THREE ESSAYS ON GAME THEORY

By

Demet Yilmazkuday

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

August, 2011

Nashville, Tennessee

Approved:

Professor Quan Wen

Professor Eric W. Bond

Professor Andrew F. Daughety

Professor Paul H. Edelman

To Hakan and Ada

ACKNOWLEDGMENTS

I am grateful to my advisor, Professor Quan Wen, and my dissertation committee members, Professors Eric W. Bond, Andrew F. Daughety, and Paul H. Edelman, for their guidance. I also would like to thank John A. Weymark for helpful suggestions. All the remaining errors are mine.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

This dissertation consists of three studies on game theory. In the first study, we examine a sequential matching problem. In particular, we study preference restrictions to implement stable matchings in the equilibrium. In the second study, we propose a solution concept for the roommate problem. In the third study, we analyze the effects of transportation costs on regional and multilateral trade agreements.

On Subgame Perfect Implementation of Stable Matchings

This study investigates subgame perfect implementation of stable matchings in a sequential matching mechanism. We identify some "ring" conditions on the domain of preferences that are necessary and sufficient for the unique, men and women optimal stable matchings to be implementable in the subgame perfect equilibrium (SPE) of the sequential mechanism. We also investigate how these ring conditions are related to the Eckhout (Econ Lett, 2000) and α^M (Suh and wen, 2008) conditions.. We introduce the No-ring-by-which-stable-matching-partners-are-swapped condition, which is weaker than the Eeckhout condition, and show that it is not only sufficient as the Eeckhout condition, but also necessary for the existence of a unique stable matching and for the unique stable matching to be in the SPE. We propose the No-ring-by-which-men-swap-optimal-partners condition and prove that it is both necessary and sufficient for the men-optimal stable matching to be in the SPE of the mechanism when men move first. Moreover, it is equivalent to the α^M condition. Hence, we prove that the α^M condition is not only sufficient but also necessary for the men-optimal stable matching to be the SPE outcome of the mechanism when men move first.

A Solution for the Roommate Problem and A Sequential Matching Mechanism

The study has two parts: In the first part, we propose a solution concept for the roommate problem, and in the second part, we focus on a sequential roommate problem. The solution concept is related to the $P-stability$ concept (Inarra et al, 2008) and is called $RP-stability$ (reduced preference profile $P-stability$). Similar to $P-stability$, an $RP-stable$ matching always exists. Moreover, whenever a Pareto improvement is possible in a $P-stable$ matching, $RP-stable$ matching provides a Pareto improvement. Furthermore, the number of matched individuals in an $RP-stable$ matching is always greater than or equal to the number of those in any $P-stable$ matching. We introduce a procedure, which is called the RP procedure, to obtain the set of $RP-stable$ matchings. In the second part, we focus on a sequential roommate problem and analyze the subgame perfect equilibrium (SPE) of this problem. First, we show that the set of all potential SPE outcomes can be identified by the RP procedure. Second, we identify a sufficient condition to guarantee the pairwise stability of the SPE outcome of the sequential game regardless of the order of individuals in solvable roommate problems, i.e., roommate problems with stable solutions.

Regional versus Multilateral Trade Agreements: A Welfare Analysis

Why are trade agreements mostly regional? By making a welfare analysis, we show that the existence of transportation costs may be a possible reason. In particular, we set up a model of N countries by considering the effects of transportation costs on the welfare of each country. We also consider the relative size of the countries in our analysis together with measures of comparative advantage. We first show that the optimal tariff rates are decreasing in transportation costs; i.e., closer countries tend to have higher tariff rates between each other, which implies a potential gain from regional agreements. After that, we make a welfare analysis by using a stationary dynamic tariff game approach and show that it is harder to make an agreement for countries if the transportation cost is higher; i.e., countries tend to sustain regional agreements to maximize their welfare.

CHAPTER II

ON SUBGAME PERFECT IMPLEMENTATION OF STABLE MATCHINGS

Introduction

A two-sided matching problem considers, in general, two sets of agents, X and Y, and the allocation of X among Y "and" Y among X. For example, a hospital-intern problem is a matching problem that allocates open positions in hospitals among interns and interns among hospitals. If an intern is matched to a hospital, then the hospital is matched to the intern. Hence, the exchange in the problem is naturally bilateral. In this paper, we focus on what is known as *the marriage problem*, in which two sets of agents are said to be men and women, and each individual can be matched (i.e., married) at most one individual of the opposite sex.[1]

A matching is *individually rational* if there is no individual who prefers to be unmatched (i.e., being single or alone) rather than being matched to the his or her partner under the matching. Two individuals *block* a matching if they prefer to be matched to each other rather than being matched to their partners under the matching. A matching is stable if it is not blocked by any individual or pair of individuals. Gale and Shapley (1962) proved the existence of a stable matching in a marriage problem through the deferred acceptance algorithm, which always finds a stable matching. They also showed that when men propose in their algorithm, every man likes the resulting matching as much as any other stable matching; this matching is called the men-optimal stable matching, μ^M. Similarly,

[1]See Roth and Sotomayor (1990) for a comprehensive review on two-sided matching problems.

when women propose in the deferred acceptance algorithm, every woman likes the resulting matching as much as any other stable matching; this matching is called the women-optimal stable matching, μ^W. Hereafter, we will state definitions and results only for men as similar definitions and results can be obtained by switching men and women's roles.

In this paper, we focus on the implementation of μ^M and the unique stable matching (when there is one) as the Subgame Perfect Equilibrium (SPE) outcomes of a sequential matching mechanism. Hereafter, we refer to this mechanism as the sequential matching mechanism (SM-mechanism). In a SM-mechanism, men and women move sequentially according to a previously specified order. An individual can choose one of three possible actions when it is his or her turn to move: (i) Accept a proposal which has been made to himself or herself at a previous stage, or (ii) propose to a potential partner who moves at a later stage, or (iii) choose to remain single. The mechanism in which men move before women is called the men-move-first mechanism.

One may expect that the SPE outcome of the SM-mechanism is a stable matching regardless of the order of individuals' moves. However, the stability of the equilibrium outcome is not guaranteed without further restrictions on the preference domain. This paper is motivated by such restrictions to implement stable matchings in the SPE. In particular, this paper aims to answer two questions: 1) Under which conditions are μ^M, μ^W, and the unique stable matching implementable in the SPE of the SM-mechanism regardless of which order the individuals move in, and 2) how are these conditions related to other conditions such as the Eeckhout condition (Eeckhout, 2000) and the alpha conditions (Suh and Wen, 2008), which are previously considered in the context of implementing μ^M, and the unique stable matching?

First, we explain our motivation for the first question and highlight the related literature. Our objective regarding the first question is to find necessary and sufficient conditions to implement μ^M and the unique stable matching. We are particularly interested in finding "ring" conditions. An ordered set of men and women is a (strict) *ring* if they can be arranged in a circle with men-women couples alternating such that each individual prefers the next person in the circle (e.g., clockwise) to the one who precedes him or her. We are interested in ring conditions (i.e., condition that are defined in terms of certain types of rings in the preference profile) because there is a close relation between stable matchings and rings in a preference profile. For example, in a more general matching problem, which is called the roommate problem, a stable matching may not exist. The relation between stable matchings and rings become more obvious in roommate problems because the existence of stable matching is directly related to the existence of some certain type of rings. We observe that the close relation between rings and stable matchings also play an important role in implementing stable matchings context. By exploring this relation, we are able to identify two ring conditions: The No-ring-by-which-stable-matching-partners-are-swapped (NRS) condition, and the No-ring-by-which-men-swap-optimal-partners (NRMO) condition. A preference profile satisfies the NRS condition if and only if there is no ring such that each man and the woman who precedes him are partners in a stable matching or each woman and the man who precedes her are partners in a stable matching . A preference profile satisfies the NRMO condition if and only if there is no ring such that each man and the woman who precedes him are partners in the men-optimal stable matching. The first objective of this paper is to show that the NRS condition is a necessary and sufficient condition for there to be a unique stable matching and for the unique stable matching to be implementable as the SPE outcome of the SM-mechanism regardless of the order of individuals. Similarly, we show that the NRMO condition is necessary and sufficient to implement μ^M as the SPE outcome of the SM-mechanism regardless of the order of individuals in the men-move-first mechanism.

Now, we explain our motivation for the second question. Our objective is to understand how our ring conditions are related to Eeckhout's condition (Eeckhout, 2000), which we refer as the Eeckhout condition hereafter, and the alpha-men (α^M) and alpha-women (α^W) conditions (Suh and Wen, 2008). A preference profile satisfies the Eeckhout condition if and only if men and women can be ordered such that each man (woman) prefers the woman (man) with the same index to all women (men) with higher indices. A preference profile satisfies the α^M condition if and only if men and women can be ordered such that each man prefers the woman with the same index to all women with higher indices, and if a man prefers a woman with a lower index to the woman with the same index as himself, then he must be preferred less by the woman with the lower index to the man with the same lower index. We focus on the Eeckhout condition and the α^M condition because, as far as we know, these are the only conditions that are previously considered in the context of implementing stable matchings. We show that the NRMO condition is an equivalent ring condition of the α^M condition. Hence, we also prove that the α^M condition is not only sufficient (Suh and Wen, 2008), but also a necessary condition for μ^M to be the SPE outcome of the men-move-first mechanism regardless of the order of moves. Furthermore, we analyze how the NRS condition and the Eeckhout condition are related. We show that the NRS condition is weaker then the Eeckhout condition.

The reason for our interest in ring conditions is the close relation between stable matchings and rings in a preference profile. There are very important studies in the literature which show particular interest in the relation between rings and stable matchings. For example, Chung (2000) identifies a condition, called the "no odd rings"condition, that is sufficient for the existence of a stable matching when preferences are not necessarily strict in the roommate problems.[2] Another study is by Irving and Leather (1986) who propose an algorithm to compute all stable matchings in a marriage problem. In this algorithm, a stable matching is obtained by breaking a marriage in another stable marriage and satisfying a subset of identified agents that form a cycle.[3] Rings play an important role in Gale's Top Trading Cycles algorithm (Shapley and Scarf, 1974) which is used to find the unique core allocation in the context of housing markets, in the YRMH-IGYT (you request my house-I get your turn) mechanism (Abdulkadiroglu and Sonmez, 1999) which is a generalization of Gale's Top Trading Cycles algorithm, and in the Top Trading Cycles and Chains mechanism (Roth, Sonmez, and Unver, 2004) which is used to solve the kidney exchange problem.

In terms of implementation of stable matchings, one close study is by Suh and Wen (2008). In their paper, Suh and Wen show that the condition proposed by Eeckhout (2000), which guarantees the uniqueness of the stable matching as Eeckhout proves, is a sufficient condition for the unique stable matching to be implementable in the SPE for any order of individuals' moves. They also prove that the uniqueness of the stable matching is not sufficient to implement the unique stable matching in the SPE for any order of individuals' moves. The α^M condition is proved to be a sufficient condition for the men-optimal stable matching to the in the SPE of any men-move first mechanism by Suh and Wen (2008). They also note that the Eeckhout condition implies the α^M and α^W conditions, but the reverse is not true.

[2] A roommate problem is a one-sided matching problem, i.e, there is only one set of agents. A marriage problem is a special case of a roommate problem. In a roommate problem, the existence of a stable matching is not guaranteed.

[3] Irving and Leather (1986), Abulkadiroglu and Sonmez (1999), and Roth, Sonmez, and Unver (2004) use the term "cycle" instead of the term "ring".

Another close study is Alcalde and Romero-Medina (2005) who study the implementation of the student-optimal stable matching in the SPE. In their study, they propose the students-sequentially-propose-and-colleges-choose mechanism and show that the student-optimal stable matching is the SPE outcome when colleges' preferences are substitutive.[4] Their mechanism is equivalent to the men-move-first mechanism with a capacity constraint imposed on women. Alcalde and Romero-Medina (2005) also argue that a symmetrical result would hold for the marriage problem (i.e., each college can accept only one student) when the role of students and colleges are exchanged. However, as Suh and Wen noted and as our example shows that their argument does not hold without further restrictions imposed on a preference domain. Alcalde and Romero-Medina (2000) and Sotomayor (2003) show that the core correspondence and/or the set of stable solutions of the college admissions are implementable in the SPE when colleges' preferences are substitutable.[5] All the mechanisms considered are two stage mechanisms such that agents on one side propose or choose a set of potential partners simultaneously in the first stage and agents on the other side choose whom to be matched simultaneously in the second stage. In the mechanism employed by Sotomayor (2003), agents of one side who move in the first stage choose a set of acceptable partners. However, in Alcalde and Romero-Medina (2000), the choice set of each agent who move in the first stage is restricted to be a singleton set. Because the mechanisms considered are different from the SM-mechanism in the sense that agents who are active, (i.e. agents who move in the first stage) do not decide sequentially, our results cannot be seen as a consequence of those obtained by Alcalde and Romero-Medina (2000)

[4] In college admissions problems, when colleges have preferences over sets of students rather than individual students, a stable solution may not exist unless colleges' preferences satisfy substitutability. Colleges' preferences satisfy substitutability if and only if for any two students $s \neq s'$ and for any subset of students A with $s, s' \in A$, if s is in $Ch_c(A)$, then s is also in $Ch_c(A \setminus \{s'\})$ where $Ch_c(A)$ is defined as the most prefered set of students by c among all subsets of A (including A).

[5] It is known that for a college admissions problem, the core of a college admissions problem coincides with the set of pairwise stable solutions under subsitutable preferences (Roth and Sotomayor, 1990).

and Sotomayor (2003).

Some other important studies that are interested in the implementation of stable solutions or the core of the problem are Alcalde (1996), Shin and Suh (1996), and Kara and Sonmez (1996). Alcalde (1996) has employed and compared two mechanisms one of which is the Gale-Shapley algorithm to analyze the possibility of implementing stable outcomes for marriage markets. Shin and Suh (1996) have considered a stable matching mechanism that implements the stable rule that selects all of the stable matchings in a strong Nash equilibrium. Kara and Sonmez (1996) have shown that all rules by which Pareto optimal and individually rational matchings are implementable as supersolutions of the stable rule.

This work is organized as follows. In section 2, we introduce the marriage model and some useful definitions and preliminary results. We present the Eeckhout and α^M conditions, and present the results of Eeckhout (2000) and Suh and Wen (2008) that are used and analyzed throughout the paper. In section 3, we introduce our ring condition. First, we introduce the NRMO condition and show the equivalency of the NRMO and α^M conditions. Second, we present our results regarding the NRS condition. In section 4, we focus on the SM-mechanism, and give some preliminary results. In Section 5, we first prove that the NRMO condition is the necessary and sufficient condition to implement μ^M in the SPE of the men-move-first mechanism regardless how individuals move. Next, we prove that the NRS condition is necessary and sufficient to have a unique stable matching and to have the unique stable matching as the SPE outcome of the Suh-Wen sequential matching game no matter how individuals move. Section 6 concludes. All proofs are in the Appendix.

The Model

The one-to-one matching problem that we consider is interpreted as a marriage problem. In a marriage problem, there are two disjoint finite sets of individuals, men and women, denoted by M and W. Each individual has preferences over the individuals of the opposite sex and being unmatched. A matching is a function $\mu : M \cup W \longrightarrow M \cup W$ such that each man is matched at most one woman or remains unmatched, and each woman is matched to at most one man or remains unmatched. Individual i is unmatched, we write $\mu(i) = i$.

We assume that there are equal number of men and women, $|M| = |W| = n$. Individuals' preferences are complete, irreflexive, transitive, and strict. Let $\succ_i$ represents individual i's preferences. Individual j is acceptable to i if i prefers to be matched to j rather than being unmatched, $j \succ_i i$. Given a preference profile $\succ = (\succ_i)_{i \in M \cup W}$, a matching μ is *individually rational if* $\mu(i)$ is acceptable to i for all $i \in M \cup W$. A pair of man and woman (m, w) *blocks* μ if $w \succ_m \mu(m)$ and $m \succ_w \mu(w)$. A matching μ is *stable* if it is individually rational and if there is no blocking pair for μ. A stable matching μ^M is the men-optimal stable matching if every man likes μ^M as much as any other stable matching. The existence of optimal stable matchings is due to Gale and Shapley (1962).

Preliminary Results

In this section, we present some preliminary results that we benefit throughout the paper. The first of these results is related to the condition considered by Eeckhout (2000).

Eeckhout Condition. It is possible to rename the individuals such that (i) for all $m_i \in M$,

$w_i \succ_{m_i} w_j$ for all $j > i$ and (ii) for all $w_i \in W$, $m_i \succ_{w_i} m_j$ for all $j > i$.

Eeckhout (2000) shows that the Eeckhout condition is a sufficient condition for there to be a unique stable matching. The following example shows that the Eeckhout condition is not necessary.

Example 1 *The preference profile $\succ$ is given by:*

m_1	m_2	m_3	w_1	w_2	w_3
w_1	w_1	w_2	m_3	m_2	m_3
w_3	w_2	w_3	m_1	m_3	m_1
w_2	w_3	w_1	m_2	m_1	m_2

The unique stable matching for $\succ$ is

$$\mu = \begin{pmatrix} m_1 & m_2 & m_3 \\ w_1 & w_2 & w_3 \end{pmatrix}.$$

The profile does not satisfy the Eeckhout Condition because there is no pair of man and woman who mutually rank each other first.

Eeckhout (2000) also shows that a preference profile that satisfies the Eeckhout condition contains no ring of a certain order. An ordered set of men and women is a (strict) *ring* if they can be arranged in a circle with men-women couples alternating such that each individual prefers the next person in the circle (e.g., clockwise) to the one who precedes him or her. The formal definition of a ring is as follows.

Definition 1 *A ring is an ordered list of men and women $(x_1, x_2, \ldots, x_k)$, $(k \geq 3)$, such that $x_{i+1} \succ_{x_i} x_{i-1}$, $\forall i = 1, \ldots, k \pmod k$.*

Observe that there must be at least two men and two women in a ring (in a marriage problem), and each individual in the ring must be the opposite sex of the preceding and succeeding individuals.

Lemma 1 (Eeckhout, 2000, Lemma 2) *A preference profile $\succ$ that satisfies the Eeckhout condition contains no ring of type $(r_k, s_k, r_l, s_l, \ldots, r_p, s_p)$, $(k \geq 2)$, where $r_i \in M$ for all $i = k, \ldots, p$, if and only if $s_i \in W$ for all $i = k, \ldots, p$ such that $k < l < \cdots < p$.*

We should highlight two points about this lemma. First, the lemma does not say that when the Eeckhout condition is satisfied, a preference profile does not contain a ring. The rings that are considered here are a certain type of rings. Second, the converse of the lemma is not true. For example, the profile considered in 1 does not contain any such ring, but it does not satisfy the Eeckhout condition.

Another condition that we consider in this paper is the α^M condition that is introduced by Suh and Wen (2008). They identify μ^M with the α^M condition.

The α^M Condition. A preference profile $\succ$ satisfies the α^M condition if it is possible to rename the individuals such that (i) for all $m_i \in M$, $w_i \succ_{m_i} w_j$ for all $j > i$ and (ii) for all $m_i \in M$, if $w_j \succ_{m_i} w_i$ for all $j < i$, then $m_j \succ_{w_j} m_i$.

Proposition 1 (Suh and Wen, 2008, Proposition 2) *Under the α^M condition, the men-optimal stable matching is μ^*, where $\mu^*(m_i) = w_i$ for all $i \in \{1, \ldots, n\}$.*

The Ring Conditions

The NRMO and α^M conditions

In this section, we identify a ring condition, which we call the No-ring-by-which-men-swap-optimal-partners (NRMO) condition. We analyze the α^M condition, which Suh and Wen (2008) introduce, and its relation to the NRMO condition. We show that the NRMO and α^M conditions are equivalent. Now, the definition of the NRMO condition is introduced.

The NRMO condition. A preference profile $\succ$ satisfies the NRMO condition if and only if there is no ring $(w_k, m_k, \ldots, w_t, m_t)$, such that $\mu^M(w_i) = m_i$ for all $i = k, \ldots, t$.

Observe that if a preference profile satisfy the α^M condition, then it is possible to order the individuals such that for any man, all women with higher indices are less preferred than the woman who has the same index as the man. Such an ordering would not be possible is there is a ring of type $(w_k, m_k, \ldots, w_t, m_t)$, such that $\mu^M(w_i) = m_i$ for all $i = k, \ldots, t$ because eventually there is at least one man who prefers a woman with a higher index to the woman with the same index due to the nature of a ring.

Lemma 2 *If a preference profile $\succ$ satisfies the NRMO condition, then there is at least one man who is matched to his top choice in μ^M.*

Proposition 2 *A preference profile $\succ$ satisfies the α^M condition if and only if it satisfies the NRMO condition.*

The Eeckhout and NRS conditions

In this section we focus on the Eeckhout condition. We introduce the No-ring-by-which-stable-matching-partners-are-swapped (NRS) condition, and show that the NRS condition is weaker then the Eeckhout condition, yet it is sufficient for there to be a unique stable matching. The NRS condition focuses on rings in which each man in the ring and the woman who precedes him or each woman in the ring and the man who precedes him are stable matching partners. The condition requires the nonexistence of such rings.

NRS Condition. A preference profile $\succ$ satisfies the NRS condition if and only if there is no ring $(r_1, s_1, \ldots, r_k, s_k)$ such that $\mu(r_i) = s_i$ where μ is a stable matching $r_i \in M$ if and only if $s_i \in W$ for all $i = k, \ldots, t$.

The preference profile in Example 1 satisfies the NRS condition. This is because for any individual $r_i \in M \cup W$, if $\mu(r_j) \succ_{r_i} \mu(r_i)$, then we have $\mu(r_j) \succ_{r_j} \mu(r_i)$. Hence, a ring which contains a group of individuals and their partners in a stable matching cannot be formed. The following example shows a preference profile which violates the NRS condition.

Example 2 *The following preference profile violates the NRS condition.*

$\succ_1$

m_1	m_2	m_3	w_1	w_2	w_3
w_3	w_3	w_2	m_2	m_1	m_3
w_1	w_2	w_3	m_1	m_2	m_1
w_2	w_1	w_1	m_3	m_3	m_2

The two stable matchings are

$$\mu^M = \begin{pmatrix} m_1 & m_2 & m_3 \\ w_1 & w_2 & w_3 \end{pmatrix} \text{ and } \mu^W = \begin{pmatrix} m_1 & m_2 & m_3 \\ w_2 & w_3 & w_1 \end{pmatrix}.$$

A ring by which the NRS condition is violated is (w_2, m_2, w_3, m_3).

Notice that the NRS condition focuses on rings that make it possible for a group of the same sex individuals to be better off by swapping their stable matching partners in the sequential matching mechanism. That is, the NRS condition suggests that rings by which a group of same sex individuals swap their stable matching partners are relevant to the uniqueness and implementation of stable matchings and all other rings are irrelevant. Observe that the Eeckhout condition is stronger than the NRS condition. This is because if there is a ring by which the NRS condition is violated, then for any ordering of individuals, we always have $r_j \succ_{r_i} s_i$ with $j > i$, for some individuals r_j, r_i, s_i. Hence, whenever the NRS condition is violated, the Eeckhout condition is also violated. Furthermore, if we let the stable matching in the definition of the NRS condition be μ^M and $s_i = m_i$, the definition becomes the definition of the NRMO condition. Hence, the NRS condition implies the NRMO condition. Next, we show that the NRS condition is a sufficient condition for there to be a unique stable matching in the following proposition.

Proposition 3 *If a preference profile $\succ$ satisfies the NRS condition, then there is a unique stable matching for $\succ$.*

A sequential Matching Mechanism

The sequential matching mechanism (SM-mechanism) considered in this paper is a $2n$ stage extensive form game of perfect information. The order of moves is given by a surjective function $\sigma : \{1, \ldots, 2n\} \longrightarrow M \cup W$, where $\sigma(k)$ denotes the individual who moves in the kth stage of the mechanism.

In stage 1, $\sigma(1)$ either proposes to a potential mate or chooses to be single.

In stage k, $1 < k < 2n$, after observing the history of the game up to stage k, $\sigma(k)$ either (i) accepts one potential mate who has proposed to him or her in a previous stage, (ii) proposes to a potential mate who will move in a subsequent stage, or (iii) chooses to remain single.

In stage $2n$, $\sigma(2n)$ either accepts one potential mate who has proposed to him or her in a previous stage, or chooses to remain single.

There is a different sequential mechanism for every order of moves. A mechanism in which all men move before women is called a *men-move-first mechanism*. A *women-move-first mechanism* is defined similarly.

Given an order of moves, the outcome of this mechanism is a matching. This is because an individual cannot accept more than one proposal, and by making a proposal or by deciding to remain single, he or she must reject all proposals that has been received. Because preferences are strict, the SPE outcome of the SM-mechanism for a given an order of moves is unique. Moreover, the SPE outcome of the game is an individually rational matching. This is because an individual i would choose to remain single rather than proposing to or accepting a proposal from someone who is not acceptable for i.

Suh and Wen (2008) show that the Eeckhout condition is also a sufficient condition to implement the unique stable matching as the SPE outcome of the SM-mechanism for any order of moves. They also show that the α^M condition is a sufficient condition to implement μ^M in the SPE of the men-move-first sequential matching mechanism regardless of which order the individuals move in.

Proposition 4 (Suh and Wen, 2008, Theorem 3) *Under the α^M condition, the SPE outcome of the men-move-first mechanism is the men-optimal stable matching μ^M.*

The men-optimal and the unique stable matchings in the equilibrium

The men-optimal stable matching in the equilibrium

In this section, we investigate the men-move-first mechanism. In particular, we focus on the conditions under which μ^M is the SPE outcome of the men-move-first mechanism. Our main purpose is to show that the NRMO condition is a necessary and sufficient condition for μ^M to be the SPE outcome of the men-move-first mechanism for any order of moves. By Propositions 2 and 4, we already know that the NRMO condition is a sufficient condition for μ^M to be the SPE of any men-move-first mechanism. Hence, we only need to prove that the NRMO condition is also necessary for μ^M to be the SPE of any men-move-first mechanism. The next proposition is used to prove the necessity and sufficiency of the NRMO condition.

Proposition 5 *There is no men-move-first mechanism such that a man is matched to a woman worse than his μ^M partner in the SPE.*

Proposition 6 *A preference profile $\succ$ satisfies the NRMO condition if and only if the SPE outcome of the men-move-first mechanism is μ^M regardless the order of moves.*

We finish this section by the following corollary and Figure 1 which summarizes our results of this section.

Corollary 1 *Given a preference profile $\succ$, the SPE outcome of the SM-mechanism for any order of moves is μ^M if and only if $\succ$ satisfies the NRMO condition if and only if $\succ$ satisfies the α^M condition.*

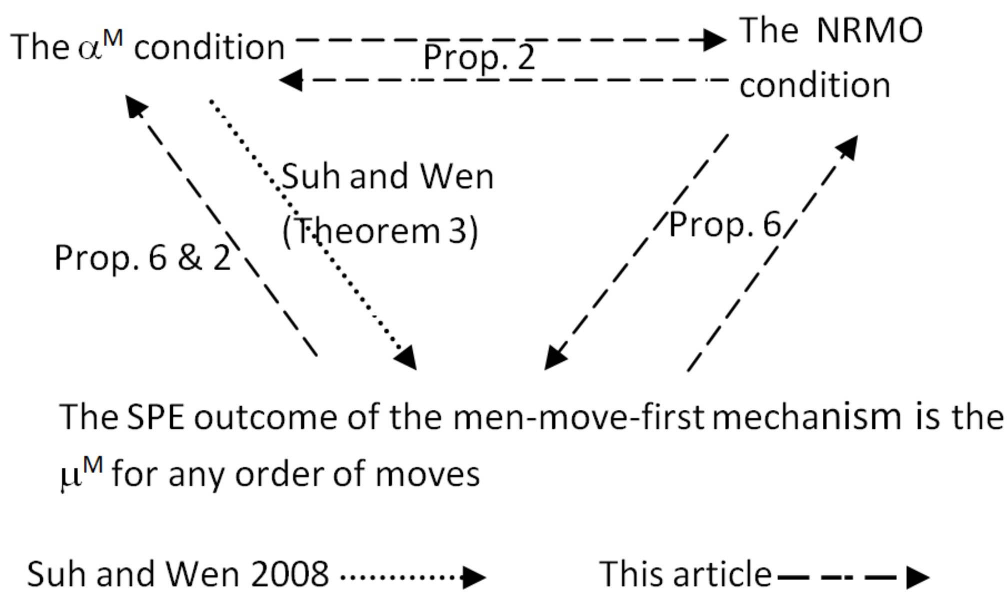

Figure 1. The NRMO and α^M Conditions

The unique stable matching in the equilibrium

In this section, we are interested in identifying a necessary and sufficient condition to guarantee the stability of the SPE outcome of any sequential mechanism. We show that the NRS condition serves the purpose of this section. First, we prove that if a preference profile satisfies the NRS condition, not only there is a unique stable matching, but this matching is also the unique SPE outcome of the SM-mechanism for any order of moves. We present two lemmas that are used throughout the paper. Before we present these lemmas, we need to introduce new notation and definitions that are useful. A preference profile $\succ$ and a SM-mechanism is given. Set $P_k^t(\sigma(i))$, $i = 1, 2, \ldots, 2n$, is defined as the set of individuals who have proposed to individual $\sigma(i)$ between the beginning of stage k until the beginning of stage t and set $T^k(\sigma(k))$ is defined as $\{\sigma(l) \in M \cup W : l > k, \sigma^{-1}(\mu(\sigma(l))) < k$, and

$\sigma(k) \succeq_{\sigma(l)} u$, for all $u \in P_1^k(\sigma(l))\}$. A strategy profile $s^* = (s_1^*, \ldots, s_{2n}^*)$, where s_j^* denotes the strategy for an individual j, and stable matching μ pair, (s^*, μ), is defined as follows:

$$
\begin{aligned}
s^*_{\sigma(1)} &= \mu(\sigma(1)) \\
&\vdots \\
s^*_{\sigma(k)} &= \begin{cases} \text{accept } v & \text{if } v \succ_{\sigma(k)} w, \text{ for all } w \in P_1^k(\sigma(k)) \cup T^k(\sigma(k)) \\ \text{propose } v & \text{if } v \succ_{\sigma(k)} w, \text{ for all } w \in P_1^k(\sigma(k)) \cup T^k(\sigma(k)) \\ \text{stay single} & \text{otherwise} \end{cases} \\
&\vdots \\
s^*_{2n} &= \begin{cases} \text{accept } v & \text{if } v \in P_1^{2n}(\sigma(2n)) \text{ and } v \succ_{\sigma(2n)} w, \text{ for all } w \in P_1^{2n}(\sigma(2n)) \\ \text{stay single} & \text{otherwise} \end{cases}
\end{aligned}
$$

Before we comment on s^*, let us explain set $T^k(\sigma(k))$. Each individual $\sigma(l)$ in $T^k(\sigma(k))$ moves after $\sigma(k)$, $\mu(\sigma(l))$ moves before $\sigma(k)$, and $\sigma(l)$ has not received a proposal from someone whom $\sigma(l)$ prefers to $\sigma(k)$ until stage k. If s^* is played, the individual who moves in stage 1, $\sigma(1)$, can only propose to his or her μ partner. In stage k, $1 < k < 2n$, the individual who moves in stage k, $\sigma(k)$, can either accept individual v's proposal or propose to individual v, if v is the most preferred individual among those who have proposed to $\sigma(k)$ and among those whose μ partners have moved before v and who like $\sigma(k)$ as much as anyone who has proposed them until stage k. Otherwise, $\sigma(k)$ stays single. The individual who moves in the last stage, $\sigma(2n)$, can either accept the best proposal in hand or stays single. Now we present the lemmas.

Lemma 3 *Given a preference profile $\succ$ and a SM-mechanism, if each individual i plays s_i^* in this mechanism, the matching induced by s^* is the stable matching is μ, for any pair (s^*, μ).*

Lemma 4 *Given a preference profile $\succ$ and a SM-mechanism, for any pair (s^*, μ), s^* is the SPE strategy, if $\succ$ does not contain a ring as $(r_i, \mu(r_i), r_{i+1}, \mu(r_{i+1}), \ldots, r_k, \mu(r_k))$ where $r_{i+1} = \mu'(\mu(r_i))$ for some matching $\mu' \neq \mu$.*

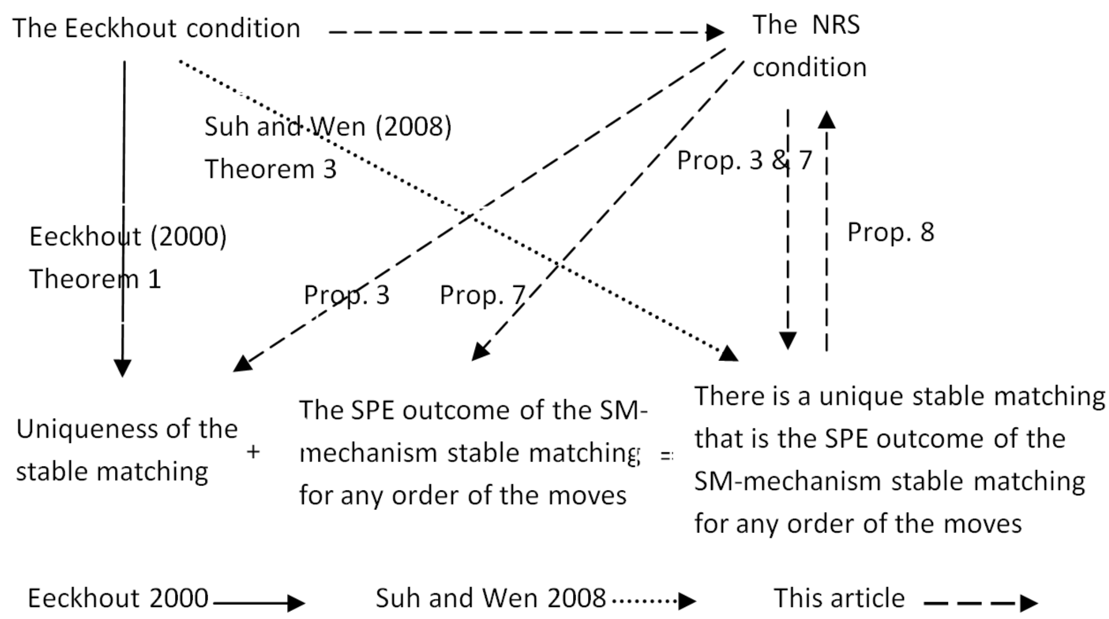

Figure 2. The Eeckhout and NRS Conditions

Proposition 7 *If a preference profile $\succ$ satisfies the NRS condition, the SPE outcome of the SM-mechanism is a stable matching for any order of moves.*

So far, only the sufficiency of the NRS condition has been considered. Our next step is to show that when there is a unique stable matching for a preference profile, the NRS condition is also a necessary condition to implement this stable matching as the SPE outcome of the sequential matching game regardless of how individuals move.

Proposition 8 *For a preference profile $\succ$, if there is a unique stable matching μ and μ is the SPE outcome of the SM-mechanism for any order of moves, then the preference profile $\succ$ satisfies the NRS condition.*

Figure 2 summarizes the results that are considered this section.

Conclusion

One may suppose that the subgame perfect equilibrium outcome of the sequential matching mechanism considered in this paper is a stable matching regardless of the order of individuals. However, as we show in our first example, the stability of the equilibrium outcome is not guaranteed under unrestricted preference domain assumption. Motivated by this fact, we search for "ring" conditions that a preference profile should satisfy to implement stable matchings. The reason for our interest in ring conditions is the relation between rings (or cycles) in a profile and stable allocations, which has gained considerable attention in many different contexts in the literature. We also investigate the conditions previously considered for implementation of stable matchings, and explain these conditions in terms of rings.

In this paper, we answer two questions: 1) Under which "ring" conditions are μ^M, μ^W, and the unique stable matching implementable in the SPE of the SM-mechanism regardless of which order the individuals move in, and 2) how are the ring conditions related to other conditions such as the Eeckhout condition (Eeckhout, 2000) and the alpha conditions (Suh and Wen, 2008) that are previously considered in the context of implementing μ^M,μ^W, and the unique stable matching?

We introduce the NRS and NRMO conditions to answer the first question. The NRS condition is used to guarantee the existence of a unique stable matching and to implement the unique stable matching. That is, we prove that the NRS condition is both necessary and sufficient for there to be a unique stable matching and for this matching to be implementable in the SPE independent of the order of individuals' moves. The NRMO condition is introduced to implement μ^M and is shown that it is a necessary and sufficient condition for μ^M to be the SPE outcome of the men-move-first mechanism.

In order to answer the second question, we first show that the NRMO condition is an equivalent ring condition of the α^M. Second, we show that the NRS condition is a weaker condition for both the existence and implementation of the unique stable matching.

As a future research project, we plan to propose a more general ring condition to implement all stable matchings in the SPE. Our conjecture for such a ring condition is the following: If a preference profile does not contain a ring such that an unstable matching can be obtained from a stable matching μ through this ring when a group of same sex individuals switch their partners in μ, then the SPE outcome of the sequential mechanism is a stable matching no matter which order individuals move in. A future research topic may be to show that such a ring condition is a necessary and sufficient condition to have a stable matching in the SPE.

Another question that we want to ask is the following: Is there any the relation between the order of individual and the equilibrium outcome? In particular, we want to focus on the order of moves of individuals and search for a certain rule to order individuals to implement specific stable matchings.

Appendix.

Lemma 2. If a preference profile $\succ$ satisfies the NRMO condition, then there is at least one man who is matched to his top choice in μ^M.

Proof. On the contrary, suppose that no man is matched to his top choice in μ^M. We show that $\succ$ does not satisfy the NRMO condition.

Take an arbitrary man and rename him as m_1. Note that $w \succ_{m_1} \mu^M(m_1) = w_1$ for some $w \in W$ by assumption. Let w_2 be m_1's top ranked woman and let $\mu^M(m_2) = w_2$. By assumption, $w \succ_{m_2} w_2$ for some $w \in W$. If m_2's top ranked woman is w_1, we have a ring (w_1, m_1, w_2, m_2) and $\succ$ does not satisfy the NRMO condition. Otherwise, let m_2's top ranked woman be w_3 and let $\mu^M(m_3) = w_3$. By assumption, $w \succ_{m_3} w_3$ for some $w \in W$. If m_3's top listed woman is in $\{w_1, w_2\}$. Then, we have a ring of type $(w_1, m_1, w_2, m_2, w_3, m_3)$ if m_3's top listed woman is w_1 or (w_2, m_2, w_3, m_3) if m_3's top listed woman is w_2. The NRMO condition is violated in either case. In an arbitrary step k of this process, we either have a ring by which the condition is violated or m_k's top ranked woman w_{k+1} is not in $\{w_1, \ldots, w_k\}$. Because there is a finite number of individuals, the process must eventually terminate a ring. Hence, the NRMO condition is violated. ■

Proposition 2. A preference profile $\succ$ satisfies the α^M condition if and only if it satisfies the NRMO condition.

Proof. First we prove the sufficiency of the NRMO condition. Suppose that $\succ$ satisfies the NRMO condition. By lemma 2, there is at least one man who is matched to his top choice in μ^M. Rename all men who are matched to their top choices in μ^M arbitrarily as $m_1, \ldots, m_k$, $1 \leq k \leq n$, and let $M_1 = \{m_1, \ldots, m_k\}$. Moreover, rename all women in $\mu^M(M_1)$ such that $\mu^M(m_i) = w_i$ for all $i \in \{1, \ldots, k\}$.

Obtain a reduced preference profile $\succ_1$ by eliminating all individuals in $M_1 \cup \mu^M(M_1)$ from the original profile $\succ$. Observe that $\succ_1$satisfies the NRMO condition because if $\succ_1$does not satisfy the NRMO condition due to a ring, then $\succ$ violates the NRMO condition by the same ring. By lemma 2, there is at least one man who is matched to his top choice in μ^M for $\succ_1$. Rename all men who are matched to their top choices in μ^M in the reduced profile $\succ_1$ arbitrarily as $m_{k+1}, \ldots, m_m$, $k+1 \leq m \leq n$. Let $M_2 = \{m_k, \ldots, m_m\}$. Also rename all women in $\mu^M(M_2)$ such that $\mu^M(m_i) = w_i$ for all $i \in \{k+1, \ldots, m\}$, and obtain a reduced preference profile $\succ_2$ by eliminating all individuals in $\bigcup_{i=1}^{2}\left(M_i \cup \mu^M(M_2)\right)$. The process eventually stops after step t when all individuals are renumbered and eliminated from $\succ_t$. After individuals are renumbered, we have $w_i \succ_{m_i} w$ for all $w \in W \setminus \{w_i\}$, for all $i \in \{1, \dots, k\}$. Moreover, any reduced profile $\succ^k$ coincides with $\succ$ in terms of preference ordering of all individuals in $(M \cup W) \setminus \bigcup_{i=1}^{k} \left(M_i \cup \mu^M (M_i)\right)$, $k = 1, \dots, t$. Thus, for all $m_i \in M$, we have $w_i \succ_{m_i} w_j$, for all $j > i$. Furthermore, if there is a woman w_j such that $w_j \succ_{m_i} w_i$ for some $j < i$, then $m_j \succ_{w_j} m_i$ by the stability of μ^M, for all $m_i \in M$. Thus, $\succ$ satisfies the α^M condition.

Now, we prove the necessity of the NRMO condition. Suppose that $\succ$ satisfies the α^M condition. By Proposition 1, $w_i = \mu^M (m_i)$ for all $i \in \{1, \dots, n\}$. Now, assume to the contrary that $\succ$ does not satisfy the NRMO condition. Consider an arbitrary ring $(w_a, m_a, w_b, m_b, \dots, w_k, m_k)$ and $w_i = \mu^M (m_i)$ for all $i \in \{a, \dots, k\}$. By the α^M condition, $a < k$ because $w_a \succ_{m_k} w_k$. By similar logic, we have $k < \cdots < b < a$; this contradicts $a < k$. ■

Proposition 3. If a preference profile $\succ$ satisfies the NRS condition, then there is a unique stable matching for $\succ$.

Proof. Suppose that there are multiple stable matchings for the preference profile $\succ$. So, the men and women optimal matchings, μ^M and μ^W, are distinct. Let $M' \subseteq M$ be the set of all men such that $\mu^M (m) \neq \mu^W (m)$ for all $m \in M'$. By optimality, $\mu^M (m) \succ_m \mu^W (m)$ for all $m \in M'$. That is, for any $m \in M'$, we have $\mu^W (m) = \mu^M (m')$ for some $m' \in M \setminus \{m\}$ with $\mu (m') \neq \mu' (m^*)$ for any $m^* \neq m$. Because M' is finite, there exists a group of men $\{m_i, \dots, m_k\} \in M'$ who form a ring $\left(\mu^M (m_i), m_i, \mu^M (m_j), m_j, \dots, \mu^M (m_k), m_k\right)$. ■

Proposition 5. There is no men-move-first mechanism such that a man is matched to a woman worse than his μ^M partner in the SPE.

Proof. Suppose to the contrary that there exists a men-move-first mechanism such that a man is matched to a woman worse than his men-optimal partner in the SPE outcome of this mechanism. Let the SPE outcome be a matching $\mu \neq \mu^M$. Consider the first man who is matched to a woman worse than his men-optimal partner in this mechanism. Let this man be m_1 and $\mu^M(m_1)$ be w_1. When it's m_1's turn to move, the reason for m_1's not proposing to w_1 is that w_1 rejects m_1 in favor of a man whom she likes more than m_1 if m_1 proposes to w_1. Consider the subgame starting with m_1's proposal to w_1. By the stability of μ^M, the man who proposes to w_1 in the subgame where m_1 proposes to w_1 likes his μ^M partner more than w_1. Hence, in the subgame where m_1 proposes to w_1, there is at least one man who proposes to a woman worse than his μ^M partner according to his equilibrium strategy, and this man moves after m_1, by assumption. Let the first man who proposes to a woman worse than his μ^M partner according to his equilibrium strategy in the subgame where m_1 proposes to w_1 be m_2. Let $\mu(m_2)$ be w_2. Note that m_2 is the *first* man who proposes to a woman worse than w_2 and is not necessarily the man who proposes to w_1. Hence, every man who moves after m_1 and before m_2 has either proposed to his μ^M partner or a more preferred woman according to his equilibrium strategy in the subgame where m_1 proposes to w_1. Thus, when it is m_2's turn to move, if w_2 has already a proposal from another man, it must be from a man to whom w_2 prefers m_2. Consider the subgame starting with m_2's proposal. Similarly, in this subgame, when it is m_2's turn to move, the reason for m_2's not proposing to w_2 is that if he does so, w_2 rejects m_2 in favor of another man, whom she likes more than m_2. By the stability of μ^M, the man who proposes to w_2 likes his μ^M partner better than w_2. Hence, there exists at least one man who proposes to a woman worse than his mate under μ^M according to his equilibrium strategy in the subgame where m_1 proposes w_1 and m_2 proposes to w_2. In any step j of this process, we pick the first man who proposes to a woman worse than his μ^M partner, and rename him as m_j. Then m_j proposes to $\mu^M(m_j) = w_j$ instead of proposing to a woman worse than w_k.

Consider an arbitrary step k of the process which starts with m_k's proposal. Observe that w_k has not received a proposal from a man whom she prefers to m_k. If w_k has already an offer in her hand, then it must be from a man to whom she prefers m_k.

Because there are finite number of men, this process must eventually end in a step, say t, in which m_t proposes to $\mu^M(m_t) = w_t$, and there is no other man who moves after m_t and proposes to a woman worse than his μ^M partner. On the other hand, we know that m_t is the first man who would propose to a woman worse than $\mu^M(m_t) = w_t$ after step $t-1$ according to his equilibrium strategy. Hence, it must be the case that if he m_t proposes to w_t, he will be rejected in favor of another man, contradiction. ■

Proposition 6. A preference profile $\succ$ satisfies the NRMO condition if and only if the SPE outcome of the men-move-first mechanism is μ^M regardless the order of moves.

Proof. By Proposition 4, the α^M condition is a sufficient condition for μ^M to be the SPE outcome. By Proposition 2, the NRMO condition is also a sufficient condition for μ^M to be the SPE outcome. Hence, we only need to show the necessity of the NRMO condition.

Suppose that $\succ$ does not satisfy the NRMO condition. Hence, $\succ$ contains a ring $(w_k, m_k, \ldots, w_t, m_t)$ such that $\mu^M(w_i) = m_i \ \forall i = k, \ldots, t$. We show that there is an order of moves such that SPE outcome of the mechanism is not μ^M.

Let $(w_1, m_1, \ldots, w_t, m_t)$ be a ring such that $\mu^M(w_i) = m_i \ \forall i = 1, \ldots, t$. If there is more than one ring of the same type, consider the ring $(w_k, m_k, \ldots, w_t, m_t)$ with $w_{i+1} \succ_{m_i} w$ for all $w \in \{w_k, \ldots, w_t\}$, for every man m_i in $\{m_k, \ldots, m_t\}$. Such a ring exists because if

$w_j \succ_{m_i} w_{i+1}$ for some $w_j \in \{w_k, \ldots, w_t\}$, then we can obtain another ring of the same type by eliminating individuals $w_{i+1}, m_{i+1}, \ldots, w_{j-1}, m_{j-1}$ from the original ring.

Let $M_1 = \{m_1, \ldots, m_t\}$ be the set of all men in the ring. Suppose that all men in $M \backslash M_1$ move before all men in M_1 and each man m_i in M_1 is followed by m_{i+1}, $i = 1, \ldots, t, (\mathrm{mod}\, t)$. In the first case that we need to consider, there exists a man $m' \in M \backslash M_1$ who proposes to a woman $w'' \neq \mu^M(m)$. So, the SPE outcome is not μ^M, and the proof is done.

In the second case that we need to consider, each man in $M \backslash M_1$ proposes to his μ^M partner. We claim that that each man m in M_1 proposes to a woman in $\mu^M(M_1)$. This is because if a man m in M_1 proposes to a woman w' in $W \backslash \mu^M(M_1)$, w' refuses m in favor of $\mu^M(w')$ because $\mu^M(w') \succ_{w'} m$ by the stability of μ^M. Hence, each man m in M_1 proposes to a woman in $\mu^M(M_1)$. Suppose that each man m_i in M_1 proposes to $w_{i+1}, i = 1, \ldots, k, (\mathrm{mod}\, k)$. We claim that no man can do better by proposing to another woman. Because $w_{i+1} \succ_{m_i} w$ for all $w \in \mu^M(M_1)$, m_i cannot do better by proposing to another woman if w_{i+1} accepts m_i's proposal. Moreover, because the only proposal in w_{i+1}'s hand is from m_i, w_{i+1} accepts m_i's offer. Hence, m_i cannot do better by proposing to another woman. So, the SPE outcome is not μ^M. The proof is complete. ■

Lemma 3. Given a preference profile $\succ$ and a SM-mechanism, if each individual i plays s_i^* in this mechanism, the matching induced by s^* is the stable matching is μ, for any pair (s^*, μ).

Proof. We use an induction proof. Suppose that each individual plays s_i^*. We show that each individual i is matched with $\mu(i)$, i.e., proposes to $\mu(i)$ and is accepted or accepts $\mu(i)$'s proposal when i plays s_i^*. Individual $\sigma(1)$ proposes to $\mu(\sigma(1))$ according to $s_{\sigma(1)}^*$. Assume that in stage k, $\sigma(k)$ plays $s_{\sigma(k)}^*$ and is matched with $\mu(\sigma(k))$. Consider step

$k+1$. We know that $\sigma(k+1)$ has not received a proposal from an individual $\sigma(j)$ with $j<k$ and $\sigma(j) \neq \mu(\sigma(k))$, because every $\sigma(j)$ with $j<k$ and $\sigma(j) \neq \mu(\sigma(k))$ is matched with $\mu(\sigma(j))$. Now, we consider all potential partners who move in subsequent stages. According to $s^*_{\sigma(k+1)}$, if $\sigma(k+1)$ prefers $\mu(\sigma(k+1))$ more than all potential partners who move in subsequent stages, then $\sigma(k+1)$ should be matched with $\mu(\sigma(k+1))$. Consider an arbitrary individual $\sigma(l)$ such that $\sigma(l) \succ_{\sigma(k+1)} \mu(\sigma(k+1))$ and $l>k$. For $\sigma(l)$ to be a potential partner of $\sigma(k+1)$, $\sigma(l)$ must be in $T^{k+1}(\sigma(k+1))$. That is, $\mu(\sigma(l))$ must move before $\sigma(k+1)$ and $\sigma(k+1) \succ_{\sigma(l)} u$ for all $u \in P_1^k(\sigma(l))$. By our initial assumption, $\mu(\sigma(l))$ must have already proposed to $\sigma(l)$, hence, $\sigma(k+1) \not\succ_{\sigma(l)} u$ for all $u \in P_1^k(\sigma(l))$. Because $\sigma(l)$ is chosen arbitrarily, we conclude that $T^{k+1}(\sigma(k+1)) \setminus \{\mu(\sigma(k+1))\}$ is empty. That is, $\sigma(k+1)$ should be matched with $\mu(\sigma(k+1))$ if he or she play $s^*_{\sigma(k+1)}$. ■

Lemma 4. Given a preference profile $\succ$ and a SM-mechanism, for any pair (s^*, μ), s^* is the SPE strategy, if $\succ$ does not contain a ring as $(r_i, \mu(r_i), r_{i+1}, \mu(r_{i+1}), \ldots, r_k, \mu(r_k))$ where $r_{i+1} = \mu'(\mu(r_i))$ for some matching $\mu' \neq \mu$.

Proof. Given a preference profile $\succ$ and a SM-mechanism, we know that the matching induced by s^* is the stable matching μ by lemma 3. We show that if s^* is not the SPE strategy, then there is a ring $(r_i, \mu(r_i), r_{i+1}, \mu(r_{i+1}), \ldots, r_k, \mu(r_k))$ where $r_{i+1} = \mu'(\mu(r_i))$ for some matching $\mu' \neq \mu$. Suppose that s^* is not the SPE strategy. Hence, there exists a subgame in which when an individual $v \in M \cup W$ is the first one to move, v deviates from s^*_v and can be matched to an individual ranked higher than $\mu(v)$ in v's preference list when all $i \in M \cup W \setminus \{v\}$ plays s^*_i. Given the order of moves of individuals, assume without loss of generality that m_1 is the person who deviates from s^* and let the matching outcome when m_1 deviates be μ'. Let $\mu(m_1) = w_1$ and $\mu'(m_1) = w_2$. Because each individual i who moves before m_1, plays s^*_i, each individual i who moves before m_1 proposes to or accepts an offer from $\mu(i)$ by lemma 3. Hence, w_2 cannot be one of those individual who move before m_1 because otherwise she would have proposed to m_2. That is, w_2 moves after m_1. Because w_2 is matched to m_1, it must be the case that the best that w_2 can do according to $s^*_{w_2}$ is to accept m_1's offer. Hence, m_1 must be the best man among those in $P_1^{\sigma^{-1}(w_2)}(w_2) \cup T^{\sigma^{-1}(w_2)}(w_2)$. This is possible only if $\mu(w_2) = m_2$ has moved before w_2 and has not proposed to w_2. Hence, there is a women in $P_1^{\sigma^{-1}(m_2)}(m_2) \cup T^{\sigma^{-1}(m_2)}(m_2)$, say w_3, such that $w_3 \succ_{m_2} w_2$, and m_2 must be the best man among those in $P_1^{\sigma^{-1}(w_3)}(w_3) \cup T^{\sigma^{-1}(w_2)}(w_3)$. If $w_3 = w_1$, this procedure terminates in this step and we have a ring (w_1, m_1, w_2, m_2) with $w_2 = \mu'(m_1)$ and $w_1 = \mu'(m_2)$, hence, the proof is done. If $w_3 \neq w_1$, there is a women in $P_1^{\sigma^{-1}(m_3)}(m_3) \cup T^{\sigma^{-1}(m_3)}(m_3)$, say w_4, such that $w_4 \succ_{m_3} w_3$. and m_3 must be the best man among those in $P_1^{\sigma^{-1}(w_4)}(w_4) \cup T^{\sigma^{-1}(w_4)}(w_4)$. If $w_4 \in \{w_1, w_2\}$, the procedure terminates in this step, and we have a ring which is $(w_1, m_1, w_2, m_2, w_3, m_3)$ with $w_2 = \mu'(m_1)$, $w_3 = \mu'(m_2)$, and $w_1 = \mu'(m_3)$ if $w_4 = w_1$, or (w_2, m_2, w_3, m_3) with $w_3 = \mu'(m_2)$, and $w_2 = \mu'(m_3)$ if $w_4 = w_2$. In either case, the proof is done.

This procedure must either terminate in some step where w_k is considered or there exists a woman w_{j+1} with $w_{j+1} = \mu'(m_j)$. If the procedure terminates, we obtain a ring $(w_i, \mu(w_i), \ldots, w_k, \mu(w_k))$ with $w_{j+1} = \mu'(\mu(w_j))$ for all $j = i, \ldots k, (\operatorname{mod} j)$. By the fact that there are finite number of individuals, the procedure must eventually terminate and we obtain a ring $(w_i, \mu(w_i), \ldots, w_k, \mu(w_k))$ with $w_{j+1} = \mu'(\mu(w_j))$ for all $j = i, \ldots k, (\operatorname{mod} j)$. Hence, the proof is done. ■

Proposition 7. If a preference profile $\succ$ satisfies the NRS condition, the SPE outcome of the SM-mechanism for any order of moves is a stable matching.

Proof. Suppose that $\succ$ satisfies the NRS condition. By Proposition 3, we know that there exists a unique stable matching μ. By lemma 4, strategy s^* is the SPE strategy for this mechanism given the pair (s^*, μ). By lemma 3, this SPE outcome is μ. Because the SPE outcome of any SM-mechanism is unique, we conclude that the unique stable matching μ for $\succ$ is the SPE outcome of the SM-mechanism. ■

Proposition 8. For a preference profile $\succ$, if there is a unique stable matching μ and μ is the SPE outcome of the SM-mechanism for any order of moves, then the preference profile $\succ$ satisfies the NRS condition.

Proof. Suppose that the preference profile $\succ$ does not satisfy the NRS condition. If there are multiple stable matchings, then the proof is complete.

Suppose that there is a unique stable matching μ. Hence, $\mu = \mu^M = \mu^W$. Because $\succ$ does not satisfy the NRS condition, there exists a ring of type $(w_l, m_l, \ldots, w_t, m_t)$ or $(m_l, w_l, \ldots, m_t, w_t)$, and $\mu(m_i) = w_i$ where μ is the unique stable matching. Without loss of generality, suppose that the ring is of type $(w_l, m_l, w_m, m_m, \ldots, w_t, m_t)$. Hence, $\succ$ does not satisfy the NRMO condition. By Proposition 6, we know that there exists an order of moves in the men-move-first mechanism such that the SPE outcome is matching μ' such that $\mu' \succ_{M'} \mu$ for a nonempty subset of men M'. Hence, the proof is complete. ■

CHAPTER III

A SOLUTION FOR THE ROOMMATE PROBLEM AND A SEQUENTIAL MATCHING MECHANISM

Introduction

In a roommate problem, there is a set of individuals and each individual has a preference over all other individuals, including oneself. A roommate problem is the problem of identifying pairs (i.e., individuals who share a room) and singletons (i.e., individuals who stay alone). The outcome is referred as a *matching.* A matching is *individually rational* if there is no individual who would stay alone rather than sharing a room with his mate in the matching. Two individuals *block* the matching if they prefer sharing a room with each other rather than sharing the room with their current mates in the matching. A matching is *(pairwise) stable* if it is individually rational and it is not blocked by a pair. A roommate problem may not have a stable solution. A roommate problem has a solution or is *solvable* if there is a stable matching, *unsolvable* otherwise. Tan (1991) shows that a roommate problem has a stable matching if and only if there is a *stable partition* without an *odd ring.* A stable partition is a partition of the set of individuals such that each set in a stable partition is either a ring, or a pair of mutually acceptable individuals, or a singleton, and the partition satisfies stability between sets and also within each set. A ring is an ordered subset of individuals such that each individual prefers the subsequent individual to the preceding individual. A ring is an odd (even) ring if there is an odd (even) number of people in the ring.

Inarra et al (2008) introduce $P - stability$ as a solution concept such that the existence of a $P - stable$ matching is guaranteed. The $P - stability$ concept is based on Tan's (1991) stable partitions. For a stable partition P, a $P - stable$ matching is one that matches each individual to his subsequent or preceding individual in any set of the stable partition P, and one individual is unmatched (i.e., stays alone) if the set is an odd ring or singleton.

The paper has two parts: In the first part, we focus on a solution concept for the roommate problem, and in the second part we study a sequential matching mechanism for the roommate problem and analyze the subgame perfect equilibrium (SPE) of this mechanism. In the first part, we propose a solution concept, which is called $RP - stability$ (reduced preference profile $P - stability$). We introduce a procedure, which is called the RP procedure, and we define an $RP - stable$ matching as the outcome of the RP procedure. Our motivation for introducing the $RP - stability$ is that the number of matched individuals in any $RP - stable$ matching is always greater than or equal to the number of those in any $P - stable$ matching. Moreover, whenever a Pareto improvement is possible based on a $P - stable$ matching, there is an $RP - stable$ matching which Pareto dominates the corresponding $P - stable$ matching.

The RP procedure starts with a stable partition[1] in a roommate problem. A *reduced preference profile* is obtained by *cutting* an odd ring in the stable partition, i.e., by eliminating each individual in an arbitrary pair of succeeding and preceding individuals in the odd ring from each other's preference list. A stable partition with respect to a reduced profile is called a *reduced stable partition*. Once an odd ring is cut and a reduced profile is obtained, the procedure continues cutting one odd ring in each reduced profile stable partition until a profile for which there is no stable partition with an odd ring is reached.

[1]The procedure chooses odd rings according to a protocol which is explained in details when the protocol is presented.

Such a reduced profile is called a *final reduced profile*. A stable partition with respect to a final reduced profile is a *final stable partition*. When a final reduced profile is reached, the procedure considers the most recent step where there is an odd ring in which there are still some individuals who have not been used to cut the odd ring, and the procedure continues by cutting that odd ring. The procedure eventually stops when there is no such step to consider. An $RP-stable$ matching is a $P-stable$ matching with respect to a final stable partition that is obtained through the procedure.

Given a profile, when there is a stable partition with an odd ring, there is always an individual i in the odd ring who is not matched to one of his neighbors (the succeeding and the preceding individuals) in any matching. The RP procedure cuts odd rings to capture the idea that an individual cannot be matched to either of his neighbors in an odd ring. In all $P-stable$ matchings in which individual i is not matched to one of his neighbors in the odd ring, i would be unmatched. But, individual i should not be unmatched if there is another individual j such that both i and j prefer being matched to each other rather than being unmatched. In such a situation, i and j would be matched in an $RP-stable$ matching and observe that matching i and j would be a Pareto improvement, and the number of matches is greater when i and j are matched.

Obviously, if there is no partition with an odd ring, then the set of $RP-stable$ matchings is equal to the set of $P-stable$ matchings. Hence, when the roommate problem is solvable, $RP-stability$ and $P-stability$ are equivalent. Moreover, because the set of $P-stable$ matchings and the set of stable matchings are equivalent when the problem is solvable (Inarra et al 2008), $RP-stability$ and stability are also equivalent concepts for solvable roommate problems. However, $P-stability$ and $RP-stability$ concepts are unrelated, i.e., neither of them implies the other, for unsolvable roommate problems.

We focus on a roommate problem when individuals move sequentially according to a previously specified order in the second part of the paper. In the mechanism, an individual can choose one of the three possible actions when it is his turn to move: (i) Accept an offer from another individual who has proposed to him at a previous stage, or (ii) propose to a potential individual who moves at a later stage, or (iii) choose to remain single. We are interested in analyzing the subgame perfect equilibrium the (SPE) outcome of the sequential matching mechanism[2].

In the second part, our first objective is to show that the RP procedure can be used to identify the set of all potential SPE outcomes. Our second objective is to identify a sufficient condition which guarantees the stability of the SPE outcome regardless of the order of individuals.

Odd rings play an important role in the sequential game as they do for the existence of stable matchings in roommate problems. Because of an odd ring, an individual may not be matched to either of his neighbors. The same individual may not be matched to his best alternative after his neighbors just because i and his next best alternative, too, are in an odd ring. Obviously, an odd ring, which contains i and his next best alternative, would only appear in a reduced stable partition, but not appear in the original stable partition. Because the RP procedure captures such odd rings, too, by considering the reduced profiles, it is able to identify all potential SPE outcomes. By analyzing the relation between the SPE and $RP-stability$, we are able to identify a sufficient condition to guarantee the stability of the SPE outcome regardless of the order of moves in solvable roommate problems.

The roommate problem was introduced by Gale and Shapley (1962). They show that a stable matching may not exist in a roommate problem, but there is a stable matching

[2]Hereafter, when we say the SPE outcome, we mean the SPE outcome of the sequential mechanism.

for every marriage problem, which is a special case of the roommate problem. There are many studies that investigate stable matchings in both roommate and marriage problems and the conditions that can guarantee the existence of a stable matching in roommate problems.[3] Tan (1990) proposes a necessary and sufficient condition for the existence of a stable matching in roommate problems with strict preferences. Chung (2000) introduces the "no odd ring" condition and proves that the no odd ring condition is a sufficient condition for the existence of a stable matching in roommate problems with weak preferences. Abeledo and Isaak (1991) show that the set of mutually acceptable pairs can be viewed as a graph, and a stable matching exists under any preference profile if and only if the acceptability graph is bipartite.

Because the existence of a stable matching is not guaranteed in a roommate problem, one may ask whether we can find matchings which always exist when we relax the stability concept. Tan (1990) proposes an algorithm that finds a *maximum stable matching*: A matching with maximum number of disjoint pairs of persons such that these pairs are stable among themselves. Abraham et al. (2005) focus on finding *almost stable matchings*: Matchings that admit fewer blocking pairs. In Klaus et al. (2008), it is shown that the set of *stochastically stable matchings* coincides with the set of *absorbing matchings*. They assert that the set of absorbing matchings is the only solution concept that is core consistent and shares the stochastic stability characteristic with the core. Inarra et al. (2008, 2010) focus on solution concepts in unsolvable roommate problems such as $P-stability$ and absorbing sets. They prove that absorbing sets always exist and if a stable matching exists, the absorbing set coincides with the set of stable matchings. Moreover, an absorbing set satisfies the *outer stability*; i.e., all matchings that are not in the solution set are dominated

[3] Gale Shapley (1962), Irving (1985), Roth and Sotomayor (1990) have also investigated the stability concept in marriage problems.

by matchings in the solution set. They also show that the solutions that Tan (1990) and Abraham et al. (2006) propose do not satisfy the outer stability property.

In this paper, we also study a sequential matching game. All existing studies focus on the implementation of stable matchings in solvable matching problems, which is reasonable, because the existence of a stable matching in roommate problems is guaranteed. Yet, these studies do not provide much intuition about the characteristics of the equilibrium outcome in a sequential game in a broader context, i.e., the roommate problem. As far as we know, this is the first study that focuses on a sequential game in a roommate problem. In terms of implementation of stable matchings, two important studies are by Alcalde and Romero-Medina (2000, 2005) and Sotomayor (2003), who study the implementation of the core correspondence and/or the set of stable solutions of the college admissions problem. Suh and Wen (2008) identify sufficient conditions to implement stable matchings in the equilibrium of a sequential marriage problem, which we mimic in this paper. Some other important studies that are interested in the implementation of stable solutions or the core of the problem are Alcalde (1996), Shin and Suh (1996), and Kara and Sonmez (1996).

In section 2, we introduce the roommate problem. We also present the concepts of stable partition and $P-stability$, and provide some preliminary results in section 2. In section 3, we introduce the RP procedure and $RP-stability$. We discuss the stable partition, $P-stability$ and $RP-stability$ concepts, and show how these concepts are related. In Section 4, we introduce the sequential game and prove our results. Section 5 concludes.

The roommate problem

The set of individuals is defined by a finite set $I = \{i_1, \ldots, i_n\}$. A roommate problem is a pair $\left(I, (\succeq_i)_{i \in I}\right)$ where for each individual $i \in I$, $\succeq_i$ denotes i's complete and transitive preference relation defined over I. A preference profile is the collection of all individuals' preferences, i.e., $\succeq = (\succeq_i)_{i \in I}$. We denote the strict preferences of $i \in I$ with $\succ_i$. A matching is a function $\mu : I \longrightarrow I$ such that each individual is matched to at most one individual. If $\mu(i) = j$, then $\mu(j) = i$, and we say that i is matched to j in μ. If $\mu(i) = i$, then i is unmatched in μ.

In this paper, we assume that individuals have strict preferences. When $j \succ_i k$, i strictly prefers j to k if $i \neq j \neq k \neq i$. Individual i prefers being alone rather than being matched to k, that is, k is *unacceptable* for i if $i = j \neq k$. Individual j is acceptable for i if $i = k \neq j$. A matching μ is *individually rational* if no individual i blocks μ, i.e., $\mu(i)$ is acceptable to i. A pair $\{i, j\} \subseteq I$ blocks μ if $j \succ_i \mu(i)$ and $i \succ_j \mu(j)$. A matching μ is stable if it is individually rational and if there is no pair $\{i, j\} \subseteq I$ that blocks μ. For a set of individuals $S \subseteq I$, and any two matchings μ and μ', we say that each individual in S prefers μ to μ' if $\mu \succ_S \mu'$. For two matchings μ and μ', μ *Pareto dominates* μ' if $\mu \succeq_I \mu'$ and $\mu \succ_S \mu'$ for a nonempty set $S \subseteq I$.

A stable matching in a roommate problem may not exist. A roommate problem is *solvable* if a stable matching exists; otherwise it is *unsolvable.* A marriage problem is a special case of a roommate problem. Gale and Shapley (1962) show that a stable matching always exists in a marriage problem. Hence, a marriage problem is considered as a solvable roommate problem.

Preliminaries

This section presents the definitions of stable partitions (Tan, 1991) and $P-$ *stability* (Inarra et al., 2008). We also present some preliminary results that we benefit throughout the paper. The first result is related to solvable roommate problems.

Theorem 1 (Roth) *In a marriage problem with strict preferences, the set of people who are single is the same for all stable matchings.*[4]

Given a roommate problem $\left(I, (\succeq_i)_{i \in I}\right)$, an ordered subset of individuals $A = \{a_1 \ldots, a_k\} \subseteq I$ is a *ring* if $k \geq 3$ and $a_{i+1} \succ_{a_i} a_{i-1} \succ_{a_i} a_i$ for all $i = 1, \ldots, k \pmod{k}$.[5] A ring is an *odd ring* if the number of individuals in the ring is odd, and is an *even ring* if the number of individuals in the ring is even. The set A is a *pair of mutually acceptable individuals* if $k = 2$ and $a_{i-1} \succ_{a_i} a_i$ for all $i = 1, 2$. Set A is a *singleton* if $k = 1$. When we say that a set is an *odd set*, we mean that the number of individuals in the set is odd, i.e., the set is either a singleton or an odd ring. A *partition* P of a set $S \subseteq I$ is a collection of disjoint subsets of S.

A *stable partition* P of I is such that

(i) For all $A \in P$, the set A is a ring or a mutually acceptable pair of individuals or a singleton, and

(ii) For any set $A = \{a_1, \ldots, a_k\} \in P$ and $B = \{b_1, \ldots, b_l\} \in P$, if $b_j \succ_{a_i} a_{i-1}$ with $b_j \neq a_{i+1}$, then $b_{j-1} \succ_{b_j} a_i$ for all $i = 1, \ldots, k$ and $j = 1, \ldots, l$.

In the definition above, we do not necessarily have $A \neq B$. The case where $A \neq B$ can be thought as the stability between sets in the stable partition. The case where $A = B$ can be thought as the stability within the set. The second condition can also be written as follows: For any $a_i \in I$, if there exists an individual a_j such that $a_j \neq a_{i+1}$ and $a_j \succ_{a_i} a_{i-1}$, then $a_{j-1} \succ_{a_j} a_i$ for all $i = 1, \ldots, n$ and $j = 1, \ldots, n$.

The following results are by Tan (1991). The first three of these results are about the characteristics of stable partitions. The last result provides a necessary and sufficient

[4] Theorem 2.22 in Roth and Sotomayor, 1990.

[5] Hereafter we omit *modulo* k.

condition to have a stable matching in a roommate problem with strict preferences. Then, we introduce the $P-stability$ concept proposed by Inarra et al. (2008).

Theorem 2 *For any roommate problem $(I, (\succ_i)_{i \in I})$,*

1. *There exists at least one stable partition.*
2. *Any two stable partitions have exactly the same odd sets.*
3. *Each even ring of a stable partition can be broken into pairs of mutually acceptable individuals preserving stability.*
4. *There is no stable matching if and only if there exists a stable partition with an odd ring.*

Definition 2 *Given a stable partition P, μ is a $P-stable$ matching if for each $A = \{a_1, \ldots, a_k\} \in P$, $\mu(a_i) \in \{a_{i+1}, a_{i-1}\}$ for all $i = 1, \ldots, k$ except for a unique j such that $\mu(a_j) = a_j$ if A is an odd set.*

Because a stable partition exists for any roommate problem, a $P-stable$ matching exists, too. By the second result of Theorem 2, without loss of generality, we only consider stable partitions which do not contain even rings throughout the paper.

In a $P-stable$ matching μ, exactly one individual in every odd set is unmatched. Hence, the number of unmatched individuals in a $P-stable$ matching is equal to the number of odd sets, i.e., the total number of odd rings and singletons. Suppose that an individual a_i in an odd ring is not matched to either of his neighbors in the ring in a $P-stable$ matching μ, whereas all other individuals in the odd ring are matched to their neighbors. In such a case a_{i-1} and a_i would prefer each other to their current mates in μ, but $\{a_{i-1}, a_i\}$ is not considered as a blocking pair in term of $P-stability$. We provide an example about stable partitions and $P-stability$.

Example 3 *A preference profile with 6 people is given below:*

1	**2**	**3**	**4**	**5**	**6**
–	–	–	–	–	–
2	3	1	2	6	5
3	1	2		4	
	4	6			

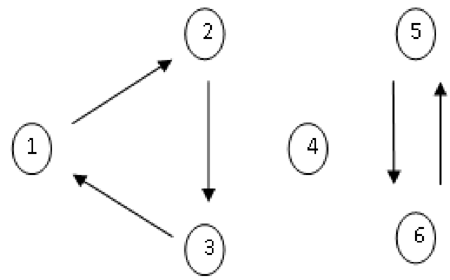

Figure 3. Stable Partition

The only stable partition is $P = \{\{1,2,3\},\{4\},\{5,6\}\}$ *where* $A_1 = \{1,2,3\}$ *is an odd ring,* $A_2 = \{4\}$ *is a singleton set, and* $A_3 = \{5,6\}$ *is a pair of mutually acceptable individuals. Figure 3 represents the stable partition* P*.The* $P - stable$ *matchings are:*

$$\mu_1 = \{\{1,2\},\{3\},\{4\},\{5,6\}\}$$
$$\mu_2 = \{\{1\},\{2,3\},\{4\},\{5,6\}\}$$
$$\mu_3 = \{\{1,3\},\{2\},\{4\},\{5,6\}\}$$

In example 3, for any $P - stable$ matching, there are two individuals in the odd ring who block the matching. For example, $\{2,3\}$ block μ_1, $\{1,3\}$ block μ_2, and $\{1,2\}$ block μ_3. Consider the $P - stable$ matching μ_3 in example 3. Pairs $\{1,2\}$ and $\{2,4\}$ are not considered as blocking pairs for μ_3 in terms of $P - stability$. Individuals 1 and 2 are in an odd ring in P, but 2 and 4 are not. Moreover, matching 2 and 4 is a Pareto improvement. Consider matching $\mu' = \{\{1,3\},\{2,4\},\{5,6\}\}$. The only blocking pair for μ' is $\{1,2\}$. Moreover, the number of individuals who are matched is greater in μ' than that in μ_3. Matching μ' is an $RP - stable$ matching. As in this example, two individuals who do not involve in an odd ring cannot form a blocking pair in terms of $RP - stability$. We introduce the $RP - stability$ concept in the next section and show how $RP - stability$ is different from $P - stability$.

The RP procedure and RP-stability

We introduce the RP procedure and $RP-stability$. We show that $P-stability$, and $RP-stability$ are not related unless the problem is solvable. Moreover, the number of matched individuals in an $RP-stable$ matching is greater than or equal to the number of matched individuals in any $P-stable$ matching. It is also true that if the problem is solvable, stability, $P-stability$, and $RP-stability$ are equivalent.

First, we introduce additional definitions. Given a profile $\succ$ and a stable partition P, when we *cut* an odd ring $A=\{a_1,\ldots,a_k\}$ by leaving a_i out, we erase a_i in a_{i-1}'s preference lists and a_{i-1} in a_i's list. When we cut an odd ring $A=\{a_1,\ldots,a_k\}$ by leaving a_i out, we say "we cut A by a_i" or "we cut $a_i \in A$".

After an odd ring is cut, we obtain a *reduced preference profile* $\succ'$. We obtain a stable partition P' with respect to the reduced preference profile $\succ'$.[6] Such a stable partition is called a *reduced stable partition.* A reduced preference profile for which no reduced stable partition has an odd ring is called a *final preference profile.* A stable partition with respect to a final profile is called a *final stable partition.*

An odd ring in a stable partition P may not appear in a reduced stable partition. An odd ring $A=\{a_1,\ldots,a_k\}$ in a stable partition is an *independent* odd ring if for all a_i in A, if $b_i \succ_{a_i} a_{i-1}$ for some $b_i \in I$, then b_i is not in an odd ring in P.[7][8] If, for some $c_i \in I$, we have $c_i \succ_{b_{i-1}} b_i$ or b_{i-1} and b_i are neighbors in an odd ring A' with $c_i \in A'$, then c_i is not in an odd ring in P. Eventually, we have an individual h_i, who is not in an odd ring in P, and there is no $l_i \in I$ with $l_i \succ_{h_{i-1}} h_i$, or h_{i-1} and h_i are neighbors in an odd ring A'' with $l_i \in A''$. An odd ring which is not independent is a *dependent* odd ring. We discuss

[6] Hereafter, we say a stable partition P' whenever it is clear that the partition P' is stable with respect to $\succ'$.

[7] Note that the subscript i is not necessarily the same number when used as the subscript of different letters. For example, for two individuals a_i and b_i, if a_i is a_1, b_i is not necessarily b_1.

[8] Note that, by the stability of P, $b_{i-1} \succ_{b_i} a_i$, and so on so forth.

some properties of independent odd rings after we introduce the RP procedure. Consider the following example:

Example 4

1	**2**	**3**	**4**	**5**	**6**
4	3	1	5	6	4
2	1	2	6	5	5
3			1		

The unique stable partition is $P = \{\{1,2,3\},\{4,5,6\}\}$. *The odd ring* $\{4,5,6\}$ *is an independent odd ring, while*$\{1,2,3\}$ *is dependent because* $4 \succ_1 3$ *and* 4 *is in an odd ring.*

We introduce a procedure, the RP procedure, to obtain $RP-stable$ matchings. A reduced profile is obtained in each step of the procedure by cutting an odd ring in a stable partition that is obtained in the previous step. In each step, if there is an independent ring in a reduced stable partition, the procedure ignores the dependent rings in that partition, i.e., the procedure cuts only the independent rings in that partition. If there is no independent ring, then the procedure cuts dependent rings.

The RP Procedure Given a profile $\succ$, let $P = \{A, \ldots, H\}$ be a stable partition.

Step 1 If there exists an independent odd ring in P, choose an arbitrary independent odd ring. Choose an arbitrary dependent odd ring otherwise. Let the chosen odd ring be A. Obtain a reduced preference profile $\succ^1$ by cutting $a_1 \in A \in P$.[9] If $\succ^1$ is a final profile (a profile for which there is no stable partition with an odd ring), obtain all $P^1 - stable$ matchings for any final stable partition P^1 and continue step k. Otherwise, continue step 2.

Step 2 Obtaining the reduced profile $\succ^2$ by cutting

$$a_1^1 \in A^1 = \left\{a_1^1, a_2^1 \ldots, a_j^1\right\} \in P^1.$$

[9] Note that in each step j of the procedure, all individuals are renumbered according to how they are partitioned in some partition that is obtained in step j. We use superscript j for individuals, sets, partitions, and profiles to indicate step j.

Repeat *step* 1.

...

Step k Let step k be the first step that a final profile $\succ^k$ is reached by cutting

$$a_1^{k-1} \in A^{k-1} \left\{ a_1^{k-1}, a_2^{k-1}, \ldots, a_j^{k-1} \right\} \in P^{k-1}.$$

Obtain all $P^k - stable$ matchings for any final stable partition P^k.

Step $k+1$ Let j be the most recent step where an odd ring A^j in which there are still some individuals who have not been used to cut the odd ring is cut. Repeat step 1 by cutting A^j.

...

Step t The process stops when we cut all individuals in any odd ring that is obtained in any step of the procedure (excluding the dependent odd rings if there is an independent odd ring in the same reduced stable partition).

Definition 3 *Given a preference profile, a matching μ is $RP - stable$ if μ is obtained by the RP procedure, i.e., is a $P - stable$ matching with respect to a final reduced profile in the RP procedure.*

We provide an example in the Appendix to show how to obtain $RP - stable$ matchings by the RP procedure. Note that an $RP - stable$ matching can be obtained more than once in different steps of the procedure. Because the purpose of the paper is not to come up with an efficient algorithm to obtain all $RP - stable$ matchings, but to define $RP - stability$, a modification to overcome this problem is not considered in this paper.[10]

We now discuss some properties of independent rings. An independent odd ring $A \in P$ can also be defined as an odd ring that appears in any reduced stable partition in

[10] McVitie and Wilson (1971) imposes further rules on the breakmarriage operation to overcome the problem of obtaining the same stable solution many times.

any step of the procedure until $a \in A \in P^j$ is cut in some step of the procedure. That is, starting from P, no matter how and which odd ring except A is cut in any reduced stable partition, A always appears as an odd ring in any reduced stable partition until $a \in A$ is cut for some $a \in A$ in some step of the procedure. No individual a_i in the independent odd ring A, would have neighbors other than a_{i-1} and a_{i+1} unless A is cut. Independent rings are "independent" of other odd rings in this sense.

Because an independent odd ring appears in any reduced stable partition no matter how other odd rings are cut (until the independent odd ring is cut), if there are more independent odd rings in some step of the procedure, it does not matter which independent ring is chosen to cut. If there is no independent odd ring, it does not matter which dependent odd ring is chosen to cut because as the procedure considers the most recent step where an odd ring in which there are still some individuals who have not been used to cut the odd ring, the procedure cuts all other dependent odd rings at any step of the procedure. Furthermore, because any two stable partitions have the same odd rings by result 2 in Theorem 2, the choice of the stable partition does not matter in any step of the procedure, either.

One should note that neither $P-stability$ nor $RP-stability$ implies the other in an unsolvable roommate problem. In Example 3, $\mu_3 = \{\{1,3\},\{2\},\{4\},\{5,6\}\}$ is a $P-stable$ matching, but not an $RP-stable$ matching. On the other hand, matching $\mu_4 = \{\{1,3\},\{2,4\},\{5,6\}\}$ is an $RP-stable$ matching, but not a $P-stable$ matching. Although, there is no relation between $P-stability$ and $RP-stability$ in unsolvable roommate problems, it is obvious that $P-stability$ and $RP-stability$ are equivalent in solvable roommate problems. Because the set of $P-stable$ matchings coincides with the set of pairwise stable matchings in solvable roommate problems (Inarra et al, 2008), pairwise stability and $RP-stability$, too, are equivalent in solvable roommate problems.

As Example 3 shows, there may be more matched individuals in an $RP-stable$ matching compared to a $P-stable$ matching. Before we prove that an $RP-stable$ matching matches more individuals than a $P-stable$ matching, we introduce some new notation. Let $A(\cdot)$ represent the set of all individuals in an odd ring in a partition, $B(\cdot)$ represent the set of all individuals who involve in a set of mutually acceptable pair in a partition or the set of matched individuals in a matching, and $C(\cdot)$ represents the set of all individuals in singleton sets in a partition or the set of all unmatched individuals in a matching. Let $O(\cdot)$ represents the set of all odds sets in a partition.

For an $RP-stable$ matching μ, let P^k be the final reduced profile that μ is obtained, that is, $\mu^k = \mu$. Moreover, P^j is a reduced stable partition in an arbitrary step j in the procedure through P^k, and μ^j is an arbitrary $P^j-stable$ matching for any $j = 1, \ldots, k-1$. We make an assumption to simplify the analysis. We assume that for any $\{a_i, a_{i-1}\} = B \in P^j$ for some pair $B \in P^j$, we keep $\{a_i, a_{i-1}\}$ as a pair in P^{j+1} if possible while preserving stability in the partition. We make such an assumption because we want individuals in even rings (if an even ring exists) to be partitioned in the same way in adjacent steps. Without this assumption, the individuals in an even ring may be partitioned differently in P^j and P^{j+1}, but the number of matched individuals is the same. Hence, our analysis is not affected. So, without loss of generality, we assume a pair in P^j is kept as a pair in P^{j+1} while preserving stability, if possible. The following lemma is used to prove our claim.

Lemma 5 *For a stable partition P and an $RP-stable$ matching μ, the number of odd sets in P^j is greater than or equal to the number of odd sets in P^{j+1} for all $j = 1, \ldots, k-1$.*

Proof. Suppose that P^{j+1} is obtained by cutting $a_i^j \in A^j \in P^j$. There are three possible cases to consider: (i) $a_i^j \in C\left(P^{j+1}\right)$ or (ii) $a_i^j \in A\left(P^{j+1}\right)$, or (iii) $a_i^j \in B\left(P^{j+1}\right)$.

Case (i): When we cut $a_i^j \in A^j$ in P^j, a_i^j involves in a singleton in the reduced stable partition P^{j+1}. Note that for all $b_i \in I \backslash \left\{a_{i-1}^j, a_i\right\}$, b_i's preference ordering in step j is preserved in step $j+1$. Moreover, no new odd set which has a_i^j and some other individuals is formed; the partition of individuals in $I \backslash A^j$ in P^{j+1} and P^j is the same (by our assumption). Hence, there is one less odd ring and one more singleton in P^{j+1}; $\left|O\left(P^j\right)\right| = \left|O\left(P^{j+1}\right)\right|$.

Case (ii): When we cut $a_i^j \in A^j$ in P^j, a_i^j involves in an odd ring in the reduced stable partition P^{j+1}. Let $A^{j+1} \in P^{j+1}$ be the odd ring such that $a_i^j \in A^{j+1}$. We claim that $A^{j+1} = \left\{a_i^j, b, b', \ldots c, c'\right\}$ with $\{i, i'\}$ is a pair in P^j for all $i = b, \ldots, c$. That is, i and i' cannot be from different sets in P^j for all $i = b, \ldots, c$. This is because there is stability among pairs in P^j, and when the stable partition P^{j+1} is obtained from P^j, preference orderings of individuals in $I \backslash \left\{a_{i-1}^j, a_i^j\right\}$ do not change. Hence, if two individuals in a pair of P^j involve in an odd set in P^{j+1}, they appear as a pair in that odd set of P^{j+1}. Now, because when an individual $i \in B\left(P^j\right)$ involves in an odd ring in P^{j+1}, his partner in the pair also involves in that odd ring; there is no individual i such that $i \in B\left(P^j\right)$ and $i \in A\left(P^{j+1}\right) \cup C\left(P^{j+1}\right)$. Hence, no odd set other than A^{j+1} is formed, i.e., $\left|O\left(P^j\right)\right| = \left|O\left(P^{j+1}\right)\right|$.

Case (iii): When we cut $a_i^j \in A^j$ in P^j, a_i^j involves in a pair in the reduced stable partition P^{j+1}. Let $\left\{a_i^j, b_i\right\} \in P^{j+1}$. Suppose that $b_i \in C\left(P^j\right) \cup A\left(P^j\right)$. Now, because b_i is in a pair in P^{j+1}, the odd set that contains b_i in P^j, say $S^j \in P^j$, does no longer an odd ring in P^{j+1}. In fact, we have either $S^j \backslash \{b_i\} = \emptyset$ or the individuals in $S^j \backslash \{b_i\}$ form pairs. Hence, no new odd set is formed, and we have $\left|O\left(P^j\right)\right| > \left|O\left(P^{j+1}\right)\right|$. If $b_i \in B\left(P^j\right)$, then an individual $c_i \in B\left(P^j\right)$ (not necessarily $c_i = b_{i-1}$) appears either in a singleton in P^{j+1} or in an odd ring in P^{j+1}. Hence, $\left|O\left(P^j\right)\right| = \left|O\left(P^{j+1}\right)\right|$.

We consider each possible case and conclude that $\left|O\left(P^{j}\right)\right| \geq \left|O\left(P^{j+1}\right)\right|$. The proof is complete. ∎

Proposition 9 *Given a stable partition P and an $RP-stable$ matching μ, the number of matched individuals in μ is greater than or equal to the number of matched individuals in any $P'-stable$ matching μ' for any stable partition P'.*

Proof. If there is no odd ring in P, then there is no odd ring in P' either by Theorem 2. Hence, μ' and μ are stable matchings. By Theorem 1, the set of matched individuals in μ and μ' are the same; we are done.

Now, consider the case where there is an odd ring in P. We know that for any stable partition P^{*}, the number of unmatched individuals in a $P^{*}-stable$ matching is equal to the number of odd sets in P^{*}Hence, the number of unmatched individuals in $\mu^{k}=\mu$ is equal to the number of odd sets in the final stable partition P^{k}. Similarly, the number of unmatched individuals in μ' is equal to the number of odd sets in P'. Hence, in order to show that the number of matched individuals in μ is greater than that in μ', it is enough to show that $O\left|(P')\right| \geq O\left|\left(P^{k}\right)\right|$.

By Theorem 2, $O(P)=O(P')$. By lemma 5,

$$\left|O(P)\right| \geq \left|O(P_1)\right| \geq \ldots \geq \left|O(P_k)\right|.$$

Hence, $\left|O(P')\right| \geq \left|O(P_k)\right|$. ∎

A sequential matching mechanism

In this section, we introduce a sequential matching mechanism and analyze the SPE outcome of this mechanism. We show that the RP procedure can be used to identify all potential SPE outcomes of the mechanism for any order of individuals' moves.

For an $RP - stable$ matching μ, let μ' be a matching which is preferred to μ by a set of individuals S. Moreover, each individual in set S is matched to an individual in $\mu(S)$ under μ' and each individual not in S is matched to his μ partner under μ'. That is, for any $RP - stable$ matching μ, let μ' be such that $\mu' \succ_S \mu$, $\mu'(S) = \mu(S)$ for some $S \subseteq I$ with $S \neq \emptyset$, and $\mu'(i) = \mu(i)$ for all $i \in I \backslash (S \cup \mu'(S))$. Matching μ' is said to be *obtained from* μ *by favoring set* S. We argue that the SPE outcome of the sequential game is either an $RP - stable$ matching or a matching that is obtained from an $RP - stable$ matching by favoring some set $S \subseteq I$.

The sequential matching mechanism considered in this paper is an n stage extensive form game of perfect information. The order of moves is given by a surjective function $\sigma : \{1, \ldots, n\} \longrightarrow I$, where $\sigma(k)$ denotes the individual who moves in the kth stage of the mechanism.

In stage 1, $\sigma(1)$ either proposes to a potential roommate or chooses to be single. Whenever $\sigma(1)$ is indifferent between proposing to a roommate and choosing to remain single, he chooses to remain single.

In stage k, $1 < k < n$, after observing the history of the game up to stage k, $\sigma(k)$ either (i) accepts one potential roommate who has proposed to her in a previous stage, or (ii) proposes to a potential roommate who will move in a subsequent stage, or (iii) chooses to remain single. Whenever $\sigma(k)$ is indifferent between proposing to a roommate and choosing to remain single, he chooses to remain single.

In stage n, $\sigma(n)$ either accepts one potential roommate who has proposed to her in a previous stage, or chooses to remain single.

There is a different sequential mechanism for every order of moves. Given an order of moves, the outcome of the mechanism is a matching. This is because an individual cannot accept more than one proposal, and by making a proposal or by deciding to remain single, she must reject all proposals that have been received. Because preferences are strict, the SPE outcome of the mechanism for a given an order of moves is unique. Observe that the SPE outcome of the game is an individually rational matching. This is because an individual i would choose to remain single rather than proposing to or accepting a proposal from someone who is not acceptable for i.

We know that odd rings play an important role for the existence of a stable matching in roommate problems. In this section, we focus on the role of odd rings in the sequential mechanism. The following example is an example of the role that odd rings play on the equilibrium path of any mechanism (the preference profile in this example is the same with the one that is used in Appendix).

Example 5 *Consider the preference profile below:*

1	**2**	**3**	**4**	**5**
2	3	1	5	2
3	1	2	2	4
	4			
	5			

The unique stable partition is $P = \{\{1,2,3\},\{4,5\}\}$*. The unique odd ring in* P *is* $A = \{1,2,3\}$*. Depending on the order of individuals' moves, an individual in* $\{1,2,3\}$ *is matched to someone not in the ring or is unmatched in the SPE. The individual in* $\{1,2,3\}$ *who is matched to someone not in the ring or is unmatched in the SPE may involve in other odd rings with other individuals. For example, consider the following order of moves:* $1,2,3,4,5$*. On the equilibrium path, we observe that* 2 *never receives a proposal from* 1 *and cannot propose to* 3*. This is because if* 1 *proposes to* 2*, he will be rejected in favor of* 3*. Then,* 1 *proposes to* 3*. Hence,* 2 *will be rejected in favor of* 1 *if he proposes to* 3*. When we cut* $2 \in \{1,2,3\}$*, we obtain another stable partition* $P^1 = \{\{1,3\},\{2,4,5\}\}$*. That is, individual* 2 *involves in another odd ring* $\{2,4,5\}$*. In this odd ring,* 4 *never receives a proposal from* 2 *and cannot propose to* 5 *by the same logic. When we cut* $4 \in \{2,4,5\}$*, we obtain* $P^2 = \{\{1,3\},\{2,5\},\{4\}\}$*. Because there is no odd ring in* P^2*, we obtain all* P^2-*stable matchings. Thus,* $\mu^2 = \{\{1,3\},\{2,5\},\{4\}\}$ *is an* $RP-$ *stable matching. Matching* μ^2 *is also the SPE outcome of the sequential mechanism.*

Consider a sequential mechanism and suppose that there is an odd ring $A = \{a, \ldots, k\}$ such that a is the first one to move. On the equilibrium path, suppose that a cannot propose to b because if he does so, b is matched to c, d is matched to e, and so on so forth. In such a situation, we say that a and b *face* an odd ring. Obviously, when two individuals face an odd ring on the equilibrium path, they cannot be matched in the SPE. Let $\{a_1, a_2\}$ be a pair in some set A of a stable partition P. By the stability of P, for any $b_i \in I$ such that $a_1 \succ_{b_i} b_{i-1}$ and for any c_i such that $a_2 \succ_{c_i} c_{i-1}$, we have $a_2 \succ_{a_1} b_i$ and $a_1 \succ_{a_2} c_i$, respectively. Hence, if a_1 and a_2 do not face an odd ring on the equilibrium path, then they are not matched in the SPE only if one of them gets an opportunity to be matched to someone more preferred. This is equal to say that if a_1 and a_2 are not matched in the SPE, then either they face an odd ring on the equilibrium path or one of them gets an opportunity to be matched to someone more preferred. By using this logic, we prove the following lemma, which is used to prove our main result in this section.

Proposition 10 *For any odd ring $A = \{a_1, \ldots, a_k\} \in P$, if A is an independent odd ring, then a_i never receives a proposal from or proposes to an individual b such that $b \succ_{a_i} a_{i-1}$ and $b \neq a_{j+1}$ when everybody plays his equilibrium strategy in the sequential mechanism for all $a_i \in A$.*

Proof. Suppose there is an independent odd ring $A = \{a_1, \ldots, a_k\} \in P$. On the contrary, suppose that for some $a_i \in A$, a_i rejects a_{i-1} in favor of an individual $b_i \neq a_{i+1}$ when a_{i-1} proposes to a_i in a sequential mechanism. Because $b_i \succ_{a_i} a_{i-1}$, we must have $b_{i-1} \succ_{b_i} a_i$ by the stability of P. Then, b_i can only be available for a_i either because b_{i-1} and b_i are in an odd ring, or b_{i-1} gets an opportunity to be matched with someone whom he prefers to b_i. Hence, there is an individual $c_i \in I$ such that either $c_i \succ_{b_{i-1}} b_i$ or $c_i \in A'$ with A' being an odd ring and $\{b_{i-1}, b_i\}$ being a pair in A'. Because A is an independent odd ring, $A' \notin P$ by the definition of an independent odd ring. The fact that $A' \notin P$ implies that c_i becomes available for someone less preferred than c_{i-1}. By a similar argument, this is possible only if c_{i-1} and c_i are in an odd ring, or c_{i-1} gets an opportunity to be matched with someone whom he prefers to c_i. Hence, there is an individual $d_i \in I$ such that either $d_i \succ_{c_{i-1}} c_i$ or $d_i \in A''$ with A'' being an odd ring and $\{c_{i-1}, c_i\}$ being a pair in A''. By repeating the same argument, in any step of the argument that is considered, there must be an individual l_i, who becomes available for someone less preferred than l_{i-1}, because either l_{i-1} and l_i are in an odd ring, or l_{i-1} gets an opportunity to be matched with someone whom he prefers to l_i. Hence, the argument continues. However, because there is a finite number of people, the argument must have finite steps, contradiction. ■

Proposition 10 implies that if there is an independent odd ring in a stable partition, then everyone except one person in the odd ring is matched to one of his neighbors in the independent odd ring, and one person is matched to someone worse than his preceding individual.

In order to prove that the SPE outcome of a sequential game is either an $RP-stable$ matching or a matching that is obtained from an $RP-stable$ matching by favoring a set of individuals, we focus on the $RP-stable$ matchings that are obtained by following a particular path in the procedure. That is, we choose an $RP-stable$ matching that can be obtained by breaking odd rings in a specific order which is consistent with the sequential mechanism considered, i.e., the order in which individuals move. Consider the following example:

Example 6 *Consider the preference profile given in example 4:*

1	**2**	**3**	**4**	**5**	**6**
4	3	1	5	6	4
2	1	2	6	4	5
3			1		

The unique stable partition is $P = \{\{1,2,3\}, \{4,5,6\}\}$*. The* $RP-stable$ *matchings are:*

$$\begin{aligned}
\mu_1 &= \{\{1,4\}, \{2,3\}, \{5,6\}\} \\
\mu_2 &= \{\{1\}, \{2,3\}, \{4,6\}, \{5\}\} \\
\mu_3 &= \{\{1,3\}, \{2\}, \{4,6\}, \{5\}\} \\
\mu_4 &= \{\{1,2\}, \{3\}, \{4,6\}, \{5\}\} \\
\mu_5 &= \{\{1\}, \{2,3\}, \{4,5\}, \{6\}\} \\
\mu_6 &= \{\{1,3\}, \{2\}, \{4,5\}, \{6\}\} \\
\mu_7 &= \{\{1,2\}, \{3\}, \{4,5\}, \{6\}\}
\end{aligned}$$

Because $\{4,5,6\}$ *is an independent ring, we start cutting* $\{4,5,6\}$*. In the RP procedure, when we cut* 4*,* $\{1,2,3\}$*, which is a dependent ring in P, disappears in the reduced profile, and we obtain the final partition* $\{\{1,4\}, \{2,3\}, \{5,6\}\}$*. Hence, we obtain* μ_1*. When we continue the procedure by cutting* 5 *or* 6*,* $\{1,2,3\}$ *appears as an independent ring (the only ring) in any reduced stable partition. Hence, we continue cutting* $\{1,2,3\}$ *by each individual in it.*

In any sequential mechanism, because $\{4,5,6\}$ *is an independent ring, there is no individual* $i \in \{4,5,6\}$ *such that* $\mu(i) \succ_i i-1$*, and exactly one individual in* $\{4,5,6\}$ *is matched to someone else other than his neighbors. Observe that if* 4 *is the one who cannot be matched to one of his neighbors, then* 4 *becomes available for* 1*, and* 4 *and* 1 *are matched in the SPE. The ring* $\{1,2,3\}$ *is not effective, i.e., no individual in the ring* $\{1,2,3\}$ *is left out because of the ring* $\{1,2,3\}$*. Hence, the way that* 1*,* 2*, and* 3 *are ordered is irrelevant. This is similar to the fact that* $\{1,2,3\}$ *disappears in the reduced profile when we cut* 4*. If either* 5 *or* 6 *is cut in a mechanism, then the order of* 1*,* 2*, and* 3 *matters. Those individuals who are matched and the one who is unmatched depend on order of* 1*,* 2*, and* 3*. This is similar to the fact that when we cut* 5 *or* 6 *in the procedure,* $\{1,2,3\}$ *appears as an independent ring (the only ring) in any reduced stable partition, and we continue cutting* $\{1,2,3\}$ *by each individual in it.*

On the equilibrium path in a sequential mechanism, there is a particular order in which the odd rings become effective. To prove our main result, we construct a particular order of cutting odd rings in the procedure that is consistent with the particular order in which the odd rings become effective in the sequential mechanism considered, by using Proposition 10. The outline of the proof is as follows: We start with showing how to construct that particular order of cutting rings. Moreover, we show that the profile that we reach by cutting the odd rings according to that particular order is a final profile that can be obtained by the procedure. Hence, if the SPE outcome is a $P-stable$ matching with respect to that final profile, it is an $RP-stable$ matching. We finish the proof by showing that if the SPE outcome is not a $P - stable$ matching with respect to that final profile, then it is a matching that is obtained from an $RP - stable$ matching by favoring a set of individuals. Now, we prove our main result.

Theorem 3 *Given a preference profile $\succ$, the SPE outcome of any sequential matching mechanism is either an $RP - stable$ matching or a matching that is obtained from an $RP - stable$ matching by favoring a set of individuals.*

Proof. Consider an order of individuals and let μ^* be the SPE outcome of the game. Let P be a stable partition. By Proposition 10, we know that if there is an independent odd ring $A = \{a_1, \ldots, a_k\}$ in P, then no individual a_j in A gets an opportunity to be matched to an individual b such that $b \succ_{a_j} a_{j-1}$ and $b \neq a_{j+1}$. Moreover, there is one individual $a_i \in A$ who cannot be matched to a_{i-1} and becomes available for other individuals who prefer a_i more than their preceding individuals in some sets of P (if there is any such individual). Hence, we eliminate a_i and a_{i-1} from each other's preference profile and obtain a reduced stable partition P^1. Observe that no individual's actions are affected on the equilibrium path when we eliminate a_i and a_{i-1} from each other's preference profile. In order to see this, observe that each individual plays only once in the sequential mechanism. Hence, once a_{i-1} cannot propose to a_i (or a_i cannot propose to a_{i-1}) on the equilibrium path because of the odd ring A, a_i can no longer be a thread for a possible partnership of a_{i-1} with another individual on the equilibrium path. That is, on the equilibrium path, when taking an action, each individual takes the fact in to account that a_i and a_{i-1} cannot be matched. Hence, there is no individual b in $I \setminus \{a_{i-2}, a_{i-1}, a_i\}$ such that a_{i-1} rejects b in favor of a_i or a_i does not propose to b because he has an offer from a_i. We can conclude that no individual's actions are affected on the equilibrium path when we eliminate a_i and a_{i-1} from each other's preference profile.

Similarly, if there is an independent odd ring in a reduced profile, we can continue the process by cutting the independent odd ring by the individual who cannot be matched to his preceding or succeeding individuals in the ring. By the same logic, as we continue the process by cutting the independent odd rings accordingly, no individual's actions are affected on the equilibrium path when we eliminate two successive individuals in the independent odd ring from each other's preference profile.

Consider an arbitrary step j of the procedure. Suppose that step j is the first step such that there is no independent odd ring in any stable partition. If there are dependent odd rings, then the procedure cuts an arbitrary dependent odd ring by an arbitrary individual in step j. The choice of the dependent odd ring does not matter because after a final stable partition is reached, the procedure considers the most recent profile with an odd ring such that not all individuals in the odd ring is used to cut the ring. Hence, all dependent rings in step j are cut one by one by each individual in the ring. As a result, any ring (whether independent or dependent) that individuals face on the equilibrium path is cut by the procedure. The particular order of cutting rings that is consistent with the sequential game must be one of those paths that the procedure reaches. That is, the final profile that is reached when we cut the odd rings in a way that is consistent with the order of individuals is one of the final profiles that is reached by the procedure.

On the path that is consistent with the order of individuals, consider the final step k, where we obtain the final profile $\succ^k$. For any individual $a_i \in I$, if b is eliminated from a_i's list, then b and a_i involve in an odd ring in some step j and cannot be matched because of the odd ring in μ^*. Hence, $\mu^*(a_i)$ is in a_i's list in $\succ^k$. Let P^k be a final stable partition with respect to $\succ^k$ and μ^k be an $P^k-stable$ matching. If $\mu^* = \mu^k$, then μ^* is an $RP-stable$ matching, and hence, the proof is done. Now, suppose that $\mu^* \neq \mu^k$. We aim to show that

μ^* is such that $\mu^* \succ_S \mu^k$, $\mu^*(S) = \mu^k(S)$ for a nonempty set $S \subseteq I$, and $\mu^*(i) = \mu^k(i)$ for all $i \in I \backslash (S \cup \mu^*(S))$.

Let S be the set of all individuals who prefer μ^* to μ^k. First, we show that $S \neq \emptyset$. Let $a_i \in I$ be such that $\mu^*(a_i) \neq \mu^k(a_i)$. If $\mu^*(a_i) \succ_{a_i} \mu^k(a_i)$, then $a_i \in S$, and $S \neq \emptyset$. If $\mu^k(a_i) \succ_{a_i} \mu^*(a_i)$, then we know that either a_i and $\mu^k(a_i)$ are in an odd ring in P^k or $\mu^k(a_i)$ gets an opportunity to be matched with someone whom he prefers to a_i. Because P^k is a final stable partition, there is no odd ring in P^k. Hence, we must have $\mu^*\left(\mu^k(a_i)\right) \succ_{\mu^k(a_i)} a_i$, which implies that $\mu^k(a_i) \in S \neq \emptyset$. Second, we show that $\mu^*(S) = \mu^k(S)$. If $\mu^k(S) \neq \mu^*(S)$, there exists an individual $a \in S$ such that $\mu^*(a) \notin \mu^k(S)$. Note that $\mu^*(a) \notin \mu^k(S)$ implies $\mu^k(\mu^*(a)) \notin S$. Because $a \in S$, $\mu^*(a) = a^* \succ_a \mu^k(a)$. By the stability of P^k, $\mu^k(a^*) \succ_{a^*} a$. Because there is no odd ring in P^k and because a^* is matched with someone worse than $\mu^k(a^*)$, $\mu^k(a^*)$ must get an opportunity to be matched with someone whom he prefers to a^* in μ^*. That is, $\mu^*\left(\mu^k(a^*)\right) \succ_{\mu^k(a^*)} a^*$. Hence, $\mu^k(a^*) \in S$, contradiction. We conclude that $\mu^*(S) = \mu^k(S)$.

Because S is the set of all individuals who prefer μ^* to μ^k, by the stability of P^k, $\mu^k(S)$ (or $\mu^*(S)$) must the set of all individuals who prefer μ^k to μ^*. Then, for all $i \in I \backslash (S \cup \mu^*(S))$, we have $\mu^k(i) = \mu^*(i)$. Thus, if $\mu^* \neq \mu^k$, then μ^* is such that $\mu^* \succ_S \mu^k$ for a nonempty set $S \subseteq I$, and $\mu^*(i) = \mu^k(i)$ for all $i \in I \backslash (S \cup \mu^*(S))$. The proof is complete. ■

A matching that is obtained from an $RP-stable$ matching by favoring a nonempty set of individuals is not necessarily $RP-stable$. Suppose that μ^* is obtained from an $RP-stable$ matching μ' by favoring a nonempty set of individuals S, and is not $RP-stable$. As it can also be seen in the proof of Theorem 3, there is no $\{a, \mu'(a)\} \in \mu'$ such that $\{a, \mu'(a)\}$ blocks μ^*. That is, only certain types of blocking pairs can be observed in the

SPE. Hence, if we can rule out only these certain types of blocking pairs, the SPE outcome μ^* is an $RP-stable$ matching. In other words, if we can rule out only the unstable matchings that are obtained from $RP - stable$ matchings by favoring some set of individuals, we can guarantee the stability of the SPE outcome. By taking this fact into account, we introduce a sufficient condition to guarantee the stability of the SPE outcome regardless of the order of individuals in a solvable roommate problem in the next section.

A sufficient condition to implement pairwise stable matchings

In this section, we focus on solvable roommate problems and on the pairwise stability concept. Our purpose is to introduce a sufficient condition to guarantee the pairwise stability of the SPE outcome of the mechanism for any order of individuals' moves in solvable roommate problems. Before we focus on the stability of the SPE, we want to explore the relation between pairwise stability, $P - stability$, and $RP - stability$.

Given a preference profile, there is no odd ring in a stable partition if and only if there exists a stable matching by Theorem 2. Hence, for a solvable roommate problem, each $P-stable$ matching is a stable matching. Moreover, for a solvable roommate problem, because there is no odd ring in any stable partition for the given preference profile, $RP - stability$ and $P - stability$, and hence, pairwise stability are equivalent by definition.

Now, we focus on implementing stable matchings in the SPE. Although a stable matching exists for any solvable roommate problem, the SPE outcome may not be stable for some order of individuals' moves as shown in the following example.

Example 7

1	**2**	**3**	**4**	**5**	**6**
6	6	5	1	2	3
4	5	6	2	3	1
5	4	4	3	1	2
1	2	3	4	5	6

There is a unique stable partition $P = \{\{1,4\}, \{2,5\}, \{3,6\}\}$ *and hence, the unique* $P-$ *stable and also* $RP-$ *stable and pairwise stable matching is* $\mu = \{\{1,4\}, \{2,5\}, \{3,6\}\}$. *Consider matching* $\mu^* = \{\{1,4\}, \{2,6\}, \{3,5\}\}$, *which is unstable because it is blocked by the pair* $\{1,6\}$. *Observe that* μ^* *can be obtained from* μ *by favoring* $S = \{2,5\}$. *Furthermore, observe that the SPE outcome of the mechanism when the individuals move as* $1,2,3,4,5,6$ *is* $\mu^* = \{\{1,4\}, \{2,6\}, \{3,5\}\}$.

As in the above example, the SPE outcome may not be $RP-stable$ and hence, may not be stable in a solvable roommate problem. Consider a solvable roommate problem. Let μ^* be in the SPE. We know by Theorem 3 that if μ^* is not $RP-stable$, then μ^* can be obtained from an $RP-stable$ matching by favoring a group of individuals. Hence, a sufficient condition to guarantee the stability of the SPE outcome in a sequential mechanism would then require that any matching that is obtained from an $RP-stable$ matching by favoring a set of individuals $S \subseteq I$ is stable. We formally present our result in the following theorem.

Theorem 4 *Consider a solvable roommate problem. Let* μ *be a stable matching. If any matching that is obtained through an* $RP-stable$ *matching by favoring a set of individuals* $S \subseteq I$ *is stable, then the SPE outcome of the sequential game is stable regardless of the order of individuals.*

In the previous section, we show that the SPE outcome for a given mechanism is either an $RP-stable$ matching or can be obtained from an $RP-stable$ matching by favoring a group of individuals. However, not all $RP-stable$ matchings are implementable in the SPE. In fact, not all pairwise stable matchings are implementable in the SPE. We provide an example of a solvable roommate problem and show that there exists a pairwise stable matching such that there is no order of moves for which this pairwise stable matching is in the SPE. By doing so, we also prove that not all $RP-stable$ matchings are implementable in the SPE because pairwise stability and $RP-stability$ are equivalent in solvable roommate problems.

Example 8 *Consider the following preference profile:*

1	**2**	**3**	**4**	**5**	**6**
4	5	6	2	3	1
5	6	4	3	1	2
6	4	5	1	2	3
1	2	3	4	5	6

The pairwise stable matchings are $\mu_1 = \{\{1,4\},\{2,5\},\{3,6\}\}$, $\mu_2 = \{\{1,5\},\{2,6\},\{3,4\}\}$, *and* $\mu_3 = \{\{1,6\},\{2,4\},\{3,5\}\}$. *In any given mechanism, whoever moves first always proposes to his or her top ranked person. Observe that for any* $i \in \{1,2,3,4,5,6\}$, *i is the last ranked person in the preference list of his top ranked person, and i is the second ranked person in the preference list of his second ranked person. Because individuals are symmetric in this sense, if we prove that* 1 *proposes to* 4 *in any mechanism with* 1 *being the first person to move, we show that whoever moves first always proposes to his top ranked person. If* 1 *proposes to* 4*, then* 6 *misses the opportunity to be matched to his top ranked choice* 1*. The best* 6 *can do is to be matched to* 2*. Thus, the worst* 2 *can do is to be matched to* 6*. Hence,* 4 *never gets the opportunity to be matched to* 2*. Then, the best* 4 *can do is to be matched to* 3*, and the worst* 3 *can do is to be matched to* 4*. Hence,* 5 *never gets the opportunity to be matched to* 3*. Because* 1 *has proposed to* 4*,* 4 *is not available for* 5*, either. Thus, the best* 5 *can do is to be matched to* 2*. Hence,* 2 *cannot do worse than matching to be* 5*, which implies that* 6 *cannot be matched to* 2*. Then, the only individual whom* 6 *can be matched is* 3*. So, if* 1 *proposes* 4*, the SPE outcome is* $\{\{1,4\},\{2,5\},\{3,6\}\}$*. We conclude that in any mechanism, whoever moves first always proposes to his top ranked person. Hence, the pairwise stable matching* μ_2 *in which everybody is matched to his or her second ranked choice, is not the SPE outcome for any order of individuals' moves.*

Conclusion

We introduce $RP-stability$ concept such that an $RP-stable$ matching allows more individuals to be matched compared to a $P-stable$ matching and is blocked by only a certain type of a blocking pair. In particular, if two individuals block an $RP-stable$

matching, then they must involve in an odd ring. We propose a procedure, which is called RP procedure, to obtain $RP - stable$ matchings.

We also employ a sequential matching mechanism to understand the characteristics of the SPE outcome of this mechanism in a roommate problem. We show that the RP procedure can be used to identify the SPE outcomes of the sequential mechanism for any order of individuals' moves. By analyzing the characteristics of the $RP - stability$ and the relation between the procedure and the sequential game, we are able to propose a sufficient condition to guarantee the stability of the SPE outcome regardless of the order of individuals in the mechanism.

We plan to investigate how the $RP - stability$ concept is related to other solution concepts that are proposed for roommate problems. We also plan to investigate if the set of $RP - stable$ matchings are absorbing sets or not and whether the set of $RP - stable$ matchings satisfies the outer stability or whether there is a random path from an unstable matching to an $RP - stable$ matching.

Appendix

Example: How to obtain $MP - stable$ matchings. Consider the preference profile below:

1	**2**	**3**	**4**	**5**	**6**
–	–	–	–	–	–
2	3	1	2	6	4
3	1	2	5	4	5
	4	6	6		

There is a unique stable partition $P = \{\{1,2,3\},\{4,5,6\}\}$. Let $A = \{1,2,3\}$, and $B = \{4,5,6\}$. Observe that A is the only independent ring. Hence, we cut $1 \in A$.

The preference profile obtained $\succ^1$ is shown below:

1	**2**	**3**	**4**	**5**	**6**
–	–	–	–	–	–
2	3	2	2	6	4
	1	6	5	4	5
	4		6		

The unique stable partition is $P^1 = \{\{1\}, \{2,3\}, \{4,5,6\}\}$. The odd ring $B = \{4,5,6\}$is the unique odd ring in P^2. We cut $4 \in B$ and obtain $\succ^2$:

1	**2**	**3**	**4**	**5**	**6**
–	–	–	–	–	–
2	3	2	2	6	5
	1	6	5	4	
	4				

The unique stable partition is $P^2 = \{\{1\}, \{2,3\}, \{4\}, \{5,6\}\}$. There is no odd ring in P^2. Hence, P^2 is a final stable partition and $\mu_1 = \{\{1\}, \{2,3\}, \{4\}, \{5,6\}\}$ is an $RP - stable$ matching. We turn back to the most recent step in which there is an odd ring such that not all individual are considered in the cutting process, yet. Hence, in the third step, we consider $\succ^1$ again. We cut $5 \in B$ and obtain $\succ^3$ as below:

1	**2**	**3**	**4**	**5**	**6**
–	–	–	–	–	–
2	3	2	2	6	4
	1	6	6		5
	4		5		

The unique stable partition is $P^3 = \{\{1\}, \{2,3\}, \{4,6\}, \{5\}\}$ and P^3 is a final stable partition. We obtain an $RP - \mathit{stable}$ matching $\mu_2 = \{\{1\}, \{2,3\}, \{4,6\}, \{5\}\}$. We continue with P^1 and cut $6 \in B$ to obtain $\succ^4$:

1	**2**	**3**	**4**	**5**	**6**
2	3	2	2	4	4
	1	6	5		
	4		6		

The unique stable partition is $P^4 = \{\{1\}, \{2,3\}, \{4,5\}, \{6\}\}$ and P^4 is a final stable partition. We obtain an $RP - \mathit{stable}$ matching $\mu_3 = \{\{1\}, \{2,3\}, \{4,5\}, \{6\}\}$. The most recent step in which there is an odd ring such that not all individual are considered in the cutting process is step 1. We continue with A and cut $2 \in A$ to obtain $\succ^5$:

1	**2**	**3**	**4**	**5**	**6**
3	3	1	2	6	4
	4	2	5	4	5
		6	6		

The unique stable partition is $P^5 = \{\{1,3\}, \{2,4\}, \{5,6\}\}$ and P^5 is a final stable partition. We obtain an $RP - \mathit{stable}$ matching $\mu_4 = \{\{1,3\}, \{2,4\}, \{5,6\}\}$. We continue with odd ring A and cut $3 \in A$ to obtain $\succ^6$:

1	**2**	**3**	**4**	**5**	**6**
$-$	$-$	$-$	$-$	$-$	$-$
2	1	1	2	6	4
3	4	6	5	4	5
			6		

The unique stable partition is $P^5 = \{\{1,2\},\{3\},\{4,5,6\}\}$. We cut $4 \in B$ and obtain $\mu_5 = \{\{1,2\},\{3\},\{4\},\{5,6\}\}$. In the following two steps we cut $5 \in B$ and obtain $\mu_6 = \{\{1,2\},\{3\},\{4,6\},\{5\}\}$, and we cut $6 \in B$ and obtain $\mu_7 = \{\{1,2\},\{3\},\{4,5\},\{6\}\}$.

All $RP - stable$ matching are given below:

$$
\begin{aligned}
\mu_1 &= \{\{1\},\{2,3\},\{4\},\{5,6\}\} \\
\mu_2 &= \{\{1\},\{2,3\},\{4,6\},\{5\}\} \\
\mu_3 &= \{\{1\},\{2,3\},\{4,5\},\{6\}\} \\
\mu_4 &= \{\{1,3\},\{2,4\},\{5,6\}\} \\
\mu_5 &= \{\{1,2\},\{3\},\{4\},\{5,6\}\} \\
\mu_6 &= \{\{1,2\},\{3\},\{4,6\},\{5\}\} \\
\mu_7 &= \{\{1,2\},\{3\},\{4,5\},\{6\}\}
\end{aligned}
$$

CHAPTER IV

REGIONAL VERSUS MULTILATERAL TRADE AGREEMENTS: A WELFARE ANALYSIS

Introduction

The lack of a forcing authority in trade relations of world countries makes it difficult to achieve a trade agreement that increases world welfare. This creates a structural problem of rules in trade agreements that will self-enforce the trading countries to achieve a more liberal trade. In this paper, we study these self-enforcing rules with asymmetric countries from the perspective of regional and multilateral trade agreements. In particular, we attempt to find an answer to the question of "Why are trade agreements mostly regional/preferential rather than multilateral?". By employing a welfare analysis, we show that the existence of transportation costs may be a possible reason.

We present an N-good-N-country partial equilibrium model by generalizing the 2-good-2-country model in Bond and Park (2002) through considering transportation costs. For simplicity, we assume that the downward-sloped demand curve and upward-sloped supply curve of each good in each country are linear in price of the good. We allow asymmetries between countries in terms of country sizes and comparative advantages. We show that international trade between any two countries is achieved through differences in supply and demand structures of the countries. After considering transportation cost between any two countries, it follows that the equilibrium autarky price (for each good) in any country depends on the individual country specific demand and supply structure, while the equilibrium price (for each good) under international trade depends on the demand and supply structures of all countries together with country-specific tariff rates and transportation costs. According to the model, national welfare is defined as the sum of consumer surplus, producer surplus and tariff incomes. We let each country to receive all the tariff income for the good that the country imports.

In order to analyze the optimal tariff setting behavior of each country, we employ the following approach in the theoretical analysis: First, we find the Nash tariff rates and best response functions. Each country maximizes its welfare given the tariff rate of other countries. We show that the Nash tariff rates are decreasing in transportation costs; i.e., closer countries tend to have higher tariff rates between each other, which lead to lower trade volume across them. This suggests that there is a potential gain from a trade agreement between closer countries in the short run. Moreover, while smaller countries set lower optimal tariff rates to larger countries, larger countries set higher optimal tariff rates to smaller countries. Then, we compare regional and multilateral trade agreements when countries involve a stationary dynamic tariff game, i.e., countries play a repeated game for tariff rates. The repeated structure of the game gives more flexibility to our model in terms of incorporating the self-enforcement agreements. In this game, each country is able to compare future payoffs out of a possible collusion (cooperation) with future payoffs out of a possible deviation from the agreement. In order to sustain collusion in a trade agreement, the trade-off between the gains from deviating from an agreed-upon tariff policy and the discounted expected future gains from collusion must be balanced in a way that the latter should keep away countries from deviating. We show that it is harder to make an agreement for any country if the transportation cost is higher; i.e., countries tend to sustain regional agreements to maximize their long-run welfare.

The relation between geography and trade agreements has been previously studied in the literature starting with Viner (1950) who has mentioned departures from the Most Favored Nation (MFN) principle between countries within Europe going as far back as the nineteenth century.[1] Nevertheless, instead of explaining these agreements through trade costs, Viner has reasoned them to "close ties of sentiment and interest arising out of ethnological, or cultural, or historical political affiliations". Other earlier studies such as Meade (1955) and Lipsey (1957) haven't mentioned about a possible effect of transportation costs on regional trade agreements either.

Recently, Wonnacott and Wonnacott (1981), Wonnacott and Lutz (1989), Krugman (1991, 1993), Summers (1991), Frankel, Stein and Wei (1995), and Bhagwati and Panagariya (1996) have attempted to find whether or not proximity between countries have made regional agreements more beneficial compared to non-regional agreements. In particular, Wonnacott and Wonnacott (1981) assign an important role to transportation costs in their analysis, but their study has been criticized by Berglas (1983) and Panagariya (1998) in the sense that the transportation costs have to be too high in order to talk about the effect of transportation costs on regional trade agreements. Besides, following Krugman (1991), other studies such as Frankel (1997), Frankel, Stein and Wei (1995), Frankel and Wei (1997) have also advocated for the effect of transportation costs on regional trade agreements. However, building on the earlier critique in Bhagwati (1993) and Bhagwati and Panagariya (1996), Panagariya (1997) has shown that transportation costs are not different than any other costs and hence should not deserve any special attention in explaining the regional trade agreements.

In terms of methodology, the common question in the literature asked is whether

[1]As Panagariya (2000) states, MFN is the centerpiece of the General Agreement on Tariffs and Trade (GATT) that governs the international trade in goods. In particular, MFN refers to the trade policy in which each World Trade Organization (WTO) member grants to all members the same advantage, privilege, favor, or immunity that it grants to any other country.

tariff reductions with nearby partners are welfare improving. Krugman (1991), Frankel Sten and Wei (1995), and Bhagwati and Panagariya (1996) have examined this particular question by considering whether exogenously given preferential tariff reductions are welfare improving in a model where there are differing levels of transport costs between trading partners. This approach has been extended by Bond (2001) who concentrates on preferential tariff reductions by self-enforcing agreements in a four-country world.

In particular, Bond (2001) assumes that the world is divided into two continents, with two countries located on each continent. There is a per unit cost on any good imported from a country on the other continent, but zero transportation cost on goods coming from the country on the same continent. He shows that Nash equilibrium tariffs on regional trading partners are higher than those on the distant partners. As Bond (2001) claims in the conclusion of his paper, one might anticipate that this fact would make it more difficult to support trade liberalization with nearby countries, because the incentive to deviate at a given agreement tariff would be higher. However, he shows that this effect is offset by the fact that the welfare level under regional free trade agreements is higher than that with a distant partner (with given internal tariffs). This is due to the fact that free trade agreements with distant partners have higher external tariffs against all countries, which leads to lower world welfare under distant free trade agreements. In sum, Bond (2001) shows that the equilibrium with regional trade agreements yields higher welfare.[2] This paper follows a similar approach with Bond (2001), but our analysis differs from his paper by considering an N-country model. In Bond's notation, instead of comparing the Nash equilibrium tariff rates within and across continents under transportation costs, we directly measure the effect of transportation costs on the Nash equilibrium tariff rates of N countries.

[2] Most recently, by working out optimal tariffs with transport costs, Zissimos (2007) also shows that trade based gains to a block with countries of the same region are higher than gains to an agreement involving (distant) countries from different regions.

Another topic that has been studied in the literature is the relation between preferential and multilateral agreements. Common questions that have been investigated are whether the trade liberalization is achieved through preferential agreements or multilateral agreements, whether preferential agreements are building blocks or stumbling blocks for multilateral agreements, and whether bilateralism or multilateralism is a better strategy for countries. Bagwell and Staiger (1999) study a competing exporters model with three countries to identify the different circumstances under which the preferential agreements can lead to multilateral agreements or block them. Other studies that investigate the relation between preferential and multilateral agreements in different settings are Bagwell and Staiger (1997a,b) (1999), Bond and Syropoulos (1995), Bond Syropoulos and Winters (2001). All of these studies assume that countries can commit tariff rates under preferential agreements, hence, only the multilateral agreements must be self-enforcing. In our paper, we do not make such an assumption about preferential agreements, i.e., preferential agreements must be self-enforcing, too. A recent study which also aims to explain the tendency toward regionalism is Freund (2000). Freund analyzes the relation between preferential and multilateral agreements in a model of imperfect competition, and argues that a multilateral tariff agreement creates the incentives to form a preferential agreement with a higher probability that the preferential agreement is self-enforcing. Limao (2007) investigates the effects of preferential trade agreements on global free trade when countries are also motivated by cooperation in non-trade issues. Limao argues that the preferential agreements motivated by cooperation in non-trade issues increase the cost of multilateral tariff reductions and, hence, decrease the likelihood of a multilateral free trade agreement. Another recent study, Saggi and Yildiz (2010), focus on the comparison of bilateralism and multilateralism through trade liberalization. They employ a competing exporters model in which tariff rates are determined endogenously. They argue that when countries have symmetric endowments, both bilateralism and multilateralism yield global free trade, but when countries are asymmetric in terms of endowments, global free trade is stable (for a large set of parameters) only through bilateral agreements. In our model, countries do not compete over the goods that they export because each country exports only one good. Hence, in our model, when two countries involve in a bilateral agreement, the non-member countries do not face discriminatory tariffs in export markets.[3]

The rest of the paper is organized as follows. Section II introduces the N-country model. Section III finds the optimal tariff rates and the best response functions. Section IV extends the analysis by considering the stationary dynamic tariff game approach. Section V concludes.

The Model

We extend the international trade model of Bond and Park (2002) by increasing the number of countries to N in order to investigate possible regional and multilateral trade agreements. In particular, our model is a N-good-N-country partial equilibrium model where transportation costs are considered. In terms of notation, $C = \{1, \ldots, N\}$ represents the set of countries and H_{ji} is related to variable H in terms of good $i \in \{1, ..., N\}$ in country $j \in \{1, ..., N\}$.

The demand for good i in country j is given as follows:

$$D_{ji} = \lambda_j \left(A - Bp_{ji} \right)$$

where we assume that the demand curve is downward-sloping and linear in price of the

[3]For a futher discussion on the relationship between preferential and multilateral liberalization see also Bhagwati et al. (1999), Saggi (2006), and Karacaovali and Limao (2007).

good.[4] Similarly, the upward-sloping supply curve of good i in country j, which is also linear in price of the good, is given as follows:

$$X_{ji} = \lambda_j \left(\alpha_{ji} + \beta p_{ji} \right)$$

where p_{ji} is the price of good i in country j, and $\lambda_j \geq 1$ is a parameter by which we measure the size of country j. Note that in a special case of $\lambda_j = 1$ for all j, we have N symmetric countries with the same size.

Each country can impose specific tariffs on its importables, with t_{jk} denoting the tariff rate imposed by country j for goods imported from country k (where $t_{jj} = 0$ for all j). Moreover, trade between any two countries is up to an exogenous symmetric iceberg transportation cost, with τ_{jk} denoting the cost from j to k (where $\tau_{jj} = 0$ for all j).

According to the supply and demand functions in each country, the autarky price of good i in country j, $p_{ji}\left(a\right)$, is given by the following expression:

$$p_{ji}(a) = \frac{A - \alpha_{ji}}{\beta + B}$$

We assume that $\alpha_{ji} \leq 0$ for all i and j; hence, the autarky price is positive for each good in any country, i.e., $p_{ji}(a) > 0$. Moreover, in order to ensure that there is a single exporter of each good k, we assume that $\alpha_{kk} - \alpha_{jk} > (\beta + B) \left(t_{jk} + \tau_{jk} \right)$. That is, because country k is the lowest cost supplier (including trade costs) of good i for all countries (i.e., $p_{ji}(a) > p_{ki}(a) + t_{jk} + \tau_{jk}$ when $i = k$ for all $j \in C \backslash \{k\}$ where $C \backslash \{k\}$ is the set of all countries excluding country k), country k is the single exporter of good i. Therefore, when

[4] See Gehrels (1956-1957) and Lipsey (1957) for early theoretical models that compare the implications of zero elasticity of demand and non-zero elasticity of demand. Also see Panagariya (2000) for a recent discussion on the implications of downward-sloped demand and upward-sloped supply.

trade is achieved, we can write the price of good k in country j by the following expression:

$$\begin{aligned} p_{jk} &= \min_i p_{ik} + t_{ji} + \tau_{ji} \\ &= p_{kk} + t_{jk} + \tau_{jk} \end{aligned}$$

Note that if $j = k$, then trade costs are zero (i.e., $t_{jk} = \tau_{jk} = 0$), so that the price of the domestically produced good is $p_{kk} = p_{jk}$ in such a case.

The market clearing condition for good i can be written as follows:

$$\sum_{m=1}^{N} \lambda_m \left(\alpha_{mk} + \beta p_{mk}\right) = \sum_{m=1}^{N} \lambda_m \left(A - B p_{mk}\right)$$

By using $p_{mk} = p_{kk} + t_{mk} + \tau_{mk}$ for all $m \in \{1, \ldots, N\}$, we can find the source (i.e., factory gate) price of good i in country k as follows when $i = k$:

$$p_{kk} = \frac{\sum_{m=1}^{N} \lambda_m \left(A - \alpha_{mk} - (\beta + B)\left(t_{mk} + \tau_{mk}\right)\right)}{\sum_{m=1}^{N} \lambda_m \left(\beta + B\right)}$$

By the assumptions of the model introduced above, p_{kk} is positive.

The volume of the imports of country j from country k is then given by the following expression:

$$M_{jk}(p_{jk}) = \lambda_j \left(A - \alpha_{jk} - (B + \beta)\, p_{jk}\right)$$

where

$$
\begin{aligned}
p_{jk} &= p_{kk} + t_{jk} + \tau_{jk} \\
&= \frac{\sum_{m=1}^{N} \lambda_m \left(A - \alpha_{mk}\right)}{\sum_{m=1}^{N} \lambda_m \left(\beta + B\right)} + \frac{\left(t_{jk} + \tau_{jk}\right) \sum_{m=1}^{N} \lambda_m - \sum_{m=1}^{N} \lambda_m \left(t_{mk} + \tau_{mk}\right)}{\sum_{m=1}^{N} \lambda_m} \\
&= \frac{\sum_{m=1}^{N} \lambda_m \left(\frac{A-\alpha_{mk}}{\beta+B} + t_{jk} + \tau_{jk} - t_{mk} - \tau_{mk}\right)}{\sum_{m=1}^{N} \lambda_m}
\end{aligned}
$$

Note that the derivative of p_{jk} with respect to t_{jk} or τ_{jk} is positive since $\lambda_j > 0$ for all j; i.e., destination prices increase in trade costs, and thus the volume of imports decreases in trade costs (and source prices) and increases in country sizes.

In order to go one step further in our analysis, we need an objective function for each country. The natural choice is, for sure, the national welfare function. Following Bond and Park (2002), national welfare is defined as the sum of consumer surplus, producer surplus and tariff incomes. As in the existing literature, we let each country to receive all the tariff income for the good that the country imports. In particular, national welfare for country j can be expressed as follows:

$$
W_j = \sum_k \left[\int_{p_{jk}}^{A/B} D_{jk}(u)\, du + \int_{-\alpha_{jk}/\beta}^{p_{jk}} X_{jk}(u)\, du \right] + \sum_k t_{jk} M_{jk}(p_{jk}) \tag{IV.1}
$$

where $\int_{p_{jk}}^{A/B} D_{jk}(u)\, du$ is the consumer surplus for good k in country j; $\int_{-\alpha_{jk}/\beta}^{p_{jk}} X_{jk}(u)\, du$ is the producer surplus for good k in country j; and $t_{jk} = 0$ if $k = j$, as before.

In order to analyze the optimal tariff setting behavior of each country, we first find the optimal tariff rates and best response functions, then we employ a stationary dynamic tariff game. The details of the approach is given in the following sections.

Analytical Results and Best Response Functions

In this section, we show the implications of each country's maximizing its welfare given the tariff rate of the other countries. We provide a closed-form solution to the maximization problem of a country. Country j chooses the optimal tariff rate for good k for all $k \in \{1, \ldots . N\}$ by

$$\max_{t_{jk}} W_j = \sum_k \left[\int_{p_{jk}}^{A/B} D_{jk}(u)\, du + \int_{-\alpha_{jk}/\beta}^{p_{jk}} X_{jk}(u)\, du \right] + \sum_k t_{jk} M_{jk}(p_{jk}).$$

The first order conditions of the above problem gives us the best response function of country j for good k. Given the tariff rates of countries other than j on good k, the best response of j for good k is:

$$t_{jk}^B = \frac{\lambda_j \left(\sum\limits_{m \in C \setminus \{j\}}^{N} \frac{\lambda_m \left(\alpha_{mk} - \alpha_{jk} \right)}{(B+\beta)} - \tau_{jk} \sum\limits_{m \in C \setminus \{j\}}^{N} \lambda_m + \sum\limits_{m \in C \setminus \{j,k\}}^{N} \lambda_m t_{mk} + \sum\limits_{m \in C \setminus \{j,k\}}^{N} \lambda_m \tau_{mk} \right)}{\left(\left(\sum\limits_{m=1}^{N} \lambda_m \right)^2 - \lambda_j^2 \right)} \tag{IV.2}$$

Therefore, the optimal tariff of each country for a good depends on the tariff rates of all other countries for the same good. The closed-form solution is discussed in the next section. According to this best response function, when a country increases its tariff for a particular good, the optimal tariff of other countries for that particular good also increases; i.e., there is tariff complementarity. Furthermore, when a country changes its tariff rate for a particular good, the tariff rates of any country on other goods are not affected.

Analytical solution: Nash Tariff Rates

In this section, we provide a closed-form solution to a country's welfare maximizing problem. By substituting demand, supply, price, and import expressions into Equation IV.1 and solving for optimal Nash tariffs for each country, we obtain the Nash tariff of country j on good k (imported from country k) as follows:

$$t_{jk}^{N} = \frac{\left\{ \begin{array}{c} q_{jk} - q_{jk} \sum_{i=1}^{|C^{*}\backslash\{j,k\}|} \left((i-1) \sum_{S_i \subset C^{n}\backslash\{j,k\}} \left(\prod_{t \in S_i} \gamma_t \right) \right) \\ + \frac{\gamma_j}{\lambda_j} \left(\sum_{m \in C^{*}\backslash\{j,k\}} \lambda_m q_{km} \left(1 + \sum_{i=1}^{|C^{*}\backslash\{j.k,m\}|} \left(\sum_{S_i \subset C^{*}\backslash\{j.k,m\}} \left(\prod_{t \in S_i} \gamma_t \right) \right) \right) \right) \end{array} \right\}}{\left\{ \begin{array}{c} 1 - \left(\sum_{i=1}^{|C^{*}\backslash\{j,k\}|} \left((i-1) \sum_{S_i \subset C^{*}\backslash\{j,k\}} \left(\prod_{t \in S_i} \gamma_t \right) \right) \right) \\ -\gamma_j \left(\sum_{i=1}^{|C^{*}\backslash\{j,k\}|} \left(i \sum_{S_i \subset |C^{*}\backslash\{j,k\}|} \left(\prod_{t \in S_i} \gamma_t \right) \right) \right) \end{array} \right\}} \quad \text{(IV.3)}$$

where

$$q_{jk} = \left(\sum_{m=1}^{N} \frac{\gamma_j \lambda_m (\alpha_{mk} - \alpha_{jk})}{\lambda_j (B + \beta)} \right) - \left(\frac{\lambda_j \tau_{jk}}{\lambda_j + \sum_{m=1}^{N} \lambda_m} \right) + \left(\sum_{m \in C\backslash\{j,k\}}^{N} \left(\frac{\gamma_j \lambda_m \tau_{mk}}{\lambda_j} \right) \right)$$

and

$$\gamma_j = \frac{\lambda_j^2}{\left(\sum_{m=1}^{N} \lambda_m \right)^2 - \lambda_j^2} \text{ for all } t$$

and S_i refers to an arbitrary subset of countries with i countries.[5] Set C^* is defined as the set of countries which are not involved in a free trade agreement. By Equation IV.3, we can find the optimal tariff rate of any country for any good, given the tariff rates of other countries. By substituting Equation IV.3, in the welfare function, the welfare function of any country can also be obtained as a function of exogenous variables.

Comparative Statics for The Case with No Agreement

In order to understand comparative statics of optimal tariffs, without loss of generality, we use the following benchmark parametrization of $\alpha_{jj} = -0.9$, $\alpha_{jk} = -1$, $B+\beta = 0.2$, $\tau_{jk} = 0$, $N = 3$, and $\lambda_j = 1$ for all j. Although the scale of these parameters is not important, they are restricted in a way that is consistent with the assumptions of the model introduced above. Nevertheless, for robustness, we consider different values of each parameter (that still satisfies the assumptions of the model), below.

The relation between optimal tariff rates and the relative size of country 1 (i.e., λ_1), ceteris paribus, is given in Figure 4. As is evident, as country 1 gets larger (i.e., as λ_1 increases), its tariff on other two symmetric countries increases, tariff rates of other two symmetric countries on country 1 decreases, and bilateral tariff between other two symmetric countries decreases as well. This result is mostly connected to increasing imports of country 1 and decreasing imports of other two symmetric countries which are important parts of welfare through tariff income.

The relation between optimal tariff rates and remoteness of country 1 (i.e., $\tau_{1j} =$

[5]For example, when we write

$$\sum_{i=1}^{n-2} \sum_{S_i \subset C \setminus \{j.k\}} \cdots$$

we mean that the summation is for any subset of $C \setminus \{j.k\}$ with i elements for all $i = 1, \ldots n-2$.

$\tau_{j1} > 0$), ceteris paribus, is given in Figure 5. As is evident, as country 1 gets remoter (i.e., as $\tau_{1j} = \tau_{j1}$ increase), its tariff on other two symmetric countries decreases, tariff rates of other two symmetric countries on country 1 decrease, and bilateral tariff between other two symmetric countries increases. In other words, as transportation costs increase between country 1 and the rest of the world, other two symmetric countries become relatively closer, and they apply higher tariff rates to each other compared to what they apply on the remote country.

The relation between optimal tariff rates and comparative advantage of each country in the good that it exports (i.e., α_{jj}'s), ceteris paribus, is given in Figure 6. As is evident, as the degree of comparative advantage increases, all symmetric countries apply higher bilateral tariff rates. This is again mostly due to increasing tariff incomes through increasing volumes of trade.

Finally, the relation between optimal tariff rates and the slope of excess supply with respect to price (i.e., $B + \beta$) , ceteris paribus, is given in Figure 7. As is evident, there is a negative relation between tariff rates of symmetric countries and the slope of excess supply, mostly because of potential gains from additional tariff income through increasing exports (and thus imports).

Figures 4-7 reflect the trade-off between many variables in terms of optimal tariffs because of their effects on the volume of trade: (i) when Figures 4 and 5 are compared, putting a tariff on a remote country has similar effects as putting a tariff on a smaller country; (ii) when Figures 6 and 7 are compared, increasing the number of countries has similar effects as decreasing comparative advantage of each country in the product that it exports; (iii) when Figures 6 and 8 are compared, increasing the number of countries has similar effects as increasing the slope of the excess supply with respect to price. In other words, just like Engel and Rogers (1996) measure the international border in terms of distance, one can measure the international border in terms of country sizes, comparative advantages, or transportation costs using the model of this paper when Nash tariffs are in charge. Such an exercise can be handled through an elasticity approach, i.e., by measuring the effects of country sizes, comparative advantages, or transportation costs on the elasticity of demand and supply.

The results in Figures 4-7 are not affected when we change transportation costs (i.e., τ_{jk}'s), comparative advantages (α_{jj}'s), absolute country sizes (λ_j's) or the slope of excess supply with respect to price (i.e., $B + \beta$), but the positions and magnitudes of comparative statics change.

Comparative Statics for The Case with Trade Agreements

In this section, we analyze how the welfare of a country changes under possible free trade agreements with different countries. Figures 4-7 reflect that countries apply higher tariff rates to countries that they have higher imports from. Although there can be many different reasons to have high import volumes, in the context of regional trade agreements, highest tariff rates would be applied on imports coming from closer countries, since the volume of trade is higher with such countries due to low transportation costs.

We consider the benchmark parametrization of $\alpha_{jj} = -0.9$, $\alpha_{jk} = -1$, $B+\beta = 0.2$, $N = 3$, and $\lambda_j = 1$ for all j, and we change the remoteness of country 1 through its transportation costs (i.e., τ_{j1} and τ_{1j}) which is used as a measure of distance. By using this parametrization, we compare the welfare of country 2 when it makes an agreement with country 1 (a remote region with a transportation cost of $\tau_{21} = \tau_{12} > 0$) with its welfare when it makes an agreement with country 3 (a close region with a transportation cost of

$\tau_{23} = \tau_{32} = 0$). Formally, welfare gains of country j from a trade agreement between country 2 and country 3 compared to an agreement between country 1 and country 2 is given by the following expression:

$$\Theta_j = W_j\left(t^A_{23}, t^A_{32}, t^B_{(-i)3}, t^B_{(-i)2}\right) - W_j\left(t^A_{12}, t^A_{21}, t^B_{(-i)1}, t^B_{(-i)2}\right)$$

where t^A_{mn} represents the agreed tariff rate between countries m and n, which is set equal to zero for simplicity, and $t^B_{-(i)n}$ represents the optimal tariff rates of other countries on good n given the agreed tariff rates.

The welfare gains of a regional trade agreement between country 2 and country 3 compared to an agreement between country 1 and country 2 (i.e., Θ_j's) are depicted in Figure 8. As country 2 is the one for which we make the comparison, we should focus on its welfare gains: as is evident, as country 1 gets remote, the welfare gain of country 2 increases. This is true mainly because of two reasons: 1) Increasing transportation costs between country 1 and country 2 decreases country 2's welfare from having an agreement with country 1, and 2) As country 1 gets remote, the tariff rate between country 2 and country 3 increases (Figure 5) and hence, the welfare of country 2 from having an agreement with country 3 increases. This result is true for alternative parameterizations as well, although the magnitude of welfare gains are different in alternative cases. When transportation costs decrease (i.e., if countries are closer to each other), the welfare gains from a regional trade agreement increase. As is also evident, the welfare gain of the world (i.e., the sum of the welfare gains of all countries) also increases when transportation cost decreases. As expected, welfare gains of country 1 are negative (and increasing with transportation costs), while welfare gains of country 3 are positive (and decreasing with transportation costs).

The Stationary Dynamic Tariff Game

It is well known that repeated interactions between parties can be used to support payoffs that Pareto dominate those obtained in the one shot game. In this section, following the lead of Bond (2001), countries play a repeated game for tariff rates. The repeated structure of the game gives more flexibility to our model in terms of incorporating the self-enforcement agreements.

In this game, each country is able to compare future payoffs out of a possible collusion (cooperation) and out of a possible deviation from a free trade agreement.[6] Country j's welfare when j makes an agreement with k is represented by $W_j\left(t^A_{jk}, t^A_{kj}, t^B_{(-i)j}, t^B_{(-i)k}\right)$ where t^A_{jk} and t^A_{kj} are the agreed tariff rates for good j and k, respectively, and $t^B_{(-i)j}$ ($t^B_{(-i)k}$) is the optimal tariff rates of all other countries on good j (resp., k). If country j cheats on country k, j sets its optimal tariff rate on good k, t^B_{jk}, given t^A_{kj}, $t^B_{(-i)j}$ and $t^B_{(-i)k}$. In this game, countries follow a grim trigger strategy, i.e., if country j cheats on country k, then both countries set their optimal tariff rates in the future and no agreement can be formed in the future.

In order to sustain collusion in a trade agreement, the trade-off between the gains from deviating from an agreed-upon tariff policy and the discounted expected future gains from collusion must be balanced in a way that the latter should keep away countries from deviating. That is,

$$\frac{1}{1-\delta_j} W_j\left(t^A_{jk}, t^A_{kj}, t^B_{(-i)j}, t^B_{(-i)k}\right) \geq \left\{ \begin{array}{c} W_j\left(t^B_{jk}, t^A_{kj}, t^B_{(-i)j}, t^B_{(-i)k}\right) \\ +\frac{\delta_j}{1-\delta_j} W_j\left(t^B_{jk}, t^B_{kj}, t^B_{(-i)j}, t^B_{(-i)k}\right) \end{array} \right\}$$

[6]Although this paper investigates free trade agreements, the model of this paper can easily be used to investigate custom unions.

where δ_j is the discount factor of country j. Hence, country j cooperates if and only if

$$\left(\frac{\delta_j}{1-\delta_j}\right)\Psi_j - \Omega_j \geq 0 \tag{IV.4}$$

or equivalently

$$\delta_j \geq \frac{\Omega_j}{\Psi_j + \Omega_j}$$

where Ψ_j is the one-period value of cooperation for country j, i.e.,

$$\Psi_j = W_j\left(t^A_{jk}, t^A_{kj}, t^B_{(-i)j}, t^B_{(-i)k}\right) - W_j\left(t^B_{jk}, t^B_{kj}, t^B_{(-i)j}, t^B_{(-i)k}\right)$$

and Ω_j is the welfare gain of country j from cheating to country k, i.e.,

$$\Omega_j = W_j\left(t^B_{jk}, t^A_{kj}, t^B_{(-i)j}, t^B_{(-i)k}\right) - W_j\left(t^A_{jk}, t^A_{kj}, t^B_{(-i)j}, t^B_{(-i)k}\right)$$

In other words, for country j to have a cooperation, the minimum discount factor of country j should be:

$$\underline{\delta}_j = \frac{\Omega_j}{\Psi_j + \Omega_j}$$

In our analysis, we calculate $\underline{\delta}_j$ values under different transportation costs to connect the model to regional trade agreements.[7]

Trade Agreements in a Stationary Dynamic Tariff Game

Consider the benchmark parametrization of $\alpha_{jj} = -0.9$, $\alpha_{jk} = -1$, $B + \beta = 0.2$, $N = 3$, and $\lambda_j = 1$ for all j, and we change the remoteness of country 1 through its transportation costs (i.e., τ_{j1} and τ_{1j}) which is used as a measure of distance. By using this

[7]Note that both the welfare function and the minimum discount factor can be written as closed form expressions because we already know the closed form expression of the best response tariff.

parametrization, we compare the minimum discount factor of country 2 when it cooperates with country 1, say, $\underline{\delta}_2^1$ (i.e., a bilateral trade agreement) with its minimum discount factor when it cooperates with country 3, say, $\underline{\delta}_2^3$ (i.e., another bilateral trade agreement), and its minimum discount factor when it cooperates with countries 1 and 3, say, $\underline{\delta}_2^{1,3}$ (i.e., a multilateral trade agreement among all countries). Such a comparison helps us understand the differences in sustainability between regional and multilateral trade agreements in a stationary dynamic tariff game.

From Country 2's point of view, the relation between the minimum discount factor of country 2 $\underline{\delta}_2$ and the relative size of country 1 λ_1 is given in Figure 9. Independent of alternative parameterizations, it is seen that when countries are symmetric in size (i.e., $\lambda_j = 1$ for all j), a multilateral trade agreement is harder to sustain for country 2, while a bilateral trade agreement is easier. However, as country 1 gets larger, ceteris paribus, it is easier for country 2 to sustain an agreement with country 1 (i.e., the larger country), while it becomes relatively harder to sustain an agreement with country 3 (i.e., the smaller country). Nevertheless, as country 1 gets large enough, the minimum discount factor of country 2 in the case of an agreement with country 1 gets very close to the one in the case of a multilateral agreement with all countries. This is mostly due to country 1 having a larger share in the world economy as it gets larger.

From Country 2's point of view, the relation between the minimum discount factor of country 2 $\underline{\delta}_2$ and the relative size of country 2 λ_2 is given in Figure 10. As seen in Figure 10, as country 2 gets larger, both a bilateral trade agreement and a multilateral trade agreement become harder to sustain for country 2. Moreover, as country 2 gets larger, a cooperation in a multilateral agreement becomes relatively more difficult to sustain compared to a cooperation in a bilateral agreement.

When countries are asymmetric in terms of their sizes, in order the check whether a collusion is sustainable between asymmetric countries in general, we need to use the maximum of the minimum discount factors of countries involving in an agreement; i.e., we need to calculate $\max\{\delta_j, \ldots, \delta_k\}$ where $j, ..., k$ are the countries involving in an agreement. Such an analysis is achieved in Figure 11. As is evident, when country 1 gets larger, a multilateral or a bilateral agreement involving country 1 (i.e., the large country) is harder to sustain, because the gains in terms of reduced tariffs get smaller for country 1.

Figure 12 investigates the possible implications of transportation costs in trade agreements: It shows how the minimum discount factor of country 2 and the remoteness of country 1 are related. As country 1 gets remote, cooperation in a regional or a multilateral agreement becomes harder to sustain for country 2. A regional agreement with country 3 is still relatively easier to sustain for country 2. In fact, the degree of sustainability of a regional agreement with country 3 is the least to be affected by the remoteness of country 1. This is due to the fact that as country 1 gets remoter, trade volume between countries 2 and 3 increases, and hence, the gains from an agreement also increases. The degree of sustainability of a multilateral agreement is still higher than that of a regional agreement even when the regional agreement is with a remote country. This is the second main result of this paper: Lower transportation costs play an important role in sustaining regional trade agreements.

In Figures 13 and 14, we see that the comparative advantage of country 2 and the slope of the excess supply, respectively, have no effect on the minimum discount factor of country 2. Nevertheless, a multilateral agreement is always more difficult to sustain compared to a bilateral agreement.

Conclusion

This paper has shown that the existence of transportation costs may be a possible reason for explaining why trade agreements are mostly regional by introducing an N-country international trade model with different country sizes and comparative advantages. In a case with no agreements, we show that the optimal tariff rates are decreasing in transportation costs; i.e., closer countries set higher tariffs to each other in equilibrium. If such high tariffs are reduced between closer countries, the welfare gain is shown to be much higher compared to reducing tariff rates with a remote country; i.e., countries tend to make regional agreements to maximize their welfare. Moreover, according to the best response functions of countries, (i) as a country gets larger, other countries apply lower optimal tariff rates to that country; (ii) as a country gets remoter from others, other countries become relatively closer, and they apply higher optimal tariff rates to each other compared to what they apply on the remote country; (iii) as the degree of comparative advantage increases across countries, all symmetric countries apply higher optimal bilateral tariff rates; (iv) there is a negative relation between tariff rates of symmetric countries and the slope of their excess supply.

After that, we make a welfare analysis by using a stationary dynamic tariff game approach and show that, for an individual country, (i) it is harder to sustain multilateral trade agreements (compared to a regional one) when country sizes are close to each other; (ii) it is more difficult to sustain a regional trade agreement with a relatively small country than a multilateral agreement when a relatively big country gets involved in the multilateral agreement; (iii) it is harder to sustain a trade agreement when the country itself gets larger; (iv) it is harder to sustain a multilateral trade agreement (compared to a regional one) when a country gets larger; (v) lower transportation costs play a significant role in sustaining regional trade agreements compared to sustaining multilateral trade agreements. The results hold under alternative cases, which further support the analysis.

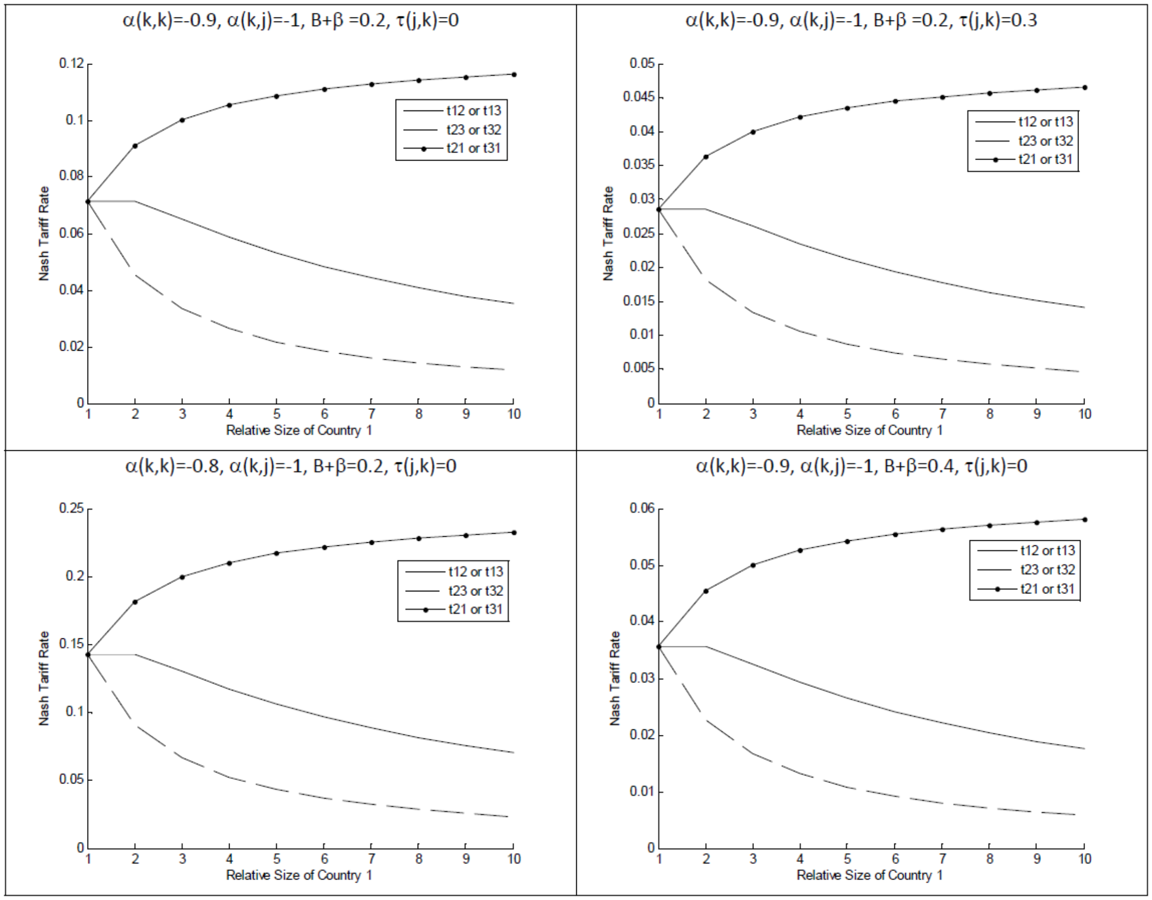

Figure 4. Nash Tariff Rates versus the Relative Size of Country 1

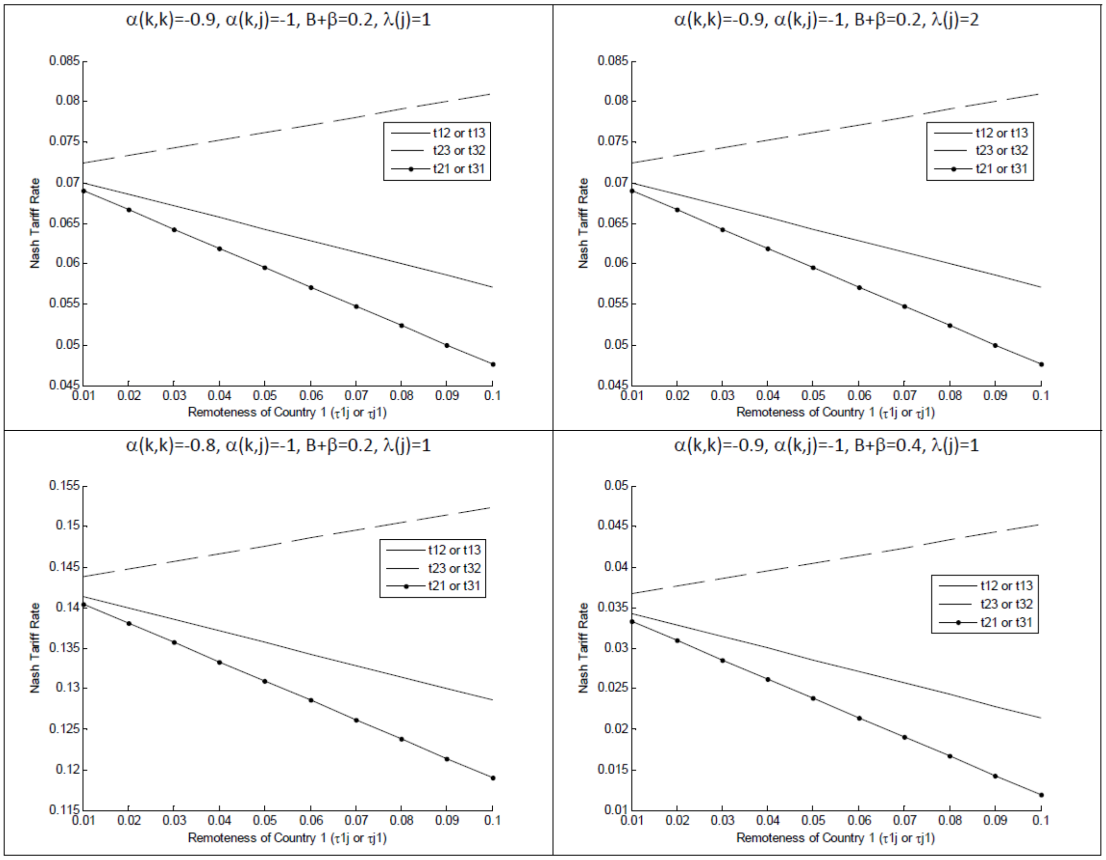

Figure 5. Nash Tariff Rates versus the Remoteness of Country 1

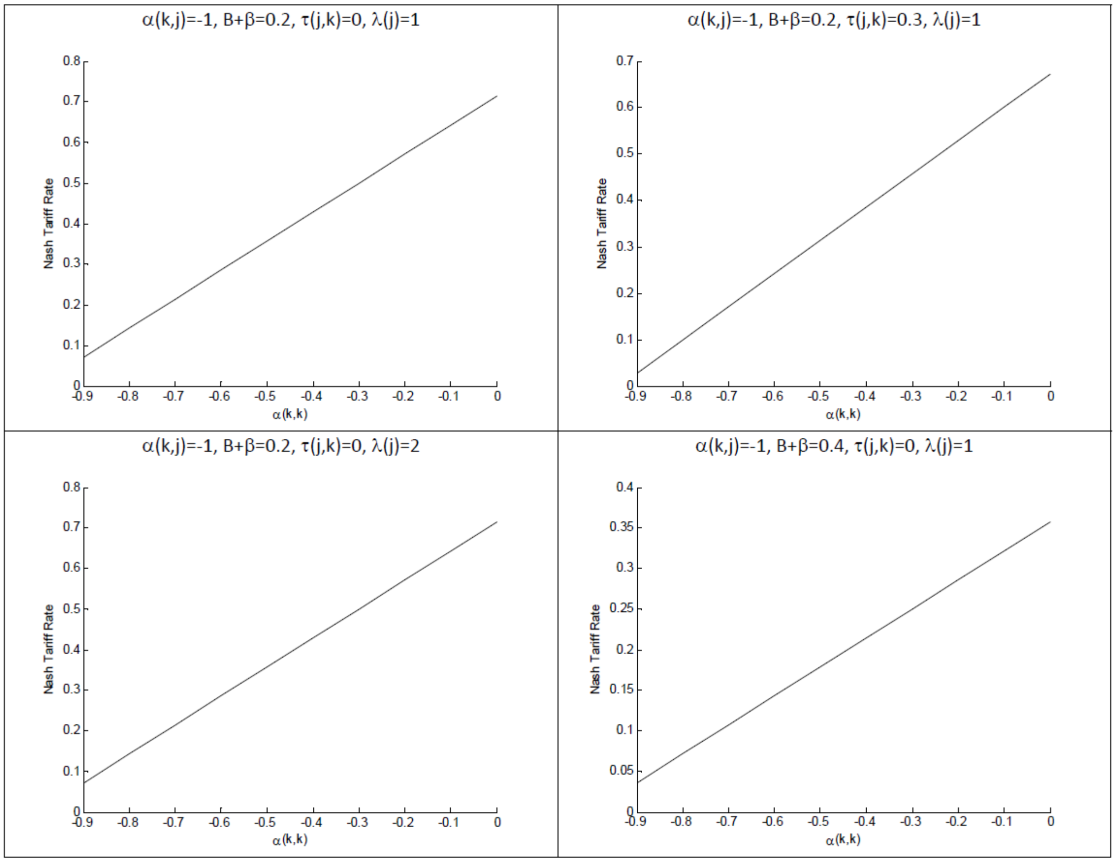

Figure 6. Nash Tariff Rates versus Comparative Advantage

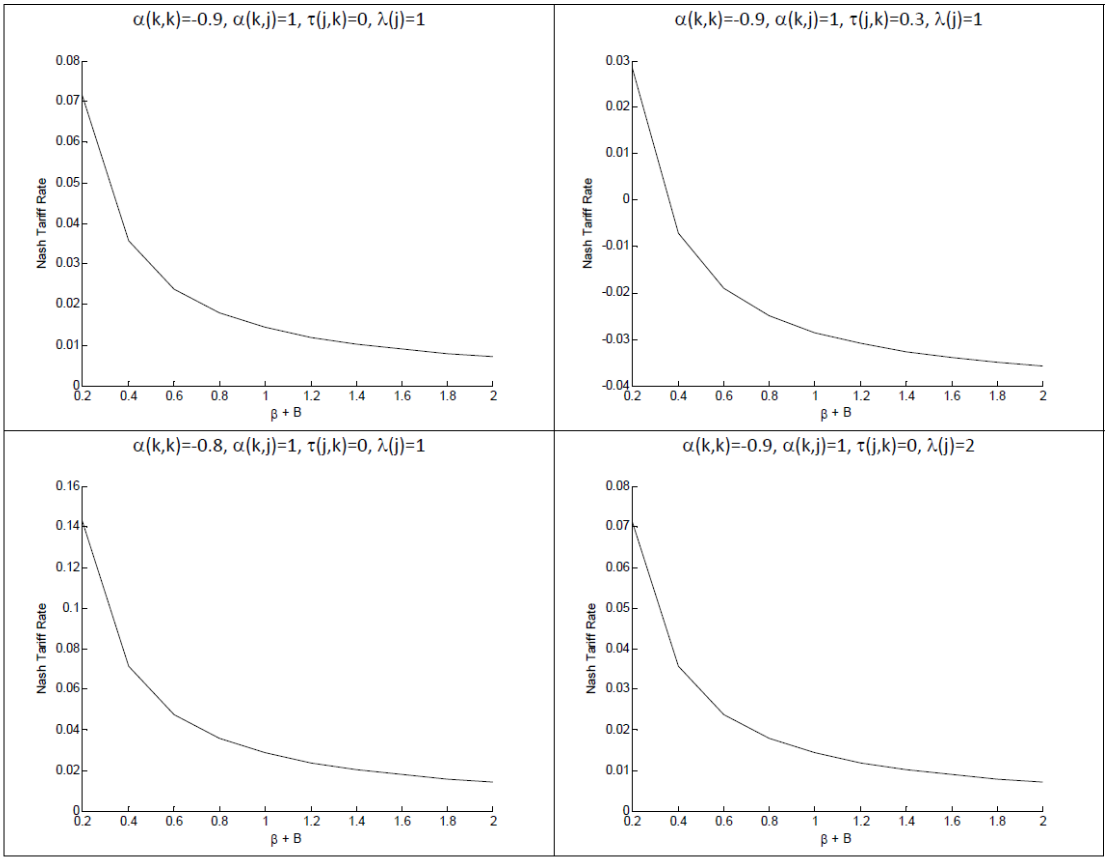

Figure 7. Nash Tariff Rates versus the Slope of Excess Supply

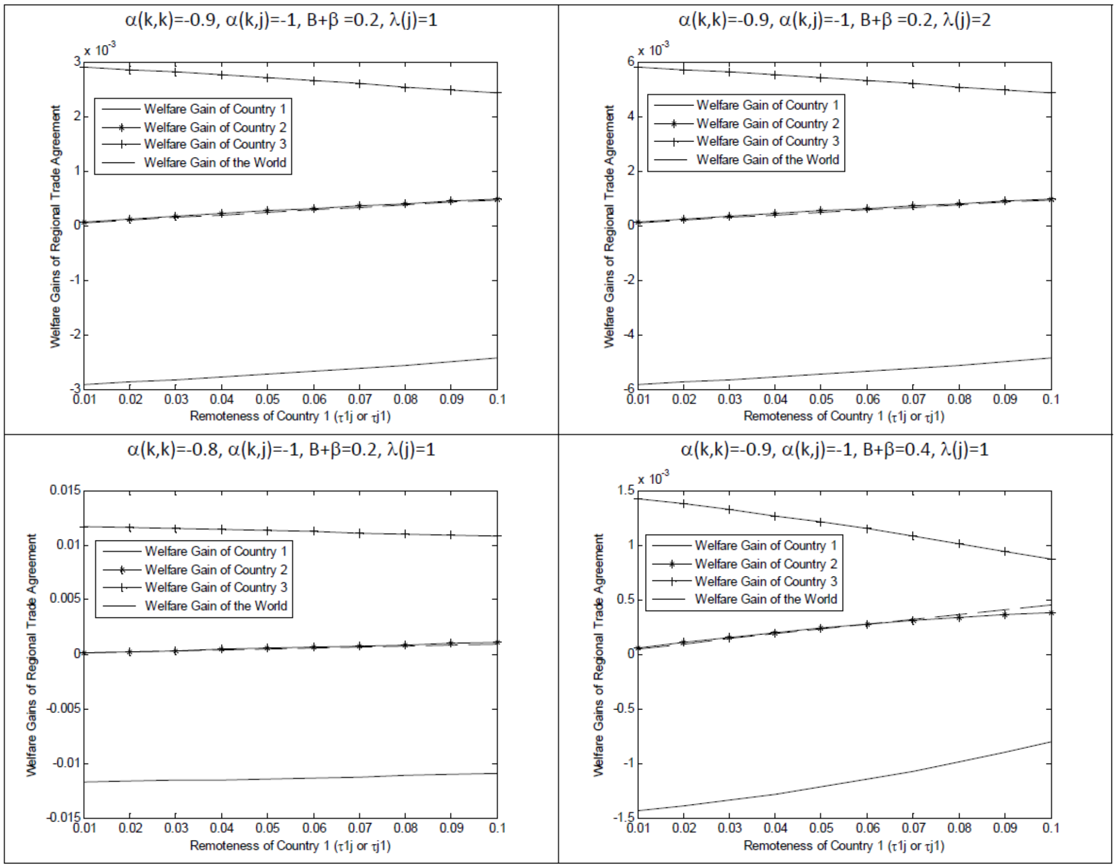

Figure 8. Welfare Gains of a Regional Trade Agreement

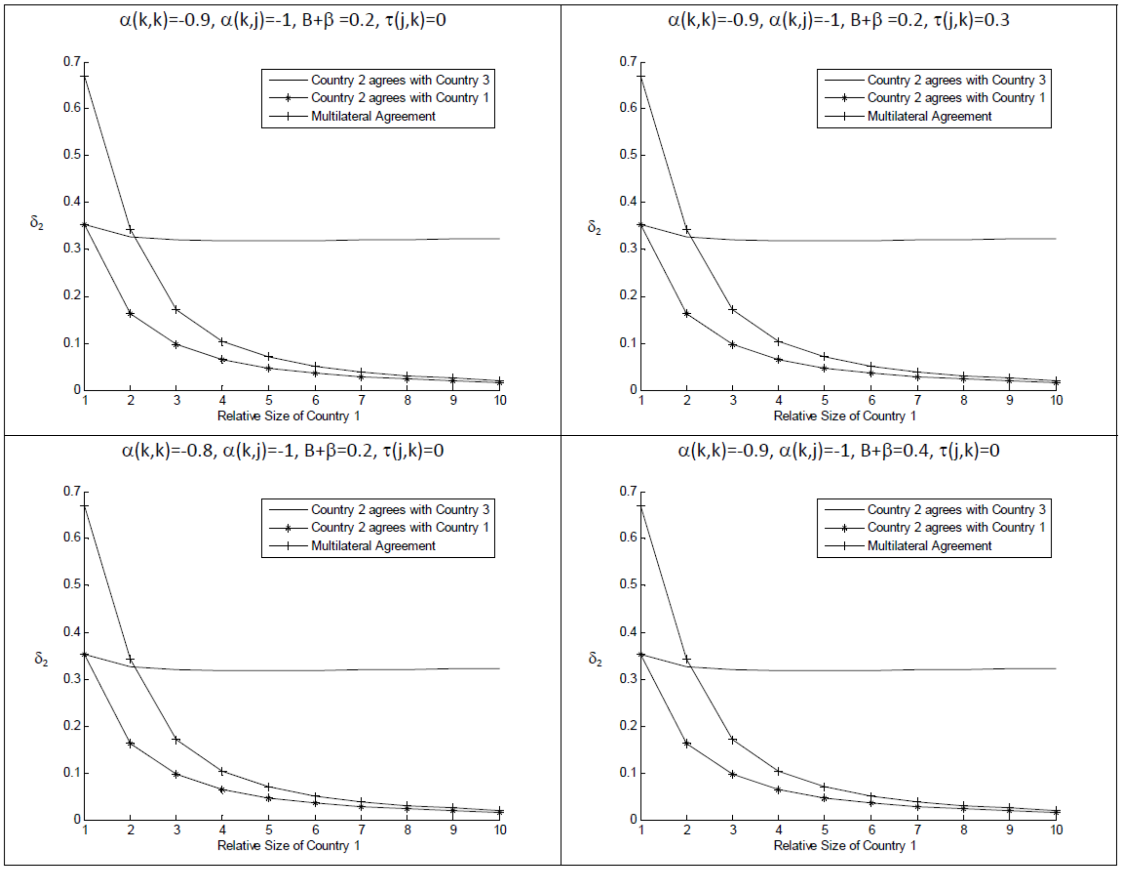

Figure 9. Minimum Discount Factor of Country 2 versus the Relative Size of Country 1

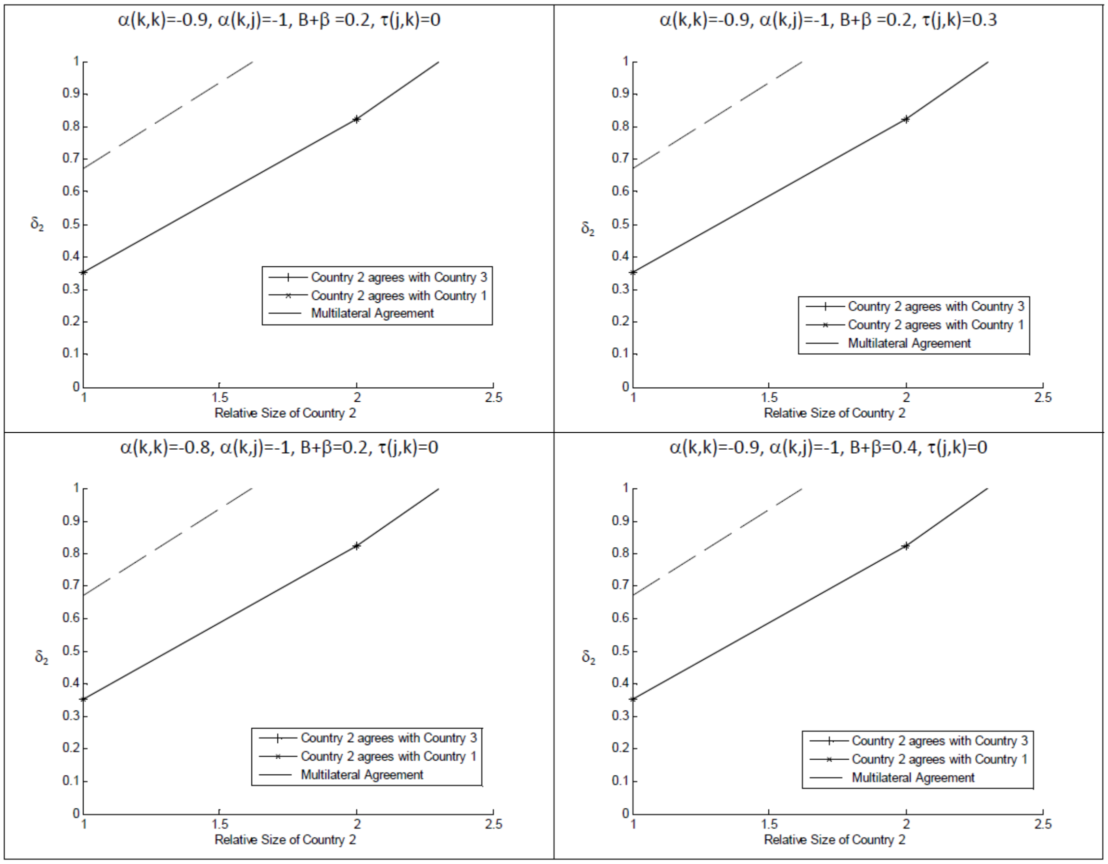

Figure 10. Minimum Discount Factor of Country 2 versus the Relative Size of Country 2

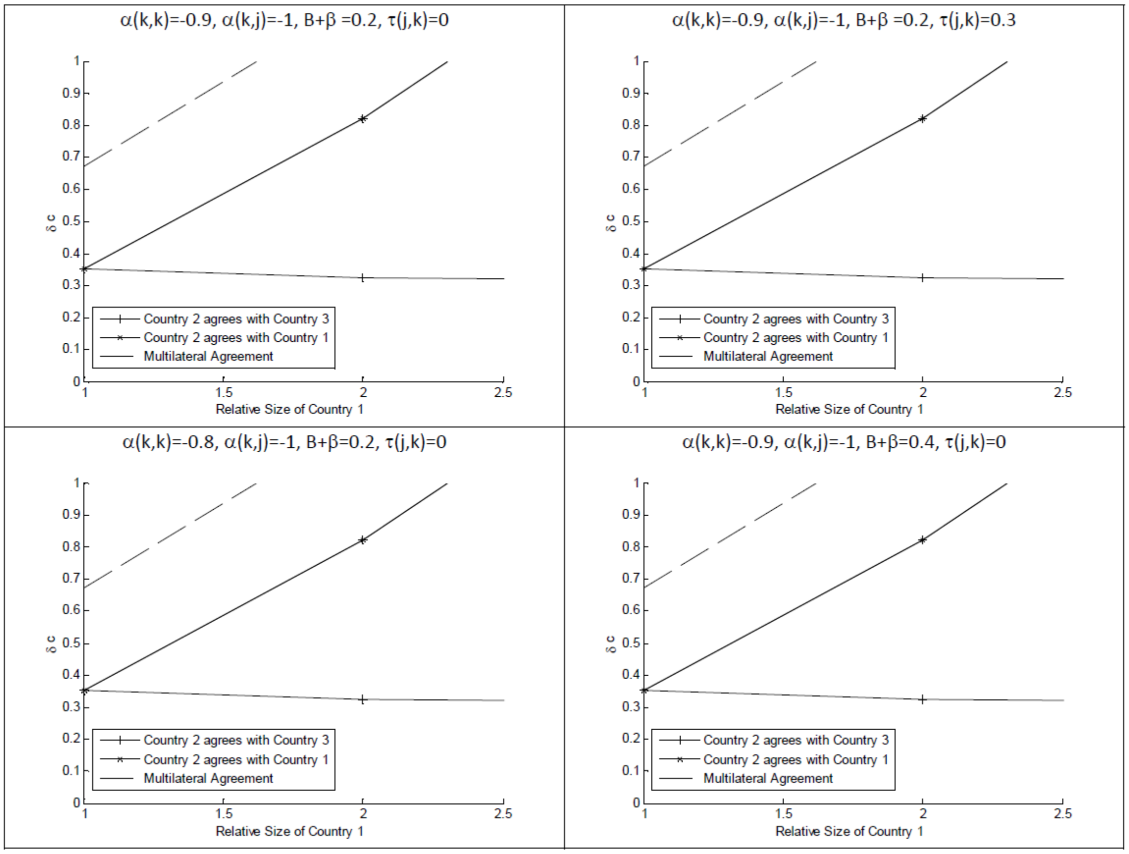

Figure 11. Minimum Discount Factor for Collusion versus the Relative Size of Country 1

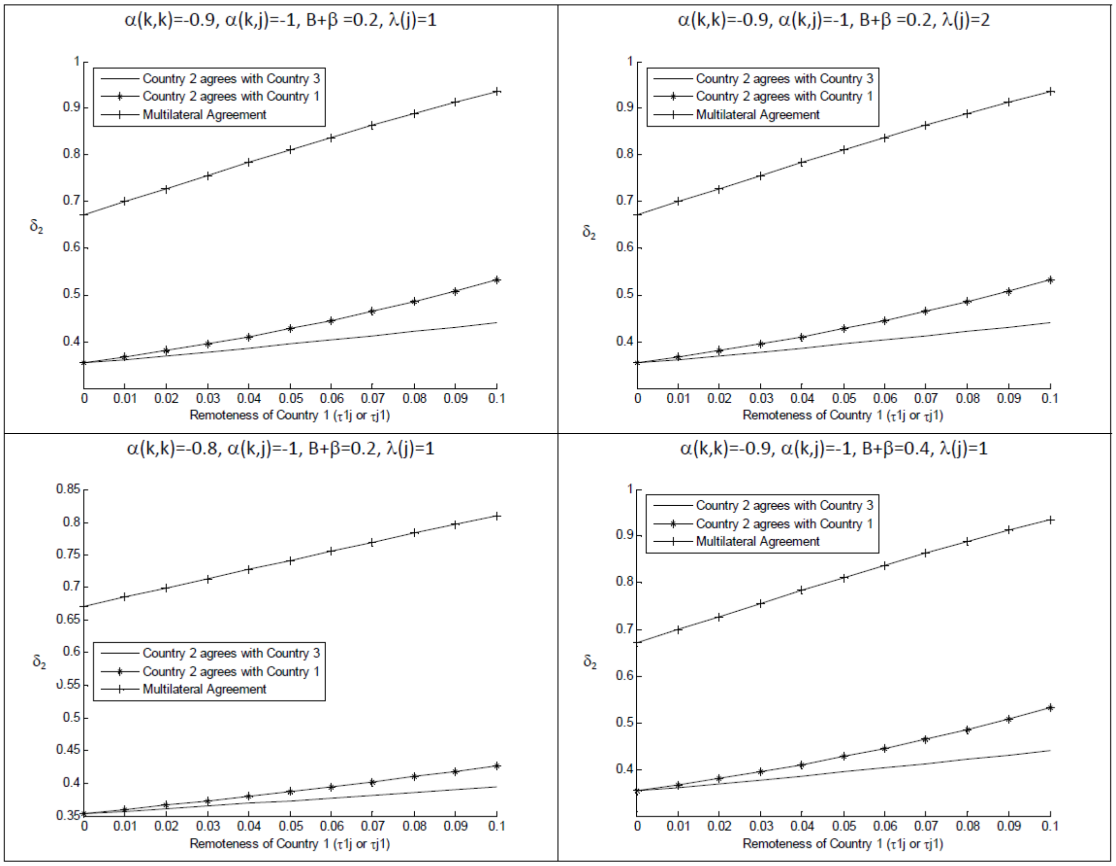

Figure 12. Minimum Discount Factor of Country 2 versus the Remoteness of Country 1

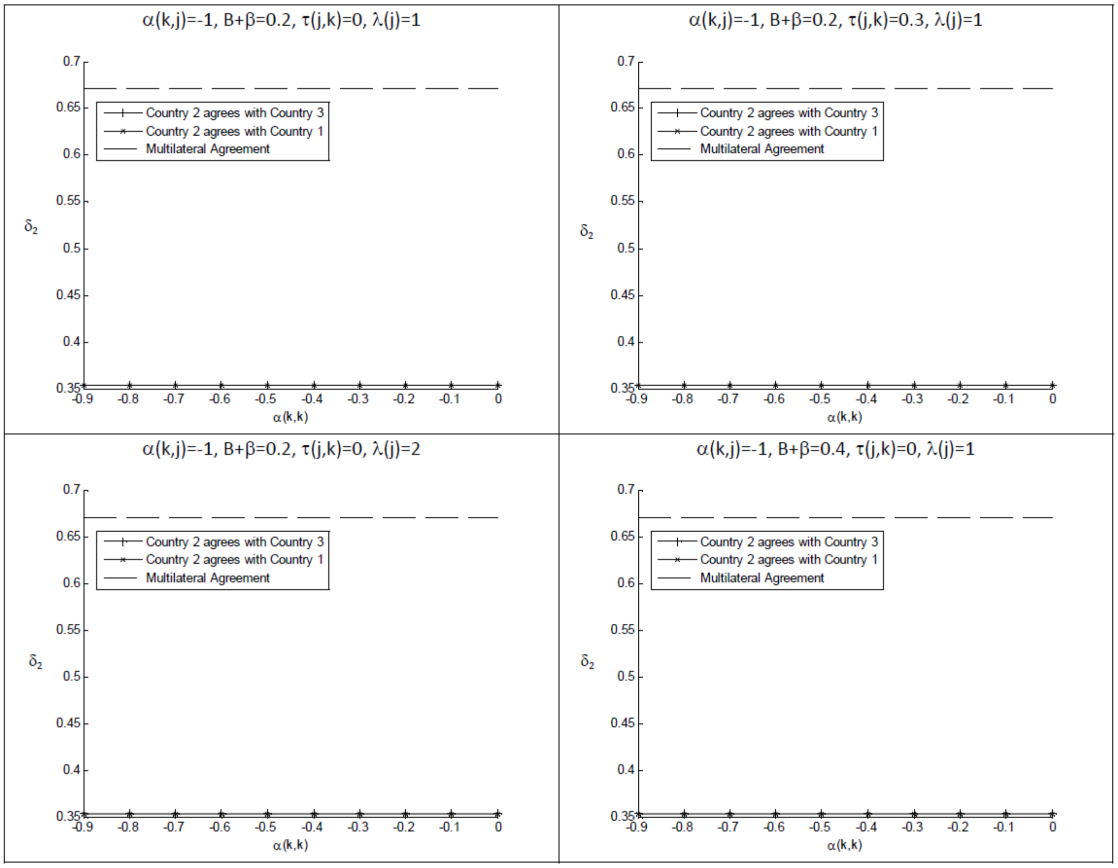

Figure 13. Minimum Discount Factor of Country 2 versus Comparative Advantage

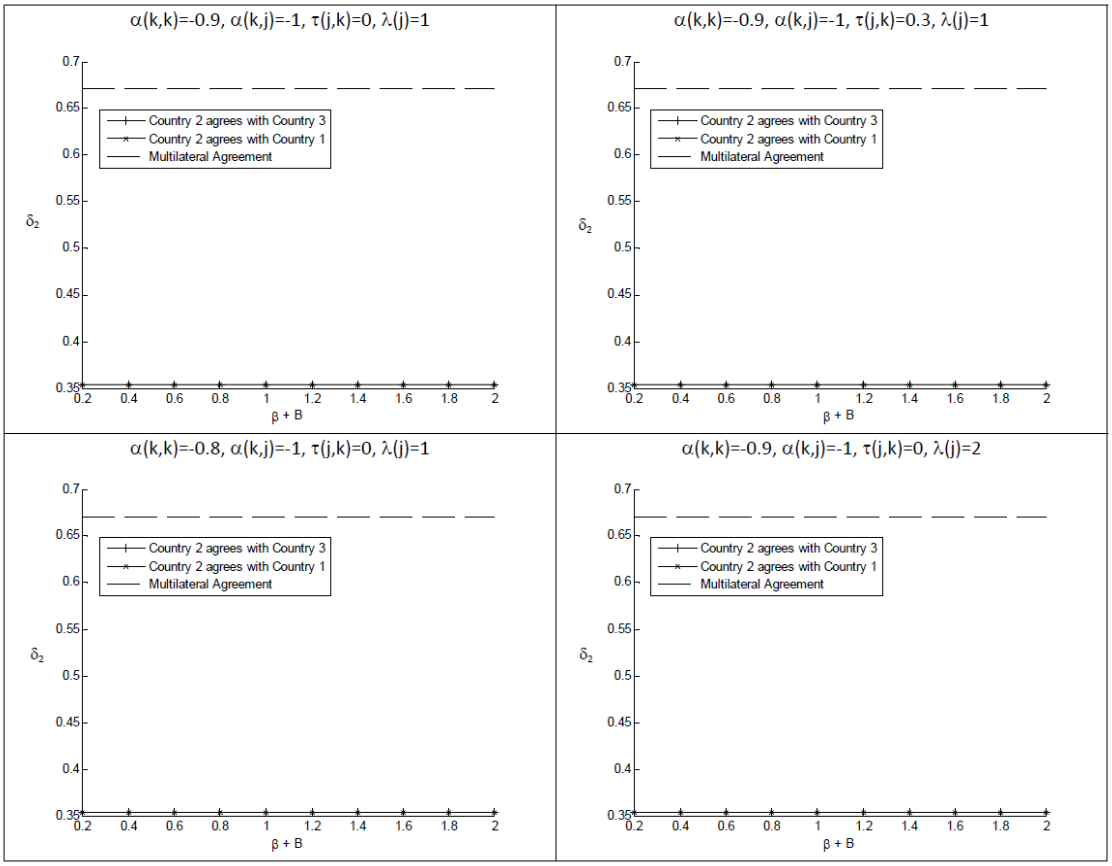

Figure 14. Minimum Discount Factor of Country 2 versus the Slope of Excess Supply

BIBLIOGRAPHY

[1] Abdulkadiroglu, A., Pathak, P., Roth, A., 2005A. The New York City High School Match. *American Econ Review*, 95(2):364-367.

[2] Abdulkadiroglu, A., Pathak, P., Roth, A., Sonmez, T., 2005B. The Boston Public School Match. *American Econ Review,* 95(2):368-371.

[3] Abdulkadiroglu, A., Sonmez, T., 1999. House Allocation with Existing Tenants. *J Econ Theory,* 88(2):233-260.

[4] Abdulkadiroglu, A., Sonmez, T., 2003. School Choice: A Mechanism Design Approach. *American Econ Review*, 93(3):729-747.

[5] Abeledo, H., Isaak, G., 1991. A characterization of graphs that ensure the existence of stable matchings. *Math. Soc. Sci., 22(1):93-96.*

[6] Abraham D., Biro P., and D., Manlove, 2005. Almost stable" matchings in the roommates problem. *Proceedings of WAOA 2005: the 3rd workshop on approximation and online algorithms*, Vol. 3879 of Lecture Notes in Computer Science:1–14.

[7] Alcalde, J., 1996. Implementation of stable solutions to marriage problems. *J Econ Theory*, 69:240–254.

[8] Alcalde J, Romero-Medina, A., 2000. Simple mechanisms to implement the core of college admissions problems. *Games Econ Behav,* 31:294–302.

[9] Alcalde, J., Romero-Medina, A., 2005. Sequential decisions in the college admissions problems. *Econ Lett,* 86:153–158.

[10] Bagwell, K. and Staiger, R. W. (1997a). Multilateral Tariff Cooperation During the Formations of Customs Unions. *Journal of International Economics*. 42: 91-123.

[11] Bagwell, K. and Staiger, R. W. (1997b). Multilateral Tariff Cooperation During the Formation of Regional Free Trade Areas. *International Economic Review*. 38(2).

[12] Bagwell, K. and Staiger, R. W. (1999). Regionalism and multilateral tariff cooperation. In Pigott, John, Woodland, Alan (Eds.), International Trade Policy and the Pacific Rim. Macmillan, London.

[13] Berglas, E. (1983). The Case for Unilateral Tariff Reductions: Foreign Tariffs Rediscovered. American Economic Review 73(5): 1142-43.

[14] Bhagwati, J. (1993). Regionalism and Multilateralism: An Overview. In Jaime de Melo and Arvind Panagariya, eds. *New Dimensions in Regional Integration*. 22-51. Cambridge: Cambridge U. Press.

[15] Bhagwati, J. and Panagariya, A. (1996). Preferential Trading Areas and Multilateralism: Strangers, Friends or Foes? In Jagdish Bhagwati and Arvind Panagariya, eds. *The Economics of Preferential Trade Agreements*. 1-78. Washington, D.C. AEI Press.

[16] Bond, E.W. (2001). Multilateralism vs. Regionalism: Tariff Cooperation and Interregional Transport Costs. In Sajal Lahiri, ed. *Regionalism and Globalization: Theory and Practice*. Routledge Press.

[17] Bond, E.W. and Park, J-H. (2002) Gradualism in Trade Agreements with Asymmetric Countries. *The Review of Economic Studies*. 69(2): 379-406.

[18] Bond, E. W. and Syropoulos, C. (1995). Trading Blocs and the Sustainability of Inter-Regional Cooperation. In M. Canzoneri, W.J. Ethier and V. Grilli, (eds.), *The New Transatlantic Economy*, London: Cambridge University Press.

[19] Bond, E. W., Syropoulos, C., and Winters, L. A. (2001). Deepening of regional integration and multilateral trade agreements. *Journal of International Economics*. 53(2): 335-361.

[20] Chen, Y., Sonmez, T., 2006. School choice: an experimental study. *J Econ Theory*, 127(1): 202-231.

[21] Chung, K., 2000. On the Existence of Stable Roommate Matchings. *Games Econ Behav, 33(2):206-230.*

[22] Deardorff, A.V., (1998). Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World? In J.A. Frankel ed. *The Regionalization of the World Economy,*.7–32. Chicago: U. Chicago Press.

[23] Eeckhout, J., 2000. On the uniqueness of stable marriage matchings. *Econ Lett,* 69:1–8.

[24] Ergin, H., Sonmez, T., 2006. Games of school choice under the Boston mechanism. *J Public Econ*, 90(1-2):215-237.

[25] Frankel, J. A. (1997). *Regional Trading Blocs in the World Trading System* Washington, DC: Institute for International Econ.

[26] Frankel, J. and Wei, S-J. (1997). The New Regionalism and Asia: Impact and Options. In A. Panagariya, M.G. Quibria and N. Rao, eds., *The Global trading System and Developing Asia*, Hong Kong: Oxford University Press.

[27] Frankel, J., Stein, E. and Wei, S. (1995). Trading Blocs and the Americas: The Natural, the Unnatural and the Supernatural. *Journal of Development Economics* 47: 61-96.

[28] Freund, C. (2000). Multilateralism and the endogenous formation of preferential trade agreements. *Journal of International Economics*. 52: 359–376.

[29] Gale, D., Shapley, L., 1962. College admissions and the stability of marriage. *Am Math Mon,* 69:9–15.

[30] Inarra, E., Larrea, C., and E., Molis, 2008. Random paths to *P-stability* in the roommate problem. *Int J Game Theory*, 36:461–471.

[31] Inarra, E., Larrea, C., and E., Molis, 2010. The stability of the roommate problem revisited. *Unpublished*.

[32] Irving, R., Leather, P., 1986. The complexity of counting stable marriages. *J Computing, 15(3):655-667.*

[33] Kara, T., Sönmez, T., 1996. Nash implementation of matching rules. *J Econ Theory,* 68:425–439.

[34] Klaus, B., Klijn, F., and M., Walzl, 2010. Stochastic Stability for Roommate Markets. *J. Econ. Theory*, in press.

[35] Krugman, P. (1991). The Move to Free Trade Zones. *Policy Implications of Trade and Currency Zones*. Symposium sponsored by Fed. Reserve Bank Kansas City: 7-41.

[36] Krugman, P. (1993). Regionalism versus Multilateralism: Analytical Notes. In Jaime de Melo and Arvind Panagariya, eds *New Dimensions in Regional Integration*. 58-84. Cambridge, UK: Cambridge U. Press.

[37] Limão, N. (2007). Are Preferential Trade Agreements with Non-trade Objectives a Stumbling Block for Multilateral Liberalization? *Review of Economic Studies*. 74(3): 821–85.

[38] Lipsey, R.G. (1957). The theory of customs unions; trade diversion and welfare. *Economica,* 24: 40-46.

[39] McVitie, D. G., Wilson, L. B., 1971. The stable marriage problem. *Communications of the ACM*, 14(7):486-490.

[40] Meade, J.E. (1955). *The Theory of Customs Unions*. Amsterdam: North-Holland.

[41] Panagariya, A. (1997). The Meade Model of Preferential Trading: History, Analytics and Policy Implications. Essays in Honor of Peter B. Kenen. B. J. Cohen, ed. *International Trade and Finance: New Frontiers for Research*. 57-88. NY: Cambridge U. Press.

[42] Panagariya, A. (1998). Do Transport Costs Justify Regional Preferential Trade Arrangements? NO. *Weltwirtschaftliches Archiv*, 134(2): 280-301.

[43] Panagariya, A. (2000). Preferential Trade Liberalization: The Traditional Theory and New Developments. *Journal of Economic Literature*, 38(2): 287-331.

[44] Roth, A., 1984. Misrepresentation and stability in the marriage problem. *J Econ Theory*, 34(2):383-387.

[45] Roth, A., Sotomayor, M., 1990. Two-sided matching: a study in game-theoretic modeling and analysis. Cambridge Univ. Press, New York.

[46] Roth, A., Sönmez, T., Ünver, U., 2004. Kidney Exchange, *The Quarterly J Econ*, 119(2):457-488.

[47] Roth, A., Sönmez, T., Ünver, U., 2005. Pairwise kidney exchange. *J Econ Theory*, 125(2):151-188.

[48] Saggi, K. and Yildiz, H. M. (2010). Bilateralism, multilateralism, and the quest for global free trade. *Journal of International Economics*. 81: 26–37.

[49] Shapley, L., Scarf, H., 1974. On Cores and Indivisibility. *J Math Econ*, 1:23-28.

[50] Shin, S., Suh, S., 1996. A mechanism implementing the stable rule in marriage problems. *Econ Lett,* 51:185–189.

[51] Sotomayor, M., 2003. Reaching the core of the marriage market through a non-revelation matching mechanism. *Int J Game Theory,* 32:241–251.

[52] Summers, L. (1991). Regionalism and the World Trading System. *Policy Implications of Trade and Currency Zones*. Symposium sponsored by Federal Reserve Bank Kansas City: 295-301

[53] Suh, S., Wen, Q., 2008. Subgame perfect implementation of stable matchings in marriage problems. *SocChoice and Welfare*, 31:163-174.

[54] Tan, J. J. M., 1990. On a maximum stable matching problem. *BIT* ,29:631–640.

[55] Tan, J. J. M., 1991. A necessary and sufficient condition for the existence of a complete stable matching. *J. Algorithms,* 12:154–178.

[56] Viner, J. (1950). *The Customs Union Issue*. NY: Carnegie Endowment for International Peace.

[57] Wonnacott, P. and Lutz, M. (1989). Is There a Case for Free Trade Areas?. In Jeffrey Schott, ed. *Free Trade Areas and U.S. Trade Policy*. 59-84. Washington, DC: Institute for International Econ.

[58] Wonnacott, P. and Wonnacott, R.J. (1981). Is Unilateral Tariff Reduction Preferable to a Customs Union? The Curious Case of the Missing Foreign Tariff. *American Economic Review,* 71(4): 704-14.

[59] Zissimos, B. (2007). Regional Trade Blocks in a Non-cooperative Network. Vanderbilt University, *mimeo.*