x}|s)$ because $s, s' \in S(T, P, M|d)$.

For any $(z_1, ..., z_i) \in \hat{Z}$ choose $y^{(z_1,...,z_i)} \in Y$ where $y^{(z_1,...,z_i)}$ is different from the candidate that is elected given any vote profile that contains $(z_1, ..., z_i)$. Construct d' such that

$$d'(z_1, ..., z_{|J|}) = \begin{cases} y^{(z_1,...,z_i)} & \text{if } (z_1, ..., z_i) \in \hat{Z} \text{ and } z_{|J|} \neq \overline{z} \\ d(z_1, ..., z_{|J|}) & \text{otherwise.} \end{cases}$$

It is straightforward to verify that under d' the collective choice is not determined before voting in the last stage takes place. Let (T, P', M) be the institution representing (r, d', t). Clearly $P'_i = P_i$ for every $i \leq |J|$. Choose any $s' \in S(T, P', M|d')$ such that $m_i(x_i, p|s') = m_i(x_i, p|\hat{s})$ for any $i \leq |J|$, $x_i \in X_i$ and $p \in P'_i$.

Arbitrarily choose $\mathbf{x} \in X$. Denote $g = \rho_{|J|}(\mathbf{x}|s')$. Clearly $\rho_{|J|+1}(\mathbf{x}|s') = \rho_{|J|+1}(\mathbf{x}|\hat{s})$ for any $\mathbf{x}$. If $g(i) \notin \hat{Z}$ then $\alpha_{|J|+1}(\mathbf{x}|s') = d'(\rho_{|J|+1}(\mathbf{x}|s')) = d(\rho_{|J|+1}(\mathbf{x}|\hat{s})) = \alpha_{|J|+1}(\mathbf{x}|\hat{s}) = \alpha_{|J|+1}(\mathbf{x}|s)$. If $g(i) \in \hat{Z}$ then $\mu(\mathbf{x}|s') = \overline{z}$, and following a similar sequence of equalities we have $\alpha_{|J|+1}(\mathbf{x}|s') = \alpha_{|J|+1}(\mathbf{x}|s)$. Thus $\boldsymbol{\alpha}(\cdot|s') = \boldsymbol{\alpha}(\cdot|s)$, implying $C(T, P, M|d) \subset C(T, P', M|d')$. It follows from Lemma 3.6.2 that (r, d', t) dominates (r, d, t). □

B.13 Proof of Proposition 3.6.4

Proof. Fix institution (T, P, M) where there is some i such that $|P_i| < |T_i|$. Let P' be a partition of T such that $P'_j = P_j$ for every $j \neq i$ and P'_i is a strict refinement of P_i. That $|P_i| < |T_i|$ implies such P'_i exists. Both complexity conditions in the proposition are satisfied by the institution (T, P', M). (T, P', M) dominates (T, P, M) by Lemma 3.5.4 because the former is obtained from the latter by refining. If (T, P, M) dominates (T, P', M) then by Theorem 3.5.9 (T, P, M) can be obtained from (T, P', M) by a sequence of operations of expanding, refining, trimming or relabeling. Note that merging does not apply because none of the player has a singleton action set. Since none of the operations decrease the number of perceptions of i, [4] it follows that $|P_i| \geq |P'_i|$, a contradiction. Thus (T, P, M) does not dominate (T, P', M). □

B.14 Proof of Proposition 3.6.5

Proof. Fix institution (T, P, M) where $|P_{i+1}| \geq 2$ for some $i \geq 2$. Let (T', P', M') be obtained from (T, P, M) by expanding, such that

1. $M'_j = M_j$ for any $j < N$ and $j \notin \{i-1, i\}$.
2. $M_{i-1} \subset M'_{i-1}$.
3. $M'_i = M_i \cup \{\overline{m}_i\}$ for some $\overline{m}_i \notin M_i$.
4. $|P'_j| = |P_j|$ for any $j \in \mathscr{N}$.
5. There is some $p' \in P'_i$ such that $h \in p'$ for any $h \in T'_i \backslash T_i$.
6. For any $m_i, m'_i \in M'_i$ there is some $h \in p'$ such that m_i and m'_i are not synonymous given h within (T', P', M').

An institution satisfying Properties $1-5$ can be easily constructed by making new messages available to players $i-1$ and i, and put message profiles containing the newly introduced messages to existing perceptions as prescribed by Property 5. If the institution (T', P', M') expanded from (T, P, M) satisfying Properties 1-5 does not satisfy Property 6, that is, there are $m_i, m'_i \in M'_i$ that are synonymous given any $h \in p'$ within (T', P', M'), then we can expand (T', P', M') by making an additional message $\overline{m}_{i-1}$ available to player $i-1$, and moreover:

- Put $g \times \overline{m}_{i-1}$ in p' for every $g \in T'_{i-1}$.
- Put $(g \times \overline{m}_{i-1}) \times m_i$ and $(g \times \overline{m}_{i-1}) \times m'_i$ in different perceptions of player $i+1$ for every $g \in T'_{i-1}$.
- Put any message profiles of length $j-1$ containing $\overline{m}_{i-1}$ arbitrarily to existing perceptions of player j for any $j > i+1$.

[4] By Lemma B.6.2, trimming does not decrease the number of perceptions of any player.

Let (T'', P'', M'') be the consequent institution. (T'', P'', M'') satisfies Properties 1-5. Let $p'' \in P''_i$ be the consequent perception enlarged from p'. It follows that m_i and m'_i are not synonymous given $g \times \overline{m}_{i-1}$ for any $g \in T''_{i-1}$ and hence they are not synonymous given p'' within (T'', P'', M''). We can keep applying this particular type of expanding until Property 6 is satisfied, without violating Properties 1-5.

Let (T', P', M') be expanded from (T, P, M) that satisfies Properties 1-6. (T', P', M') dominates (T, P, M) by Lemma 3.5.3. If (T, P, M) dominates (T', P', M') then the proof of Theorem 3.5.9 implies there are institutions (T^1, P^1, M^1), (T^2, P^2, M^2), (T^3, P^3, M^3) and (T^4, P^4, M^4) such that

1. (T^1, P^1, M^1) is obtained from (T', P', M') by merging.
2. (T^2, P^2, M^2) is obtained from (T^1, P^1, M^1) by refining.
3. (T^3, P^3, M^3) is obtained from (T^2, P^2, M^2) by a sequence of trimming operations.
4. (T^4, P^4, M^4) is obtained from (T^3, P^3, M^3) by relabeling.
5. (T, P, M) is obtained from (T^4, P^4, M^4) by expanding.

$(T^1, P^1, M^1) = (T', P', M')$ because everyone's action set is non-singleton. That $|P'_j| = |P_j|$ for every j implies $(T^2, P^2, M^2) = (T^1, P^1, M^1)$ because strict refining increases the number of perceptions for some player, which will not be decreased by trimming, relabeling or expanding. Thus $(T^2, P^2, M^2) = (T', P', M')$. $p' \cap T^3_i \in P^3_i$ because trimming does not decrease the number of perceptions by Lemma B.6.2. Since there do not exist $m_i, m'_i \in M'_i$ which are synonymous given p' within (T', P', M'), $M^3(p \cap T^3) = M'_i$. Thus for any $h' \in p' \cap T^3$ we have $|M^3(h')| = |M'_i| = |M_i| + 1$, implying there is $h \in T_i$ such that $|M(h)| = |M_i|+1$ because relabeling and expanding do not decrease the number of children of any message profile. This leads to a contradiction because $|M(h)| = |M_i|$ for any $h \in T_i$. Hence (T, P, M) does not dominate (T', P', M'). $\square$

Bibliography

Adachi, Hiroyuki (2003), "A search model of two-sided matching under nontransferable utility." *Journal of Economic Theory*, 113, 182 – 198.

Alcalde, José (1996), "Implementation of stable solutions to marriage problems." *Journal of Economic Theory*, 69, 240 – 254.

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

Alcalde, José and Antonio Romero-Medina (2005), "Sequential decisions in the college admissions problem." *Economics Letters*, 86, 153 – 158.

Amann, Erwin and Wolfgang Leininger (1996), "Asymmetric all-pay auctions with incomplete information: The two-player case." *Games and Economic Behavior*, 14, 1 – 18.

Baik, Kyung Hwan, In-Gyu Kim, and Sunghyun Na (2001), "Bidding for a group-specific public-good prize." *Journal of Public Economics*, 82, 415 – 429.

Barbieri, Stefano and David A. Malueg (2015), "Private-information group contests."

Barbieri, Stefano, David A. Malueg, and Iryna Topolyan (2013), "The best-shot all-pay (group) auction with complete information."

Blackwell, David (1951), "Comparison of experiments." In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 93–102, University of California Press, Berkeley, Calif.

Bloch, Francis and Effrosyni Diamantoudi (2011), "Noncooperative formation of coalitions in hedonic games." *International Journal of Game Theory*, 40, 263–280.

Bloch, Francis and Harl Ryder (2000), "Two-sided search, marriages, and matchmakers." *International Economic Review*, 41, 93–115.

Blumrosen, Liad and Michal Feldman (2013), "Mechanism design with a restricted action space." *Games and Economic Behavior*, 82, 424 – 443.

Blumrosen, Liad, Noam Nisan, and Ilya Segal (2007), "Auctions with severely bounded communication." *J. Artif. Int. Res.*, 28, 233–266.

Burdett, Ken and Melvyn G. Coles (1997), "Marriage and class." *Quarterly Journal of Economics*, 112, 141–168.

Chowdhury, Subhasish M., Dongryul Lee, and Iryna Topolyan (2016), "The max-min group contest: Weakest-link (group) all-pay auction." *Southern Economic Journal*, n/a–n/a.

Chowdhury, Subhasish M. and Iryna Topolyan (2015), "The group all-pay auction with heterogeneous impact functions."

Condorcet, Marquis de (1785), *Essai sur l'application de l'analyse á la probabilité des décisions rendues á la pluralité des voix*. Paris.

Cornes, Richard and Roger Hartley (2005), "Asymmetric contests with general technologies." *Economic Theory*, 26, 923–946.

Costinot, Arnaud and Navin Kartik (2007), "On optimal voting rules under homogeneous preferences."

Dekel, Eddie and Michele Piccione (2000), "Sequential voting procedures in symmetric binary elections." *Journal of Political Economy*, 108, pp. 34–55.

Diamantoudi, Effrosyni, Eiichi Miyagawa, and Licun Xue (2015), "Decentralized matching: The role of commitment." *Games and Economic Behavior*, 92, 1 – 17.

Duggan, John and Cesar Martinelli (2001), "A bayesian model of voting in juries." *Games and Economic Behavior*, 37, 259 – 294.

Eeckhout, Jan (1999), "Bilateral search and vertical heterogeneity." *International Economic Review*, 40, 869–887.

Eeckhout, Jan (2000), "On the uniqueness of stable marriage matchings." *Economics Letters*, 69, 1–8.

Esteban, Joan and Debraj Ray (2001), "Collective action and the group size paradox." *American Political Science Review*, 95, 663–672.

Esteban, Joan and Debraj Ray (2008), "On the salience of ethnic conflict." *American Economic Review*, 98, 2185–2202.

Ewerhart, Christian (2014), "Unique equilibrium in rent-seeking contests with a continuum of types." *Economics Letters*, 125, 115 – 118.

Ewerhart, Christian and Federico Quartieri (2013), "Unique equilibrium in contests with incomplete information."

Fadel, Ronald and Ilya Segal (2009), "The communication cost of selfishness." *Journal of Economic Theory*, 144, 1895 – 1920.

Feddersen, Timothy and Wolfgang Pesendorfer (1998), "Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting." *The American Political Science Review*, 92, pp. 23–35.

Fu, Qiang, Jingfeng Lu, and Yue Pan (2015), "Team contests with multiple pairwise battles." *American Economic Review*, 105, 2120–40.

Gale, David (2000), *Strategic Foundations of General Equilibrium: Dynamic Matching and Bargaining Games*. Cambridge University Press, Cambridge.

Gale, David and Lloyd Shapley (1962), "College admissions and the stability of marriage." *American Mathematical Monthly*, 69, 9–15.

Gale, Douglas and Hamid Sabourian (2006), "Markov equilibria in dynamic matching and bargaining games." *Games and Economic Behavior*, 54, 336 – 352.

Haeringer, Guillaume and Myrna Wooders (2011), "Decentralized job matching." *International Journal of Game Theory*, 40, 1–28.

Hillman, Arye L. and John G. Riley (1989), "Politically contestable rents and transfers." *Economics and Politics*, 1, 17–39.

Kirkegaard, Rene (2013), "Handicaps in incomplete information all-pay auctions with a diverse set of bidders."

Konrad, Kai A. and Dan Kovenock (2009), "The alliance formation puzzle and capacity constraints." *Economics Letters*, 103, 84 – 86.

Konrad, Kai A. and Wolfgang Leininger (2007), "The generalized stackelberg equilibrium of the all-pay auction with complete information." *Review of Economic Design*, 11, 165–174.

Kos, Nenad (2012), "Communication and efficiency in auctions." *Games and Economic Behavior*, 75, 233 – 249.

Kos, Nenad (2014), "Asking questions." *Games and Economic Behavior*, 87, 642 – 650.

Lauermann, Stephan and Georg Nöldeke (2014), "Stable marriages and search frictions." *Journal of Economic Theory*, 151, 163 – 195.

Lizzeri, Alessandro and Nicola Persico (2000), "Uniqueness and existence of equilibrium in auctions with a reserve price." *Games and Economic Behavior*, 30, 83 – 114.

Marschak, Jacob and Roy Radner (1972), *Economic theory of teams*. Cowles Foundation for Research in Economics at Yale University, Yale University Press, New Haven, London.

McLennan, Andrew (1998), "Consequences of the condorcet jury theorem for beneficial information aggregation by rational agents." *The American Political Science Review*, 92, pp. 413–418.

McNamara, J. M. and E. J. Collins (1990), "The job search problem as an employer-candidate game." *Journal of Applied Probability*, 27, 815–827.

Michael R. Baye, Casper G. de Vries, Dan Kovenock (1996), "The all-pay auction with complete information." *Economic Theory*, 8, 291–305.

Mookherjee, Dilip and Masatoshi Tsumagari (2014), "Mechanism Design with Communication Constraints." *Journal of Political Economy*, 122, 1094 – 1129.

Münster, Johannes (2007), "Simultaneous inter- and intra-group conflicts." *Economic Theory*, 32, 333–352.

Münster, Johannes (2009), "Group contest success functions." *Economic Theory*, 41, 345–357.

Niederle, Muriel and Leeat Yariv (2009), "Decentralized matching with aligned preferences." NBER Working Papers 14840, National Bureau of Economic Research, Inc.

Nisan, Noam and Ilya Segal (2006), "The communication requirements of efficient allocations and supporting prices." *Journal of Economic Theory*, 129, 192 – 224.

Nitzan, Shmuel (1991), "Collective rent dissipation." *The Economic Journal*, 101, 1522–1534.

Nitzan, Shmuel and Kaoru Ueda (2009), "Collective contests for commons and club goods." *Journal of Public Economics*, 93, 48 – 55.

Nitzan, Shmuel and Kaoru Ueda (2011), "Prize sharing in collective contests." *European Economic Review*, 55, 678 – 687.

Osborne, Martin J. and Ariel Rubinstein (1990), *Bargaining and Markets*. Academic Press, Inc., San Diego, CA.

Pais, Joana (2008), "Incentives in decentralized random matching markets." *Games and Economic Behavior*, 64, 632 – 649.

Romero-Medina, Antonio and Matteo Triossi (2013), "Acyclicity and singleton cores in matching markets." *Economics Letters*, 118, 237 – 239.

Roth, Alvin E. (2008), "What have we learned from market design?" *Economic Journal*, 118, 285–310.

Roth, Alvin E. and Marilda A. Liveria Sotomayor (1990), *Two-Sided Matching: A Study In Game-theoretic Modeling and Analysis*. Cambridge University Press, Cambridge.

Roth, Alvin E and John H Vande Vate (1990), "Random Paths to Stability in Two-Sided Matching." *Econometrica*, 58, 1475–80.

Rubinstein, Ariel and Asher Wolinsky (1990), "Decentralized trading, strategic behaviour and the Walrasian outcome." *Review of Economic Studies*, 57, 63–78.

Ryvkin, Dmitry (2010), "Contests with private costs: Beyond two players." *European Journal of Political Economy*, 26, 558 – 567.

Segal, Ilya (2007), "The communication requirements of social choice rules and supporting budget sets." *Journal of Economic Theory*, 136, 341 – 378.

Segal, Ilya R. (2010), "Nash implementation with little communication." *Theoretical Economics*, 5, 51–71.

Siegel, Ron (2009), "All-pay contests." *Econometrica*, 77, 71–92.

Siegel, Ron (2014), "Asymmetric all-pay auctions with interdependent valuations." *Journal of Economic Theory*, 153, 684 – 702.

Skaperdas, Stergios (1998), "On the formation of alliances in conflict and contests." *Public Choice*, 96, 25–42.

Smith, Lones (2006), "The marriage model with search frictions." *Journal of Political Economy*, 114, 1124–1144.

Suh, Sang-Chul and Quan Wen (2008), "Subgame perfect implementation of stable matchings in marriage problems." *Social Choice and Welfare*, 31, 163–174.

Topolyan, Iryna (2013), "Rent-seeking for a public good with additive contributions." *Social Choice and Welfare*, 42, 465–476.

Van Zandt, Timothy (2007), "Communication complexity and mechanism design." *Journal of the European Economic Association*, 5, pp. 543–553.

Wilson, Andrea (2014), "Bounded memory and biases in information processing." *Econometrica*, 82, 2257–2294.