EL COLEGIO DE MÉXICO, A.C.
CENTRO DE ESTUDIOS ECONÓMICOS

THREE ESSAYS ON SCHOOL CHOICE

TESIS PRESENTADA POR:

MIRIAM LIZBETH GONZÁLEZ CAVAZOS

PROMOCIÓN 2016-2019

CIUDAD DE MÉXICO ENERO DE 2020

EL COLEGIO DE MÉXICO, A.C.
CENTRO DE ESTUDIOS ECONÓMICOS

THREE ESSAYS ON SCHOOL CHOICE

TESIS PRESENTADA POR:

MIRIAM LIZBETH GONZÁLEZ CAVAZOS

PARA OPTAR POR EL GRADO DE

DOCTORA EN ECONOMÍA

PROMOCIÓN 2016-2019

DIRECTOR DE TESIS

DR. DAVID RENÉ MICHEL CANTALA

CIUDAD DE MÉXICO ENERO DE 2020

EL COLEGIO DE MÉXICO

CENTRO DE ESTUDIOS ECONÓMICOS

CONSTANCIA DE APROBACIÓN

Doctorante: Miriam Lizbeth González Cavazos

Tesis: Three Essays on School Choice

Director de Tesis: Dr. David René Michel Cantala

Aprobada por el Jurado Examinador:

Dr. David René Michel Cantala	Presidente	______________________
Dr. Leobardo Plata Pérez	Primer Vocal	______________________
Dr. Damián Emilio Gibaja Romero	Vocal Secretario	______________________
Dra. Adriana Gama Velázquez	Suplente	______________________

Ciudad de México, 24 de enero de 2020

Dedicado a mis padres,
Rosy y Héctor.

Acknowledgments

Durante el desarrollo de este trabajo conté con el apoyo de varias personas e instituciones, las cuáles ayudaron a la culminación del mismo. Por medio de las siguientes líneas quisiera reconocer y agradecer a cada uno de ellos.

En primer lugar, quisiera reconocer el apoyo brindado por mi director de tesis, el Doctor David Cantala, quien me guió durante el desarrollo de mi tesis doctoral; gracias por su paciencia, orientación e impulso.

Quisiera agredecer los comentarios y sugerencias otorgadas por Jordi Massó durante mi estancia de investigación en la Universitat Autònoma de Barcelona. Así como el apoyo recibido por parte del Doctor Edwin Van Gameren para el desarrollo del tercer capítulo de este trabajo, cuyos comentarios ayudaron en gran medida a la culminación de dicho capítulo.

Un reconocimiento a mis lectores, Dr. Leobardo Plata y Dr. Damián Gibaja, quienes hicieron comentarios acertivos sobre los tres capítulos de la tesis y ayudaron a su perfeccionamiento. Gracias también a la Dra. Adriana Gama, miembro suplente del comité, quien durante mi estancia en El Colegio de México siempre fue un apoyo y una inspiración para mí.

Un agradecimiento a las instituciones que me apoyaron de alguna manera durante mi formación. Agradezco a El Colegio de México, y en particular al Centro de Estudios Económicos por brindarme la herramientas necesarias para poder concluir con éxito mis estudios doctorales. A CONACYT por otorgarme una beca durante los tres años de los estudios doctorales, y por el apoyo económico para poder realizar una estancia de investigación en el Departament d'Economia i d'Història Econòmica de la Universitat Autònoma de Barcelona, de igual manera, a la Fundación Colmex por brindarme apoyo para poder realizar dicha estancia. Gracias a la UAB por otorgarme todos los medios para poder realizar mi investigación en sus instalaciones. También un agradecimiento a los Amigos Mexicanos de la Universidad Hebrea de Jerusalem por otorgarme una beca para asistir a la Escuela de Jerusalem de Teoría Económica.

Un reconocimiento especial a mis padres, por su cariño y apoyo incondicional. En especial a mi madre, a quien admiro profundamente y quien es una fuente de inspiración para mí por su fortaleza y su forma de ver la vida. A mis hermanas, Rosa y Gloria, de quienes, a pesar de la distancia, siempre he sentido su apoyo.

Gracias a mis amigos, Judith, Yaritza, Adriana, Selene, Elizabeth, Luis y Víctor, quienes me impulsaron en cada paso y con quienes me desahogaba en los momentos de frustación, gracias por su cariño y paciencia. A José Alberto, quien siempre me ha impulsado a lograr mis objetivos y vencer mis miedos.

Contents

Chapter 1

Introduction

This thesis is a collection of three essays on assignment problems in School Choice market. Each essay is self-contained, however, the three essays are closely related.

The problems analyzed in the two first essays lie at mechanism design, and the last essay uses econometric tools. Our main concern is to study the allocation of school seats to students in the context of School Choice market, where the students have preferences over the schools and each school is endowed with a priority relation over the set of students. In the first essay, we propose a mechanism to improve the efficiency to any stable matching and establish conditions that make the outcome essentially stable. The second essay focuses on a school choice problem where the priorities of the schools are considered like preferences and all the students have to be assigned to a school. The last essay shows an empirical analysis of the mechanism used in Mexico City to assigned students to high school.

The first essay analyzes the School Choice problem with consent for any stable matching. We propose a mechanism, the adjusted $\mathcal{E}$ mechanism (A$\mathcal{E}$M), that takes any stable matching, and improves its efficiency using the SEADAM approach. First, we compute the full set of stable matchings, find these matchings and the truncated preferences to reach them. Then, we take every stable matching with its truncated preferences and run the SEADAM. Our algorithm follows the SEADAM structure, and inherits its properties (Pareto efficiency and essential stability) to the subproblem generated by the truncated preferences used to find the initial stable matching, but this properties do not keep in the original problem unless the outcome of the A$\mathcal{E}$M is the same as in SEADAM when all students consent. Finally, we search for some conditions that assure us that the matching gotten by the adjusted $\mathcal{E}$ mechanism is essentially stable. We ask for two conditions: (1) the improvable students has to improve at SEADAM, and (2) the unimprovable students are the same at every stable matching. If these two conditions are satisfied then the matching reached by the $A\mathcal{E}M$ is essentially stable.

In the second essay, we focus on solving a School Choice problem, where we assign school seats to students, adding the condition that all students get a school seat. This condition, different from the original problem requires to develop new stability and efficiency concepts. We work using the priorities of schools as preferences. Then, we propose an extension to the classic school choice market and new concepts of matching, r-stable matching and r-optimal matching. With these definitions, we develop an algorithm (the market extension algorithm) to find an r-stable matching and prove that the set of r-stable matchings is non-empty. Then, we prove there are r-optimal matchings for both sides of the market and that they could be found when one side of the market makes the proposals in both rounds of the market extension algorithm. Finally, we analyze special cases where the r-stable matching is unique, under which conditions the mechanism is manipulable or not, and an example where the algorithm limits the number of preferred schools listed by the students.

The third essay studies the School Choice problem in Mexico City to allocate high school seats to applicants. The Comisión Metropolitana de Instituciones Públicas de Educación Media Superior (COMIPEMS) implements the assignment taking into consideration three factors: (1) the applicants preferences for schools, (2) the score obtained in the admission standardized test to high school level, and (3) the capacity of every school. So, we use the data from the high school student assignment process in Mexico to answer three research questions: if the time spent to arrive at school directly affects the probability to graduate from high school, if the home-school distance has effect on the enrollment decision, and whether the applicants consider the time that they will spend in public transportation at the moment of selecting their preferred schools when they apply to COMIPEMS admission test. With this information we analyze how the students are doing their schools' election and whether it matters to applicants the home-school distance to the moment of doing this election. Also, we count with information of the high schools where the applicants are assigned and we analyze if this time spent to arrive at their school have some implications in enrollment and graduation. The preference for a distant school decreases the probability of enrollment and graduation. Nevertheless, if they will have to spend more time in public transport, they will opt for less demanding schools as their top preferences.

Chapter 2

Efficiency-Adjusted Stable Matchings

2.1 Introduction

In the two sided matching literature à la Gale and Shapley (1962) there is a trade-off between Pareto efficiency and stability/fairness. In the marriage market model the difference is formalized by the property called weak Pareto: the stable matching obtained by the deferred acceptance (DA) algorithm is the most preferred stable matching for the proposing side of the market (Roth, 1982), still it can be Pareto dominated by another non-stable matching. In the school choice problem introduced by Abdulkadiroğlu and Sönmez (2003), the matching obtained by the DA which is always fair does not necessarily coincide with the one obtained by the top trading cycle (TTC) which is always Pareto efficient.

Kesten (2010) shows light on the trade-off. He establishes that one can reach Pareto efficiency from the DA outcome by iteratively asking to a bossy agent to consent for waiving her justified claim on an object for hurting other agent without benefiting her. The starting point of the procedure is the DA outcome, if interrupters are found during the run of the DA algorithm, the Efficiency-Adjusted Deferred Acceptance Mechanism (EADAM) picks the last one and deletes from her preferences the object for which she is an interrupter. The DA algorithm is run with the new preferences. The procedure is iterated until there is no interrupter; then the outcome matching is Pareto efficient.

Tang and Yu (2014) propose the simplified efficiency-adjusted deferred acceptance mechanism (SEADAM). This mechanism is an alternative to the EADAM with the difference that it does not change the preferences of the students. They define the underdemanded schools and show that their matches are the unimprovable students. The DA outcome is also the starting point of SEADAM which definitely matches unimprovable students with underdemanded schools, withdraw these agents from the market and run the DAA with the remaining agents. The procedure is iterated until all schools are underdemanded, the outcome of SEADAM coincides with the one of EADAM.

As we notice in these two mechanisms the authors start with the DA outcome, which is somehow arbitrary. We analyze what happens when we start from any stable matching and improve efficiency using the Kesten's algorithm with the original preferences and notice that it always reaches the EADAM outcome. Then, following Martínez, Massó, Neme and Oviedo (2004) we use the truncated preferences necessary to obtain every one of the stable matchings and the SEADAM. We choose to work with SEADAM instead EADAM because it reduces the number of rounds needed in running the algorithm and in the end the outcome is the same. Finally, this paper proposes a mechanism for any stable matching, the adjusted $\mathcal{E}$ mechanism (A$\mathcal{E}$M), that takes any stable matching, that could be different from the DA matching, and improves its efficiency using the SEADAM approach.

As we mention before, we take the mechanism developed by Martínez, Massó, Neme and Oviedo (2004) to compute the full set of stable matchings, find these matchings and the truncated preferences to reach them. They start with the matchings obtained by DA when the students propose (the students optimal stable matching) and when the schools propose (the schools optimal stable matching), look for the differences between these two matchings and iteratively withdraw from the students preferences the schools assigned by students optimal stable matching different from the schools optimal stable matching; run the DA mechanism with the new preferences and verify if this matching is stable under the original preferences. The procedure is repeated until the matching found with the truncated preferences is the same as the schools optimal stable matching. Then, we choose any stable matching with its truncated preferences and run the SEADAM. The outcome of the A$\mathcal{E}$M is not always the same as in SEADAM and is not always Pareto efficient under the original preferences.

Troyan, Delacrétaz and Kloosterman (2018) propose a new stability criterion called *essentially stable* matching. They ask if all matchings are equally unfair, and argue that the answer is no because the stability criterion excludes many matchings unnecessarily, they take these matchings into account. They prove that the SEADAM outcome is essentially stable. On the contrary, when we use any stable matching and their correspondent truncated preferences we can ensure that the A$\mathcal{E}$M is essentially stable only when we get the SEADAM outcome, namely only when we start with the DA matching.

Finally, we search for some conditions that assure that the matching gotten by the adjusted $\mathcal{E}$ mechanism is essentially stable. These conditions are taken over the set of improvable and unimprovable students of the stable matching. First, we prove that the set of improvable students is the same in every stable matching when the unimprovable students at DA are also at any stable matching. Then, we ask that the unimprovable students are the same at every stable matching and that the improvable students actually improve from their assignment under DA. If these two conditions are satisfied then the matching reached by the $A\mathcal{E}M$ is essentially stable.

The paper is organized as follows. We introduce the basic model of school choice in Section 2. Section 3 presents the SEADAM and the essentially stable property. In Section 4, we introduce the adjusted $\mathcal{E}$ mechanism, its properties and the main results. Section 5 concludes. The appendix presents two used mechanisms.

2.2 The model

Let $N = \{1, 2, ..., n\}$ denote the set of students and $S = \{1, 2, ..., m\}$ the set of schools. Let $q = (q_s)_{s \in S}$ where the integer $q_s \geq 1$ denotes the number of seats at school s.

Each student $i \in N$ has strict preferences (complete, transitive and antisymmetric binary relation) over the set of schools S, denoted by P_i. A *preference profile* is an n-tuple of preferences, denoted by $P = (P_1, ..., P_n)$. For each school $s \in S$, there is a strict priority order (complete, transitive and antisymmetric binary relation) over the set of students N, denoted by $\succ_s$. Define the priority profile $\succ = (\succ_s)_{s \in S}$. We assume that all students are acceptable at any school. Then, a school choice problem consists of a quatern $(N, S, P, \succ)$, and the preference-priority profile is denoted by the pair $(P, \succ)$.

A *matching* is a function $\mu : N \longrightarrow S \cup \{\emptyset\}$ such that: (i) for all $i \in N$, $\mu(i) \in S$ and (ii) for all $s \in S$, $\mid \{i \in N | \mu(i) = s\} \mid \leq q_s$, we denote by $\mu(i)$ the assignment of student i under μ, and denote by $\mathcal{M}$ the set of all matchings. We say that student i *desires* school s at matching μ if $sP_i\mu(i)$. The matching μ *violates the priority* of i for s if $\mu(j) = s$, $sP_i\mu(i)$, and $i \succ_s j$, that is, i would rather be matched to s, but j who has lower priority over s than i is matched to s. A student i *claims a seat at school* s if: (i) $sP_i\mu(i)$, and (ii) either $\mid \mu(s) \mid < q_s$ or $i \succ_s j$ for some $j \in \mu(s)$. If no student claims a seat at any school, then we say μ is **stable**[1] at $(P, \succ)$. We say that (i, s) is a *blocking pair* of matching μ if (i) the priority of i for s is violated, or (ii) i desires s while s still has unassigned seats.

Consider the profile of priorities $\succ = (\succ_s)_{s \in S}$, the set of students and the set of schools (N and S), a *mechanism* φ is a function that associates a matching to every preference profile P; we denote the matching under the profile of preferences P as $\varphi[P]$. A mechanism φ is stable if for each profile P, $\varphi[P]$ is stable at P. Following Martínez, Massó, Neme and Oviedo (2004), we elaborate a mechanism where preferences are iteratively modified, thus we make explicit the set of stable matchings under the preference at hand P, and denote the set $\mathcal{E}(P)$.

We denote R_i the preferences of student i over the set of matchings. Student i strictly prefers matching μ to matching μ' if $\mu(i)P_i\mu'(i)$, she is indifferent between μ and μ' if $\mu(i) = \mu'(i)$. Thus, at any matching students only care about their assignment. To simplify notations, from now on R_i and P_i denotes respectively preferences and strict preferences of student i both over matchings and schools. A matching μ is *Pareto-efficient* at R if there is no other matching μ' for which all students are at least as well off at μ than at μ', and at least one student better off; that is, $\mu(i)R_i\mu'(i)$ for all $i \in N$ and $\mu(j)P_j\mu'(j)$ for some $j \in N$. A mechanism φ is *Pareto-efficient* if for any preference profile P, $\varphi[P]$ is Pareto efficient at P.

[1]This definition of stability is equivalent to: A matching μ is *stable* at P and $\succ$ if it is fair and non-wasteful. Where, a matching μ is *fair* if no student's priority for any school is violated and it is *non-wasteful* if any school $s \in S$ that is desired by some student at μ, satisfies $\mid \{i \in N | \mu(i) = s\} \mid = q_s$.

2.3 SEADAM and Essentially Stable Matchings

In a school choice problem, the student-proposing DA selects the optimal stable matching for students, that is the stable matching that each student likes at least as much as any other stable matching. Nevertheless, the DA outcome is not necessarily Pareto efficient for students, and to Pareto improve on the DA outcome, Kesten (2010) proposes the school choice problem with consent, in which each student is asked whether he allows other students to violate her priorities.

In order to reach this objective, Kesten (2010) proposes the efficiency-adjusted deferred acceptance mechanism (EADAM). Formally, given a problem to which the DA algorithm is applied, some student i applies to school s and is tentatively accepted, but his tentative acceptance at s initiates a chain of rejections that eventually leads s to reject student i herself. By applying to school s, student i gains nothing, but potentially interrupts a desirable settlement among other students, i is called an *interrupter* at s and the pair (i, s) is an *interrupting pair*. If the outcome of DA is inefficient, then there is at least one interrupting pair, even though the converse is not true.

For any school choice problem with consenting students, Kesten's EADAM operates as follows[2] (i) run DA for the school choice problem; (ii) identify iteratively the last step of the DA procedure in which consenting interrupter(s) are rejected, and then identify all interrupting pairs of this step that contain a consenting interrupter and, for each pair, remove the respective school from the interrupter's preference; (iii) rerun DA with the new preference profile; these steps are repeated until there are no more consenting interrupters.

Theorem 1. *(Kesten 2010) The EADAM Pareto dominates the DA as well as any other stable mechanism. If all students consent, then for all P the EADAM outcome is Pareto Efficient at P.*

Following Kesten(2010), Tang and Yu (2014) propose a simpler description of EADAM. The approach they take is to directly examine consenting incentives, that is, they use the consent of a student only when her assignment is not Pareto improvable anymore. And,

[2]See the complete algorithm and an example in Appendix.

they call their mechanism the simplified efficiency-adjusted deferred acceptance mechanism (SEADAM)

Definition 1. *The student i is **not Pareto improvable** (or **unimprovable**) at μ if for any matching ν that Pareto dominates μ, $\nu(i) = \mu(i)$.*

$U(DA[P])$ denotes the set of unimprovable students found under DA at P, and $I(DA[P])$ the set of improvable students found under DA at P.

Definition 2. *A school s is **underdemanded** at matching μ if no student strictly prefers s to her assignment under μ.*

By Lemma 1 of Tang and Yu (2014), all students matched with underdemanded schools at the DA outcome are not Pareto improvable.

Theorem 2 (Tang and Yu, 2014)**.** *The $SEADAM$ is Pareto efficient when all students consent.*

Example 1 (Underdemanded schools)**.** *Consider the sets of students $N = \{i_1, i_2, i_3, i_4\}$ and of schools $S = \{s_1, s_2, s_3, s_4\}$, let $q_s = 1$ for all $s \in S$ and assume all students consent. Let the profile P of preferences and the profile $\succ$ of priorities be as follows:*

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}
s_3	s_4	s_3	s_4
s_1	s_3	s_4	s_2
s_4	s_1	s_1	s_1
s_2	s_2	s_2	s_3

$\succ_{s_1}$	$\succ_{s_2}$	$\succ_{s_3}$	$\succ_{s_4}$
i_2	i_4	i_4	i_1
i_1	i_1	i_2	i_3
i_3	i_2	i_1	i_2
i_4	i_3	i_3	i_4

The DA outcome is:

Step	s_1	s_2	s_3	s_4
1			$\underline{i_1}, i_3$	$\underline{i_2}, i_4$
2		i_4	i_1	$i_2, \underline{i_3}$
3		i_4	$i_1, \underline{i_2}$	i_3
4	i_1	i_4	i_2	i_3

In the DA procedure, student i_1 is the only interrupter and the interrupter pair is (i_1, s_3). The interruption of i_1 blocks a trading between i_2 and i_3 and makes the DA outcome inefficient, that is, the students i_2 and i_3 could exchange their schools s_3 and s_4 and could be better, but the student i_1 blocks this trading.

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}
s_3	$\boxed{s_4}$	$\boxed{s_3}$	s_4
$\underline{s_1}$	$\underline{s_3}$	$\underline{s_4}$	$\underline{s_2}$
s_4	s_1	s_1	s_1
s_2	s_2	s_2	s_3

In this case, the schools s_2 and s_1 are underdemanded. Note that the interrupter i_1 is matched with s_1, which is an underdemanded school at the DA.
Furthermore, notice the students i_1, i_4 are matched with underdemanded schools, and this two students are also Pareto unimprovable, this because their assignment under DA and under simplified EADAM is the same.

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}
s_3	$\cdot s_4$	$\cdot s_3$	s_4
$\cdot\boxed{\underline{s_1}}$	$\underline{s_3}$	$\underline{s_4}$	$\cdot\boxed{\underline{s_2}}$
s_4	s_1	s_1	s_1
s_2	s_2	s_2	s_3

The DA outcome is underlined, the EADAM when all students consent is indicated by a center dot and the underdemanded schools are showed in a box. □

Consider any school choice problem $(N, S, P, \succ)$ with all students consenting. The SEADAM operates as follows:

- **Round 0**: Run DA for the problem $(N, S, P, \succ)$
- **Round** k, $k \geq 1$: This round consists of two steps[3]:
    1. Identify the schools that are underdemanded at the round-$(k-1)$ DA matching, settle the matching at these schools, and remove these schools and the students matched with them.

[3] We remove the step 2 from the original mechanisms because we consider that all students consent.

2. Rerun DA (the round-k DA) for the subproblem that consists of only the remaining schools and students.
   Stop when all schools are removed.

Henceforth, $SEADAM[P]$ denotes the outcome of the $SEADA$ mechanism with the preferences profile P. Tang and Yu (2014) also prove that SEADAM produces the same matching as Kesten's EADAM does.

In order to illustrate SEADAM, we present an example.

Example 2 (SEADAM). *The sets of students is $N = \{i_1, i_2, i_3, i_4, i_5\}$, and of schools is $S = \{s_1, s_2, s_3, s_4, s_5\}$; let $q_s = 1$ for all $s \in S$ and assume all students consent. Let the profile P of preferences and the profile $\succ$ of priorities be as follows:*

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}	P_{i_5}
s_3	s_4	s_3	s_1	s_1
s_1	s_3	s_1	s_5	s_5
s_4	s_5	s_5	s_4	s_2
s_5	s_1	s_4	s_3	s_3
s_2	s_2	s_2	s_2	s_4

$\succ_{s_1}$	$\succ_{s_2}$	$\succ_{s_3}$	$\succ_{s_4}$	$\succ_{s_5}$
i_2	i_5	i_4	i_3	i_1
i_1	i_1	i_5	i_4	i_2
i_3	i_2	i_2	i_1	i_3
i_4	i_3	i_1	i_2	i_5
i_5	i_4	i_3	i_5	i_4

Round 0: Run the DA.

Step	s_1	s_2	s_3	s_4	s_5
1	$\underline{i_4}, i_5$		$\underline{i_1}, i_3$	i_2	
2	$\underline{i_3}, i_4$		i_1	i_2	i_5
3	i_3		i_1	i_2	$i_4, \underline{i_5}$
4	i_3		i_1	$i_2, \underline{i_4}$	i_5
5	i_3		$i_1, \underline{i_2}$	i_4	i_5
6	$\underline{i_1}, i_3$		i_2	i_4	i_5
7	i_1		i_2	i_4	$\underline{i_3}, i_5$
8	i_1	i_5	i_2	i_4	i_3

Round 1: Step 1: The underdemanded school is s_2, we settle the matching (i_5, s_2) and remove school s_2 and student s_2 from the problem.

Step 2: We rerun the DA algorithm with the following preference profile:

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}
s_3	s_4	s_3	s_1
s_1	s_3	s_1	s_5
s_4	s_5	s_5	s_4
s_5	s_1	s_4	s_3

$\succ_{s_1}$	$\succ_{s_3}$	$\succ_{s_4}$	$\succ_{s_5}$
i_2	i_4	i_3	i_1
i_1	i_2	i_4	i_2
i_3	i_1	i_1	i_3
i_4	i_3	i_2	i_4

Step	s_1	s_3	s_4	s_5
1	i_4	$\underline{i_1}, i_3$	i_2	
2	$\underline{i_3}, i_4$	i_1	i_2	
3	i_3	i_1	i_2	i_4

Round 2: Step 1: The underdemanded schools are s_4 and s_5, we settle the matching at these schools, and remove these schools and the students i_2 and i_4 from the problem.

Step 2: We rerun the DA algorithm with the following profile:

P_{i_1}	P_{i_3}
s_3	s_3
s_1	s_1

$\succ_{s_1}$	$\succ_{s_3}$
i_1	i_1
i_3	i_3

Step	s_1	s_3
1		$\underline{i_1}, i_3$
2	i_3	i_1

Round 3: Step 1: The underdemanded school is s_1, we settle its matching and remove this school ant the student i_3 from the problem.

Step 2: Rerun the DA algorithm with the profile:

P_{i_1}
s_3

$\succ_{s_3}$
i_1

Step	s_3
1	i_1

Round 4: Step 1: The underdemanded school is s_3, then we remove this school and its matching i_1. Then, all schools have been removed and the algorithm stops. The matching gotten is:

$$SEADAM[P] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 & i_5 \\ s_3 & s_4 & s_1 & s_5 & s_2 \end{pmatrix}$$

$\square$

Troyan, Delacrétaz and Kloosterman (2018) propose to relax the concept of stability using the definition of *vacuous claims*. Given a matching, student i is said to have a (justified) *claim* to school s if she prefers s to her assignment and she has higher priority than another student who is assigned to s. In this case, (i, s) denotes i's claim to a seat at s.

Definition 3. *Consider a matching μ and a claim (i, s). The* ***reassignment chain*** *initiated by claim (i, s) is the list $i^0 \to s^0 \to i^1 \to s^1 \to ... \to i^K \to s^K$ where, $i^0 = i$, $\mu^0 = \mu$, $s^0 = s$ and for each $k \geq 1$:*

- *i^k is the lowest-priority student in $\mu^{k-1}(s^{k-1})$,*
- *μ^k is defined as: $\mu^k(j) = \mu^{k-1}(j)$ for all $j \neq i^{k-1}$, i^k, $\mu^k(i^{k-1}) = s^{k-1}$ and student i^k is unassigned,*
- *s^k is student i^k's most preferred school where she can claim a seat at μ^k*

Terminates at the first K such that $\mid \mu^K(s^K) \mid < q_{s^K}$.

We say that a claim is **vacuous** if the initial claim is unfounded, that is, if the student i who started the reassignment chain, is removed from the school s that she claimed initially by some student with higher priority. For a reassignment chain started by a claim (i, s), if there exists $k \neq 0$ such that $i^k = i$, we say that the reassignment chain returns to i. If the reassignment chain returns to i, then i will ultimately be removed from the school s that she claimed initially by some student with higher priority. When this is the case, we say that claim (i, s) is vacuous.

Definition 4. *A matching μ is* ***essentially stable*** *if all claims at μ are vacuous.*

If there exists at least one claim at μ that is not vacuous, μ is *strongly unstable*. A mechanism φ is said to be essentially stable if $\varphi(P)$ is an essentially stable matching for all P. If φ is not essentially stable, then we say it is strongly unstable.

Theorem 3. *(Troyan, Delacrétaz and Kloosterman, 2018) The final matching produced by the SEADAM is essentially stable.*

2.4 Adjusted $\mathcal{E}$ Mechanism

We propose the adjusted $\mathcal{E}$ mechanism which adapts the simplified efficiency adjusted mechanism of Tang and Yu (2014) to any stable matching $\mu \in \mathcal{E}(P)$.

First, we recall the algorithm to compute the full set of stable matchings developed by Martínez, Massó, Neme and Oviedo (2004). This algorithm works as follows[4]:

- **Round 0**: Take the original profile of preferences (P), compute the student-optimal matching, denoted by $\mu_N[P]$ and the school-optimal matching (when the schools propose) denoted by $\mu_S[P]$

- **Round** k, $k \geq 1$: This round consists of four steps:

  1. Truncate students' preferences, this is done by deleting of the preference profile of student i the school which was assigned by the DA in the last round when the assigned school is different from the assigned by $\mu_S[P]$. That is, we find a truncated preference for every student which gets a school different that the obtained under the school-optimal matching.
  2. Use the DA with these truncated preferences and find new matchings, one for each truncated preference.
  3. Compare the assignments before and after the truncation, if the assignment after the truncation is preferred by the school then it is stable, dismiss the unstable matchings.

[4]We present in Appendix the algorithm in the way that the authors do it.

4. We compare by pairs the stable matchings and discard the stable matchings that do not assign the same school to students who do not truncate their preferred schools in any of them.

   These steps are repeated until the school-optimal assignment is reached.

Then, with the set of stable matchings in the original problem $(N, S, P, \succ)$ and the truncated preferences used to find them, we apply the SEADAM using all the truncated preferences and find the Pareto efficient matching respect these truncated preferences. From now on, $\mu_N[P]$ denotes the outcome of DA when students propose with the profile of preferences P, in the same way $\mu_S[P]$ denotes the outcome of DA when schools propose with P as the profile of preferences.

Now we can proceed to show how the adjusted $\mathcal{E}$ mechanism[5] works to find the outcome starting from all stable matchings:

- **Stage 0**: Run the Martínez-Massó-Neme-Oviedo algorithm and find the set of stable matchings denoted by $\mathcal{E}(P)$.

- **Stage** k, $1 \leq k \leq\mid \mathcal{E}(P) \mid$: This stage consists of the following rounds:

  * **Round 0**: Identify and fix the truncated preferences P^k that were used to find the stable matching $\mu_N[P^k]$.

  * **Round** l, $l \geq 1$. This round consists of two steps:

    1. Identify the schools that are underdemanded at the round $l-1$ DA outcome with the profile (P^k), settle the matching at these schools, and remove these schools and the students matched with them.
    2. Rerun DA (the round l DA) for the subproblem that consists of only the remaining schools and students under P^k. Stops when all schools are removed.

The algorithm ends in a finite number of steps because the number of schools and students are finite.

[5] We name the Adjusted $\mathcal{E}$ Mechanism by its initials $A\mathcal{E}M$.

Let's denote the outcome of the adjusted $\mathcal{E}$ mechanism by $A\mathcal{E}M[P]$, that is, the set of all adjusted stable matchings gotten using the adjusted $\mathcal{E}$ mechanism, and denote one particular adjusted stable matching by $A\mathcal{E}M[P^k]$ with $1 \leq k \leq\mid \mathcal{E}(P) \mid$. Note that the outcome $A\mathcal{E}M[P^k]$ refers to the element of the family of results obtained by $A\mathcal{E}M[P]$ when the preferences P^k and the stable matching $\mu_N[P^k]$ are used. From this moment, we will abuse the notation and refer to $A\mathcal{E}M[\mu_N(P^k)]$ only as $A\mathcal{E}M[P^k]$. Also, $U(\mu_N[P^k])$ denotes the set of unimprovable students found under DA at P^k, and $I(\mu_N[P^k])$ the set of improvable students found under DA at P^k.

Notice that with this mechanism we can find different outcomes, depending of the truncated preferences that we use, that is, we propose a mechanism that results in a family of matchings, one for every stable matching $\mu \in \mathcal{E}(P)$. This family is $A\mathcal{E}M[P] = \{A\mathcal{E}M[P^1], A\mathcal{E}M[P^2], ..., A\mathcal{E}M[P^m]\}$ with $m =\mid \mathcal{E}(P) \mid$.

Example 3 (Adjusted $\mathcal{E}$ Mechanism)**.** *Consider the sets of students $N = \{i_1, i_2, i_3, i_4\}$ and of schools $S = \{s_1, s_2, s_3, s_4\}$ where each school has only one seat, and assume all students consent. The profiles of preferences and of priorities are as follows:*

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}
s_2	s_2	s_3	s_1
s_3	s_1	s_2	s_2
s_1	s_3	s_4	s_3
s_4	s_4	s_1	s_4

$\succ_{s_1}$	$\succ_{s_2}$	$\succ_{s_3}$	$\succ_{s_4}$
i_1	i_4	i_2	i_3
i_2	i_3	i_1	i_1
i_4	i_2	i_3	i_4
i_3	i_1	i_4	i_2

Stage 0: We find the set of stable matchings using the Martínez, Massó, Neme, Oviedo algorithm. In this case, there are only two stable matchings, the students-optimal stable matching and the schools-optimal stable matching.

Stage 1: Round 0: First, we identify and fix the preferences used to find the students-optimal stable matching $\mu_N[P]$, in this case we use the original preferences $P^1 = P$:

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}
s_2	s_2	s_3	s_1
s_3	s_1	s_2	s_2
s_1	s_3	s_4	s_3
s_4	s_4	s_1	s_4

$$\mu_N[P^1] = \mu_N[P] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ s_3 & s_1 & s_4 & s_2 \end{pmatrix}$$

Round 1: Step 1: Then, we identify the set of underdemanded schools and settle their matching and fix her assignment: in this case the underdemanded school is s_4 and her assignment $A\mathcal{E}M[P^1](s_4) = i_3$, which is the unimprovable school. Now, we remove this school and the student i_3 from the preferences and priorities of students and schools respectively:

P_{i_1}	P_{i_2}	P_{i_4}
s_2	s_2	s_1
s_3	s_1	s_2
s_1	s_3	s_3

$\succ_{s_1}$	$\succ_{s_2}$	$\succ_{s_3}$
i_1	i_4	i_2
i_2	i_2	i_1
i_4	i_1	i_4

Step 2: We rerun the DA algorithm and find the matching $\mu'_N[P]$[6]:

$$\mu'_N[P] = \begin{pmatrix} i_1 & i_2 & i_4 \\ s_3 & s_2 & s_1 \end{pmatrix}$$

Round 2: Step 1: In this case the underdemanded schools are s_1 and s_3 and their assignments are $A\mathcal{E}M[P^1](s_1) = i_4$ and $A\mathcal{E}M[P^1](s_3) = i_1$, which are the unimprovable schools. Now, we remove these schools and the students i_1 and i_4 from the preferences and priorities of students and schools respectively:

P_{i_2}
s_2

$\succ_{s_2}$
i_2

Step 2: We rerun the DA algorithm and find the matching $\mu''_N[P]$[7]:

$$\mu''_N[P] = \begin{pmatrix} i_2 \\ s_2 \end{pmatrix}$$

[6] μ'_N is the DA matching at the subproblem described in the Stage 1 Round 1 Step 1.
[7] μ''_N is the DA matching at the subproblem described in the Stage 1 Round 2 Step 1.

Round 3: Step 1: The underdemanded schools is s_2 and her assignment is $A\mathcal{E}M[P^1](s_2) = i_2$. We remove this school and notice that all schools have been removed. Then, the Stage 1 ends and the $A\mathcal{E}M[P^1]$ is the following:

$$A\mathcal{E}M[P^1] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ s_3 & s_2 & s_4 & s_1 \end{pmatrix}$$

Stage 2: Round 0: We identify and fix the preferences used to find the stable matching $\mu_N[P^2]$, in this case we use the preferences $P^2 = (P'_{i_1}, P_{i_2}, P_{i_3}, P_{i_4})$:

P'_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}
s_2	s_2	s_3	s_1
s_1	s_1	s_2	s_2
s_4	s_3	s_4	s_3
	s_4	s_1	s_4

$$\mu_N[P^2] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ s_1 & s_3 & s_4 & s_2 \end{pmatrix}$$

Round 1: Step 1: We identify the set of underdemanded schools and settle their matching and fix her assignment: in this case the underdemanded school is s_4 and her assignment $A\mathcal{E}M[P^2](s_4) = i_3$, which is the unimprovable school. Now, we remove this school and the student i_2 from the preferences and priorities of students and schools respectively:

P'_{i_1}	P_{i_2}	P_{i_4}
s_2	s_2	s_1
s_1	s_1	s_2
	s_3	s_3

$\succ_{s_1}$	$\succ_{s_2}$	$\succ_{s_3}$
i_1	i_4	i_2
i_2	i_2	i_1
i_4	i_1	i_4

Step 2: We rerun the DA algorithm and find the matching $\mu'_S[P]$[8]:

$$\mu'_S[P] = \begin{pmatrix} i_1 & i_2 & i_4 \\ s_1 & s_3 & s_2 \end{pmatrix}$$

Round 2: Step 1: In this case the underdemanded school is s_3 and her assignment is $A\mathcal{E}M[P^2](s_3) = i_2$, which is the unimprovable school. Now, we remove this school and the student i_2 from the preferences and priorities of students and schools respectively:

[8] μ'_S is the DA matching at the subproblem described in the Stage 2 Round 1 Step 1.

P'_{i_1}	P_{i_4}		$\succ_{s_1}$	$\succ_{s_2}$
s_2	s_1		i_1	i_4
s_1	s_2		i_4	i_1

Step 2: We rerun the DA algorithm and find the matching $\mu''_S[P]$[9]:

$$\mu''_S[P] = \begin{pmatrix} i_1 & i_4 \\ s_2 & s_1 \end{pmatrix}$$

Round 3: Step 1: The underdemanded schools are s_1 and s_2 and their assignments are $A\mathcal{E}M[P^2](s_1) = i_4$ and $A\mathcal{E}M[P^2](s_2) = i_1$. We remove this school and notice that all schools have been removed. Then, the Stage 2 ends and the $A\mathcal{E}M[P^2]$ is the following:

$$A\mathcal{E}M[P^2] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ s_2 & s_3 & s_4 & s_1 \end{pmatrix}$$

When we calculate the $SEADAM[P]$ we found that $SEADAM[P] = A\mathcal{E}M[P^1]$, but different from $A\mathcal{E}M[P^2]$. This observation makes us wonder if this always happen.
We know by Theorem 3 that $SEADAM[P] = A\mathcal{E}M[P^1]$ is essentially stable, so we analyze what happen with $A\mathcal{E}M[P^2]$. Suppose student i_2 claims the seat at school s_2, then student i_1 becomes unmatched and she asks to be assigned to s_3, her next most-preferred school. That is the end of our chain because s_3 is the school that i_2 gave up to claim s_2. In this case the original claim is founded (no vacuous), then the $A\mathcal{E}M[P^2]$ assignment is strongly unstable.

□

Observation 1. *The outcome of the adjusted $\mathcal{E}$ mechanism using the original preferences profile $P = P^1$ is the same that the outcome of $SEADAM$ at P, that is, $A\mathcal{E}M[P^1] = SEADAM[P]$.*

This occurs because $A\mathcal{E}M$ replays $SEADAM$ for every truncated preference profile. The $A\mathcal{E}M$ finds in Stage 0 all the stable matchings, then in the Round 0 of Stage 1 we fix the preference profile to $P^1 = P$, and notice that from Round 1 to Round l of stage 1 the $A\mathcal{E}M$ consists of the same steps that the $SEADAM$. Then, the outcome of $A\mathcal{E}M$ at $P^1 = P$ is the same as in $SEADAM[P]$.

[9] μ''_S is the DA matching at the subproblem described in the Stage 2 Round 2 Step 1.

The adjusted $\mathcal{E}$ mechanism replays the $SEADAM$ in every subproblem $(N, S, P^k, \succ)$, with $1 \leq k \leq |\mathcal{E}(P)|$, that is, $SEADAM[P^k] = A\mathcal{E}M[P^k]$. Then, the $SEADAM$ inherits its properties to every element of $A\mathcal{E}M[P]$ in every subproblem, in particular the Pareto efficiency and the essentially stable properties. Although, this does not happen at the original problem $(N, S, P, \succ)$, when we use the outcome of $A\mathcal{E}M[P^k]$ with $k \neq 1$.

Proposition 1. *The $A\mathcal{E}M[P^k]$ is Pareto efficient at P^k.*

Proof. The proof consists of two parts: (i) first, we prove that $A\mathcal{E}M[P^k] = SEADAM[P^k]$, and (ii) then, we prove that $A\mathcal{E}M[P^k]$ is Pareto efficient at P^k.

- We prove the first part by induction. Both mechanisms $A\mathcal{E}M$ and $SEADAM$ has an analogous process, then we can compare every element of the family $A\mathcal{E}M[P]$ with the outcome of $SEADAM$ at every subproblem $(N, S, P^k, \succ)$ with $1 \leq k \leq |\mathcal{E}(P)|$. By Observation 1, $A\mathcal{E}M[P^1] = SEADAM[P]$. We assume that $A\mathcal{E}M[P^k] = SEADAM[P^k]$ and prove for $k+1$. It is enough notice that the Stage $k+1$ of $A\mathcal{E}M$ consists of the same steps as the $SEADAM$ but with the preferences P^{k+1}, then the outcome of $A\mathcal{E}M[P^{k+1}]$ is equal to the outcome of $SEADAM$ with the preferences profile P^{k+1}. Then, $A\mathcal{E}M[P^k] = SEADAM[P^k]$ for all $1 \leq k \leq |\mathcal{E}(P)|$.

- The second part of the proof is a consequence of the Theorem 2 and the first part of this proof: $A\mathcal{E}M[P^k] = SEADAM[P^k]$ and $SEADAM[P^k]$ is Pareto efficient at P^k. Therefore, the matching $A\mathcal{E}M[P^k]$ is Pareto efficient at P^k.

■

Although, the Pareto efficiency property of the $A\mathcal{E}M[P^k]$ keeps at the subproblem with truncate preferences P^k, this does not happen at the original problem respect the original preferences P.

Proposition 2. *The $A\mathcal{E}M[P^k]$ is not Pareto efficient at P.*

Proof. The proof is by counterexample. Consider the sets of students $N = \{i_1, i_2, i_3, i_4\}$ and of schools $S = \{s_1, s_2, s_3, s_4\}$ where all students consent and each school has only one seat. The profiles of preferences and of priorities are:

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}
s_3	s_4	s_4	s_2
s_1	s_2	s_2	s_4
s_4	s_3	s_1	s_1
s_2	s_1	s_3	s_3

$\succ_{s_1}$	$\succ_{s_2}$	$\succ_{s_3}$	$\succ_{s_4}$
i_2	i_1	i_3	i_4
i_3	i_2	i_4	i_1
i_1	i_3	i_2	i_2
i_4	i_4	i_1	i_3

And consider the mechanism applied to the stable matching $\mu[P^3]$ with the truncated preferences $P^3 = (P'_{i_1}, P'_{i_2}, P_{i_3}, P_{i_4})$:

P'_{i_1}	P'_{i_2}	P_{i_3}	P_{i_4}
s_1	s_4	s_4	s_2
s_4	s_2	s_2	s_4
s_2	s_1	s_1	s_1
		s_3	s_3

$$\mu_N[P^3] = \mu_N[P] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ s_3 & s_4 & s_1 & s_2 \end{pmatrix}$$

The outcome of $A\mathcal{E}M[P^3]$ is[10]:

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}
$\cdot s_3$	$\underline{s_4}$	s_4	$\underline{s_2}$
$\underline{s_1}$	s_2	s_2	s_4
s_4	s_3	$\cdot s_1$	s_1
s_2	s_1	$\underline{s_3}$	s_3

$$A\mathcal{E}M[P^3] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ s_1 & s_4 & s_3 & s_2 \end{pmatrix}$$

We notice that the students i_1 and i_3 could make an exchange of schools[11] s_1 and s_3 and get a more preferred school under this matching, which we call ν, than with the assignment made by $A\mathcal{E}M[P^3]$, all this under the original preferences P.
Then, the matching $A\mathcal{E}M[P^3]$ is not Pareto efficient at P, because $\nu(i_1)\ P_{i_1}\ A\mathcal{E}M[P^3](i_1)$ and $\nu(i_3) P_{i_3} A\mathcal{E}M[P^3](i_3)$. Therefore, the adjusted $\mathcal{E}$ mechanism is not Pareto efficient at P. ■

Now we know that our mechanism is Pareto efficient at P^k but not necessary at P, we analyze if the outcome satisfies another important property, stability. The outcome of $A\mathcal{E}M[P^1]$ is not stable, as in the case of the $SEADAM$, but the former mechanism inherits the essentially stable property from the latter.

[10] Underlined in the original preferences profile
[11] Marked by a center dot in the original preferences

Observation 2. *$A\mathcal{E}M[P^1]$ is essentially stable.*

This is a direct consequence of the Observation 1 and Theorem 3. The matching $A\mathcal{E}M[P^1]$ is equal to $SEADAM[P]$ and the matching produced by $SEADAM$ is essentially stable, then $A\mathcal{E}M[P^1]$ is essentially stable.

But this is the only case when we can assure that this happens. As we can see in the Example 3, $SEADAM[P]$ and $A\mathcal{E}M[P^2]$ are different, even more $A\mathcal{E}M[P^2]$ is strongly unstable. Looking that, we can search for some special conditions over the school choice problem that allow us to ensure that the matchings are essentially stable, and this only occurs when the outcome $A\mathcal{E}M[P^k]$ is equal to the one of $SEADAM$.

Theorem 4. *If $A\mathcal{E}M[P^k] = SEADAM[P]$ for any $1 \leq k \leq \mid \mathcal{E}(P) \mid$, then $A\mathcal{E}M[P^k]$ is essentially stable. Otherwise, the mechanism is strongly unstable.*

Proof. The first statement, if $A\mathcal{E}M[P^k] = SEADAM[P]$ then $A\mathcal{E}M[P^k]$ is essentially stable, is a direct consequence of Theorem 3 and the equivalence $A\mathcal{E}M[P^k] = SEADAM[P]$.
The second one, if $A\mathcal{E}M[P^k] \neq SEADAM[P]$ then $A\mathcal{E}M[P^k]$ is strongly unstable, is proved by the Example 3. In that case, when the preferences are P^2 the outcome is not equal to $SEADAM$ and is not essentially stable because we found that the original claim is founded (no vacuous). This shows that the property is satisfied depending of the preferences. Then, the $A\mathcal{E}M[P^k]$ is strongly unstable.

■

The properties of Pareto efficiency and essential stability are surely satisfied when the outcome $A\mathcal{E}M[P^k]$ is equal to $SEADAM[P]$. These two properties are very desirable for any matching because it assure us that the students do not want to exchange their school with another student and do not make founded claims. So, we look for characteristics in the students that warrant us that the outcome of the $A\mathcal{E}M[P^k]$ is equal to $SEADAM$.

In the Round l Step 1 of the adjusted $\mathcal{E}$ mechanism we look for the underdemanded schools and their assignments under DA (the unimprovable students) and remove them. This gives a hint that these students could be the key to get or not the $SEADAM$ outcome.

Example 4 (Unimprovable students with different assignment at different stable matchings)**.** *Consider the set of students $N = \{i_1, i_2, i_3, i_4\}$, and the set of schools $S = \{s_1, s_2, s_3, s_4\}$ where each school has only one seat, and assume all students consent. The profile P of preferences and the profile $\succ$ of priorities are as follows:*

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}
s_4	s_3	s_4	s_2
s_2	s_4	s_2	s_3
s_3	s_1	s_1	s_1
s_1	s_2	s_3	s_4

$\succ_{s_1}$	$\succ_{s_2}$	$\succ_{s_3}$	$\succ_{s_4}$
i_1	i_2	i_3	i_4
i_2	i_1	i_4	i_2
i_4	i_3	i_2	i_1
i_3	i_4	i_1	i_3

In this case, there are four stable matchings, these ones and the matchings $A\mathcal{E}M[P^k]$ obtained by the adjusted $\mathcal{E}$ mechanisms for $k = \{1, 2, 3, 4\}$, where $P^1 = P$, are shown next:

$$\mu_N[P^1] = \mu_N[P] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ s_2 & s_4 & s_1 & s_3 \end{pmatrix} \quad A\mathcal{E}M[P^1] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ s_4 & s_3 & s_1 & s_2 \end{pmatrix}$$

$$\mu_N[P^2] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ s_2 & s_4 & s_3 & s_1 \end{pmatrix} \quad A\mathcal{E}M[P^2] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ s_4 & s_3 & s_2 & s_1 \end{pmatrix}$$

$$\mu_N[P^3] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ s_2 & s_1 & s_3 & s_4 \end{pmatrix} \quad A\mathcal{E}M[P^3] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ s_4 & s_1 & s_2 & s_3 \end{pmatrix}$$

$$\mu_N[P^4] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ s_1 & s_2 & s_3 & s_4 \end{pmatrix} \quad A\mathcal{E}M[P^4] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ s_1 & s_3 & s_4 & s_2 \end{pmatrix}$$

The $A\mathcal{E}M[P^k]$ assignments with $k = \{2, 3, 4\}$ are different to $SEADAM$. Also, these outcomes are Pareto efficient at P, but they are not essentially stable (this is justified below). Consider $A\mathcal{E}M[P^2]$ first. Suppose student i_4 claims the seat at school s_3, then student i_2 becomes unmatched and she ask to be assigned to s_4 her next most-preferred school where she has higher priority than the student i_1. Now student i_1 is unmatched, then she asks for s_2 where she has high enough priority to be assigned. Student i_3 is now unmatched, and asks for s_1. That is the end of our chain because s_1 is the school that i_4 gave up to claim s_3. In this case the original claim is founded (no vacuous), then the $A\mathcal{E}M[P^2]$ assignment is strongly unstable.

In the $A\mathcal{E}M[P^3]$ case we have something similar. The chain of claims starts with the student i_2 claiming the seat at school s_4, then student i_1 becomes unmatched and she ask for school s_2 where she has high enough priority to be assigned. Student i_3 is now unmatched, and asks for s_1 and the reassignment chain ends because s_1 is the school that i_2 gave up to claim s_4. Then, the $A\mathcal{E}M[P^3]$ assignment is strongly unstable because the claim is no vacuous.

Finally, the chain of claims for $A\mathcal{E}M[P^4]$ stars with the student i_1 claiming the seat at school s_2, then student i_4 becomes unmatched and she asks for s_3, then student i_2 is unmatched and asks for school s_4, becoming unmatched student i_3 who asks for a seat in school s_1 ending the chain of claims because s_1 is the school that i_1 gave up to claim s_2. Therefore, $A\mathcal{E}M[P^4]$ is strongly unstable. Notice that this chain of claims is not the only one for $A\mathcal{E}M[P^4]$, but it is enough that one of them is vacuous to say that the matching is strongly unstable. □

First, we state a lemma where we show the consistency of the set of improvable students, that is, we assure that the set of improvable students will be the same under every stable matching as long as the set of unimprovable students is the same in every stable matching.

Lemma 1. *For every school choice problem $(N, S, P, \succ)$ when all students consent, if for all $i \in U(DA[P])$ and for all $1 \leq k \leq\mid \mathcal{E}(P) \mid$, $i \in U(\mu_N[P^k])$, then $I(\mu_N[P^k]) = I(DA[P])$ for all $\mu_N[P^k] \in \mathcal{E}(P)$.*

Proof. The set of students is a partition between unimprovable and improvable students at DA, that is $I(DA[P]) = N \setminus U(DA[P])$. We prove that, if the set of unimprovable students is the same in all the stable matchings, if i is improvable under P also will be improvable under P^k.

For all $i \in U(DA[P])$ and for all $1 \leq k \leq\mid \mathcal{E}(P) \mid$ is satisfied that $i \in U(\mu_N[P^k])$ for all $\mu_N[P^k] \in \mathcal{E}(P)$, then $U(DA[P]) \subseteq U(\mu_N[P^k])$, that is if i is unimprovable under P then it is unimprovable under every P^k.

Also, by lattice properties of the stable matchings, the set of schools that are preferred to the stable matching is going increased according the matching is going less preferred by the students, this causes that the set of underdemanded schools could be reduced, then the set of unimprovable students could be reduced too, $U(\mu_N[P^k]) \subseteq U(DA[P])$.

Then, the unimprovable are the same under every P^k, $U(\mu_N[P^k]) = U(DA[P])$. Thus,

$N \setminus U(\mu_N[P^k]) = N \setminus U(DA[P])$, therefore $I(\mu_N[P^k]) = I(DA[P])$ the set of improvable students is the same under every stable matching, that is the set of students that could trade theirs schools is the same under every P^k. ■

Theorem 5. *For every school choice problem $(N, S, P, \succ)$ when all students consent, if for the stable matching $\mu_N[P^k] \in \mathcal{E}(P)$ with $1 \leq k \leq\mid \mathcal{E}(P) \mid$:*

1. *For all $i \in U(DA[P])$, $i \in U(\mu_N[P^k])$*
2. *For all $i \in I(DA[P])$, $DA[P](i) \neq SEADAM[P](i)$*

Then every $A\mathcal{E}M[P^k] \in A\mathcal{E}M[P]$ produces the same matching as $SEADAM[P]$ does, that is $A\mathcal{E}M[P^k] = SEADAM[P]$ for all $1 \leq k \leq\mid \mathcal{E}(P) \mid$.
Furthermore, the adjusted $\mathcal{E}$ mechanism generate a family of essentially stable matchings.

That is, whenever we have a school choice problem and every student consent, when the unimprovable students are the same in every stable matching, and the improvable students under the original preferences actually improve from DA at P, then the matchings reached under the $SEADAM$ with the preferences profile P and under the $A\mathcal{E}M[P^k]$ are the same.

Proof. The set of students is a partition between unimprovable and improvable students at DA, first we prove the case of the unimprovable students and then the case of improvable students, if both are satisfied the theorem holds:

- **Case 1**: Consider $i \in U(DA[P])$ (the student i is unimprovable at $DA[P]$), by definition, i satisfies $DA[P](i) = SEADAM[P](i)$. Also, by definition of unimprovable student and by the lattice structure of the stable matchings , $DA[P](i) = \mu_N[P^k](i)$, then $\mu_N[P^k](i) = SEADAM[P](i)$.
  Furthermore, since i is improvable at $\mu_N[P^k]$, then $\mu_N[P^k](i) = A\mathcal{E}M[P^k](i)$ with $1 \leq k \leq \mid \mathcal{E}(P) \mid$.
  Therefore, by transitivity, $A\mathcal{E}M[P^k](i) = SEADAM[P](i)$, $\forall i \in U(DA[P])$.

- **Case 2**: Consider $i \in I(DA[P])$, and suppose $SEADAM[P](i) \neq DA[P](i)$, then $SEADAM[P](i) P_i DA[P](i)$ $\forall i \in I(DA[P])$. Also, we prove in the Lemma 1 that the set of students that could trade their schools is the same under every P^k, $I(\mu_N[P^k]) = I(DA[P])$.

We prove this case by contradiction. Assume that there is a student i for whom $A\mathcal{E}M[P^k](i) \neq SEADAM[P](i)$ with $1 \leq k \leq\mid \mathcal{E}(P) \mid$, and suppose $SEADAM[P](i)$ P_i $A\mathcal{E}M[P^k](i)$, the inverse $A\mathcal{E}M[P^k](i)$ P_i $SEADAM[P](i)$ has a symmetric process. Consider $s_{i_1} \equiv SEADAM[P](i_1)$ and $s_{i_2} \equiv A\mathcal{E}M[P^k](i_1)$, in general $s_{i_l} \equiv SEADAM$ $[P](i_l)$ and $s_{i_{l+1}} \equiv A\mathcal{E}M[P^k](i_l)$ (and $s_{i_1} \equiv A\mathcal{E}M[P^k](i_L)$) where $i_l \in \{i_1, i_2, ..., i_{L-1}, i_L\}$ $= \{i \mid i \in I(\mu_N[P^k])$, these are all the improvable students under the stable matching $\mu_N[P^k]$, and they prefer the matching under $SEADAM$ than under $A\mathcal{E}M$.

Then as $s_{i_l}P_{i_l}s_{i_{l+1}}$, and i does not get s_{i_l} under $A\mathcal{E}M[P^k]$, should happen one of two alternatives:

- Case 2.1: i_l never applies to s_{i_l} in the $A\mathcal{E}$ mechanism, but the preferences P^k are a truncation of P, then the only possibility for this is that s_l has been truncated from P_{i_l}. This imply that $s_{i_l} = \mu_N[P^r](i_l)$ for some $r < k$, because the Martínez, Massó, Neme and Oviedo algorithm[12]. As $s_{i_l} = SEADAM[P](i_l)$ and $s_{i_l} = \mu_N[P^r](i_l)$, then $SEADAM[P](i_l) = \mu_N[P^r](i_l)$, also $DA[P](i_l)$ R_{i_l} $\mu_N[P^r](i_l)$ because DA when the students propose is optimal for them, therefore $DA[P](i_l)$ R_{i_l} $SEADAM[P](i_l)$ which contradicts the Pareto efficiency of $SEADAM$ (condition 2 of the theorem: $SEADAM[P](i)P_iDA[P](i), \forall i \in I(DA[P])$).

- Case 2.2: s_{i_l} rejects i_l for some $i_j \in \{i_1, i_2, ...i_L\}$ during the $A\mathcal{E}$ mechanism. We suppose s_{i_1} rejects i_1 for i_L, then $i_L \succ_{s_{i_1}} i_1$, this happens because in some previous step s_{i_L} rejects i_L for i_{L-1}, what implies $i_{L-1} \succ_{s_{i_L}} i_L$, because in a previous step $s_{i_{L-1}}$ rejects i_{L-1} for i_{L-2}, then $i_{L-2} \succ_{s_{i_{L-1}}} i_{L-1}$, ..., because in previous step s_{i_3} rejects i_3 for i_2, then $i_2 \succ_{s_{i_3}} i_3$, because s_{i_2} rejects i_2 for i_1, what implies $i_1 \succ_{s_{i_2}} i_2$, all this during the $A\mathcal{E}$ process.

  Then i_1 applied first to s_{i_2} before to s_{i_1}, that is $s_{i_2}P^k_{i_1}s_{i_1}$, what implies that $s_{i_2}P_{i_1}s_{i_1}$ because P^k is a truncation of P and the schools in the preferences list of i_1 do not change the order.

  This contradicts our premise $s_{i_l}P_{i_l}s_{i_{l+1}}$, with $l = 1$, $s_{i_1}P_{i_1}s_{i_2}$.

Therefore, like $A\mathcal{E}M[P^k](i) \neq SEADAM[P](i)$ is not true, then
$A\mathcal{E}M[P^k](i) = SEADAM[P](i), \forall i \in I(DA[P])$ happens.

[12] All the truncated schools are schools that were assigned in some previous stable matching, the inverse is not true in general.

So, as Case 1 satisfies $A\mathcal{E}M[P^k](i) = SEADAM[P](i)$, $\forall i \in U(DA[P])$, and Case 2 satisfies $A\mathcal{E}M[P^k](i) = SEADAM[P](i)$, $\forall i \in I(DA[P])$, then $A\mathcal{E}M[P^k](i) = SEADAM[P](i)$, $\forall i \in N$ holds.

The proof of the essentially stable property is direct. The matching $A\mathcal{E}M[P^k]$, for all $1 \leq k \leq\mid \mathcal{E}(P) \mid$, is the same as in $SEADAM[P]$. Also, by Theorem 3, the matching produced by $SEADAM$ is essentially stable. Therefore, the $A\mathcal{E}M[P]$ is essentially stable.

■

Consider the examples 3 and 4. In the first one, notice that the improvable student i_1 has not gotten better assignment under $A\mathcal{E}M[P^1]$ than under $\mu_N[P] = DA[P]$, even more $A\mathcal{E}M[P^1](i_1) = DA[P](i_1)$, and by Proposition 1 $A\mathcal{E}M[P^1] = SEADAM[P]$. Then by transitivity $SEADAM[P](i_1) = DA[P](i_1) = s_3$. In the example 4, notice that the unimprovable student i_3 under DA is not unimprovable under $\mu_N[P^k]$ for $k = \{2, 3, 4\}$, that is $i_3 \notin U(\mu_N[P^k])$ for $k = \{2, 3, 4\}$.

The Examples 3 and 4 show when the assignments are different to $SEADAM$, in these cases two conditions are not satisfied: in Example 3, the improvable students are not improving; and in Example 4, the unimprovable students are different for every stable matching. Then, we ask all the unimprovable students at DA are the same in every stable matching $\mu_N(P^k)$, and all improvable students actually improve their assignment under P, this with the purpose of assure that every $A\mathcal{E}M[P^k]$ gets the same outcome that $SEADAM[P]$.

Let's show an example where these two conditions are satisfied. Also, we use this example to show the essentially stable property.

Example 5 ($A\mathcal{E}M$)**.** *Let's return to Example 2, where* $N = \{i_1, i_2, i_3, i_4, i_5\}$*,* $S = \{s_1, s_2, s_3, s_4, s_5\}$*, each school has only one seat and all students consent.*

Stage 0: Find all the stable matchings using the Martínez-Massó-Neme-Oviedo algorithm. In this case, there are only three stable matchings.

Stage 1: Round 0: We use the original preferences $P^1 = P$ *and the matching*$\mu_N[P]$*):*

$$\mu_N[P^1] = \mu_N[P] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 & i_5 \\ s_1 & s_3 & s_5 & s_4 & s_2 \end{pmatrix}$$

Round 1: The underdemanded school is s_2 and her assignment $A\mathcal{E}M[P^1](s_2) = i_5$, which is the unimprovable school. We remove this school and the student i_5 from the preferences and priorities of students and schools respectively and run the DA algorithm and find the matching $\mu'_N[P]$:

$$\mu'_N[P] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ s_3 & s_4 & s_1 & s_5 \end{pmatrix}$$

Round 2: The underdemanded schools are s_4 and s_5, we remove these schools and the students $A\mathcal{E}M[P^1](s_4) = i_2$ and $A\mathcal{E}M[P^1](s_5) = i_4$ from the preference-priority profile, and run the DA algorithm with this subproblem and find the matching:

$$\mu''_N[P] = \begin{pmatrix} i_1 & i_3 \\ s_3 & s_1 \end{pmatrix}$$

Round 3: The underdemanded school is s_1 and her assignment $A\mathcal{E}M[P^1](s_1) = i_3$, we remove this school and the student i_3 from the preference-priority profile, and run the DA algorithm with this subproblem and find the matching:

$$\mu'''_N[P] = \begin{pmatrix} i_1 \\ s_3 \end{pmatrix}$$

Round 4: The underdemanded school is s_3 and is assigned to $A\mathcal{E}M[P^1](s_3) = i_2$, then the $A\mathcal{E}M[P^1] = SEADAM[P]$ is:

$$A\mathcal{E}M[P^1] = SEADAM[P] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 & i_5 \\ s_3 & s_4 & s_1 & s_5 & s_2 \end{pmatrix}$$

Stage 2: Round 0: We identify and fix the preferences used to find the stable matching $\mu_N[P^2]$, in this case we use the preferences $P^2 = (P_{i_1}, P_{i_2}, P'_{i_3}, P_{i_4}, P_{i_5})$:

P_{i_1}	P_{i_2}	P'_{i_3}	P_{i_4}	P_{i_5}
s_3	s_4	s_3	s_1	s_1
s_1	s_3	s_1	s_5	s_5
s_4	s_5	s_4	s_4	s_2
s_5	s_1	s_2	s_3	s_3
s_2	s_2		s_2	s_4

$$\mu_N[P^2] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 & i_5 \\ s_1 & s_5 & s_4 & s_3 & s_2 \end{pmatrix}$$

Round 1: The underdemanded school is s_2 and her assignment $A\mathcal{E}M[P^2](s_2) = i_5$, which is the unimprovable school. We remove this school and the student i_5 from the preferences and priorities of students and schools respectively and run the DA algorithm to find the matching:

$$\mu'_N[P^2] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ s_3 & s_4 & s_1 & s_5 \end{pmatrix}$$

Round 2: The underdemanded schools are s_4 and s_5, we remove these schools and the students $A\mathcal{E}M[P^2](s_4) = i_2$ and $A\mathcal{E}M[P^2](s_5) = i_4$ from the preference-priority profile, and run the DA algorithm with this subproblem and find the matching:

$$\mu''_N[P^2] = \begin{pmatrix} i_1 & i_3 \\ s_3 & s_1 \end{pmatrix}$$

Round 3: The underdemanded school is s_1 and her assignment $A\mathcal{E}M[P^2](s_1) = i_3$, we remove this school and the student i_3 from the preference-priority profile, and run the DA algorithm with this subproblem and find the matching:

$$\mu'''_N[P^2] = \begin{pmatrix} i_1 \\ s_3 \end{pmatrix}$$

Round 4: The underdemanded school is s_3 and is assigned to $A\mathcal{E}M[P^2](s_3) = i_2$, then the $A\mathcal{E}M[P^2]$ is:

$$A\mathcal{E}M[P^2] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 & i_5 \\ s_3 & s_4 & s_1 & s_5 & s_2 \end{pmatrix}$$

Stage 3: Round 0: We identify the preferences used to find the stable matching $\mu_N[P^3]$, *in this case we use the preferences* $P^3 = (P_{i_1}, P'_{i_2}, P'_{i_3}, P_{i_4}, P_{i_5})$*:*

P_{i_1}	P'_{i_2}	P'_{i_3}	P_{i_4}	P_{i_5}
s_3	s_4	s_3	s_1	s_1
s_1	s_3	s_1	s_5	s_5
s_4	s_1	s_4	s_4	s_2
s_5	s_2	s_2	s_3	s_3
s_2			s_2	s_4

$$\mu_S[P] = \mu_N[P^3] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 & i_5 \\ s_5 & s_1 & s_4 & s_3 & s_2 \end{pmatrix}$$

Round 1: The underdemanded school is s_2 *and her assignment* $A\mathcal{E}M[P^3](s_2) = i_5$*, which is the unimprovable school. Now, we remove this school and the student* i_5 *from the preferences and priorities of students and schools and run the DA algorithm to find the matching:*

$$\mu'_N[P^3] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ s_3 & s_4 & s_1 & s_5 \end{pmatrix}$$

Round 2: The underdemanded schools are s_4 *and* s_5*, we remove these schools and the students* $A\mathcal{E}M[P^3](s_4) = i_2$ *and* $A\mathcal{E}M[P^3](s_5) = i_4$ *from the preference-priority profile, and run the DA algorithm with this subproblem and find the matching:*

$$\mu''_N[P^3] = \begin{pmatrix} i_1 & i_3 \\ s_3 & s_1 \end{pmatrix}$$

Round 3: The underdemanded school is s_1 *and her assignment* $A\mathcal{E}M[P^3](s_1) = i_3$*, we remove this school and the student* i_3 *from the preference-priority profile, and run the DA algorithm with this subproblem and we get:*

$$\mu'''_N[P^3] = \begin{pmatrix} i_1 \\ s_3 \end{pmatrix}$$

Round 4: The underdemanded school is s_3 *and is assigned to* $A\mathcal{E}M[P^3](s_3) = i_2$*, then the* $A\mathcal{E}M[P^3]$ *is:*

$$A\mathcal{E}M[P^3] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 & i_5 \\ s_3 & s_4 & s_1 & s_5 & s_2 \end{pmatrix}$$

Notice that the three mechanisms, $A\mathcal{E}M[P^1] = SEADAM[P]$, $A\mathcal{E}M[P^2]$ and $A\mathcal{E}M[P^3]$ have the same outcome.

By Theorem 3, the final matching produced by $SEADAM[P]$ is essentially stable, therefore the other two $A\mathcal{E}M[P^2]$ and $A\mathcal{E}M[P^3]$ are essentially stable too, because the outcome is the same as in $SEADAM[P]$.

Also, notice that this example satisfies the two conditions mentioned in the Theorem 5, unlike the examples 3 and 4. Condition 1: the unimprovable student is the same in every stable matching, $i_5 \in U(\mu_N[P^k])$ for $k = \{1, 2, 3\}$. Condition 2: the improvable students $\{i_1, i_2, i_3, i_4\}$ improve from DA, $SEADAM[P](i_1) = s_3\ P_{i_1}\ DA[P](i_1) = s_1$, $SEADAM[P](i_2) = s_4\ P_{i_2}\ DA[P](i_2) = s_3$, $SEADAM[P](i_3) = s_1\ P_{i_3}\ DA[P](i_3) = s_5$ and $SEADAM[P](i_4) = s_5\ P_{i_4}\ DA[P](i_4) = s_4$. □

In Theorem 5, we prove that conditions one and two are sufficient to produce the same matching than $SEADAM$, but this conditions are not necessary to produce the $SEADAM$ outcome. The following example shows that:

Example 6 (No necessary condition)**.** *The sets of students is $N = \{i_1, i_2, i_3, i_4, i_5, i_6\}$, and of schools is $S = \{s_1, s_2, s_3, s_4, s_5, s_6\}$ where each school has only one seat and all students consent. The profiles P of preferences and $\succ$ of priorities are as follows:*

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}	P_{i_5}	P_{i_6}
s_1	s_3	s_3	s_5	s_4	s_4
s_3	s_1	s_2	s_4	s_1	s_5
s_4	s_5	s_5	s_1	s_6	s_1
s_5	s_4	s_4	s_3	s_5	s_3
s_2	s_6	s_1	s_2	s_3	s_6
s_6	s_2	s_6	s_6	s_2	s_2

$\succ_{s_1}$	$\succ_{s_2}$	$\succ_{s_3}$	$\succ_{s_4}$	$\succ_{s_5}$	$\succ_{s_6}$
i_3	i_2	i_6	i_2	i_1	i_5
i_2	i_4	i_4	i_1	i_2	i_3
i_4	i_3	i_1	i_4	i_3	i_1
i_5	i_1	i_2	i_5	i_6	i_6
i_6	i_5	i_3	i_3	i_4	i_4
i_1	i_6	i_5	i_6	i_5	i_2

In this case, there are five stable matchings:

$$\mu_N[P^1] = \mu_N[P] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 & i_5 & i_6 \\ s_3 & s_1 & s_2 & s_4 & s_6 & s_5 \end{pmatrix}$$

$$\mu_N[P^2] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 & i_5 & i_6 \\ s_4 & s_1 & s_2 & s_3 & s_6 & s_5 \end{pmatrix}$$

$$\mu_N[P^3] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 & i_5 & i_6 \\ s_4 & s_1 & s_5 & s_2 & s_6 & s_3 \end{pmatrix}$$

$$\mu_N[P^4] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 & i_5 & i_6 \\ s_4 & s_5 & s_1 & s_2 & s_6 & s_3 \end{pmatrix}$$

$$\mu_S[P] = \mu_N[P^5] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 & i_5 & i_6 \\ s_5 & s_4 & s_1 & s_2 & s_6 & s_3 \end{pmatrix}$$

And applying the adjusted $\mathcal{E}$ mechanism to the truncated preferences $P = P^1$, P^2, P^3, P^4, and P^5, we find that for $k = \{1, 2, 3, 4, 5\}$:

$$SEADAM[P] = A\mathcal{E}M[P^k] = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 & i_5 & i_6 \\ s_1 & s_3 & s_2 & s_5 & s_6 & s_4 \end{pmatrix}$$

By Theorem 3, the final matching produced by $SEADAM[P]$ is essentially stable, therefore the other four $A\mathcal{E}M[P^k]$ are essentially stable too, because the outcomes are the same as in $SEADAM[P]$.

Also, notice that this example satisfies the condition 2 in the Theorem 5, all the improvable students $\{i_1, i_2, i_4, i_6\}$ are improving. Nevertheless, it does not satisfy the condition 1 because even when the unimprovable student i_5 is unimprovable in every stable matching, the unimprovable student i_3 is not, $i_3 \in U(\mu_N[P^k])$, for $k = \{1, 2\}$ and $i_3 \notin U(\mu_N[P^k])$, for $k = \{3, 4, 5\}$. Then, the condition 1 is not satisfied for all unimprovable students even though the outcome is the same of $SEADAM$. This shows that this condition is sufficient but not necessary to get the same outcome that in $SEADAM$.

Besides, the outcome is essentially stable because is the same as in $SEADAM$, therefore the two conditions are not necessary to accomplish essentially stable matchings. $\square$

2.5 Conclusion

We revisit SEADAM and ask if when we are looking to improve efficiency it is necessary to start with the DA outcome.

We propose a mechanism, the adjusted $\mathcal{E}$ mechanism, that takes any stable matching and improves its efficiency. Our algorithm follows the SEADAM structure, and inherits its properties (Pareto efficiency and essential stability) to the subproblem generated by the truncated preferences used to find the initial stable matching, but this properties do not keep in the original problem unless the outcome of the A$\mathcal{E}$M is the same as in $SEADAM$ when all students consent. Then, we look at which characteristics have to show the original problem to accomplish this goal. We ask for two conditions: (1) the improvable students has to improve at SEADAM, and (2) the unimprovable students are the same at every stable matching.

From this analysis we can conclude that the only mechanism Pareto efficient and essentially stable is the one who starts with the DA matching, that is, we can use any stable matching and improves its efficiency, but the only outcome that surely satisfies these two desirable properties is the SEADAM.

Our article opens the debate to future investigations regarding the existence of a family of mechanisms that obtain all the Pareto efficient and essentially stable matchings. We look for these mechanisms and find that the only way to find a Pareto efficient and essentially stable matching starting from any stable matching and using the SEADAM approach is to put restrictions on the market, and it is the same as the SEADAM outcome. Finding such mechanisms remains an interesting question.

2.6 Appendix

A.1 EADAM

We present the efficiency-adjusted deferred acceptance mechanism. Consider any school choice problem $(N, S, P, \succ)$ with consenting students. The EADAM operates as follows:

- *Round 0*: Run the DA algorithm

- *Round 1*: If there are no interrupting pairs, then stop. Otherwise, find the last step at which an interrupter is rejected from the school at which she is an interrupter. Identify all interrupting pairs of that step. For each identified interrupting pair (i, s), remove school s from the preferences of student i. Rerun the DA with the new preference profile.

- *Round k*, $k \geq 2$: In general, find the last step (of the DA run in Round $k-1$) at which an interrupter is rejected from the school at which she is an interrupter. If there are no interrupting pairs, then stop. Otherwise, identify all interrupting pairs of that step. For each interrupting pair (i, s), remove school s from the preferences of student i. Rerun the DA with the new preference profile.

Because the number of schools and students are finite, the algorithm eventually terminates in a finite number of steps. To illustrate EADAM, an example is shown.

Example 7 (EADAM)**.** *The sets of students is $N = \{i_1, i_2, i_3, i_4, i_5\}$, and of schools is $S = \{s_1, s_2, s_3, s_4, s_5\}$; let $q_s = 1$ for all $s \in S$ and assume all students consent. Let the profile P of preferences and the profile $\succ$ of priorities be as follows:*

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}	P_{i_5}
s_3	s_4	s_3	s_1	s_1
s_1	s_3	s_1	s_5	s_5
s_4	s_5	s_5	s_4	s_2
s_5	s_1	s_4	s_3	s_3
s_2	s_2	s_2	s_2	s_4

$\succ_{s_1}$	$\succ_{s_2}$	$\succ_{s_3}$	$\succ_{s_4}$	$\succ_{s_5}$
i_2	i_5	i_4	i_3	i_1
i_1	i_1	i_5	i_4	i_2
i_3	i_2	i_2	i_1	i_3
i_4	i_3	i_1	i_2	i_5
i_5	i_4	i_3	i_5	i_4

Round 0: Run the DA [13]

[13] The underlined student is the student with higher priority in the school s_i.

Step	s_1	s_2	s_3	s_4	s_5
1	$\underline{i_4}, i_5$		$\underline{i_1}, i_3$	i_2	
2	$\underline{i_3}, i_4$		i_1	i_2	i_5
3	i_3		i_1	i_2	$i_4, \underline{i_5}$
4	i_3		i_1	$i_2, \underline{i_4}$	i_5
5	i_3		$i_1, \underline{i_2}$	i_4	i_5
6	$\underline{i_1}, i_3$		i_2	i_4	i_5
7	i_1		i_2	i_4	$\underline{i_3}, i_5$
8	i_1	i_5	i_2	i_4	i_3

Round 1: The last step in which an interrupter is rejected is Step 7, where the interrupting pair is (i_5, s_5). We remove school s_5 from the preferences of student i_5 and rerun the DA algorithm with the following preference profile.

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}	P_{i_5}
s_3	s_4	s_3	s_1	s_1
s_1	s_3	s_1	s_5	s_2
s_4	s_5	s_5	s_4	s_3
s_5	s_1	s_4	s_3	s_4
s_2	s_2	s_2	s_2	

$\succ_{s_1}$	$\succ_{s_2}$	$\succ_{s_3}$	$\succ_{s_4}$	$\succ_{s_5}$
i_2	i_5	i_4	i_3	i_1
i_1	i_1	i_5	i_4	i_2
i_3	i_2	i_2	i_1	i_3
i_4	i_3	i_1	i_2	i_5
i_5	i_4	i_3	i_5	i_4

Step	s_1	s_2	s_3	s_4	s_5
1	$\underline{i_4}, i_5$		$\underline{i_1}, i_3$	i_2	
2	$\underline{i_3}, i_4$	i_5	i_1	i_2	
3	i_3	i_5	i_1	i_2	i_4

Round 2: The last step in which an interrupter is rejected is Step 2, where the interrupting pair is (i_4, s_1). We remove school s_1 from the preferences of student i_4 and rerun the DA algorithm.

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}	P_{i_5}
s_3	s_4	s_3	s_5	s_1
s_1	s_3	s_1	s_4	s_2
s_4	s_5	s_5	s_3	s_3
s_5	s_1	s_4	s_2	s_4
s_2	s_2	s_2		

$\succ_{s_1}$	$\succ_{s_2}$	$\succ_{s_3}$	$\succ_{s_4}$	$\succ_{s_5}$
i_2	i_5	i_4	i_3	i_1
i_1	i_1	i_5	i_4	i_2
i_3	i_2	i_2	i_1	i_3
i_4	i_3	i_1	i_2	i_5
i_5	i_4	i_3	i_5	i_4

Step	s_1	s_2	s_3	s_4	s_5
1	i_5		$\underline{i_1}, i_3$	i_2	i_4
2	$\underline{i_3}, i_5$		i_1	i_2	i_4
3	i_3	i_5	i_1	i_2	i_4

Round 3: There are no interrupting pairs, hence EADAM stops. The DA result is underlined and the EADAM assignment when all students consent is indicated by a center dot:

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}	P_{i_5}
$\cdot s_3$	$\cdot s_4$	s_3	s_1	s_1
$\underline{s_1}$	$\underline{s_3}$	$\cdot s_1$	$\cdot s_5$	s_5
s_4	s_5	$\underline{s_5}$	$\underline{s_4}$	$\cdot \underline{s_2}$
s_5	s_1	s_4	s_3	s_3
s_2	s_2	s_2	s_2	s_4

$\square$

A.2 The full set of stable matchings

First, we define the truncated preferences. We say that the preference $P_i^{(i,s)}$ is the (i,s)-**truncation of** P_i if:

1. All sets containing s are unacceptable to i according to $P_i^{(i,s)}$; that is, if $s \in S$ then $\emptyset P_i^{(i,s)} S$.
2. The preferences P_i and $P_i^{(i,s)}$ coincide on all sets that do not contain s; that is, if $s \notin S_1 \cup S_2$ then $S_1 P_i S_2$ if and only if $S_1 P_i^{(i,s)} S_2$.

3. The preferences P_i and $P_i^{(i,s)}$ coincide on all sets that contain s; that is, if $s \in S_1 \cap S_2$ then $S_1 P_i S_2$ if and only if $S_1 P_i^{(i,s)} S_2$.

4. All sets "artificially" made unacceptable in $P_i^{(i,s)}$ are preferred to the original unacceptable sets; that is, is S_1 and S_2 are such that $s \in S_1$ and $S_1 P_i \emptyset P_i S_2$ then $S_1 P_i^{(i,s)} S_2$.

We denote by $P^{(i,s)}$ the preference profile obtained by replacing P_i in P by $P_i^{(i,s)}$.

Now, we show an adaptation of the Martínez, Massó, Neme and Oviedo algorithm to compute the set of all stable matchings. We consider the algorithm in the context of the school choice problem, with the preference profile P. The algorithm proceeds as follows:

- **Round 0**: Set $T^0(P) := P$, $\mathcal{E}^0(P) := \{\mu_N\}$, and $k := 0$.

- **Round** k, $k \geq 1$: This round consist of four steps:

  1. Define $\tilde{T}(T^k(P)) = \{P^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})(i,s)} \mid s \in \mu_N^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})}(i) \setminus \mu_S(i)$, $P^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})} \in T^k(P)$, and $i \in N\}$[14].

  2. If $\tilde{T}(T^k(P)) = \emptyset$ set $T^{k+1}(P) = \emptyset$, $\mathcal{E}^{k+1}(P) = \mathcal{E}^k(P)$, else for each truncation $P^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})(i,s)} \in \tilde{T}(T^k(P))$ obtain $\mu_N^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})(i,s)}$.

  3. Define $T^*(T^k(P)) = \{P^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})(i,s)} \in \tilde{T}(T^k(P)) \mid \mu_N^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})(i,s)}(s) \succ_s \mu_N^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})}(s)\}$[15]. Order the set $T^*(T^k(P))$ in an arbitrary way and let $\prec^{k+1}$ denote this ordering.

  4. Define $\hat{T}(T^k(P)) = \{P^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})(i,s)} \in T^*(T^k(P)) \mid \forall P^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})(i',s')} \in T^*(T^k(P))$ such that $P^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})(i,s)} \prec^{k+1} P^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})(i',s')}$, $s' \in \mu_N^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})(i,s)}(i')\}$.
  Set $T^{k+1}(P) := \hat{T}(T^k(P))$, the set of stable matchings $\mathcal{E}^{k+1}(P) := \mathcal{E}^k(P) \cup \{\mu_N^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})(i,s)} \mid P^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})(i,s)} \in T^{k+1}(P)\}$, and $k := k+1$. Stops when $T^k(P)$ is empty.

[14]Where $P^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})}$ is the preference profile obtained from P after successively truncating the corresponding preference(s), and $\mu_N^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})}$ denotes its corresponding students optimal stable matching.

[15]Where $\mu_N^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})}$ denotes the schools optimal stable matching of truncated preferences $P^{(i_{j_1},s_{l_1})\dots(i_{j_k},s_{l_k})}$.

References

[1] Atila Abdulkadiroğlu and Tayfun Sönmez. *House Allocation with Existing Tenants.* Journal of Economic Theory. Vol 88, 1999; 233-260.

[2] Atila Abdulkadiroğlu and Tayfun Sönmez. *School Choice: A Mechanism Design Approach.* The American Economic Review. Vol 93, No.3, 2003; 729-747.

[3] Atila Abdulkadiroğlu and Tayfun Sönmez. *Matching Markets: Theory and Practice.* Advances in Economics and Econometrics Theory and Applications. Tenth World Congress, Vol 2, 2013; 3-47.

[4] Lars Ehlers and Bettina Klaus. *Efficient Priority Rules.* Games and Economic Behavior, Vol 55, 2006; 372-384.

[5] Haluk I. Ergin. *Efficient Resource Allocation on the Basis of Priorities.* Econometrica, Vol 70, No.6, 2002, 2489-2497.

[6] D. Gale and L.S. Shapley. *College Admissions and the Stability of Marriage.* The American Mathematical Monthly, Vol 69, No.1, 1962; 9-15.

[7] Onur Kesten. *School Choice with Consent.* The Quarterly Journal of Economics, 2010; 1297-1348.

[8] Fuhito Kojima. *Impossibility of Stable and Nonbossy Matching Mechanisms.* Economics Letters, Vol 107, 2010; 69-70.

[9] Szilvia Pápai. *Strategyproof Assignment by Hieralchical Exchange.* Econometrica, Vol 68, No.6, 2000; 1403-1433.

[10] Alvin E. Roth. *The Economics of Matching: Stability and Incentives.* Mathematics of Operations Research, Vol 7, No.4, 1982; 617-628.

[11] Alvin E. Roth and Marilda A. Oliveira Sotomayor. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis.* New York, Cambridge University Press, 1990.

[12] Qianfeng Tang and Jingsheng Yu. *A New Perspective on Kesten's School Choice with Consent Idea*. Journal of Economic Theory, Vol 154, 2014, 543-561.

[13] Qianfeng Tang and Yongchao Zhang. *Weak Stability and Pareto Efficiency in School Choice*. 2017. Working Paper. Available at SSRN: https://ssrn.com/abstract=2972611.

[14] Peter Troyan, David Delacrétaz and Andrew Kloosterman. *Essentially Stable Matchings*. July 18, 2016. Available at SSRN: https://ssrn.com/abstract=2812217 or http://dx.doi.org/10.2139/ssrn.2812217.

[15] Ruth Martínez, Jordi Massó, Alejandro Neme and Jorge Oviedo. *An Algorithm to Compute the Full Set of Many-to-Many Stable Matchings*. Mathematical Social Sciences, Vol 47, Issue 2, 2004; 187-210

Chapter 3

School Choice: A Market Extension

3.1 Introduction

The literature traditionally assumes that in the original School Choice model introduced in Abdulkadiroğlu and Sönmez (2003), the students submit a complete list of all available schools. However, contrary to what is studied in these markets, in many actual applications, the students submit only a reduced number of schools, and individual rationality can be violated. Many studies such as the developed by Roth and Vande Vate (1991), Ehlers (2008), Ashlagi and Klijn (2010) and Mongell and Roth (1990) show that the agents tend to underrepresent and truncate their list of real preferences in order to get a better assignment. This kind of behavior is notorious in situations such as the marriage market, medical market, college admissions, school choice, and even in sororities market.

The agents underrepresent their preferences thinking that this kind of behavior helps them to achieve a preferred assignment, and sometimes the algorithm used restricts the number of options that the agents can list. Some examples of these actions are presented in the following papers: Coles and Shorrer (2014) demonstrate that in large, uniform markets using the Men-Proposing Deferred Acceptance Algorithm, each woman's best response to truthful behavior by all other agents is to truncate her list substantially; Kojima and Pathak (2009) found truthful reporting by every participant is an approximate equilibrium under the student-optimal stable mechanism in large markets. Also, Mongell and Roth (1990) observe that the students tend to list only one or two sororities, as a consequence once the assignment process ends, between 1.5 and 30 percent of the students are not assigned; and Romero-Medina (1982) analyze the admission mechanism used in Spanish universities and found that the system is open to strategic manipulation because students are not allowed to express the whole list of available options. This behavior is also present in the school choice market of Mexico City[1], where on average students rank nine schools (they are allowed to rank 20) and a little more than the 20 percent remain without assigned school.

In this paper, we focus on solving the problem of unassigned students in the classic school choice problem, where we assign school seats to students, and we propose two different solutions to assign all of them. The first one adds to the original problem the condition that all students have to be assigned, this condition requires to develop new stability and efficiency concepts. The second one proves that report all the available schools as acceptable since the first step of our procedure is a dominant strategy and all students are assigned. We assume that the cardinality of the school seats is greater than the students' cardinality, and the preferences and priorities of both sets are strict. The aim of the school choice market is to assign a school seat to each student, and the students have incentives to finish school.

As we mentioned before, the first proposal to solve the problem is to develop several concepts and an algorithm (the Market Extension Algorithm) to find an assignment for every student. First, we develop the following concepts: *market extension*, where the market extension is carried out using only the students and school seats that have not been assigned by any stable matching, and asking students to extend their list of acceptable schools, if they refuse to do it we consider a random order among the remaining schools; *matching* in the extended market, that takes the assignment done by every market where the students have any school seat assigned; *r-stable*, if every assignment is stable in their market; and *r-optimal*, where every set, of students and schools, have a most preferred r-stable matching for all elements in the set. Second, we propose the Market Extension Algorithm (MEA) to find an r-stable matching and prove that the set of r-stable matchings is non-empty. The MEA when the students propose gets the r-optimal matching for the students and in analogous way we find the r-optimal matching for the schools with the market extension algorithm

[1]This market is described in the next chapter: "Does the home-school distance impact high school achievement? an analysis for Mexico City".

when the schools make the proposals. Third, the r-optimal matching for the students is the most preferred for all them, and the r-optimal matching for the schools is the least preferred for the students.

The second proposal to solve the problem uses the original concepts and proves that the dominant strategy for all students under the Market Extension Algorithm (MEA) when the students propose is to reveal a complete rank of all the schools as acceptable since the beginning. This method also achieves the aim of assigning all the students.

In addition to these general results, we present an example where the algorithm limits the number of preferred schools listed by the students, and we find that this restriction expands the incentives to misrepresent their preferences. Also, we analyze under which conditions the mechanism is manipulable or not, finding that it is manipulable by the students only when the schools are proposing. Finally, we analyze special cases where the r-stable matching is unique. The condition asked to this is satisfied is that the preferences and the extensions of the students are homogeneous (they order the schools in the same way).

The paper is organized as follows. We introduce the basic model of school choice in Section 2. Section 3 gives some important definitions as the market extension. Section 4 presents the model extension and some properties of the matching obtained (r-stable and r-optimal), along with the principal results. Section 5 presents some particular cases and manipulability. Section 6 concludes. The appendix presents the DA algorithm used in school choice market.

3.2 The model

Let $N = \{1, 2, ..., n\}$ denote the set of students and $S = \{1, 2, ..., m\}$ the set of schools. Let $q = (q_s)_{s \in S}$ where the integer $q_s \geq 1$ denotes the number of seats at school s. We assume that the total number of seats is no less than the number of students, $n \leq \sum_{s \in S} q_s$.

Each student $i \in N$ has strict preferences (complete, transitive and antisymmetric binary relation) over the set $S \cup \{i\}$, denoted by P_i, where the associated "at-least-as-good-as" relation is denoted by R_i. These preferences have a fixed length, but students have the freedom to truncate it, that is, to declare as unacceptable the schools that they consider. A *preference profile* is an n-tuple of preferences, denoted by $P = (P_1, ..., P_n)$. For each school $s \in S$, there is a strict priority order (complete, transitive and antisymmetric binary relation) over the set of students N, denoted by $\succ_s$. Define the priority profile $\succ = (\succ_s)_{s \in S}$. We consider the priorities of the schools a sort of preferences, in the sense that the priorities reflect an order over the students based on some criteria such as the district, exam score, past grades, etc., or a combination of these. Then, a school choice problem (or a market) consists of a quintuple $(N, S, P, \succ, q)$, and the preference-priority profile is denoted by the pair $(P, \succ)$.

A *matching* is a function $\mu : N \longrightarrow S \cup N$ such that: (i) if $\mu(i) \notin S$, then $\mu(i) = i$ for all $i \in N$, and (ii) for all $s \in S$, $|\{i \in N | \mu(i) = s\}| \leq q_s$. We say that a matching μ in a school choice problem $(N, S, P, \succ, q)$ is *blocked by a student* $i \in N$ if $iP_i\mu(i)$. The matching μ is *blocked by a pair* (i, s) if $s \neq \mu(i)$, $sP_i\mu(i)$, and $i \succ_s \mu(s)$. A matching μ is *stable* at $(P, \succ)$ if it is blocked neither by a student nor by a pair. We denote $\mathcal{E}(N, S, P, \succ, q)$ the set of all stable matchings at $(N, S, P, \succ, q)$. Gale and Shapley (1962) prove that this set is not empty.

We say that a matching μ is at least as good as μ' for all students $\mu P_N \mu'$, if $\mu(i) P_i \mu'(i)$ for at least one student and $\mu(i) R_i \mu'(i)$ for all students; and $\mu R_N \mu'$ denote that $\mu P_N \mu'$ or every student is indifferent between μ and μ'. We consider that the schools have priorities over the set of students, but when matchings are gotten the schools have preferences over them in the same way that the students have. Then, we say that a matching ν is at least as good as ν' for all schools $\nu P_S \nu'$, if $\nu(s) P_s \nu'(s)$ for at least one school and $\nu(s) R_s \nu'(s)$ for all schools; and $\nu R_S \nu'$ denote that $\nu P_S \nu'$ or every school is indifferent between ν and ν'.

Definition 5. *For every market $(N, S, P, \succ, q)$, a stable matching μ is N-optimal if for any other stable matching μ', $\mu R_N \mu'$. Similarly, a stable matching ν is S-optimal if for any other stable matching ν', $\nu R_S \nu'$.*

When all students have strict preferences and all schools have strict priorities, always exists a stable N-optimal matching and a stable S-optimal matching. Also, the matching μ_N obtained by the deferred acceptance algorithm with the students proposing is the stable N-optimal matching; and the stable S-optimal matching is the matching μ_S obtained by the deferred acceptance algorithm when the schools propose (Gale and Shapley, 1962). When this happens, for all matching stable μ, $\mu R_S \mu_N$ and $\mu R_N \mu_S$ (Knuth, 1976). By other hand, McVitie and Wilson (1970) and later Roth (1984) prove that the set of school seats filled and the set of students who are assigned are the same at every stable matching.

Fix a profile of priorities $\succ = (\succ_s)_{s\in S}$. A *mechanism* φ is a function that associates an assignment to every preference profile P. A mechanism φ is stable if for each profile P, $\varphi[P]$ is stable at P. The DA mechanism is stable (Gale and Shapley, 1962).

3.3 Market Extension

There are school choice problems where the stable matchings do not assign all the students, this could be a problem, especially in countries where the education is mandatory. We look for a way to solve this inconvenient. With this aim, we define the extended market M_1.

We call *acceptable schools* for the student i under P, denoted by S_i, to the set of schools preferred to i, that is $S_i = \{s \in S \mid sP_i i\}$. And the set of *schools (totally) assigned by* μ under P by $S_\mu = \{s \in S \mid |\{i \in N | \mu(i) = s\}| = q_s\}$.[2] In the same way, we define, $N_\mu = \{i \in N \mid \mu(i) \in S\}$, the set of students assigned by μ under P. Also, we have a quota filled by the matching μ, we denote this quota as $q_\mu = (q_s^\mu)_{s\in S}$, where $q_s^\mu = |\{i \in N | \mu(i) = s\}|$.

We denote the market $M_0 = M$ by the quintuple $(N^0, S^0, P^0, \succ^0, q^0) = (N, S, P, \succ, q)$, and by $\mathcal{M}^0$ the set of matchings at the market M_0.

The schools that have not been assigned by the matching $\mu^0 \in \mathcal{M}^0$, that are those for which the quota has not been filled, will be denoted by $S^1 = S^0 \setminus S_{\mu^0}$, the students that have not been assigned by the same matching will be denoted by $N^1 = N^0 \setminus N_{\mu^0}$, and the new quota is $q^1 = (q_s^1)_{s\in S^1}$ where $q_s^1 = q_s^0 - q_s^{\mu^0}$.

[2]We say that a school is totally assigned (or simply assigned) by μ when all the seats at this school are filled.

The market M_0 and the matching μ^0 define a new subproblem $M_1 = (N^1, S^1, P^1, \succ^1, q^1)$, which only depends of them and how are the preferences over S^1 and the priorities over N^1. Also, the sets N^1 and S^1 only depend of the students and schools assigned by the matching $\mu^0 \in \mathcal{M}^0$. The preferences and priorities in the new problem M_1 has to satisfy some rules in order to do not affect the original preferences P^0 and to get the aim of assign every student.

Definition 6. *Consider the market $M_0 = (N^0, S^0, P^0, \succ^0, q^0)$ and the matching $\mu^0 \in \mathcal{M}^0$. If $N^1 \neq \emptyset$ we define an* ***extension*** *at M_0 and μ^0 as $M_1 = (N^1, S^1, P^1, \succ^1, q^1)$, where the preference-priority profile $(P^1, \succ^1)$ is defined as follows:*

- *For all $s \in S^1$, $\succ_s^1 = \succ_s^0|_{N^1}$,*
- *For all $i \in N^1$, P_i^1 is an order under the elements of $S^1 \cup \{i\}$ where $s_j P_i^1 s_k$ if and only if $s_j P_i^0 s_k$ for all $s_j, s_k \in S^1$, and*
- *If there is at least a student $i \in N^1$ for which at least a school $s \in S^1$ is unacceptable[3], that is $s \in S^1 \setminus S_i^1$, then we consider a random order among these unacceptable schools, such that $sP_i^1 i$ for all $s \in S^1 \setminus S_i^1$.*

That is, the preferences and priorities are over the set of students and schools that have not been assigned by the matching μ^0. The priorities are the same that the priorities in the original problem but removing the students assigned by μ^0. The preferences are the same that the preferences in the original problem removing the schools assigned by μ^0, and all the students have to include all the remaining schools in their set of acceptable schools, if they refuse to do it or do it partially we consider a random order among the remaining schools.

Students may choose not to extend, in that case the extension is forced, which causes a violation of the preferences and generates non-individually rational assignments with respect to the original preferences, therefore we will treat that extension as a kind of consent. That is, when students reject their right to extend preferences, they will be consenting to be assigned a school that is among their unacceptable.

In order to illustrate this concept we present the following example.

[3]The acceptable schools for the student i at the extension are denote by S_i^1

Example 8 (extension)**.** *The sets of students is $N^0 = N = \{i_1, i_2, i_3, i_4, i_5, i_6\}$ and of schools is $S^0 = S = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7\}$, and let $q_s = 1$ for all $s \in S$. Let the profile $P^0 = P$ of preferences and the profile $\succ^0 = \succ$ of priorities be as follows:*

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}	P_{i_5}	P_{i_6}
s_1	s_1	s_2	s_2	s_2	s_2
s_2	s_2	s_1	s_1	s_1	s_1
s_7	i_2	s_7	i_4	i_5	i_6
s_5	s_4	i_3	s_4	s_5	s_6
s_6	s_3	s_6	s_7	s_6	s_3
i_1	s_6	s_5	s_3	s_7	s_5
s_3	s_7	s_4	s_5	s_4	s_7
s_4	s_5	s_3	s_6	s_3	s_4

$\succ_{s_1}$	$\succ_{s_2}$	$\succ_{s_3}$	$\succ_{s_4}$	$\succ_{s_5}$	$\succ_{s_6}$	$\succ_{s_7}$
i_1	i_3	i_4	i_3	i_1	i_1	i_4
i_2	i_4	i_2	i_1	i_3	i_2	i_6
i_3	i_1	i_3	i_2	i_2	i_3	i_2
i_4	i_6	i_5	i_5	i_6	i_6	i_3
i_6	i_2	i_6	i_6	i_4	i_5	i_5
i_5	i_5	i_1	i_4	i_5	i_4	i_1

When we apply the DA mechanism when the students propose, the matching found is: $\mu^0(i_1) = s_1$, $\mu^0(i_2) = i_2$, $\mu^0(i_3) = s_2$, $\mu^0(i_4) = i_4$, $\mu^0(i_5) = i_5$, $\mu^0(i_6) = i_6$.
Then, $S^1 = S^0 \setminus S_{\mu^0} = \{s_3, s_4, s_5, s_6, s_7\}$, $N^1 = N^0 \setminus N_{\mu^0} = \{i_2, i_4, i_5, i_6\} \neq \emptyset$, $q^1 = (1, 1, 1, 1, 1)$, and the extension of the market M_0 is the market $M_1 = (N^1, S^1, P^1, \succ^1, q^1)$, where $(P^1, \succ^1)$ could be the following, in the case that all students extend voluntarily:

$P^1_{i_2}$	$P^1_{i_4}$	$P^1_{i_5}$	$P^1_{i_6}$
s_4	s_4	s_5	s_6
s_3	s_7	s_6	s_3
s_6	s_3	i_7	s_5
s_7	s_5	s_4	s_7
s_5	s_6	s_3	s_4
i_2	i_4	i_5	i_6

$\succ^1_{s_3}$	$\succ^1_{s_4}$	$\succ^1_{s_5}$	$\succ^1_{s_6}$	$\succ^1_{s_7}$
i_4	i_2	i_2	i_2	i_4
i_2	i_5	i_6	i_6	i_6
i_5	i_6	i_4	i_5	i_2
i_6	i_4	i_5	i_4	i_5

Notice that these preferences and priorities are satisfying the three conditions for to be an extension:

- *Keep the order of the priorities $\succ^0_s$ only over the set N^1.*
- *All students are extending their acceptable schools to the remaning schools and keep the order of the preferences P^0_i over the set $S^1 \cup \{i\}$.*

- *In this case the third condition is satisfied by vacuity.*

With this, the new problem M_1 is correctly extended from M_0 and μ^0. □

Now, with this problem extension, we can obtain a matching $\mu^1 \in \mathcal{M}^1$, where $\mathcal{M}^1 := \mathcal{M}(\mu^0)$ denote all the matchings at the problem M_1.

Definition 7. *Given (M_0, M_1) and $\mu^k \in \mathcal{M}^k$, with $k = 0, 1$ we define $(\mu^0, \mu^1) : N \longrightarrow S \cup N$ if:*

i) For all $i \in N$, $(\mu^0, \mu^1)(i) = \mu^k(i)$ if $\mu^k(i) \neq i$,

ii) For all $s \in S \setminus S^1$, $|\{i \in N^0 \mid \mu^0(i) = s\}| = q_s$, and

iii) For all $s \in S^1$, $\sum_{k=0}^{1} |\{i \in N^k \mid \mu^k(i) = s\}| \leq q_s$

The function (μ^0, μ^1) assign a school s to each student i in the market in which the assignment is different to i for the first time. These assignments satisfy the conditions to make them matchings.

Lemma 2. *Given (M_0, M_1) and $\mu^k \in \mathcal{M}^k$, $k = 0, 1$. Then, (μ^0, μ^1) is* ***matching*** *in $\mathcal{M}^1$.*

Proof. The function $(\mu^0, \mu^1) : N \longrightarrow S \cup N$ needs to satisfy two conditions to consider it matching:

(i) If $(\mu^0, \mu^1)(i) \notin S$, then $(\mu^0, \mu^1)(i) = i$ for all $i \in N$. This condition is satisfied by vacuity, because $(\mu^0, \mu^1)(i) \in S \cup \{i\}$ and $(\mu^0, \mu^1) \neq i$, then $(\mu^0, \mu^1) \in S$.

(ii) For all $s \in S$, $|\{i \in N | (\mu^0, \mu^1)(i) = s\}| \leq q_s$. This condition is satisfied because the number of students assigned to school s at all the matchings (μ^0, μ^1) are equal to the sum of the number of students assigned to school s at every matching μ^0, μ^1. For the schools $s \in S \setminus S^1$ the sum is equal to the quote of every school, by the condition *ii)* in Definition 7, and for the schools $s \in S^1$ it is satisfied $\sum_{k=0}^{1} |\{i \in N^k \mid \mu^k(i) = s\}| \leq q_s$ by condition 3 in this definition. Then, $\sum_{k=0}^{1} |\{i \in N^k \mid \mu^k(i) = s\}| \leq q_s$ for all $s \in S$.

■

3.4 Properties

Definition 8. *Given* (M_0, M_1) *and* $\mu^k \in \mathcal{M}^k$, $k = 0, 1$. (μ^0, μ^1) *is r-**stable** if every* μ^k *is stable in the correspondent market* $(N^k, S^k, P^k, \succ^k, q^k)$.

The set of r-stable matchings are denoted by $\mathcal{E}_r((N^k, S^k, P^k, \succ^k, q^k)_{k=0}^1)$.

The simplest way to show that there is at least one r-stable matching in the market is given an algorithm whose outcome is an r-stable matching to every preference-priority profile. The algorithm of market extension design to reach this aim is an adaptation of the deferred acceptance algorithm when the students propose, and is the following.

Market Extension Algorithm

- *Round 1*: Given $(N, S, P, \succ, q)$, we define $N^0 = N$, $S^0 = S$, $P^0 = P$, $\succ^0 = \succ$, and $q^0 = q$ and run the deferred acceptance algorithm with the students making the proposes in the market $M_0 = (N^0, S^0, P^0, \succ^0, q^0)$, getting the matching μ^0. Then, we define the market $M_1 = (N^1, S^1, P^1, \succ^1, q^1)$ as indicated in the definition of market 1-extended.

- *Round 2*: We run the deferred acceptance algorithm in the market $M_1 = (N^1, S^1, P^1, \succ^1, q^1)$. The algorithm stops.

The algorithm stops in the round 2 because when we extend the market all the students add all the remaining schools to the acceptable sets, then in second round are assigned the remaining students.

Theorem 6. *Given the markets* (M_0, M_1), *the set of r-stable matchings* $\mathcal{E}_r((N^k, S^k, P^k, \succ^k, q^k)_{k=0}^1)$ *is not empty.*

Proof. We have the markets M_0, M_1, for market M_0 we always have at least one stable matching (Gale and Shapley, 1962). This is true also for the market M_1. Therefore, there is at least one stable matching at every market, by definition of r-stable matching, we have at least one r-stable matching.

■

Example 9 (Market Extension Algorithm)**.** *Let's continue with the Example 1, we obtained using the deferred acceptance algorithm the following assignments:* $\mu^0(i_1) = s_1$*,* $\mu^0(i_3) = s_2$*. Then, the extended market is* $M_1 = (N^1, S^1, P^1, \succ^1, q^1)$*, where* $S^1 = \{s_3, s_4, s_5, s_6, s_7\}$*,* $N^1 = \{i_2, i_4, i_5, i_6\}$*,* $q^1 = (1, 1, 1, 1, 1)$*, and* $(P^1, \succ^1)$ *is:*

$P^1_{i_2}$	$P^1_{i_4}$	$P^1_{i_5}$	$P^1_{i_6}$
s_4	s_4	s_5	s_6
s_3	s_7	s_6	s_3
s_6	s_3	i_7	s_5
s_7	s_5	s_4	s_7
s_5	s_6	s_3	s_4
i_2	i_4	i_5	i_6

$\succ^1_{s_3}$	$\succ^1_{s_4}$	$\succ^1_{s_5}$	$\succ^1_{s_6}$	$\succ^1_{s_7}$
i_4	i_2	i_2	i_2	i_4
i_2	i_5	i_6	i_6	i_6
i_5	i_6	i_4	i_5	i_2
i_6	i_4	i_5	i_4	i_5

We run the deferred acceptance algorithm at this market M_1 *and we obtain the following assignments:* $\mu^1(i_2) = s_4$*,* $\mu^1 i_4 = s_7$*,* $\mu^1(i_5) = s_5$*, and* $\mu^1(i_6) = s_6$*.*
Notice that our principal goal is reached, all students have assignment already. Then, the school s_7 *stays with the seat empty.*
Then, the matching is:

$$(\mu^0, \mu^1)(i_1) = \mu^0(i_1) = s_1, (\mu^0, \mu^1)(i_2) = \mu^1(i_2) = s_4$$

$$(\mu^0, \mu^1)(i_3) = \mu^0(i_3) = s_2, (\mu^0, \mu^1)(i_4) = \mu^1(i_4) = s_7$$

$$(\mu^0, \mu^1)(i_5) = \mu^1(i_5) = s_5, (\mu^0, \mu^1)(i_6) = \mu^1(i_6) = s_6$$

Also, because all the matchings have been founded by deferred acceptance algorithm, these are stables in each market, that is, the matching (μ^0, μ^1) *is* r*-stable.* □

Notice that does not matter which stable matching $\mu^0 \in \mathcal{M}^0$ is used in the market M_0, the sets N^1 and S^1 are formed by the same elements. This is because in a market with strict preferences and priorities, the set of unmatched students and schools are the same in every stable matching (Roth, 1984).

We consider that the schools can do the proposes too during the market extension algorithm (MEA). In order to distinguish who is making the proposes during the algorithm in every market M_k, the students or the schools, we denote by μ_N^k the matching gotten by the MEA in market k when the students propose, and by μ_S^k the matching gotten by the MEA in market k when the schools propose with $k = 0, 1$. Notice that μ_N^k and μ_S^k are stables in the market M_k, therefore the succession of this matchings with $k = 0, 1$ is r-stable.

Lemma 3. *Let (μ_N^0, μ_N^1) be the matching gotten by the market extension algorithm where the students propose, then the matching (μ_N^0, μ_N^1) is r-stable. Analogously, (μ_S^0, μ_S^1) is the matching gotten by the market extension algorithm where the schools propose, then the matching (μ_S^0, μ_S^1) is r-stable.*

Proof. To proof that (μ_N^0, μ_N^1) is r-stable it is necessary that every one of the element of the succession is stable. This is satisfied because every element μ_S^0 and μ_S^1 is the outcome of the deferred acceptance algorithm when the students propose, and these are stable at its corresponding market M_k, $k = 0, 1$. Therefore, (μ_N^0, μ_N^1) is r-stable. The proof for the matching (μ_S^0, μ_S^1) is analogous.

■

We say that a matching μ^k is at least as good as μ'^k for all the students N^k at the market M_k, $\mu^k P_{N^k}^k \mu'^k$, if $\mu^k(i) P_i^k \mu'^k(i)$ at least for an $i \in N^k$ and $\mu^k(i) R_i^k \mu'^k(i)$ for all the students, for $k = 0, 1$; and $\mu^k R_{N^k}^k \mu'^k$ denotes that it is satisfied $\mu^k P_{N^k}^k \mu'^k$ or that all students in N^k are indifferent between μ^k and μ'^k. Notice that the indifference between μ^k and μ'^k implies that $\mu^k = \mu'^k$ because we are using strict preferences. Similarly, we use $P_{S^k}^k$ and $R_{S^k}^k$ to denote the preferences of the schools over the matchings in the market M_k.

Denote by $\mathcal{R}_N = (R_{N^0}^0, R_{N^1}^1)$ the preferences of the students over the matchings in the markets M_0 and M_1. Analogously, $\mathcal{R}_S = (R_{S^0}^0, R_{S^1}^1)$ denotes the preferences of the schools over the matchings in the markets M_0 and M_1.

For the markets M_0 and M_1, we say that an r-stable matching (μ^0, μ^1) is r-optimal for the students, if for every student assigned at M_k, with $k = 0, 1$, (μ^0, μ^1) is at least as good as any other r-stable matching in the same market. We define the optimal r-stable matching for the schools in the same way.

Definition 9. *Given the markets (M_0, M_1), an r-stable matching (μ^0, μ^1) is* $\mathbf{N_r}$-***optimal*** *(r-optimal for the students) if $(\mu^0, \mu^1)\mathcal{R}_N\ (\mu'^0, \mu'^1)$ for any other r-stable matching (μ'^0, μ'^1). Similarly, an r-stable matching (ν^0, ν^1) is* $\mathbf{S_r}$-***optimal*** *(r-optimal for the schools) if for any other r-stable matching (ν'^0, ν'^1) is fulfilled $(\nu^0, \nu^1)\mathcal{R}_S\ (\nu'^0, \nu'^1)$.*

At (M_0, M_1) where the preferences of the students are strict, each one of them have only one favorite r-stable matching, this is the one where the assigned school is the most preferred achievable school for the student in every market.[4] This is doing by the stable matching N-optimal in every market. The case of the stable matching S-optimal is analogous.

Theorem 7. *The matching (μ_N^0, μ_N^1) obtained by the market extension algorithm when the students propose is the r-stable N_r-optimal matching. And the r-stable S_r-optimal is the matching (μ_S^0, μ_S^1) gotten by the market extension algorithm when the schools propose.*

Proof. The matching μ_N^0 obtained by the deferred acceptance algorithm when the students propose in the market M_0 is the stable N-optimal matching (Gale and Shapley, 1962). Using the same result we have that μ_N^1 is N-optimal in M_1. Then $\mu_N^k R_{N^k}^k \mu^k$ in every market M_k, for $k = 0, 1$ and any other μ^k in the market M_k. Therefore, $(\mu_N^0, \mu_N^1)\mathcal{R}_N(\mu^0, \mu^1)$, that is the matching (μ_N^0, μ_N^1) is the N_r-optimal.
Analogously, it proves the case of the matching (μ_S^0, μ_S^1) which results the S_r-optimal.

■

Students and schools have opposite interests, that is, the r-stable matching which is optimal for one side of the market is the worst r-stable matching for the other side. The r-stable N_r-optimal matching is the worst stable matching for the schools, that is, assigns to every school with its least preferred set of achievable students. In the same way, the r-stable S_r-optimal assigns to every student with her least preferred achievable school.

Theorem 8. *Given the extended markets (M_0, M_1). If (μ_N^0, μ_N^1) and (μ_S^0, μ_S^1) are the N_r-optimal and S_r-optimal matchings, respectively. Then, any other r-stable matching (μ^0, μ^1) satisfies:*

$$(\mu_N^0, \mu_N^1)\mathcal{R}_N(\mu^0, \mu^1)\mathcal{R}_N(\mu_S^0, \mu_S^1)$$

[4] A school is achievable for a student if it is assigned to the student at some stable matching.

Proof. The proof is divided in two parts, (1) $(\mu_N^0, \mu_N^1)\mathcal{R}_N(\mu^0, \mu^1)$, and (2) $(\mu^0, \mu^1)\mathcal{R}_N(\mu_S^0, \mu_S^1)$:

1. μ_N^k is the k-stable N_k-optimal matching at the market M_k and always is the most preferred by all the students $i \in N^k$, this happens for markets M_0 and M_1, that is $\mu_{N^k}^k R_{N^k}^k \mu^k$ for $k = 0, 1$. Therefore, $(\mu_N^0, \mu_N^1)\mathcal{R}_N\ (\mu^0, \mu^1)$.

2. μ_S^k is the k-stable S_k-optimal matching, that is this is the most preferred by all the schools at the market M_k, $\mu_{S^k}^k R_{S^k}^k \mu^k$ for $k = 0, 1$. By Knuth (1976), for every market M_0 and M_1, the matching most preferred by the schools is the least preferred by the students, that is $\mu_{S^k}^k R_{S^k}^k \mu^k$ if and only if $\mu^k R_{N^k}^k \mu_{S^k}^k$. Therefore, $(\mu^0, \mu^1)\mathcal{R}_N\ (\mu_S^0, \mu_S^1)$.

Then, $(\mu_N^0, \mu_N^1)\mathcal{R}_N(\mu^0, \mu^1)\mathcal{R}_N(\mu_S^0, \mu_S^1)$. ■

Each student, whose preferences are P_i, is faced with the problem of deciding what preference ordering Q_i to state. We consider that each school with priorities $\succ_s$ state it as her priority ordering and we fix the priority profile $\succ$, then the only agents with strategies will be the students. The mechanism used to find a matching $\mu[Q]$ is a function of the stated preference profile (or strategy) $Q = \{Q_{i_1}, ..., Q_{i_n}\}$. If Q represents the choices of all students, then $Q = (Q_{-i}, Q_i)$ focus on the decision facing one of them, where Q_i is the choice of student i and Q_{-i} is the set of choices of all students other than i, and $Q' = (Q_{-i}, Q_i')$ differ from Q only in student i's choice.

A particular strategy choice Q_i^* by student i is a *best response* by i to Q_{-i} if student i likes $\mu[Q_{-i}, Q_i^*]$ at least as well as any of the outcomes $\mu[Q_{-i}, Q_i]$ that would have resulted from any other strategy Q_i she could have chosen. A *dominant strategy* for a student i is a strategy Q_i^* that is a best response to all possible sets of strategy choices Q_{-i} by the other students.

Commonly, the students do not know how exactly the assignment mechanism works, that is why they think that the best response to the strategies of the other students is to list a shorter rank of acceptable schools. Other times, the mechanism limits the number of schools that they could to list as their acceptable schools. Nevertheless, the dominant strategy that the students have is to rank all the schools as acceptable.

Theorem 9. *Given a market $(N, S, P, \succ, q)$, the strategy of student i, Q_i, where all schools are acceptable, it is the dominant strategy under the market extension algorithm when the students propose, that is $(\mu_N^0, \mu_N^1)[P_{-i}, Q_i]\ R_i\ (\mu_N^0, \mu_N^1)[P_{-i}, P_i]$ for any P_{-i}.*

The dominant strategy of student i to any preference profile of the others students P_{-i} is rank all the schools as acceptable, that is, the best that a student can do when the market extension algorithm when the students propose is used to make the assignment, is to establish as their rank of preferences P_i the complete list where all the schools appear as acceptable.

Proof. The proof is by comparing the outcomes of MEA with the original preferences P and with the strategy Q_i. Take the problem where the preferences of the student i are P_i and the preferences of the other students are P_{-i}, and the students have to decide if they state this preferences or not. We start the proof in the last round of the market extension (round 2) where all students state their real preference profile $P = \{P_{-i}, P_i\}$ and extend their preferences the remaining students in round 2 as the market extension algorithm indicates. We take a student $i \in N$ which is assigned in the last round of the MEA, and we compare the outcome of the matching $(\mu_N^0, \mu_N^1)[P]$ and the outcome of the matching $(\mu_N^0)[P_{-i}, Q_i]$ for this student i, where Q_i is the strategy where the student i reveals all the list of the remaining schools as acceptable to her at round 1, and $(\mu_N^0, \mu_N^1)[P](i) = s_1$. Notice, it is more probable that some school s_0 which is more preferred than the school s_1 by i have (more) free seats in the round 1 than in the round 2, and she can be assigned at this school because i has less competition in round 1 than in round 2 where more students add this school in their acceptable schools. Then, the school assigned by $(\mu_N^0)[P_{-i}, Q_i]$ to i is at least as good as the school s_1. It is indifferent to s_1 if is equal to s_1 because the students' preferences are strict, or is preferred by i to s_1, where $(\mu_N^0)[P_{-i}, Q_i](i) = s_0$, that is $s_0 R_i s_1$. Therefore, $(\mu_N^0)[P_{-i}, Q_i](i) = s_0\ R_i\ (\mu_N^0, \mu_N^1)[P](i) = s_1$, that is, the dominant strategy of student i is to give since the beginning the preferences $P_i = Q_i$ where she lists all schools as acceptable. ■

The following example serves to illustrate the previous theorem.

Example 10. *The sets of students is $N^0 = N = \{i_1, i_2, i_3, i_4, i_5, i_6\}$ and of schools is $S^0 = S = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7\}$, and let $q_s = 1$ for all $s \in S$. Let the profile $P^0 = P$ of preferences and the profile $\succ^0 = \succ$ of priorities be as follows:*

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}	P_{i_5}	P_{i_6}
s_1	s_1	s_2	s_2	s_2	s_2
s_2	s_2	s_1	s_1	s_1	s_1
s_7	i_2	s_7	s_4	i_5	i_6
s_5	s_4	i_3	i_4	s_5	s_6
s_6	s_3	s_6	s_7	s_6	s_3
i_1	s_6	s_5	s_3	s_7	s_5
s_3	s_7	s_4	s_5	s_4	s_7
s_4	s_5	s_3	s_6	s_3	s_4

$\succ_{s_1}$	$\succ_{s_2}$	$\succ_{s_3}$	$\succ_{s_4}$	$\succ_{s_5}$	$\succ_{s_6}$	$\succ_{s_7}$
i_1	i_3	i_4	i_3	i_1	i_1	i_4
i_2	i_4	i_2	i_1	i_3	i_2	i_6
i_3	i_1	i_3	i_2	i_2	i_3	i_2
i_4	i_6	i_5	i_5	i_6	i_6	i_3
i_6	i_2	i_6	i_6	i_4	i_5	i_5
i_5	i_5	i_1	i_4	i_5	i_4	i_1

When we apply the DA mechanism when the students propose, the matching found is: $\mu^0(i_1) = s_1$, $\mu^0(i_2) = i_2$, $\mu^0(i_3) = s_2$, $\mu^0(i_4) = s_4$, $\mu^0(i_5) = i_5$, $\mu^0(i_6) = i_6$.
Then, $S^1 = S^0 \setminus S_{\mu^0} = \{s_3, s_5, s_6, s_7\}$, $N^1 = N^0 \setminus N_{\mu^0} = \{i_2, i_5, i_6\} \neq \emptyset$, $q^1 = (1,1,1,1)$, *and the extension of the market* M_0 *is the market* $M_1 = (N^1, S^1, P^1, \succ^1, q^1)$, *where* $(P^1, \succ^1)$ *could be the following, in the case that all students extend voluntarily:*

$P^1_{i_2}$	$P^1_{i_5}$	$P^1_{i_6}$
s_3	s_5	s_6
s_6	s_6	s_3
s_7	i_7	s_5
s_5	s_3	s_7
i_2	i_5	i_6

$\succ^1_{s_3}$	$\succ^1_{s_5}$	$\succ^1_{s_6}$	$\succ^1_{s_7}$
i_2	i_2	i_2	i_6
i_5	i_6	i_6	i_2
i_6	i_5	i_5	i_5

We run the deferred acceptance algorithm at this market M_1 *and we obtain the following assignments:* $\mu^1(i_2) = s_3$, $\mu^1(i_5) = s_5$, *and* $\mu^1(i_6) = s_6$ *and the school* s_7 *stays with the seat empty. Then, the matching is:*

$$(\mu^0, \mu^1)(i_1) = \mu^0(i_1) = s_1, (\mu^0, \mu^1)(i_2) = \mu^1(i_2) = s_3$$

$$(\mu^0, \mu^1)(i_3) = \mu^0(i_3) = s_2, (\mu^0, \mu^1)(i_4) = \mu^0(i_4) = s_4$$

$$(\mu^0, \mu^1)(i_5) = \mu^1(i_5) = s_5, (\mu^0, \mu^1)(i_6) = \mu^1(i_6) = s_6$$

Now, we consider the student i_2 *takes all schools as acceptable, that is, the set of strategies and priorities where all stay with the original preferences except the student* i_2, $Q = \{P_{-i_2}, Q_{i_2}\}$ *is:*

P_{i_1}	Q_{i_2}	P_{i_3}	P_{i_4}	P_{i_5}	P_{i_6}
s_1	s_1	s_2	s_2	s_2	s_2
s_2	s_2	s_1	s_1	s_1	s_1
s_7	s_4	s_7	s_4	i_5	i_6
s_5	s_3	i_3	i_4	s_5	s_6
s_6	s_6	s_6	s_7	s_6	s_3
i_1	s_7	s_5	s_3	s_7	s_5
s_3	s_5	s_4	s_5	s_4	s_7
s_4	i_2	s_3	s_6	s_3	s_4

$\succ_{s_1}$	$\succ_{s_2}$	$\succ_{s_3}$	$\succ_{s_4}$	$\succ_{s_5}$	$\succ_{s_6}$	$\succ_{s_7}$
i_1	i_3	i_4	i_3	i_1	i_1	i_4
i_2	i_4	i_2	i_1	i_3	i_2	i_6
i_3	i_1	i_3	i_2	i_2	i_3	i_2
i_4	i_6	i_5	i_5	i_6	i_6	i_3
i_6	i_2	i_6	i_6	i_4	i_5	i_5
i_5	i_5	i_1	i_4	i_5	i_4	i_1

Also, assume that the students that do not get an assignment after the first round of the MEA extend their list in the same way than in the previous case.

When we apply the market extension algorithm, we find the next assignment:

$$(\nu^0, \nu^1)(i_1) = \nu^0(i_1) = s_1, (\nu^0, \nu^1)(i_2) = \nu^0(i_2) = s_4$$

$$(\nu^0, \nu^1)(i_3) = \nu^0(i_3) = s_2, (\nu^0, \nu^1)(i_4) = \nu^1(i_4) = s_7$$

$$(\nu^0, \nu^1)(i_5) = \nu^1(i_5) = s_5, (\nu^0, \nu^1)(i_6) = \nu^1(i_6) = s_6$$

Notice that student i_2 is at least as well under this matching as under the matching gotten in the previous case, actually i_2 is strictly better, that is $(\nu^0, \nu^1)(i_2) = s_4 \; P_{i_2} \; (\mu^0, \mu^1)(i_2) = s_3$, then Q_{i_2} is the best response to the original preferences of the other students. This will happen no matters which are the preferences of the other students, then the strategy Q_{i_2} is the dominant strategy.

3.5 Manipulability and Particular Cases

Now, we suppose that the students, who know the process to find the assignment, want to find a way to benefit, that is, they try to manipulate in some way to be more satisfied with the new assigned school than with the original one.

In the beginning we think that the students always try to reveal their complete list of preferred schools as fast as possible, using this strategy they avoid that any other student get ahead and acquire a seat in some more preferred school than other that could be assigned to her. An example where this conjecture is not true is presented following.

Example 11. *The market M_0 with two students $N^0 = \{i_1, i_2\}$ and two schools $S^0 = \{s_1, s_2\}$ where each school has only one seat, and the preferences-priorities profile:*

$P^0_{i_1}$	$P^0_{i_2}$
s_2	s_1
i_1	i_2
s_1	s_2

$\succ^0_{s_1}$	$\succ^0_{s_2}$
i_1	i_2
i_2	i_1

With the market extension algorithm when the schools propose we have that the final matching is: $\mu^0(i_1) = s_2$ and $\mu^0(i_2) = s_1$.

Instead if students provide a complete list of preferences extending their acceptable schools to all the available schools, the matching gotten by the same algorithm is different. That is, if:

$P'^0_{i_1}$	$P'^0_{i_2}$
s_2	s_1
s_1	s_2
i_1	i_2

$\succ^0_{s_1}$	$\succ^0_{s_2}$
i_1	i_2
i_2	i_1

Then, the matching gotten is: $\mu'^0(i_1) = s_1$ and $\mu'^0(i_2) = s_2$, which is a worse option for the students than the previous given by μ. Therefore, the students prefer give the original and real preferences P^0 and not the extended ones P'^0. □

In the previous example we see that the students may not have incentives to extend their preferences from the start. If the students have real preferences P'^0, they gotten a bad assignment, but they could agree and misrepresent their preferences, manipulate the algorithm and give the preferences P^0 to get a more preferred school than the assigned by μ'.

The market extension algorithm is manipulable by the students when the schools are proposing, but when the students propose at the DA algorithm, Roth (1985) proves that makes it a dominant strategy for every $i \in N$ to state her true preferences. This result could be extended to our model, in the markets (M_0, M_1) the Roth's result is satisfied.

Another important variant inside the school choice problem is when the students are limited to submit a list containing only a limited number of schools. This happens in several countries around the world when use a centralize mechanism to assigned seats at high schools and college. For instance in New York the students have allowed to list 12 options for the entrance to high school, in Spain and Hungary they are limited to list eight and four options respectively when they apply to college. In Mexico City is not different, the COMIPEMS[5] allows 20 school options. This variant could be adapted to our model too.

When a limit on the number of school options exists, the DA algorithm can no longer guarantee that the dominant strategy for students is state her true preferences. Calsamiglia, Haeringer and Klijn (2010) show that students exhibit a higher proportion of misrepresentation of real preferences in the constrained case. So, we analyze what happen in our case of extension of preferences with a limited number of school options.

Example 12. *The sets of students is $N^0 = N = \{i_1, i_2, i_3, i_4, i_5, i_6\}$ and of schools is $S^0 = S = \{s_1, s_2, s_3, s_4, s_5, s_6\}$, and let $q_s = 1$ for all $s \in S$. Let the profile $P^0 = P$ of real preferences and the profile $\succ^0 = \succ$ of priorities be as follows:*

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}	P_{i_5}	P_{i_6}
s_1	s_2	s_1	s_1	s_2	s_2
s_2	s_1	s_2	s_2	s_1	s_1
i_1	i_2	s_4	i_4	s_4	i_6
s_3	s_4	i_3	s_4	i_5	s_6
s_5	s_5	s_5	s_5	s_5	s_4
s_6	s_6	s_3	s_4	s_3	s_5
s_4	s_3	s_6	s_6	s_6	s_3

$\succ_{s_1}$	$\succ_{s_2}$	$\succ_{s_3}$	$\succ_{s_4}$	$\succ_{s_5}$	$\succ_{s_6}$
i_1	i_4	i_6	i_2	i_4	i_2
i_4	i_2	i_2	i_3	i_5	i_1
i_2	i_6	i_1	i_5	i_6	i_3
i_6	i_1	i_4	i_6	i_3	i_6
i_5	i_5	i_3	i_1	i_1	i_4
i_3	i_3	i_5	i_4	i_2	i_5

[5] COMIPEMS is the organization in charge to assign high school seats to applicants

Suppose that the mechanism restricts the number of school options to 2, then student i_5 *and* i_3 *have to get out of their acceptable schools the school* s_4*. When we apply the* MEA *when the students propose, the matching found is:* $\mu^0(i_1) = s_1$, $\mu^1(i_2) = s_4$, $\mu^1(i_3) = s_3$, $\mu^0(i_4) = s_2$, $\mu^1(i_5) = s_5$, $\mu^1(i_6) = s_6$.
Now, consider that the student i_3 *misrepresents her preferences to achieve a better school than in the previous case, and presents the preferences* P'_{i_3}*:*

$P^1_{i_3}$
s_1
s_4
i_3
s_2
$\vdots$

When we apply the MEA*, the matching found is:* $\mu^0(i_1) = s_1$, $\mu^1(i_2) = s_5$, $\mu^0(i_3) = s_4$, $\mu^0(i_4) = s_2$, $\mu^1(i_5) = s_3$, $\mu^1(i_6) = s_6$*. Notice that the student* i_3 *is better if she state the preferences* P'_{i_3} *than with the real preferences when are restricted.* □

Eeckhout (2000) analyzes the case when there is a unique matching and the conditions that have to satisfy the preferences. This kind of analysis leads us to ask when there is a preference extension, which conditions in the preferences are enough for two r-stable matchings are equal; and which conditions have to be satisfied for two matchings with different extensions are the same.

We say that two preferences are homogeneous for two students $i, i' \in N$ if the order in their preferences is the same, and we denote this by $P_i \sim P_{i'}$. Analogously, two preferences in a extended market are homogeneous $P^1_i \sim P^1_{i'}$ for the students $i, i' \in N^1$ if they have the same order in their extended preferences.

Considering this concepts, we have that in the school choice problem, the N_r-optimal and the S_r-optimal matchings are the same when the students preferences and the extensions of these preferences are always the same for all students. That is, if all students have the same order in their preferences and they extend their preferences in the same way, then does not matter who is doing the proposes, the outcome is the same.

Proposition 3. *If $P_i^k \sim P_{i'}^k$ for all $i, i' \in N$ and $k = 0, 1$, then $(\mu_N^0, \mu_N^1) = (\mu_S^0, \mu_S^1)$.*

Proof. The proof is by contradiction. Suppose $(\mu_N^0, \mu_N^1) \neq (\mu_S^0, \mu_S^1)$, then by Theorem 3, $(\mu_N^0, \mu_N^1)\mathcal{R}_N(\mu_S^0, \mu_S^1)$, that is $\mu_N^k(i)P_{N^k}^k\mu_S^k(i)$ for $k = 0, 1$, all students and the markets (M_0, M_1)(in this case is not possible the indifference because we are taken strict preferences and $(\mu_N^0, \mu_N^1) \neq (\mu_S^0, \mu_S^1)$). We have two cases where exists at least one i for which are different the assignments of this two matchings:

1. This student takes a seat already taken under μ_N^k, $k = 0, 1$: We take a student i_j and suppose $\mu_N^k(i_j) = s_j$ and $\mu_S^k(i_j) = s_l$ where $s_jP_{i_j}^k s_l$, this implies that $s_jP_i^k s_l$ for all $i \in N^k$ because the preferences are homogeneous, but if i_j gets s_l under μ_S^k then i_l where $\mu_N^k(i_l) = s_l$ has to change school under μ_S^k, that is $\mu_S^k(i_l) = s_j$. Thus, $\mu_S^k(i_l)P_{i_l}^k\mu_N^k(i_l)$, which contradicts $\mu_N^k(i)P_{N^k}^k\mu_S^k(i)$ for all $i \in N^k$. Therefore, μ_S^k is not worse than μ_N^k.

2. This student takes an empty seat under μ_N^k, with $k = 0, 1$: We take any student $i \in N^k$ and suppose $\mu_N^k(i) = s$ and $\mu_S^k(i) = s'$ where s' is an empty seat, this is impossible because the school seats assigned by any stable matching are the same (Roth, 1984).

Therefore, $(\mu_N^0, \mu_N^1) = (\mu_S^0, \mu_S^1)$. ■

With this proposition we find a case where no matters who is making the proposes, the students or the schools, the assignment made by both matchings is the same. This is not the only case where this happens, that is, the condition is sufficient but not necessary.

Example 13 (Unique matching)**.** *The sets of students is $N^0 = N = \{i_1, i_2, i_3, i_4, i_5\}$ and of schools is $S^0 = S = \{s_1, s_2, s_3, s_4, s_5, s_6\}$, and let $q_s = 1$ for all $s \in S$. Also suppose the students preferences are homogeneous $P_i \sim P_i'$ for all $i, i' \in N$. Let the profile $P^0 = P$ of preferences and the profile $\succ^0 = \succ$ of priorities be as follows:*

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}	P_{i_5}
s_3	s_3	s_3	s_3	s_3
s_1	s_1	s_1	s_1	s_1
i_1	i_2	i_3	i_4	i_5
s_2	s_2	s_2	s_2	s_2
s_4	s_4	s_4	s_4	s_4
s_5	s_5	s_5	s_5	s_5
s_6	s_6	s_6	s_6	s_6

$\succ_{s_1}$	$\succ_{s_2}$	$\succ_{s_3}$	$\succ_{s_4}$	$\succ_{s_5}$	$\succ_{s_6}$
i_2	i_1	i_5	i_3	i_4	i_2
i_1	i_3	i_2	i_4	i_3	i_1
i_5	i_4	i_1	i_2	i_1	i_5
i_4	i_5	i_4	i_1	i_5	i_3
i_3	i_2	i_3	i_5	i_2	i_4

We calculate first the market extension algorithm when the students propose. And the matching found in the first round is: $\mu_N^0(i_2) = s_1$, $\mu_N^0(i_5) = s_3$.
Then the extended market is $M_1 = (N^1, S^1, P^1, \succ^1, q^1)$, *with* $S^1 = \{s_2, s_4, s_5, s_6\}$, $N^1 = \{i_1, i_3, i_4\}$, $q^1 = (1,1,1,1)$, *and* $(P^1, \succ^1)$ *as follows:*

$P^1_{i_1}$	$P^1_{i_3}$	$P^1_{i_4}$
s_2	s_2	s_2
s_4	s_4	s_4
s_5	s_5	s_5
s_6	s_6	s_6
i_1	i_3	i_4

$\succ^1_{s_2}$	$\succ^1_{s_4}$	$\succ^1_{s_5}$	$\succ^1_{s_6}$
i_1	i_3	i_4	i_1
i_3	i_4	i_3	i_3
i_4	i_1	i_1	i_4

After the extension we found $\mu_N^1(i_1) = s_2$, $\mu_N^1(i_3) = s_4$, $\mu^1(i_4) = s_5$, *and the schools* s_6 *does not have assignment.*
On the other hand, we run the market extension algorithm when the schools make the proposes. With the original problem the schools s_1 *and* s_6 *propose to* i_2, *school* s_2 *proposes to* i_1, *school* s_3 *to* i_5, *school* s_4 *to* i_3 *and school* s_5 *to* i_4, *but only the schools* s_1 *and* s_3 *are accepted because they are the only ones that are acceptable for all the students, then we have the assignment:* $\mu_S^0(i_2) = s_1$, $\mu_S^0(i_5) = s_3$. *Therefore, the market extension is the same that in the previous case where the students propose* M_1. *Then, in the second round the schools* s_2 *and* s_6 *propose to* i_1, *school* s_4 *proposes to* i_3 *and school* s_5 *to* i_4, *in this case the assignment is:* $\mu_S^1(i_1) = s_2$, $\mu_S^1(i_3) = s_4$ *and* $\mu_S^1(i_4) = s_5$.
Therefore, we have the same assignment under the market extension algorithm when the students propose than when the schools propose:

$$(\mu_N^0, \mu_N^1)(i_1) = (\mu_S^0, \mu_S^1)(i_1) = s_2$$

$$(\mu_N^0, \mu_N^1)(i_2) = (\mu_S^0, \mu_S^1)(i_2) = s_1$$

$$(\mu_N^0, \mu_N^1)(i_3) = (\mu_S^0, \mu_S^1)(i_3) = s_4$$

$$(\mu_N^0, \mu_N^1)(i_4) = (\mu_S^0, \mu_S^1)(i_4) = s_5$$

$$(\mu_N^0, \mu_N^1)(i_5) = (\mu_S^0, \mu_S^1)(i_5) = s_3$$

This means that there is only one matching. □

Now, we consider the case of two different extensions, and we notice that we can not ensure that the assignments are equals, this will happen only when the original preferences and the extended preferences are homogeneous for all students and the second extension matches with the students' order of the unaccepted schools at least until all students have been assigned. This happens because at the time of making the unacceptable schools acceptable without knowing the real order that the students give them, the real order can be changed and cause the student to accept a school that is less desired by him.

Observation 3. *If the preferences of all students are homogeneous, $P_i^0 \sim P_{i'}^0$, and we make two different extensions, where all students extend their preferences in an homogeneous way, $P_i^1 \sim P_{i'}^1$, and $P_i^{1'} \sim P_{i'}^{1'}$ for all $i, i' \in N$. Then $(\mu_S^0, \mu_S^1) \neq (\mu_S'^{0'}, \mu_S'^{1'})$.*

We present an example to clarify the observation:

Example 14. *Let's continue with the Example 6, we obtained using the market extension algorithm when the schools propose the following assignments: $\mu_S^0(i_2) = s_1$, $\mu_S^0(i_5) = s_3$. Then, with the extended market $M_1 = (N^1, S^1, P^1, \succ^1, q^1)$, we found $\mu_S^1(i_1) = s_2$, $\mu_S^1(i_3) = s_4$ and $\mu_S^1(i_4) = s_5$, and the seat in school s_6 stays empty.*
Now, we find with the market extension algorithm when the schools propose another matching using another extension, different from the used in Example 6.
In round 1 with the original preferences and priorities, the assignment is the same: $\mu_S'^0(i_2) = s_1$, $\mu_S'^0(i_5) = s_3$, but now we extend the market in a different way, $M_1' = (N^1, S^1, P^{1'}, \succ^1, q^1)$, with $S^1 = \{s_2, s_4, s_5, s_6\}$, $N^1 = \{i_1, i_3, i_4\}$, $q^1 = (1,1,1,1)$, and $(P^{1'}, \succ^1)$ as follows:

$P_{i_1}^{1'}$	$P_{i_3}^{1'}$	$P_{i_4}^{1'}$
s_2	s_2	s_2
i_1	i_3	i_4
s_4	s_4	s_4
s_5	s_5	s_5
s_6	s_6	s_6

$\succ_{s_2}^1$	$\succ_{s_4}^1$	$\succ_{s_5}^1$	$\succ_{s_6}^1$
i_1	i_3	i_4	i_1
i_3	i_4	i_3	i_3
i_4	i_1	i_1	i_4

That is, the students only extend in one school their acceptable schools, then the extension is made total randomly, for example:

$P^{1'}_{i_1}$	$P^{1'}_{i_3}$	$P^{1'}_{i_4}$
s_2	s_2	s_2
s_6	s_6	s_6
s_4	s_4	s_4
s_5	s_5	s_5
i_1	i_3	i_4

Then the schools s_2 and s_6 propose to student i_1, school s_4 proposes to i_3 and school s_5 to i_4. Then the assignment after this extension is $\mu'^{1'}_S(i_1) = s_2$, $\mu'^{1'}_S(i_3) = s_6$ and $\mu'^{1'}_S(i_4) = s_4$, the school s_5 stays with the seat free. Notice this assignment is different than in the Example 6 with the fist extension. □

3.6 Conclusions

We solve a school choice problem where the students tend to underrepresent their preferences. Then, we propose two solutions for this problem where usually schools are not assigned to all students. First, we propose an extension to the classic school choice market and new concepts of matching, r-stable matching and r-optimal matching. Theorem 6 shows that in a extended market always there is at least one r-stable matching. Theorem 7 proves there are r-optimal matchings for both sides of the market and that they could be found when one side of the market makes the proposals in both steps of the market extension algorithm. That is, in the market of students and schools, it always exists an r-optimal matching for the students when they propose in the algorithm and one r-optimal matching for the schools when they make the proposals. Theorem 8 proves that the r-optimal matching for the students is the most preferred by them and the r-optimal matching for schools is the least preferred by the students. This leads to find, in future research, a lattice structure in r-stable matchings.

Second, we prove in Theorem 9 that the dominant strategy for the student i is to state since the beginning as acceptable all the schools, because the assigned school obtained with this strategy when the remaining students state any preference profile is at least as good as any other assignment gotten with any other strategy.

Also, we show two examples about manipulation that lead us to an understanding of the strategies that students could use in different cases. In the fist example, we state that the students have incentives to present their real list of preferred schools and not the extended list from the beginning. And, in second example, we see that the students have incentives to misrepresent their preferences in order to get a most preferred school in the first step of the algorithm instead of a least preferred school in the extension.

Finally, through analysis of particular cases, we state Proposition 3, where we find sufficient conditions to the r-optimal for both sides of the market are equal; and Observation 3 states that if the students preferences are homogeneous then the matchings gotten by different extensions are not equal.

3.7 Appendix

If we interpret the school priorities like preferences, we apply the following version of the Gale-Shapley deferred acceptance algorithm to the school choice market:

- *Step 1*: Each student proposes to her first choice. Each school tentatively assigns its seats to its proposers one at a time following their priority order. Any remaining proposers are rejected.

- *Step k*: Each student who was rejected in the previous step proposes to her next choice. Each school considers the students it has been holding together with its new proposers and tentatively assigns its seats to these students one at a time following their priority order. Any remaining proposers are rejected.
  The algorithm terminates when no student proposal is rejected and each student is assigned her final tentative assignment.

References

[1] Atila Abdulkadiroğlu and Tayfun Sönmez. *Matching Markets: Theory and Practice.* Advances in Economics and Econometrics Theory and Applications. Tenth World Congress, Vol 2, 2013; 3-47.

[2] Atila Abdulkadiroğlu and Tayfun Sönmez. *School Choice: A Mechanism Design Aproach.* The American Economic Review. Vol 93, No. 3, June 2003; 729-747.

[3] Ahmet Alkan. *Nonexistence of Stable Threesome Matchings.* Mathematical Social Sciences, 16, 1986; 207-209.

[4] Itai Ashlagi and Flip Klijn. *Manipulability in Matching Markets: Conflict and Coincidence of Interests.* Social Choice and Welfare, Vol 39, 2012; 23-33.

[5] Caterina Calsamiglia, Guillaume Haeringer and Flip Klijn. *Constrained School Choice: An Experimental Study.* American Economic Review. Vol 100, September 2010; 1860-1874.

[6] Peter Coles and Ran Shorrer. *Optimal Truncation in Matching Markets.* Games and Economic Behavior. Vol 87, September 2014; 591-615.

[7] Jan Eeckhout. *On the Uniqueness of Stable Marriage Matchings.* Economics Letters 69, University of Pennsylvania, 2000; 1-8.

[8] Lars Ehlers. *Truncation Strategies in Matching Markets.* Mathematics of Operations Research, Vol 33, No. 2, 2008; 327-335.

[9] Aytek Erdil and Haluk Ergin. *What's the Matter with Tie-Breaking? Improving Efficiency in School Choice.* American Economic Review, 98, Vol. 3, 2008; 669-689.

[10] D. Gale y L.S. Shapley. *College Admissions and the Stability of Marriage.* The American Mathematical Monthly, Vol 69, No.1, 1962; 9-15.

[11] Dan Gusfield. *The Structure of the Stable Roommate Problem: Efficient Representation and Enumeration of All Stable Assignments.* SIAM Journal on Computing, 17, 1988; 742-769.

[12] Dan Gusfield and Robert W. Irving. *The Stable Marriage Problem: Structure and Algorithms*. Cambridge: MIT Press, 1989.

[13] Fuhito Kojima and Parag A. Pathak. *Incentives and Stability in Large Two-Sided Matching Markets*. American Economic Review. Vol 99, June 2009; 608-627.

[14] Donald E. Knuth. *Mariages Stables*. Montreal:Les Presses de l'Université de Montreal, 1976.

[15] Susan Mongell and Alvin E. Roth. *Sorority Rush as a Two-Sided Matching Mechanism.* The American Economic Review, 81, No.3, 1991; 441-464.

[16] D.G. McVitie and L.B. Wilson. *Stable Marriage Assignments for Unequal Sets*. BIT, 10, 1970; 295-309.

[17] Antonio Romero-Medina. *implementation of Stable Solutions in a Restricted Matching Market*. Review of Economic Design, Vol 3, 1998; 137-147.

[18] Alvin E. Roth. *The Economics of Matching: Stability and Incentives*. Mathematics of Operations Research, Vol 7, No.4, 1982; 617-628.

[19] Alvin E. Roth. *The Evolution of the Labor Market for Medical Interns and Residentes: A Case Study in Game Theory*. Journal of Political Economy, 92, 1984; 991-1016.

[20] Alvin E. Roth and John H. Vande Vate. *Incentives in Two-Sided Matching with Random Stable Matchings*. Economic Theory, Vol 1, No. 1, 1991; 31-44.

[21] Alvin E. Roth and Marilda A. Oliveira Sotomayor. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. New York, Cambridge University Press, 1990.

Chapter 4

Does the Home-School Distance Impact High School Achievement? An Analysis for Mexico City

4.1 Introduction

In Mexico, 6.2 percent of the Gross Domestic Product (GDP) is assigned to educational institutions of any level (Organization for Economic Co-operation and Development, 2014), this proportion is a little bit more than the mean of all countries in the OECD. Nevertheless, in standardized tests, as the Programme for International Student Assessment (PISA), Mexico always gets results far below the mean. In 2015, Mexico was ranked in position 56 (average between the three areas evaluated by PISA, science, mathematics, and reading) of 70 countries that participated in the evaluation. This result can be explained by the impact of some factors, such as the investment in education per student, the proportion of teachers per student, the training of professors. Another way to address these problems in education is analyzing if the students are taking advantage of the existing resources.

A way to study the effect of the available resources over the student outcomes is analyzing the schools where they are enrolled. In some states of Mexico, like in many other educational systems in the world, the mechanism used to assign the places in the educational institutions is centralized, that is, a government institution assigns the free places in every public educational institution no matter their kind (technical, charter, federal, state, etc.). Since 1996, in Mexico City, the organism in charge of this mechanism for the middle school is the Metropolitan Commission of Public Institutions of High School Education (COMIPEMS)[1].

COMIPEMS implements the assignment taking into consideration three factors: (1) the applicants preferences for schools, where students make a list of schools from highest to lowest in order of preference (at least one and at most twenty); (2) the score obtained in the standardized admission test; and (3) the capacity of every school, that is, the number of available seats.

Despite the list of preferred schools given by the applicants, nothing ensures that they get a seat in their first choice school. This is because the mechanism takes into consideration the priority that they have in every school according to their score in the admission exam. That is why mechanisms, like the one used by COMIPEMS (developed in the market of school choice), have become a debated topic among researchers in the field. Some researchers consider that the applicants could benefit from attending a school which is their first choice because this situation could maximize the commitment that they feel about their academic labor, and this could play a very important role in their academic outcomes. Although other researchers consider this does not have an effect over the scores gotten by the students in standardized exams, or in their enroll and graduation of universities.

Recently, many papers are using COMIPEMS data, in some of those papers the authors seek to answer different questions that could be analyzed with this database. They find family networks are important when the applicants have to decide where to go for studying high school (Dustan, 2017a), as well as for the effect in stratification of the increasing demand of elite schools (Estrada, 2016), or the effect of transit improvement on school choice in Mexico City (Colin ,2015), but these authors do not talk about changes in student outcomes or preferences arising by the distance between their home and the school or how these are related. It is precisely this effect that we investigate in this paper.

[1]In Spanish, Comisión Metropolitana de Instituciones Públicas de Educación Medio Superior.

In order to do so, we use the data from the high school student assignment process in Mexico to estimate the effect of enrollment in a school closer to home on the probability to graduate from high school; and the effect of assignment in a nearby school on the probability to enroll in it. Also, we want to know if, when the applicants have to submit their ranked preferred schools, they consider the factor home-school distance to make the decision. Specifically, we answer three research questions: The first question asks if the home-school distance has some causal effect on the academic outcomes of high school students, in particular, we want to know if the time spent to arrive at school directly affects the probability to graduate from high school. The second question asks if the home-school distance has effect on the enrollment decision, that is, whether the home-school travel time affects the probability of enrolling at the assigned school. For the third question, we go back to the start of the students' high school careers, and ask ourselves if the home-school distance has an effect on the preferred schools, that is, whether the applicants consider the time that they will spend in public transportation at the moment of selecting their twenty preferred schools when they apply to COMIPEMS admission test.

For the first two questions, we merge data from the applicants registered for COMIPEMS admission test with data from the schools where the applicants enroll. A logit regression can then we used to answer the first and second questions. The third question uses only information known at the time of the COMIPEMS test, and therefore an instrumental variable approach accounting for the simultaneity and endogeneity of decisions is used to answer the second research question.

We analyze these three questions in parallel to have a broad picture of how distance is affecting these three levels (preferences, enrollment and graduation), that is, we want to compare for example if the student is considering the distance at the time of choosing and if this will be really important for your educational performance. If distance is relevant to graduate from high school and to enroll in it then it should be an important variable when choosing which school you want to attend.

Section 2 continues with a more extensive review of the empirical literature regarding school choice, school quality and graduation probabilities. The empirical strategy used to respond the three questions is outlined in more detail in Section 3. The COMIPEMS admission process and the data derived from their records and from the high schools where the students were enrolled is summarized in Section 4. The descriptive statistics and the econometric results are presented in Section 5, followed by a section with the conclusions obtained from the analysis.

4.2 Literature Review

This section reviews the literature regarding the impact of school choice - in particular, the quality of the school a student has been admitted to on the outcome (graduation from high school), and factors driving the latter. In particular, we review the evidence regarding the relevance of the home-school distance for the outcome, and how it connects to the initial decision-making processes and school preferences.

Hastings et al. (2012) find that the opportunity of admission to a high quality school increase the students' motivation even before enrolling and that this has positive effects on test scores and attendance. Abdulkadiroğlu et al. (2014a) find that the students who have been moved from a state school to a charter school with best academic level are more motivated and show higher performance than those who were assign by lotteries. Also, Hastings and Weinstein (2008) find that if the parents have access to information in a simpler way the probability that they move their children to a more competitive school raises, especially if the school is close to home, and this increases the scores reached by their children.

Nevertheless, other studies have showed negative or nonsignificant effects on the students' outcomes when they attend to a more competitive school. This is the case of Ortega (2015), who uses a quasi experimental design and instrumental variables to estimate the causal impact of the supply of and the enrollment in a more competitive school in Mexico City, finding that the students who are barely offered admission to a more competitive option have lower probability of graduating relative to applicants who scored just below the admission cut-off (and therefore went to a lower quality school). Also in Mexico City, Dustan et al.

(2017b) find admission to an elite school increases the probability of high school drop out by 9.4 percentage points. The additional risk of drop out is a result of the higher academic level and greater opportunity costs of attendance. These results agree with those found by Abdulkadiroğlu et al. (2014b) where the applicants who attended a selective school in New York and Boston get little gains, even for those who are close to the threshold. Also, Dobbie and Fryer (2014) find that attending an elite school has little effect on the reading, writing and mathematics exams, and can have negative effects in college enrollment as well as the college graduation. Nevertheless, it can produce positive effects in the long run thanks to the social contacts created during high school.

The results obtained in this kind of analysis could depend on different factors, e.g. the importance that the parents give to the educative quality when they have to choose the institution where their children will attend, or on the socioeconomic characteristics. In Hastings et al. (2006) the effect of enrolling in the most preferred school increases when the importance that parents give to the school quality increases, this impacts mostly white and high-income students. When the parents do not give so much importance to the quality, the utility increases when this depends of other factors like the proximity to school or same racial mates. Burgess et al. (2014) find that families with high socioeconomic status have access to a higher quality school than those with low socioeconomic status. Also, in Chile, Elacqua(2012) finds that public schools are more likely to serve student from low socioeconomic status than the private ones. The disadvantaged students are more segregated among private voucher schools than those that attend a public school. In Burgess and Briggs (2010), the authors find that the poor kids do not have high possibilities of attending a high quality school, this may be due to several factors, one of them is the home-school distance, this is because one of the priorities to be accepted in an English school is the proximity to it.

We can observe that the researchers have found that the effects of attending a high-quality school may be due to different reasons like the motivation and interest, the school mates (peer effects), the school resources, the distance, etc. Some analyses find that the distance have some effect on the outcomes, that is the case of Cullen et al. (2006). They try to explain, without success, why they find little impact on the scholarly achievements (exam scores, the conclusion of school, the attendance, etc.) when the student goes to a first choice school. Özek (2009) analyzes the effect of inter and intra-district school choice in Pinellas District (Florida), finding negative effects in the student outcomes. The students that change their assigned school and go out of the district show worse performance in standardized exams than those that stay in their district school. Other studies have appeared in the literature talking about the effect of distance on the decisions when the applicants rank the school options. For example, Burgess et al. (2014) argue that the parents consider many factors before choosing a school for their children, one of the most important is the home-school distance. In fact they talk about a trade-off between distance and quality, that is, the parents could sacrifice the quality of the school in favor of proximity.

An important example of how the distance affects is a natural experiment in Stockholm, when a reform was implemented in the city. In 2000, the government changed the way to assign seats in high schools. Before the reform, the applicants who wanted to get a place in high school were assigned to the nearest school, but from the autumn of 2000 onward the admission depends on the scores obtained at the previous level. Those who proposed the reform argue that this change increases the efficiency and help the schools to be responsive to the parents preferences. On the other hand, the opponents argue that this reform increases segregation. Söderström and Uusitalo (2010) analyze the segregation level generated after the reform, finding that the segregation actually increases. Andersson et al. (2012) analyze the same reform in Sweden, but they study how much the distance increases by fifteen years old students to arrive to school, finding that the distance increases except for foreigners, students in social programs and minorities, unless they have families with high educational level, concluding that the free school choice is decreasing equality.

In 2008 a similar reform took place in Linz (Austria), the assignment mechanism changed from district allocation to free school choice, that means every student can apply to any school that they prefer. Altrichter et al. (2011) analyze this reform and find an increase in segregation, but the sample is too small to find significant results. One case of segregation in Mexico is treated in Estrada (2016), he shows that the stratification in Mexico City is by abilities not by family income, because the admission to high school is by exam instead of proximity to school. The stratification depends on how strong the correlation between family income and ability is, and on the relation between family income and demand. The two natural experiments in Stockholm and Linz and the study of Estrada show changes in segregation, but the authors do not talk about changes in student outcomes (graduation and enrollment)or preferences arising by distance traveled, that is, the effect that we will analyze in this paper.

Recently some articles about Mexico City have been written, considering the distance issue. Some studies treat the distance as a factor to consider by the parents when they have to enlist the school preferences. For example, Dustan and Ngo (2018) find that the demand of elite schools raises when the transit accessibility increases for high-achieving students with highly-educated parents. Dustan (2017a) finds that the applicants in Mexico City prefer schools where the older siblings had attended, irrespective of the proximity to the school. The probability of choose the sibling high school or some other similar is large and positive, even when the siblings are too far apart in age to attend school together. Another way to analyze distance is presented in Colin (2015), who studies the effect of new public transportation (metrobús) in Mexico City, he finds that one out of every five students is assigned to an under-matched school. This occurs generally to low-income students, while the geographic areas where it is more likely to happen are the areas far away from the city center. The advances in public transportation help to less students are under-matched (when a student enrolls in a school below her school qualifications), because the applicants choose schools closer to public transportation and make a trade-off between quality and distance.

Hence, the literature review suggests that home-school distance may affect achievements of students while enrolled, but also that in the earlier stage of school choice, distance matters. It is well known that the distances in Mexico City and the Metropolitan area are huge, and therefore it is relevant to analyze the impact of the home-school distance on schooling achievements in Mexico City. As stated before, our principal concern is if the distance to the school and the time used to arrive have some effect on the academic outcomes gotten by the students. Another concern could be ask us if the applicants are considering the distance factor (or the time to travel this distance) at the moment of choosing their most preferred schools when they apply for admission.

4.3 Empirical Strategy

To answer the first question, whether the home-school travel time has an effect on the academic outcomes, we take as outcome the high school graduation. Let Y_{ij} denote a dichotomous variable which takes a value equal to 1 if student i graduates from school j and 0 if not, where school j is the school that she attends. We base our analysis on Hastings et al. (2006), but unlike them - they use a linear regression because their independent variable is continuous - we propose a logit regression to explain the effect of the travel time because it accounts for the binary nature of the dependent variable (probability of graduation) and explain the probability of each of the possible outcomes.

$$P(Y_{ij} = 1) = \alpha_0 + \beta_0 X_i + \gamma_0 Z_i + \theta_0 T_{ij} + \epsilon_{ij}$$

Where Y_{ij} is the variable of graduation, X_i are socioeconomic variables, Z_i individual characteristics, and T_{ij} is the minimum home-school time of travel. Our socioeconomic and individual controls include: father's age, father's education, older siblings, number of siblings, female, GPA in middle school, study hours, number of times that she has taken the exam, score at admission test, ranked assigned option and number of options listed.

The graduation could be affected, not only by personals or socioeconomic background, but also by school characteristics. We take the minimum score that is required to attend to the school as a control of school characteristics. Then, we consider this variable in consideration and propose the following logit regression:

$$P(Y_{ij} = 1) = \alpha_1 + \beta_1 X_i + \gamma_1 Z_i + \theta_1 T_{ij} + \delta_1 S_j + \epsilon_{ij}$$

Where the same variables are used as above but now adding S_j, the minimum score required for attending to school j.

For the second question, we want to analyze if the home-school travel time has an effect on the enrollment. As in the previous case, we propose a logit regression to explain the effect of the travel time on the probability of enrollment. Let Y_{ij} denote a dichotomous variable which takes a value equal to 1 if student i enrolls at school j and 0 if not, where school j is the assigned to the student i.

$$P(Y_{ij}=1)=\alpha_0+\beta_0X_i+\gamma_0Z_i+\theta_0T_{ij}+\epsilon_{ij}$$

Where Y_{ij} is the variable of enrollment, and X_i, Z_i and T_{ij} are the same variables as in the previous model. The enrollment, in the same way as graduation, could be affected by school characteristics, then we take the minimum score that is required to attend to the school S_j as a control of school characteristics. Hence, we propose the following logit regression:

$$P(Y_{ij}=1)=\alpha_1+\beta_1X_i+\gamma_1Z_i+\theta_1T_{ij}+\delta_1S_j+\epsilon_{ij}$$

For the third research question, whether the home-school travel time has an effect on the preferred schools, we analyze the relevance of the socioeconomic and personal variables for the school choice, while being primarily interested in the impact of the potential travel time from home to school on this preference:

$$Y_{ij}=\alpha_2+\beta_2X_i+\gamma_2Z_i+\theta_2T_{ij}+\epsilon_{ij}$$

Where Y_{ij} an indicator of the school preference, X_i and Z_i are socioeconomic and individual characteristics respectively, and T_{ij} is a measure of the distance from home to the preferred school.

For the indicator of the school preference, Y_{ij}, we have several options to summarize the list of maximum 20 schools registered when applying into one single measure. We opt for four indicators of the school quality the student prefers: the minimum score to be accepted at the first, second and third schools in the student's list of preferences, and the average of these three scores. The measure of the home-school distance, T_{ij}, is calculated using the school or schools that are used when constructing the school indicators, hence, we consider the average of the three first options as a relevant indicator of the quality sought for by the student because on average the students are assigned to their 3rd or 4th school of preference.

It is straightforward to see that the proposed equation suffers a problem of bidirectional causality, because the school(s) that are used to construct the quality indicator Y_{ij} and the distance measure T_{ij} are the same school(s), and moreover, are determined at the same moment in time (when the list of 20 preferred schools is submitted to COMIPEMS). In other words, quality preferences (Y_{ij}) and travel distance preferences (T_{ij}) are jointly determined. We specify the equation as we do, with quality as the main outcome, because ideally school quality is what should matter most in the decision process - as it determines future opportunities for the students. Nonetheless, running a regression without accounting for the joint decision-making process will only give us biased estimates of the effect of preferences regarding distance on the preferred school quality. We resolve the issues by following an instrumental variable approach.

We look for some variable that causes exogenous variation in travel time and does not have an effect on score, but there are no variables in our data base that convincingly satisfy these requirements, either because they are determined at the same time as the score of the preferred school as the case of the travel time, or because it directly determines the score of the school. As a more viable alternative, we propose to construct an instrumental variable, the average travel time of all students that put the school j in their top three options, A_j. This variable tells us something about the time student i is expected to travel from her home to school, but it does not have effect on the minimum score required by the school that student i prefers, that is the average time of all the applicants to this school gives us a measure of the school location and this help us to identify the time that the student i will do in case that she is enrolled in that school, but it is not a variable to determine the score required to be accepted on the school (the quality measure of the school). Then, our first and second stages are determined by:
First stage:

$$T_{ij} = \alpha_3 + \beta_3 X_i + \gamma_3 Z_i + \theta_3 A_j + \epsilon_{ij}$$

Second stage:

$$Y_{ij} = \alpha_2 + \beta_2 X_i + \gamma_2 Z_i + \theta_2 T_{ij} + \epsilon_{ij}$$

4.4 Data

As noted in the introduction, our primary source of information is formed by the registry of the Metropolitan Commission of Public Institutions of High School Education (COMIPEMS). In this Section, we first address the underlying admission process (5.1) followed by a brief introduction of the information available to us (5.2). We finish with some more details on the construction of the relevant variables for the empirical model, in particular, the quality and distance indicators (5.3).

4.4.1 COMIPEMS Admission Process

The Mexico City metropolitan area, which comprises the Federal District and 22 municipalities from the neighboring State of Mexico, has the largest education market in the country. Every year close to $300,000$ students aged 15 years (in large majority) apply to public high schools through a centralized process, this process is a merit-based student-assignment mechanism.

In February 1996, nine subsystems that offer programs of public high school in Mexico City metropolitan area signed a collaboration agreement, and they called a contest of assignment that significantly modified the traditionally followed procedures (COMIPEMS, 2012). The substance of the agreement settled the realization of a single registry of all high school aspirants and the evaluation of their abilities and knowledge by a single standardized test, known as the High School Education Entrance Competition (CIEMS)[2].

Before that, the teenagers who wanted to enter at high school level had to apply to different academic institutions simultaneously and, then choose the most-preferred among those who had accepted. This produced inefficiency in the process, some schools were left with vacant seats and many applicants were not admitted to any of the schools to which they applied. Also, the lower-income applicants were at disadvantage, because they had to put aside many schools except the most-preferred since they had fewer financial resources to cover the costs of applying to many different schools. Through CIEMS, the applicants are assigned to public schools based on an algorithm that uses the score of the admission test, a hierarchical list of at most twenty preferred schools, the capacity of the schools and the school's minimum score.

The steps of the student-assignment process are as follows: (1) the announcement is

[2]In Spanish, Concurso de Ingreso a la Educación Media Superior, since 2014 is called Concurso de Asignación a la Educación Media Superior, CAEMS (High School Education Assignment Competition).

published, late January every year, in the local newspapers and the COMIPEMS and SEP web page[3]; (2) then the students in final (third) year of middle school must pre-register on the COMIPEMS web page before the end of February, in this step they receive material explaining the process; (3) by the beginning of March, the registration must be complete, included the hierarchical list of preferred schools and the required documents, applicants receive a study guide for the test and an identification voucher necessary for taking the test; (4) at the end of June the students take a comprehensive achievement test with 128 multiple-choice questions, the exam covers verbal and mathematical skills and knowledge in the fields of history, geography, mathematics, physics, chemistry, literature, biology and ethics.

Once the exams are graded by a computer, the system generate information that is incorporated to a database and to a computational program that process it. After the database is integrated, the system dismiss the applicants who get less than 31 points[4], those who did not graduate from middle school or did not attend the exam. After that, the assignment process starts. And, the results are published at the end of July, after the computer algorithm has done the assignment.

In the first round of the assignment process, each of the nine subsystems sets a maximum capacity at each high school, the applicants are ranked by the number of points gotten in the entrance examination (from highest to lowest). In that order, each applicant is assigned a place in the educational option of their highest preference that has available places. This is, the applicant who has the highest number of points is assigned to his first choice, and the second highest gets his most preferred school among the schools with available seats, and so on. If several applicants have the same score and compete for the last seat in one school, a school subsystem representative has to choose if every applicant in this situation is admitted or all of them are rejected. The process ends when every applicant is assigned except those who got a score below the minimum request for his options.

In the second round of the assignment process, if there are applicants who got the ad-

[3]SEP is the acronym for Secretaría de Educación Pública, in English Ministry of Public Education.
[4]This restriction was removed in 2013, after our study period.

mission requirements but have not been assigned to one of their options, they have the opportunity of select some school with free seats. This round takes several days and the selection is made by the applicants in the order assigned by their score in the entrance exam, with students with higher scores having a priority on selection. Finally and importantly, applicants are only allowed to enroll at their assigned school where they must complete the paperwork in order to formally enroll the school.

4.4.2 Data Sets

The database used to answer the research questions is a merge between data of the registered applicants who apply for the admission exam to COMIPEMS (CIEMS) and data from the schools where the applicants enroll.

The main data is extracted from COMIPEMS files. We consider the applicants that registered for the admission exam between 2005 and 2009. We take these years because we have school data where the applicants enroll until 2012, and we can know if they are already graduated from high school or not at this time.

The database contains information regarding the applicants' academic background, score in the admission exam, the ranked preferred schools, number of options listed, the assigned school, number of times that she has taken the exam, and rank of assigned option. In addition, the database includes information about the high schools, such as modality (technical high school, general high school and technical professional education), address and minimum score required for admission.

Moreover, the database contains the responses to a questionnaire that the student filled in before taking the exam (when submitting the papers), the student fills a background questionnaire including information on gender, date of birth, zip code, middle school attended and grade point average. Also, the survey provides information about the family structure (siblings), parental education and age, family income proxies, indigenous origin and other socioeconomic characteristics; and personal characteristics of the student, such as study habits, recreational activities, abilities, academic aspirations, etc.

Also, we added data obtained from the administrative records from different high school subsystems, for the 2008-2012 years when the applicants from 2005-2009 should be graduated. Those subsystems have student records about enrollment and graduation. This database is merged with the COMIPEMS database using the national identity document (CURP)[5] and/or the COMIPEMS registration number.

4.4.3 Constructed Variables

Using the information described before we create several variables that result useful for our analysis. Students report household income in the background questionnaire selecting one of 15 income brackets. We assume that each discrete income category corresponds to the upper limit, for the last bracket, which does not have a maximum limit, we make an assumption about this upper value and consider 30,000 mexican pesos as this value. In the same way we proceed with the father's age and years of education .

Students' zip codes are available in the COMIPEMS data, allowing us to geographically locate students according to their zip code centroids. With this information and with the exact address of the high school, we calculate first an approximate home-school distance and later the time, measured in hours, that the student spends traveling from home to high school, using different types of conveyances. This helps us to determinate if the student lives close or not of the schools, we consider that a student lives close of the school if the home-school time is less or equal to one hour. We choose one hour because that is the average time that the full sample do to arrive to the assigned school. This variable will be used only for an exploratory (descriptive) analysis, while for the explanatory (econometric) models the travel time will be used as a continuous variable.

Besides, we use the minimum score to be accepted at schools as a gross measure of quality school. In our database the highest minimum score to be assigned is 100 and the lowest is 31. For the descriptive analysis we say that a school is considered "elite" if it requires at least 83 points (the superior quartile) to be accepted at it. The binary closeness and elite

[5]The national identity document in Mexico is the Clave Única de Registro de Población, CURP (in English: Unique Population Registry Code)

variables are constructed with the aim of making descriptive statistics easier to explain by dividing the sample into two set of students, those who are assigned to a close school and those who are not, for the case of close and non-close; and those who are assigned to a elite school and those who are not in this case. But, for both cases, the continuous variants are used as the main explanatory variables in the econometric.

With the information of the high schools where the applicants enroll, we take the year of graduation and create the dummy variable of graduation, that takes the value of 1 if the student graduates before or at 2012, and 0 if she does not. Also, we have the enrollment variable for the students that were assigned, that takes the value of 1 if they effectively enroll and 0 if they decide not to enroll the school that was assigned to them based on the COMIPEMS exam. These variables along with the score of the first, second and third option and the average score of the top-three options are our outcome variables.

4.5 Results

Before we discuss the estimation results we first briefly discuss the descriptive statistics of dependent and explanatory variables (Section 6.1). Next, Section 6.2 discusses the results for the model relating to the effect of the home-school travel time on the high school graduation. In Section 6.3, we present the results for the enrollment model where we show the effect of the home-school travel time on the enrollment. Finally, in Section 6.4 we show the results related to the relation between the home-school travel time and the school election of the students.

4.5.1 Descriptive Statistics

Table 4.1 presents the summary of descriptive statistics for three different samples. In column 1, we report the average values of some socioeconomic and individual characteristics for all applicants in the five cohorts (2005-2009). About 78 percent of the applicants were assigned to a school, the remaining were not because they do not get the minimum score, do not present the test or do not graduate from middle school. In columns 2 and 3, we display the average values of the same variables in two restricted samples, in column 2, we summarize the variables of the students who were assigned to a school, and in column 3, we describe the variables of the students who effectively decided to enroll the assigned school.

Notice that those who enroll seem to be slightly better off than the assigned group, while the assigned group is generally closer (more similar) to the full sample. The most noticeable differences are in the variables COMIPEMS score and minimum score required at the assigned school, where we notice that the students enrolled are those who get a greater score in the admission score and they are enrolling in the demanding schools.

In column 2, we observe that 50 percent are females, the average GPA in middle school is around 8 (out of 10), the average father education in years is 9.4 (that is, in average they finish the middle school) and almost 60 percent have older siblings. The average score in the entrance test is 66, they listed about 9 school options, get their 3 or 4 listed school, and the score[6] of the assigned school is about 56.

In column 3, we see that those who enroll seem to be slightly better off than the assigned group: more educated father, fewer siblings, higher GPA, fewer indigenous. Also, the students who decide to enroll themselves in the assigned schools are those who were assigned to a more demanding school; the family monthly income is higher for these students, they tend to study a little bit more (5.2 hours per week) and aspire to reach a fourth level of studies corresponding to a bachelor degree.

In Table 4.2, we present the summary for other four different samples. In column 1 we summarize the variables of the students who were assigned to a close school, that is, the students who live an hour or less from school, and column 2 presents the variables of the students who live more than an hour from school. Column 3 reports the means and standard deviation of the same variables for the students assigned to an elite school, that is, the schools who ask for at least 83 points to accept some student, and column 4 reports these variables for students assigned to a non-elite school (which ask less than 83 points). Notice that columns 1 and 2 do not report the values for the time to the assigned school and columns 3 and 4 do not report the values for the minimum score required at the assigned

[6] We refer to the minimum score required to be accepted at school just as score.

school, this because these are values used to make the classification of students assigned to a close school and to an elite school respectively.

In column 1, we observe that 50 percent are females, the average GPA in middle school is around 8 (out of 10), the average father education in years is 9.8 (that is, a little more than finishing the middle school) and 60 percent have older siblings. Regarding their study habits and personal characteristics, we notice that they study around 5 hours a week and aspire to reach a fourth level of studies corresponding to a bachelor's degree. The average score in the entrance test is 66, they listed about 9 school options, get their 3 or 4 listed school, and the score of the assigned school is about 56.

Notice that there are no clear differences, in most of the variables shown, in averages between the samples of students assigned to schools close to or far away from their homes (columns 1 and 2). The characteristics to highlight are: the indigenous origin, where we notice that the students assigned to a faraway school have a greater possibility of having indigenous origin; the family monthly income is lower for these students, they tend to have a higher GPA in the middle school and their fathers tend to have less years of education. Also notice that the students assigned to a faraway school are very similar to the students assigned to a non-elite school (column 4) in all the personal and socioeconomic characteristics, the differences only lie in the exam outcomes (COMIPEMS score and rank of assigned option), and the students assigned to a faraway school are more dedicated, they study a little more and get a higher GPA in middle school.

Unlike columns 1 and 2, we see that the summary of statistics differs in almost all the variables between the students who were assigned to an elite school and those who were not (columns 3 and 4). Those that are assigned to an elite school tend to have higher GPA in middle school and come from a less vulnerable family (with higher monthly income, more educated father, less probability of indigenous origin). For example, the elite school sample has a greater family income and fathers with greater years of schooling than the other samples.

In the elite school sample, column 3, we notice that only one percent is indigenous, just under half are female and the family is less numerous. Talking about school habits and life projects, they study almost two hours more per week than the others and aim to achieve about one level more of study (get a master degree). Regarding their school preferences, applicants in this sample listed, on average, 9 school options, but were assigned to their first or second option, unlike the rest of the applicants who were assigned to their third or fourth option in spite of having the same number of school options, and the time spend to arrive at school is about a little less than an hour, almost the same than for the non-elite sample.

In Table 4.3, we show the descriptive statistics of outcome variables. The proportion of graduation is higher for the students assigned to elite schools than for those who did not by almost ten percent. Also, they show about 20 points more than the average in their first option or three most preferred options. Similarly, more than 90 percent of students assigned to an elite school tend to enroll. Remember that, as is explained in Section 5.3, close and elite are dichotomous variables that represent the main explanatory variables: school quality and home-school travel time.

Notice that the impact of distance on graduation seems to be small, but that, in contrast with the expectations we derived from the literature review, a school further away from home may be beneficial for graduation chance, probably related to the fact that the schools were slightly less demanding (as revealed by the required scores). Regarding enrollment, it seems that the students assigned to a nearby school tend to enroll in greater proportion than those who were assigned to a remote school. Being assigned to an Elite school increments enrollment rates to over 95%, students assigned to a Non-Elite school as much less likely to decide to go there, only 70%. Also, it seems that attending a nearby school goes with a preference for better schools, while those who attend a school further away were less ambitious in their first preferred schools. By default, aiming high correlates with ending up in an elite school.

4.5.2 Distance and Graduation Probability

In this section, we present the results for the first model, the logit model to show the effect of the home-school travel time on the high school graduation.

Table 4.4 reports the logit results for four specifications for the enrolled sample of years 2005 to 2009, notice that the sample considers only the enrolled students whose descriptive statistics are shown in column three of Table 1, the difference in the sample size is due to the missing values. All columns show the marginal effect and the standard errors in parenthesis. The first column shows the first specification which controls for family and individual variables, such as father age and education, number of siblings, if the applicant have older siblings, gender, GPA in middle school, study hours at week and number of times that she applies the admission exam. In column two, we show the second specification which adds exam outcomes the number of options listed, the rank of the assigned school, exam score and time from home to the assigned option. Column three shows specification which controls by the minimum score to be admitted in the assigned school, this exam outcome is not added in the previous specification because we want to see separately the effect of this variable on the travel time, that is, observe how change the effect of the travel time on graduation when the school is demanding. Finally, Column four adds preferences variables (score of the first option and average score of the three first options), which intends to relate the graduation with the aspirations of students before they apply the COMIPEMS test.

In the first column, we observe that father age and education have a positive and significant connection with the probability of graduation. Also, the females, the applicants with older siblings, the more dedicated students and with higher middle school GPA are more likely to graduate. Specifically, one extra point in higher middle school GPA increase 20 percentage points the likelihood of graduating. On the other hand, the number of siblings and the number of times that the applicant applies the admission exam are negative and significant.

The Column two shows a little bit stronger relation of the variables of the first specification with graduation than in the first model. And for the added variables we have that the number of options listed is negative and significant and the rank of assigned school is also negatively associated with graduation, which is intuitive since it is likely that students are more motivated when they are assigned to a more preferred school. The COMIPEMS score is negatively associated with graduation, this is counterintuitive, but the marginal effect is really low, so we conclude that, although statistically significant, in practice the impact of a one-point higher score in the exam is negligible.

In the third column, upon adding the minimum score that was requested by the school where the student got assigned to, we show that almost all the variables maintain similar estimations (positive or negative) on graduation, except for the variables related to the exam outcomes. Noteworthy is that, the COMIPEMS score has a change in sign and does not have the counterintuitive effect we found in column 2. This has much more sense intuitively than the previous cases because the score in the admission exam is a measure of the student's abilities to get good grades. Including only the COMIPEMS score (column 2) implied that students with higher scores - who therefore were assigned to the better and more demanding schools - seemed to be worse off regarding graduation probabilities - not because of their own capacities but because it made them becoming assigned to more difficult school. Accounting for the school difficulty by the inclusion of the minimum score required (column 3) shows that, given the school level, a better COMIPEMS exam predicts better chances to graduate.

The fourth model adds a couple of variables related to the preferences of the students over the schools, the score of the first option and the average of the first three options. We observe that the relation of these two variables with the graduation - given the applicant's ability (COMIPEMS score) and the school quality (minimum required) - is ambivalent. On the one hand, the first option seems to indicate that when the students aspire a more demanding school as their most-preferred one the probability of graduation is greater, but when we consider the average of the three first schools that are enlisted the effect turns into negative, that is the likelihood of graduating is greater when the students choose less competitive schools. It could be happening because applicants who rank more demanding schools have a greater probability of being assigned at one of them and these schools have a negative effect on graduation or because they were assigned to a less-preferred school and the motivation decrease.

Henceforth, the home-school travel time is a continuous variable measured in hours, unlike the close variable which is the dichotomous variant. Notice this variable and the score of the assigned option have a negative and significant association with graduation. That is the likelihood of graduating decreases when the student has to travel more time. In the same way, this likelihood decreases when the school when she is admitted is more selective (a school with a rather high minimum required score). Also, notice that the relation of graduation with the home-school travel time is reduced by adding the score of the school as a control and returns almost to its same level when we add the scores of the three first chosen schools. This negative relation is intuitive because usually, the students stop spending this time studying and use it on transportation. It is a small effect but not negligible, as time is measured in hours, one hour extra reduces the graduation probability by 0.64 percentage points. This coincides with the findings of Özek (2009) who shows the students have worse performance in standardized exams when they go to another school out of their district. On the other hand, Anderson et al. (2012) find that when the students have the opportunity of travel more to attend to a quality school they will do, Colin (2015) coincides with them, the students are willing to travel more in order to attend a quality school, but he does not analyze if this is beneficial for them. The results showed here find that this behavior could be counterproductive for them because even when the effect is small, it has some influence over the graduation of high school.

4.5.3 Distance and Enrollment

The rather small effect of travel time on graduation from the assigned school shown in the previous section are conditional on enrollment to the assigned school. In this section, we present the results for the second model, the logit model to show the effect of the home-school travel time on the high school enrollment.

In Table 4.5, we present the results of the logit regression for four specifications for the sample of applicants that gets a seat in some high school for 2005 to 2009. The sample considers only the assigned students whose descriptive statistics are shown in column two of Table 1, the difference in the sample size is due to the missing values. All columns show the marginal effect and the standard errors in parenthesis. As in the previous section, the first column shows the specification which controls for the same family and individual variables. In column two, we show the specification which adds exam outcomes the number of options listed, the rank of the assigned school, the COMIPEMS score and the travel time from home to the assigned school. Column three shows specification which controls by the minimum score required at the assigned school, and Column four adds preferences variables, which intends to relate the enrollment with the aspirations of students before they apply the COMIPEMS test.

In the first column, we observe that father's age and father's education have a positive and significant connection with the probability of enrollment, this effect is a little bit stronger that the effect on the graduation. The females, the applicants with more siblings and with older siblings, are less likely to enroll. Specifically, if the student is female, the probability of enrolling decreases 6.1 percentage points, this effect is contrary to what exists on graduation where is positive, that is, the females are less likely to enroll but more likely to graduate. For one more sibling the probability of enrolling decreased 1.5 percentage points and if the applicant has older sibling the probability also decrease what differs from the effect on graduation where having older siblings increases the probability of graduation. The middle school GPA still has a positive effect on enrollment, but it is not that big as in the graduation, one extra point in middle school GPA increase 6.6 percentage points the likelihood of enrollment compare with the 20 percentage points on graduation. Also, the more dedicated students and those who are taking the exam more than once, tend to be more likely to enroll. For the students who have taken the exam on several occasions, the probability of enrolling increases while the probability of graduation decreases, this indicates that although these applicants have a great interest in continuing to study high school they are less likely to finish their studies.

Column two shows that the relation of family and personal variables with enrollment is a little bit weaker than the relation with graduation. Also, the relation compare with the first model on Column one is weaker, even more, some of these variables have a change in the sign of the effect, such as father's education and older siblings. This behavior indicates that the exam outcome variables have an important effect on the decision to enroll or not. For the exam variables we have that the number of options listed and the rank of assigned school is positively associated with enrollment. The COMIPEMS score is positively associated with enrollment, this is intuitive, in practice, the impact of a one-point higher score increase the likelihood of enrolling in .62 percentage points which small but significant. Notice that the relation of these variables with the enrollment is opposite to the relation with the graduation, unlike time, that is, if the applicant lists more options, is assigned to a less-preferred school or gets a higher test score increases their chances of enrolling but this does not imply that they increase their chances of graduating.

Column three shows that all the variables maintain similar estimations on enrollment and variables such as father's age and COMIPEMS score have a negligible effect on the decision. But, accounting for the school difficulty by the inclusion of the minimum score required shows that students consider it of greater importance for the decision. The likelihood of enrolling increased by .71 percentage points with each extra point requested by the school to be accepted, that is, the applicants prefer to enroll in a demanding school. Comparing this relation with the relation between this variable and graduation, we find that although applicants prefer to enroll in a demanding school, they are less likely to graduate.

Column four shows estimations for the model that adds variables related to the preferences of the students over the schools, the minimum score required by the first option and the average minimum score of the top three options. We observe that as in the case of the relation of these variables with graduation, the relation of these two variables with the enrollment is ambivalent, but opposite to the relation analyzed in the previous section. The first option seems to indicate that when the students aspire a more demanding school as their most-preferred one the probability of enrolling is smaller, even when the probability of graduation is greater, but when we consider the average of the three first schools that are listed the effect turns into positive, and the effect on graduation becomes negative, that is the likelihood of enrolling is greater when the students choose more competitive schools, but the likelihood of graduation is smaller in this case.

Notice the home-school travel time of the assigned option has a negative and significant association with enrollment, and this variable has the same association with graduation unlike other exam outcome variables, that is the home-school travel time has the same effect on graduation than on enrollment. The likelihood of enrolling and graduating decreases when the student has to travel more time. On the contrary, the effect of the demanding school es opposite on graduation and enrollment, the likelihood of enrollment increases when the school where she is admitted is more selective (a school with a rather high minimum required score), and the likelihood of graduation decreases. The negative relation between enrollment and home-school travel time is intuitive because usually, the students analyze the time that they will spend on transportation and consider that their motivation could decrease when they have to travel a lot. As time is measured in hours, one extra hour decrease the enrollment probability by 1.42 percentage points. In summary, students consider that the longer they have to travel to get to school, the less likely they are to enroll in the assigned school, but if they also consider the quality of the school, then the more demanding the school is, the greater the probability that they enroll, that is, they will travel more to attend school when the school is more demanding.

4.5.4 Distance and School Quality Preferences

This section is dedicated to the findings in the relation between the home-school travel time and the school election of the students, we want to know if there is a trade-off between these two variables.

Before we investigate the role of home-school distance, we first want to understand better what kind of students request a place in the most demanding schools. For this we use the sample (for the years 2005 to 2009) with the assigned applicants whose descriptive statistics are in column two of Table 1. Table 4.6 shows the results of the linear regression with only socioeconomic and personal variables, considering four different dependent variables: the score to be admitted to the first option of the student, the score to the second option, score to third option and the average score of the three first options. We notice that the father's education and age have a positive and significant effect on the score of the four variables, the effect of the education is stronger than the father's age, that is one more year of education of the father adds more to the student's aspirations than one more year of age of the father. The number of siblings has a negative effect on the decision of the applicant, each sibling reduces the aspirations with 1.22 score points, but notice that the students that have older siblings have a positive effect, that is the number of siblings reduces the aspirations but if these siblings are older the effect is not so big; moreover, the older siblings have a greater effect in the option 2 and 3 than in the first, this could be a sign of using the information provided by the oldest.

The personal characteristics such as female, GPA and study hours per week have a positive and significant effect on the score of the three first options. It is not surprising that for every extra point in the GPA of middle school the applicant asks for a more demanding high school which asks more than almost 4 points extra in the admission exam, but the effect decrease for the second and third options which might show a sign of safeness, that is the students prefers put a low competitive school and be admitted than top rank the highly competitive schools and do not be admitted. Girls, compared to boys, always aim for a more demanding school, that is, they present a more risky attitude than boys who seem to go for a "safe" option rather quickly. Note that girls also are more likely to graduate (Table 4): the strategy of aiming higher (and probably being assigned higher) does not seem to harm graduation from the higher-quality school.

Applicants that have applied the admission exam more than one time and those who list more options than the average tend to request options with a higher entrance score. Even more, having done the exam before, comes with higher aims for second and third option than students who do the exam for the first time. Also, notice that the applicants request for more demanding schools in their first option than in their third, even more, they reduce their aspirations in 1 point or more in every option.

In Table 4.7 we analyze the effect of the home-school travel time on the score of the selected school. We show if the time to arrive at school is considered at the moment of choosing the schools and listing the 20 most preferred ones. The personal and family characteristics behave in the same way as in the previous model, then we directly analyze the effect of time.

As we mention in Section 4, the proposed equation suffers a problem of bidirectional causality because the quality indicator and the distance measure are determined at the same moment in time. Then, we use as instrument the average travel time of all students that put the school j in their top three options, that is the average time of all applicants to school j, this gives us a measure of the school location and this help us to identify the time that the student will do in case that she is enrolled in that school. We make some tests in order to show that this proposal is a good instrument for our IV regressions. These tests are showed in Table 4.8. We observe that the tests of Durbin-Wu-Hausman which tests if the explanatory variable (home-school travel time, T_{ij}) is exogenous indicate that this variable is endogenous and that the instrument is indeed necessary. The F-test and the Wald-F are bigger enough to assure that we have a strong instrument for every IV regression.

In the first IV regression, we use the time to first choice school to explain the score that is required to be accepted at this school, that is how this time is a determinant of the selection of the first choice. Notice that the effect is negative and significant, that is for an hour more of travel the student chooses a school with a 2.40 less of score. Observe that when we run the same regression but for the second and third choice, the effect is almost the same, 2.80 and 2.44 points respectively. Then, as the applicants are choosing less demanding schools at the second and third options, these schools are faraway too.

The previous analysis lead us to think that the students are considering the time factor at the moment of choosing, but they are not conceiving these two variables as a trade-off but as an entity, a nearby and demanding school or a faraway and less competitive one. That is, if they will have to spend more time in public transport, they will opt for less demanding schools as their top preferences. Now, this behavior could be counterproductive for them because as is shown in the first model, the likelihood of graduation of high school decreases when the student enrolls in a faraway school.

4.6 Conclusions

In this paper, we study the impact of travel time on preferences, enrollment and graduation, while controlling for other factors. We analyze the applicants' school choices and the implications of several variables on it, such as father education, number of siblings, gender, middle school GPA and study hours a week. We also analyze the implications of these personal and family characteristics as well as variables related to the assignment and the preferences of the students before the test was taking on the likelihood of the eventual graduation three years (or more) later and on the likelihood of the enrollment in the school assigned by COMIPEMS. In order to answer those questions, we use as primary source of information the database of COMIPEMS, which contains information collected at the moment of registering for the exam and after the assignment process. Moreover, we count with information of the high schools where the applicants are enrolled that permits us to know if students graduated. We use IV for preferences because both decisions (preferences and travel time) are simultaneous, while for graduation and enrollment logit regressions suffice because all the explanatory information is predetermined.

Our main findings indicate that, although travel time has only a small impact on graduation, in earlier stages, it negatively affected much more strongly the enrollment decision to the assigned school (given the exam results), and the preferences or aspirations before presenting the exam. The preference for a distant school decreases the probability of graduation. Also, the likelihood of enrolling decreases when the student has to travel more time, that is, the students will travel more to attend school when the school is more demanding. And, if the students will have to spend more time in public transport, they will opt for less demanding schools as their top preferences. Specifically, one extra hour of home-school travel time reduces the graduation probability by 0.64 percentage points, reduces the probability of enrollment by 1.42 percentage points and the student chooses a school with a 2.40 less of score. These results have implications for the students' graduation that they should consider at the moment of election.

We find that father's education, have older siblings, have higher notes in middle school and dedicate more hours to study make that the students have higher aspirations, as reflected in the preferred schools listed in their COMIPEMS application. Also, these students have greater probability of enrolling and graduating, though the effect is greater on graduation than on enrolling. On the other hand, if the student has more siblings, he reduces both his aspirations and his chances of enrollment and graduation, although the effect is greater on graduation. The father's age and be female have a positive effect on graduation and aspirations but negative on enrollment, that is, these students have greater aspirations and more chances to accomplish it once they are enrolled but are less likely to actually enroll.

The number of times that the applicant has taken the exam has a positive effect on aspirations and enrollment but negative on graduation, that is, the more times the exam has been taken the more demanding schools are requested and the more likely they are to be enrolled, but the less likely they are to graduate. The number of options listed has no effect on graduation, but positive effect on aspiration and enrollment, that is, the more number of options has been listed the more demanding schools are requested and the more likely they are to be enrolled.

COMIPEMS score has a positive effect on both graduation and enrollment, the effect on enrollment is smaller than on graduation, that is, the students with higher COMIPEMS score tend to be more likely to enroll and graduate. When the assigned school is less preferred than others then the effect on both graduation and enrollment is negative and small, this seems to indicate that the motivation decreases. Regarding the effect of the "quality" preferences on the likelihood of graduation and enrollment, it seems to have a very low effect on both. Accounting for the school difficulty by the inclusion of the minimum score required shows that the applicants prefer to enroll in a demanding school, but the probability of graduation decreases.

We conclude that it is very important to consider the distance when choosing the school the applicant wants to attend, since as observed in our analysis, this variable has a negative effect on the three variables analyzed, graduation, enrollment and quality of schools. The first thing would be to have detailed information about the educational options and their location, and then analyze if the student is really willing to transport themselves to school considering all that this entails. On the part of the providers of the educational service, in this case the government, it is important to consider the option of bringing quality education to the most remote areas of the city.

While the effects presented here may be particular to Mexico City, additional work is needed to examine if these effects are present in different student assignment systems. Future work in this particular context lead us to analyze the effect of distance is considering the information that the applicants have at the moment of election, using the fact that some of them have older siblings who already present the test one or more years before.

Table 4.1: Descriptive Statistics. Full, Assigned and Enrolled Samples

	Full Sample	Assigned	Enrolled
Family Monthly Income	5,263.6	5271.0	5419.0
(pesos per month)	(4,229.8)	(4176.6)	(4225.2)
Indigenous	.037	.036	.033
	(.188)	(.186)	(.179)
Father's age	45.0	45.0	45.1
	(7.1)	(7.1)	(7.1)
Father's education (years)	9.62	9.64	9.80
	(3.53)	(3.53)	(3.55)
Older siblings	.601	.592	.586
	(.490)	(.491)	(.493)
Number of siblings	2.14	2.12	2.07
	(1.43)	(1.42)	(1.38)
Female	.497	.502	.490
	(.500)	(.500)	(.500)
GPA in middle school	8.00	8.03	8.09
	(.84)	(.85)	(.87)
Study hours (per week)	4.95	5.07	5.22
	(3.19)	(3.22)	(3.25)
Aspired level of studies	3.88	3.93	3.99
	(1.12)	(1.09)	(1.06)
Number of times has taken	1.16	1.17	1.18
the exam	(.42)	(.42)	(.43)
COMIPEMS score	61.94	66.02	68.75
	(20.65)	(18.45)	(18.72)
Number of options listed	9.28	9.43	9.47
	(3.71)	(3.71)	(3.73)
Rank of assigned option		3.70	3.55
		(3.29)	(3.22)
Minimum score required		55.67	59.27
at the assigned school		(19.11)	(19.50)
Time assigned school		1.00	.97
		(1.08)	(1.05)
Observations	1,503,782	1,186,126	873,013

Full sample of all students who applied for the test in the first column.
Restricted sample to students accepted in an elite and non-elite school in columns 2 and 3 respectively. Some characteristics have missing values.
The aspired level of studies is in a scale of 1 to 5, where 1 is high school, 2 is technical professional, 3 is senior technician, 4 is bachelor degree and 5 is postgraduate studies.

Table 4.2: Descriptive Statistics. Close, Non-Close, Elite and Non-Elite Schools

	Close school	Non-Close school	Elite school	Non-Elite school
Family Monthly Income	5,371.4	5,071.7	8,185.3	4916.08
(pesos per month)	(4,289.3)	(3,928.9)	(5,934.5)	(3,757.5)
Indigenous	.033	.040	.011	.039
	(.180)	(.198)	(.103)	(.194)
Father's age	45.2	44.6	46.5	44.8
	(7.2)	(7.0)	(7.0)	(7.0)
Father's education (years)	9.78	9.37	12.06	9.34
	(3.57)	(3.44)	(3.77)	(3.39)
Older siblings	.593	.590	.507	.602
	(.491)	(.491)	(.500)	(.489)
Number of siblings	2.08	2.20	1.61	2.19
	(1.41)	(1.43)	(1.09)	(1.44)
Female	.502	.503	.477	.505
	(.500)	(.500)	(.500)	(.500)
GPA in middle school	7.99	8.12	8.66	7.96
	(.84)	(.86)	(.80)	(.82)
Study hours (per week)	5.07	5.09	6.65	4.88
	(3.21)	(3.22)	(3.32)	(3.15)
Aspired level of studies	3.93	3.91	4.60	3.84
	(1.09)	(1.09)	(.64)	(1.11)
Number of times has taken	1.18	1.15	1.10	1.18
the exam	(.43)	(.40)	(.33)	(.43)
COMIPEMS score	66.14	65.74	94.90	62.67
	(18.29)	(18.74)	(9.12)	(16.20)
Number of options listed	9.48	9.32	9.17	9.46
	(3.73)	(3.67)	(3.88)	(3.70)
Rank of assigned option	3.79	3.52	1.46	3.96
	(3.34)	(3.18)	(.94)	(3.36)
Minimum score required	55.76	55.49		
at the assigned school	(19.08)	(19.18)		
Time assigned school			.95	1.00
			(.95)	(1.09)
Observations	780,836	391,295	123,050	1,063,076

Full sample of all students who applied for the test in the first column.
Restricted sample to students accepted in a close and non-close school in columns 2 and 3 respectively. Some characteristics have missing values.
The aspired level of studies is in a scale of 1 to 5, where 1 is high school, 2 is technical professional, 3 is senior technician, 4 is bachelor degree and 5 is postgraduate studies.

Table 4.3: Descriptive Statistics. Outcome Variables

	Full Sample	Close School	Non-Close School	Elite school	Non-Elite school
Graduation	.598	.591	.615	.670	.587
	(.490)	(.492)	(.487)	(.470)	(.492)
Enrollment	.733	.749	.702	.956	.707
	(.442)	(.433)	(.457)	(.204)	(.455)
Score first choice	74.44	75.20	72.99	92.22	72.40
	(19.47)	(19.84)	(18.62)	(6.19)	(19.43)
Average score of 3 first options	71.91	72.73	70.10	88.04	70.00
	(17.31)	(17.61)	(16.48)	(6.79)	(17.15)
Observations	1,503,782	780,836	391,295	123,050	1,063,076

Full sample of all students who applied for the test in the first column.
Restricted sample to students accepted in a close, non-close, elite and non-elite school in columns 2, 3, 4 and 5 respectively.
Some characteristics have missing values.

Table 4.4: Logistic Regression on Graduation

	[1]	[2]	[3]	[4]
Time assigned school		-.0067***	-.0060***	-.0064***
		(.0006)	(.0006)	(.0006)
Score to first option				.0009***
				(.0001)
Average score of three first options				-.0014***
				(.0001)
Minimum score required at the assigned school			-.0055***	-.0053***
			(.0001)	(.0001)
Number options listed		-.0010***	.0001	-.0000
		(.0002)	(.0002)	(.0002)
Rank assigned option		-.0054***	-.0051***	-.0036***
		(.0002)	(.0002)	(.0003)
COMIPEMS score		-.0007***	.0035***	.0037***
		(.0000)	(.0001)	(.0001)
Father's age	.0009***	.0011***	.0014***	.0015***
	(.0001)	(.0001)	(.0001)	(.0001)
Father's education (years)	.0026***	.0036***	.0056***	.0057***
	(.0002)	(.0002)	(.0002)	(.0002)
Number of siblings	-.0085***	-.0094***	-.0112***	-.0114***
	(.0005)	(.0005)	(.0005)	(.0005)
Older siblings	.0314***	.0297***	.0325***	.0326***
	(.0015)	(.0015)	(.0015)	(.0015)
Female	.0280***	.0270***	.0334***	.0339***
	(.0013)	(.0014)	(.0014)	(.0014)
GPA in middle school	.2025***	.2072***	.2146***	.2151***
	(.0009)	(.0009)	(.0010)	(.0010)
Study hours (per week)	.0039***	.0044***	.0051***	.0052***
	(.0002)	(.0002)	(.0002)	(.0002)
Number of times has taken the exam	-.0495***	-.0501***	-.0466***	-.0464***
	(.0015)	(.0015)	(.0015)	(.0015)
Observations	627,987	620,443	620,443	617,415
Pseudo-RSquared	0.1054	0.1065	0.1180	0.1182

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Marginal effects are average marginal effect over the sample, and standard errors in parenthesis. Column 1 family and individual characteristics, column 2 adds the exam results, column 3 adds score at assigned school and column 4 adds information about preferences. All specifications include COMIPEMS exam year dummies.

Table 4.5: Logistic Regression on Enrollment

	[1]	[2]	[3]	[4]
Time assigned school		-.0141***	-.0153***	-.0142***
		(.0004)	(.0004)	(.0004)
Score to first option				-.0009***
				(.0000)
Average score of three first				.0029***
options				(.0001)
Minimum score required			.0071***	.0062***
at the assigned school			(.0000)	(.0000)
Number options listed		.0037***	.0021***	.0028***
		(.0001)	(.0001)	(.0001)
Rank assigned option		.0022***	.0023***	-.0051***
		(.0002)	(.0002)	(.0002)
COMIPEMS score		.0062***	.0009***	.0006***
		(.0000)	(.0000)	(.0000)
Father's age	.0028***	.0006***	.0002**	.0001
	(.0001)	(.0001)	(.0001)	(.0001)
Father's education (years)	.0060***	-.0012***	-.0037***	-.0044***
	(.0001)	(.0001)	(.0001)	(.0001)
Number of siblings	-.0157***	-.0078***	-.0055***	-.0047***
	(.0004)	(.0004)	(.0004)	(.0004)
Older siblings	-.0051***	.0111***	.0083***	.0079***
	(.0011)	(.0011)	(.0011)	(.0011)
Female	-.0610***	-.0232***	-.0314***	-.0330***
	(.0010)	(.0010)	(.0010)	(.0010)
GPA in middle school	.0664***	.0181***	.0143***	.0123***
	(.0006)	(.0007)	(.0007)	(.0007)
Study hours (per week)	.0064***	.0020***	.0013***	.0010***
	(.0002)	(.0002)	(.0001)	(.0001)
Number of times has taken	.0473***	.0312***	.0252***	.0264***
the exam	(.0012)	(.0012)	(.0012)	(.0012)
Observations	845,717	836,831	836,831	832,843
Pseudo-RSquared	0.0282	0.0682	0.1029	0.1063

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Marginal effects are average marginal effect over the sample, and standard errors in parenthesis. Column 1 family and individual characteristics, column 2 adds the exam results, column 3 adds score at assigned school and column 4 adds information about preferences. All specifications include COMIPEMS exam year dummies.

Table 4.6: Linear Regression of Score Options

	Score First Option	Score Second Option	Score Third Option	Average Top Three
Father's age	.223***	.225***	.227***	.225***
	(.003)	(.003)	(.003)	(.002)
Father's education (years)	1.12***	1.10***	1.12***	1.11***
	(.005)	(.006)	(.006)	(.005)
Number of siblings	-1.22***	-1.22***	-1.24***	-1.23***
	(.015)	(.015)	(.015)	(.013)
Older siblings	.380***	.490***	.419***	.426***
	(.042)	(.044)	(.044)	(.037)
Female	.828***	1.045***	1.011***	.959***
	(.038)	(.039)	(.039)	(.033)
GPA in middle school	4.05***	3.72***	3.49***	3.75***
	(.023)	(.024)	(.024)	(.021)
Study hours (per week)	.447***	.451***	.460***	.453***
	(.006)	(.006)	(.006)	(.005)
Number of times has taken the exam	.025	.741***	1.152***	.661***
	(.044)	(.045)	(.046)	(.039)
Number of options listed	.603***	.626***	.652***	.619***
	(.005)	(.005)	(.005)	(.004)
Constant	15.70***	14.37***	12.78***	14.41***
	(.241)	(.247)	(.249)	(.212)
Observations	957,111	955,865	951,401	951,401

Standard Errors in parenthesis. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. All specifications include COMIPEMS exam year dummies.

Table 4.7: Effect of Home-School Travel Time in School Choice

	Score First Option	Score Second Option	Score Third Option	Average Top Three
Time to First Option	-2.40***			
	(.055)			
Time to Second Option		-2.80***		
		(.052)		
Time to Third Option			-2.44***	
			(.055)	
Average Time to 3 First				-2.55***
Options				(.046)
Father's age	.204***	.204***	.209***	.206***
	(.003)	(.003)	(.003)	(.002)
Father's education (years)	1.10***	1.07***	1.10***	1.09***
	(.005)	(.006)	(.006)	(.005)
Number of siblings	-1.15***	-1.13***	-1.17***	-1.15***
	(.015)	(.016)	(.016)	(.013)
Older siblings	.390***	.467***	.388***	.410***
	(.043)	(.044)	(.045)	(.038)
Female	.675***	.854***	.844***	.788***
	(.039)	(.040)	(.040)	(.034)
GPA in middle school	4.38***	4.10***	3.81***	4.09***
	(.025)	(.026)	(.026)	(.022)
Study hours (per week)	.444***	.447***	.458***	.450***
	(.006)	(.006)	(.006)	(.005)
Number of times has	-.097*	.589***	1.031***	.531***
taken the exam	(.045)	(.046)	(.047)	(.040)
Number of options listed	.582***	.603***	.635***	.599***
	(.005)	(.005)	(.005)	(.004)
Constant	17.04***	16.09***	14.29***	15.93***
	(.247)	(.254)	(.256)	(.217)
Observations	946,995	945,773	941,378	941,378

Standard Errors in parenthesis. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. All specifications include COMIPEMS exam year dummies.

Table 4.8: Instrument's Tests

	Score First Option	Score Second Option	Score Third Option	Average Top Three
F-test	8,825.13	9,941.11	9,280.90	10,810.45
Durbin-Wu-Hausman	0.0000	0.0000	0.0000	0.0000
Wald-F	53,599.9	53,510.8	56,703	66,057.9
Observations	946,995	945,773	941,378	941,378

First line shows the first stage F value.
Second line shows p-values for Durbin-Wu-Hausman statistics.
Line three shows the value of F in the Wald test, the p-values are 0.0000.

References

[1] Atila Abdulkadiroğlu, Joshua Angrist, Peter Hull and Parag Pathak. *Charters without Lotteries: Testing Takeovers in New Orleans and Boston.* NBER Working Paper No. 20792 National Bureau of Economic Research, 2014.

[2] Atila Abdulkadiroğlu, Joshua Angrist and Parag Pathak. *The Elite Illusion: Achievement Effects at Boston and New York Exam Schools.* Econometrica, Vol 82, No. 1, 2014; 137-196.

[3] Atila Abdulkadiroğlu and Tayfun Sönmez. *School Choice: A Mechanism Design Approach.* The American Economic Review. Vol 93, No.3, 2003; 729-747.

[4] Herbert Altrichter, Johann Bacher, Martina Beham, Gertrud Nagy and Daniela Wetzelhütter. *The Effects of a Free School Choice Policy in Parents? School Choice Behavior.* Studies in Educational Evaluation, 37, 2011, 230?238.

[5] Eva Andersson, Bo Malmberg and John Östh. *Travel-to-School Distances in Sweden 2000-2006: Changing School Geography with Equality Implications.* Journal of Transport Geography, 23, 2012, 35?43.

[6] Simon Burgess and Adam Briggs. *School Assignment, School Choice and Social Mobility.* Economics of Education Review, 29, 2010, 639?649.

[7] Simon Burgess, E. Greaves, A. Vignoles and D. Wilson. *What Parents want: School Preferences and School Choice.* The Economic Journal, 125 (Septiembre), 2014; 1262-1289.

[8] Jorge Ubaldo Colin Pescina. *The Effect of Transit Improvements on School Choice; The Case of Public High Schools in Mexico City.* Doctoral dissertation, Columbia University Academic Commons, 2015

[9] COMIPEMS. *Informe del Concurso de Ingreso a la Educación Media Superior de la Zona Metropolitana de la Ciudad de México (1996-2010).* Ceneval Publication, 2012.

[10] J.B. Cullen, B. Jacob and S. Levitt. *?The Effect of School Choice on Participants: Evidence from Randomized Lotteries.* Econometrica, Vol. 74, No. 5, 2006, pp. 1191-1230.

[11] W. Dobbie and R. Fryer. *Exam High Schools and Academic Achievement: Evidence from New York City.* NBER Working Paper 17286 National Bureau of Economic Research, 2014.

[12] Andrew Dustan. *Family Networks and School Choice.* Working paper, 2017.

[13] Andrew Dustan, Alain De Janvry and Elisabeth Sadoulet. *Flourish or Fail?: The Risky Reward of Elite High School Admission in Mexico City.* Journal of Human Resources, 52(3), 2017, 756-799.

[14] Andrew Dustan and Diana K.L. Ngo. *Commuting to Educational Opportunity? School Choice Effects of Mass Transit Expansion in Mexico City.* Economics of Education Review, Vol 63, 2018, 116-133.

[15] Gregory Elacqua. *The impact of school choice and public policy in segregation: Evidence from Chile.* International Journal of Educational Development, 32(3), 2012, 444-453.

[16] Ricardo Estrada. *The Effect of the Increasing Demand for Elite Schools on Stratification.* EUI Working Paper MWP 2016/2.

[17] D. Gale and L.S. Shapley. *College Admissions and the Stability of Marriage.* The American Mathematical Monthly, Vol 69, No.1, 1962; 9-15.

[18] Daniel Gómez, Rómulo A. Chumacero and Ricardo D. Paredes. *I Would Walk 500 Miles (If It Paid): Vouchers and School Choice in Chile.* Economics of Education Review, Vol 30, No.5, 2011; 1103-1114.

[19] J. Hastings, T. Kane and D. Staiger. *Preferences and Heterogeneous Treatment Effects in a Public School Choice Lottery.* NBER Working Paper 12145, National Bureau of Economic Research, 2006.

[20] J. Hastings, C. Neilson and S. Zimmerman. *?The Effect of School Choice on Intrinsic Motivation and Academic Outcomes.* NBER Working paper No. 18324 National Bureau of Economic Research, 2012.

[21] J. Hastings and J. Weinstein. *Information, School Choice and Academic Achievement: Evidence from Two Experiments.* The Quarterly Journal of Economics, 2008; 1373-1414.

[22] M.E. Ortega. *School Choice and Educational Opportunities: The Upper-Secondary Student-Assignment Process in Mexico City.* Doctoral dissertation, Harvard Graduate School of Education, 2015.

[23] Umut Özek. *The effects of Open Enrollment on School Choice and Student Outcome.* National Center for Analysis of Longitudinal Data in Education Research. Working Paper 26, May 2009.

[24] M. Söderström and R. Uusitalo. *School Choice and Segregation: Evidence from an Admission Reform.* Scandinavian Journal of Economics, 112(1), 2010, 55?76.

List of Tables