A study on the relationship between student allocation and proficiency in Belo Horizonte, Brazil *

Delgado, V.M.S.; Rios-Neto, L. E. G. and Davis Jr., C. A.

Resumo O artigo conduz um estudo sobre o sistema de cadastro escolar para matrículas da cidade de Belo Horizonte - Brasil (ou *mecanismo* de cadastro, como mais utilizado pela literatura). O mecanismo de cadastro escolar de Belo Horizonte completou, em 2013, 20 anos desde a sua criação. Tal mecanismo executa um procedimento de registro dos estudantes e sugere em quais escolas públicas os estudantes que vivem em Belo Horizonte devem se matricular, o procedimento é com base na proximidade do aluno da escola e dos limites do distrito escolar. O objetivo do artigo é verificar se o equilíbrio proposto pela administração de Belo Horizonte é estável e se há alguma maneira de melhorar os resultados de proficiência dos alunos do 5º ano do Ensino Fundamental, fazendo-se uma correspondência entre as alocações de alunos e o bem-estar. Para comparar a alocação de estudante observada proposta pela administração municipal com alocações ótimas, empregamos o algoritmo proposto por Gale & Shapley (1962) [9], bem como o algoritmo de *Top Trading Cycles* (TTC) proposto por Shapley & Scarf (1974) [12]. Aplicações recentes dessa literatura podem ser encontradas em Abdulkadiroğlu & Sönmez (2003) [2], Fernandes (2007) [6] and Abdulkadiroğlu, Pathak & Roth (2005) [1].

Palavras-Chave: Matching de Mercados de Dois Lados, Desenhos de Mecanismos, Educação, Gale-Shapley, Top Trading Cycles

Delgado, V.M.S.
Departamento de Economia at Universidade Federal de Ouro Preto (UFOP), Mariana-MG, Brazil, e-mail: victor.delgado@icsa.ufop.br

Rios-Neto, L. E. G.
Departamento de Demografia at Universidade Federal de Minas Gerais (UFMG), Belo Horizonte-MG, Brazil, e-mail: eduardo@cedeplar.ufmg.br

Davis Jr., C. A.
Departamento de Ciência da Computação at Universidade Federal de Minas Gerais (UFMG), Belo Horizonte-MG, Brazil, e-mail: clodoveu@dcc.ufmg.br

* Os autores agradecem à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) pelo suporte financeiro para esse trabalho.

Abstract This paper conducts a study of the school registration system in Belo Horizonte - Brazil (called most frequently as *mechanism* by the literature). The registration mechanism completed, in 2013, 20 years since its inception. The mechanism performs a registration procedure and determines enrollment places in public elementary schools for students living in Belo Horizonte, based on geographic proximity and school district boundaries. The main purpose is to check whether the equilibrium proposed by Belo Horizonte's administration is stable and if there is any way to improve the results of students' proficiency on the 5^{th} grade through connecting allocation and welfare. To compare the actual student allocation proposed by the municipal administration with optimal allocations we employed the matching algorithm proposed by Gale & Shapley (1962) [9] and also the *Top Trading Cycles* (TTC) algorithm proposed by Shapley & Scarf (1974) [12]. Recent applications of this literature could be found in Abdulkadiroğlu & Sönmez (2003) [2], Fernandes (2007) [6] and Abdulkadiroğlu, Pathak & Roth (2005) [1].

Key words: Two-sided Matching, Mechanism Design, Education, Gale-Shapley, Top Trading Cycles

1 Introduction

Belo Horizonte is the 6^{th} largest city in Brazil (measured by population, around 2.5 million pop.) and receives applications to enroll approximately 20,000 students per year in its public school system, most of them in the first year of Brazilian elementary school (students of 5 to 6 years old). In 1993, the city started an administrative procedure to allocate students according to the location of the student's residence. Before that year, the process was inordinate and decided by the choices of parents, who tried to enroll their children at their most preferred school. At the time for enrollment, parents stood for hours or days in queues at their most wanted school to try to ensure a seat for their children, since the selection was made by order of arrival.

If all schools were homogeneous with a supply of seats that exceeded the demand, there would be no problem with the former decentralised mechanism. However, in the 1990s Belo Horizonte had a *boom* in demand for schools in the public system, and that caused several allocation problems. Fonseca & Zuppo (1994) [8] pointed three main reasons for that situation: 1) First, the 1990s economic crises, that brought students from the private system to public schools;[2] 2) Lack of supply in the public system, since Brazil went through the 1980s with deprived investments in public education, causing limitations in the beginning of the 1990s; 3) Belo Horizonte's enrollment mechanism previous to 1993 was very precarious, not allowing school officials to clear the queues quickly.

[2] It is important to point out that in Brazil the private system is very prominent: roughly 30% of elementary schools are private. This system is demanded by families of the upper middle class and high income.

We offer one additional explanation to the ones presented by Fonseca & Zuppo [8]: 4) Queues did not really happen at all schools, only at those that were the most desired. If that was not the case, part of the students would not be able to register at all, and that did not happened, all students could enroll. Although limited, the former Belo Horizonte's mechanism was able to provide all the necessary seats. The fact was that the queue mechanism led to an unnecessary delay in the enrollment procedures, wasting too much of the parents' and students' time.

In this paper we describe, in short, how Belo Horizonte's schooling administration managed to change their enrollment mechanism.[3] We are especially interested in investigating what are the consequences of the current enrolment mechanism in the students' proficiency. This question becomes more important as we consider that the current mechanism allocates students according to a proximity criterion (distance between school and home). In the next section we present the current enrollment procedures and briefly introduce the Shapley-Scarf [12] and Gale-Shapley [9] algorithms. In Section 3 we show a model for proficiency based on distance, we give a simple example of this construction. Section 4 presents our main results and finally in Section 5 we conclude with a discussion of our results.

2 The School Enrollment Mechanism of BH (BHM) and popular Algorithms

In 1993 BH changed its enrollment mechanism to avoid queues and problems described in [8] and in [10]. This new mechanism was based on a georeferenced database already built for BH urban planning and other applications (see [4] for a description of using GIS by the BH city computing agency (PRODABEL)).[4] In this section we describe BH's current school enrollment mechanism, giving some important aspects in the construction of the distance criterion. We also give a brief presentation of Gale-Shapley (GS) [9] and Shapley-Scarf algorithms, known as *Top Trading Cycles* (TTC) [12]. These two algorithms are important to our modeling, particularly the Gale-Shapley algorithm, which was used to explore new allocation possibilities and to check if the actual allocation is stable, in the sense that no student would want to change his/her school and no school would want this student more than any student that is already allocated for it (to see this definition in more conceptual form see subsection 2.3).

The São Paulo's students registration mechanism appeared in 1995 and was inspired by the Belo Horizonte. It was studied by [6] in a paper that explores the properties of such a system and the stability of the matching. Compared to this work, our

[3] Belo Horizonte is also called BH by Brazilians. For now on we are going to use this abbreviation to designate the city and BHM to designate the Belo Horizonte's *mechanism*.

[4] Fonseca et al. (2000) [7] and Davis & Fonseca (2007) [5] give more recent descriptions of geocoding systems and use BH as an example.

paper attempts to simulate an existing situation before the current mechanism and try to link the simulated allocations with the student proficiency performance.[5]

2.1 The BHM School Enrollment Mechanism

BHM was implemented in 1993. This is one of the first enrollment mechanism to appear in the country, which had previously used a decentralized enrollment process for schools, the former mechanism generated queues that transferred the burden of time waiting to the citizens. Now the BH's enrollment process, working since 1993, can be summarized in six steps:

1. In the month of June, a municipal education secretariat resolution establishes the deadlines for enrollment and other guidelines.[6]
2. Between August and September a pre-enrollment season takes place at the postal service agencies (only for residents living in BH, but not in the neighboring towns). Parents interested in enrolling their children in public schools must go to an accredited postal agency, carrying the child's birth certificate and one address confirmation document, typically an electrical bill.
3. At the end of step 2 (the pre-enrollment process), the number of applicants is verified. With the information of who is demanding the public school system and their addresses, the prospective students' homes are geocoded (i.e., geographically located) and stored by PRODABEL (the municipal information technology company). After student locations are known, each one is allocated to the school that is responsible for the school district that encloses their home location. [7]
4. The education committee sends a letter to the parents identifying the school for which the registration is indicated (according to the address informed in the electrical bill).
5. In the third week of December the actual enrollment process takes place for the following school year. If the indication from the pre-enrollment mechanism is accepted by parents, the student has a guaranteed place in the indicated school. However, at this stage, some parents may decide to request enrollment in another

[5] Another applications of Gale-Shapley algorithm in Brazilian context of educational matching could be found in [3], which study the Brazilian universities admission tests applying the GS-algorithm. For Brazilian graduate courses [13] studied the selection process of graduate courses in economics, centralized by ANPEC - *Associação Nacional dos Centros de Pós-Graduação em Economia*, National Association of Centres for Graduate in Economics, giving suggestions to improve the clearing process.

[6] The Brazilian school year has 200 schooldays and occurs concomitantly with the calendar year. Classes start at February, with a one or two-week interruption in July, and ends in early December.

[7] School districts usually contain a single public school. If the student demand is greater than the number of vacancies at that school, younger students are prioritized (birth year, month and day), and exceeding students are allocated to the nearest school among those schools that are responsible for the neighboring districts. All students coming from another district are prioritized over those of the district in which the vacancy is being requested. This process continues until all students are allocated.

school of their choice. This can occur, among other problems, in the case of parents who moved, or parents that think there is an error in the address information used in the process. At any time, the school office may consult the registration committee to verify the address provided. In case of incorrect address, the committee will indicate the student for a nearby schools where there are vacancies.[8]

6. Latecoming students, who have not entered the pre-enrollment mechanism in a timely process, are allocated to schools where there is a vacancy.

The mechanism described above is still in use, with only small corrections since 1993. The main evolution is a greater computerization of the process. Nowadays the data provided in postal agencies are immediately transferred to PRODABEL's headquarters. It is important to highlight that the main objective of the BHM is administrative, to avoid queues and not to minimize the distances covered by the students. Although having children studying in nearby schools at a safe distance is a desired characteristic, this is not the only desirable result. Other mechanisms are more explicit about distance, like Singapore's mechanism, reported by [14]. Other mechanisms include more criteria, see survey by [2]. More recently, school enrollment systems in the US include the stability condition more explicitly [1]. That condition says that a mechanism must comply with the property of stability, in short, no school or student may prefer each other outside the mechanism.

To investigate how BHM is related to student achievement we made three mainly assumptions in this article: 1) The schools are not perceived by the parents as *homogeneous* in quality (this deviates from the official declaration of municipal education secretariat); 2) Before the current *mechanism*, the queuing system that existed approached the TCC algorithm, with parents choosing the best schools by order of proficiency and schools giving priorities at random (which simulates their order in queue); 3) We suppose that the actual mechanism proceeds *somewhat like* the GS algorithm considering the distance as a criterion (in fact, we examine the observed allocation and compare with the GS suggested allocation and try to extract from it some information about the relationship between the suggested allocation and possible gains in learning).[9]

2.2 The Top Trading Cycles (TTC) Algorithm for BH

The TTC was proposed by [12], the example contained in the original article is about the housing market, but Abdulkadiroğlu & Sönmez [2] ingeniously translated that algorithm to a school registration scenario. Here we give a description of this algorithm based in [2] and adapted to BH case.

1. Each school has a *counter* to keep track of the number of seats. This *counter* registers the number of seats that are still available, being set initially to the number

[8] We comment more about intentional deviation in section Sect. 3.

[9] We explain why we call this a *somewhat like* GS *mechanism* in Section 4.

of seats of the school. Each student *points* to his/her favorite school (according to his/her true preferences); we could imagine that *points* corresponds to enter in the queue for that school. Each school *points* to the student of highest priority. Since our algorithm simulates the situation before BH's current registration mechanism, and considering that at that time the school was indifferent between any student, we established a random priority ranking, the school gives highest priority to who comes first in the queue.[10] Since the number of students and schools are finite, the algorithm has at least one cycle. Every student in a cycle is assigned a seat at the school she *points* to, and he/she is removed after that. The counter of each school in a cycle is reduced by one and if it reaches zero, the school is removed.

2. Generic step k. Each student who is still in the mechanism (with no seat yet) *points* to his/her preferred school among the remaining schools, and each remaining school *points* to the student with highest priority (the next student in the school queue) among the remaining. There is at least one cycle. Again, every student in a cycle is assigned to a seat at the school she *pointed* to, and removed after that. The counter of each school in a cycle is reduced by one and if it reaches zero, the school is also removed. Students who did not participate in a cycle and the schools that have vacancies remain in the mechanism. The algorithm stops when there is no student remaining.

2.2.1 A Simple Example

Suppose that we have three schools in S, $S = \{s_1, s_2, s_3\}$, each school with two seats. The vector of seats of each school is Q where $Q = \{q_1, q_2, q_3\}$ with $q_s = 2$ for any $s \in S$. We have four students $I = \{i_1, i_2, i_3, i_4\}$. Suppose that we have the following preferences for each student in the set I:

$$i_1 : s_3,\ s_2,\ s_1$$
$$i_2 : s_1,\ s_2,\ s_3$$
$$i_3 : s_3,\ s_1,\ s_2$$
$$i_4 : s_2,\ s_1,\ s_3$$

Consider now that each school is faced with completely random queues, given priorities for each school in set S:

[10] It is possible to think that our queuing mechanism proposed here has no time restriction for the student who enters the queue. This is not a very realistic assumption. Suppose, for example, a student i who tries to enroll in school s. However, suppose also that in s she is in the last place and does not get the desired seat. The *i's* second choice is the school *s'*, however, as she had already waited in the queue at her first school s, i cannot stand in front of a student j who prefers school *s'* in the first place. This is equivalent to placing a restriction, the random place in the queue that i can get at *s'* does not come before the position in the queue of students who prefer the school *s'* in the first place. Our current algorithm implementation does not do this because the preferences of the schools are totally random and thus, depending on luck, i can be faster than j, even if i prefers the school *s'* at the second place.

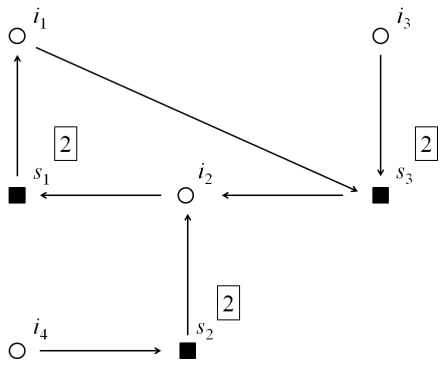

Fig. 1 TTC algorithm **step 1**. Each students points to his/her favorite school, and school points back to the its first student in line. In this step the pairs (i_1, s_3) and (i_2, s_1) are formed and schools s_1 and s_3 have their counters reduced by one seat. As school s_2 do not form a cycle, it remains with no students. Students i_3 and i_4 that don't participate in a cycle have to wait for the next step.

$$s_1 : i_1,\ i_3,\ i_4\ i_2$$
$$s_2 : i_2,\ i_4,\ i_3\ i_1$$
$$s_3 : i_2,\ i_3,\ i_1\ i_4$$

In the TTC, the student i_1 *points* to her first option, which is s_3, on the other hand, the school s_3 points to its first student in list, i_2, but i_2 does not want to go to s_3 (we could think that he is not in queue for s_3) but s_1. The school s_1 points to i_1 and this closes a cycle $(i_1, s_3, i_2, s_1, i_1)$. In Figure 1 the unfilled circles represent the students and the filled squares represent the schools, with number of seats inside the box right and above. In the first step we have just one cycle, $(i_1, s_3, i_2, s_1, i_1)$. Students i_1 and i_2 leave the algorithm and no school is complete yet, but s_1 and s_3 have their counters reduced by one.

At the second step school s_1 points to the second student in the priority list (given randomly by the queue) which is i_3. i_3 still pointing to s_3 and s_3 points him back, a new cycle is formed: (i_3, s_3, i_3). Another cycle is formed in this step. As i_2 is not in the mechanism anymore, school s_2 have to point to the next option, i_4. As i_4 still pointing to s_2, another cycle is formed: (i_4, s_2, i_4). After this second step, school s_3 is full and leaves the mechanism, students i_3 and i_4 also leave and the algorithm ends (see Figure 2).

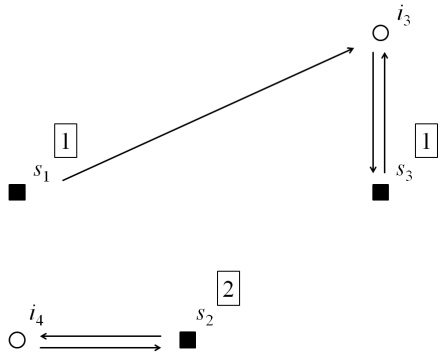

Fig. 2 TTC algorithm **step 2**. Students that didn't get one seat in the previous step still pointing to the most preferred school, since this school still vacant. i_3 and i_4 *points* to school s_3 and s_2 respectively. Now s_3 *points* to student i_3 (the next in line) and s_2 *points* to i_4 and two new cycle are formed. School s_1 still needing one more student and points to the next in line, i_3, but its seat will not be filled.

It took only two steps to allocate all the students in this example. In fact as there is always at least one cycle, the number of steps is never greater than the number of students. Abdulkadiroğlu & Sönmez [2] presents a *proof* for this and also a longer example with more schools and students.

If we represent students in rows and schools in columns we could represent the final allocation in this example as one matrix A_1 where 1 represents the student and 0 represents no allocation or a "empty" seat. In the final allocation schools s_1 and s_2 remains with one non occupied seat and all students are enrolled:

$$A_1 = \begin{array}{c} \\ i_1 \\ i_2 \\ i_3 \\ i_4 \end{array} \begin{array}{c} \begin{array}{ccc} s_1 & s_2 & s_3 \end{array} \\ \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \end{array}$$

In order to simulate the case of BH before the current school registration mechanism, we suppose that students have a non-random preferences for schools. The students preferences $\wp_i$ are given by lexicographic preferences between x_s and d_s where s is the index for schools in S $(s \in S)$, x_s is one value of a categorical variable X that indicates the quality of the school s and d_s is the distance from $i's$ students home to school s. For each student i we have the following definition for preferences:

$$\wp_i : \begin{cases} (x_s, d_s) \succeq (x_{s'}, d_{s'}) \; if \; x_s \geq x_{s'} \; \; or; \\ (x_s, d_s) \succeq (x_{s'}, d_{s'}) \; if \; x_s = x_{s'} \; and \; d_s \leq d_{s'} \end{cases} \quad \forall i \in I \quad and \quad \forall s, s' \in S. \quad (1)$$

For $s \neq s'$. This lexicographical preferences state that inside one given range of the average proficiency x, the student prefer the nearest school. So, we need to define the variable X:

$$X = \begin{cases} 3 & if \quad \overline{y}_s \geq 250 \\ 2 & if \quad 250 > \overline{y}_s \geq 200 \\ 1 & if \quad 200 > \overline{y}_s \geq 150 \\ 0 & if \quad \overline{y}_s < 150 \end{cases} \quad 500 \geq \overline{y}_s \geq 0, \quad \forall s \in S. \quad (2)$$

The y is the value of the proficiency, which is given by a regional exam that classifies the school by the students achievement, and the $\overline{y}_s$ is the mean of proficiency for school s. The scale of variable Y varies between 0 and 500, more typically, it stays inside the interval of 50 to 350 for the 5$^{\text{th}}$ grade.

These preferences can be justified by the desired characteristics in a school, parents know how to assess the quality of a school from some information that is public. There is a national exam built on national parameters for proficiency. This results are generally widely publicized by the IDEB - *Índice de Desenvolvimento da Educação Básica*, Basic Education Development Index - and parents can consult what is the school qualification in nation (or city) rank and that is publicly available. In particular, we are using here proficiency results of PROEB - *Programa de Avaliação*

da Rede Pública de Educação Básica - Public Education Evaluation Programme of Minas Gerais.

Consider, as one example, the set of four schools, $S = \{s_1, s_2, s_3, s_4\}$, let's say that for a particular student i the pair of average proficiency and distance is given by the list $L = \{(3,3),(3,2),(2,4),(2,1)\}$.[11] The preferences of student i over the set of schools is $\wp_i : s_2 \succ s_1 \succ s_4 \succ s_3$. No matter the distance, any school with $x = 3$ is preferred than a school with smaller value of x. The closest school of the set, s_4, is only the third in the preferences of i.[12]

2.3 The Gale-Shapley (GS) Algorithm

In 1962, professor David Gale and prof. Lloyd Shapley [9] proposed the algorithm to produce a stable matching. This theory is used to produce matchings between two finite and disjoint sets, let's say a set of men (M) and women (W). We call each particular man and woman by lowercase m and w, following by a subscript number when necessary.[13] Roth & Sotomayor (1992) [11] defines a *stable matching* as a matching that don't have *blocking pairs*. Two or more persons form a *blocking pair* if they are discontent in the current allocation, suppose a marriage between the man m_1 and woman w_2 and another marriage between m_2 and w_1, if m_1 prefers w_1 and, by the other hand, w_1 prefers m_1, this two individuals can break the current marriage (*coalition*) and constitute a new marriage. More formally, following the definition of [11] we have:

Definition 1. A *matching* μ is a one-to-one correspondence from the set $M \cup W$ onto itself of order two (that is, $\mu^2(x) = x$) such that if $\mu(m) \neq m$ then $\mu(m) \in W$ and if $\mu(w) \neq w$ then $\mu(w) \in M$. We refer to $\mu(x)$ as the *mate* of x.

We also should say that all marriages are *individually rational* that, broadly speaking, means that no person is forced to marry, the individual always have the option of staying single and no one could *individually* block a pair which he/she is part of. If he or she is married, it means that the current pair is at least better than be-

[11] We could state the list with $\bar{y}_s$ rather than x, so $L = \{(260,3),(255,2),(240,4),(245,1)\}$. We could think that $\bar{y}$ is a grade for each school, as grades *A*, *B*, *C*, etc., in fact, the limits used in Equation 2, are already widely used by Brazilian educational researchers.

[12] Other three types of preferences arrangements were simulated: one that considerate only distances, one for quality and other that assumed a function that combined quality and distance. The latter resulted in similar results to the lexicographical preferences presented here. It is obvious that a number of other characteristics could plug-in in a utility function for parents and thus change the preferences here obtained, including some of very subjective variables. Such preferences would be difficult to work without any direct consultation on the real preferences of the parents, and we obtained no resources for this deepest level of research, however, it may be in the scope of future research.

[13] Later we will define matching between schools (S) and students (I), like we have seen in subsection 2.2. We will also explore further implications of this changings in the end of this section.

ing single (in this case we could say that a single person marry with himself/herself). So, we could state another definition following [11]:

Definition 2. A matching μ is *stable* if it is not blocked by any individual or any pairs of agents.

By the *individual rationality* assumption, no one could block a own pair. Since we made this assumption, we need to define what is a *blocking individual* and what is a *blocking pair* to to achieve the concept of a stable matching:

Blocking Definitions:

Definition 3. In the pair (m,w), a individual m is a *blocking individual* if m strictly prefers w' over w $(w' \succ w)$ and w' weakly prefers m over any other man m' which she is allocated to $(m \succeq m')$.

Definition 4. The pair (m,w) is a *blocking pair* if, at any other allocation, m and w strictly prefers one to each other. So, m prefers w over any other woman w' who he is currently allocated to $(w \succ w')$. And w prefers m over any other man m' who she is currently allocated to $(m \succ m')$.

The GS algorithm (also called Deferred-Acceptance one-to-one) consists in men proposing to women and have a finite number of steps since no man can proposed again to a woman that already rejected him once. All this apply in reverse with women proposing to men.

Generally speaking, the algorithm work as follows: in the first step all men propose to their first choice, each man individually propose a matching to his first ranked woman. If the woman receive only one proposal, she must hold (*defer*) the current offer until receive a best ranked option. If one woman receive more than one offer at the same time (i.e. at the same step), she must decline all the least preferred offers and accept, temporarily, the best man who proposed to her.

In the second step, all men rejected in the first step have to propose to the their second option, if one woman receive more than one proposal she must hold the best option and reject all the others. In this step a woman that was holding an offer could receive a new one, in this case, she must consider what is the best offer and decline the other (this could be the one she was holding before).

The algorithm stops when all women receives at least one offer and have no more than $n^2 - 2n + 2$ steps, where n is the number of individuals in the biggest set. [14]

When applied to students and schools the algorithm is sometimes called Deferred-Acceptance algorithm many-to-one because many students can be matched to one school (each student can be "married" just with only one school, but, in the other

[14] As the GS is widely known we only give here the general definition to the many-to-one procedure, the one-to-one matching is more simpler and could be find in the original article of Gale & Shapley [9, pp.12–13] or in Roth & Sotomayor [11, pp.27–30] book.

hand, one school can "marry" with many students, in fact, this number is equal to the number of vacancies).[15]

The GS algorithm for students and schools could be described like this:

1. Each student requires a seat in the first preferred school. The school that has received a number of students proposals greater than the total seats available selects the students according to their priority (all the best options) and rejects the remaining exceeding students.
2. Generic step k $(k \geq 2)$. Each student rejected in the previous step proposes to his next most preferred school. Each school considers the students that it was "holding" in relation to the new proposals that it perhaps have received. If any new received student have higher priority than another that it was holding, the new proposal will be accepted and the student with the lowest priority will be rejected. If there is no longer any student proposing to one school with all positions filled, the algorithm terminates and schools should enroll all the proposals received. Otherwise the algorithm continues until a school that already has all positions filled no longer receive any offers.

No student may propose again to a school that he/she has already been rejected and schools consider their priority lists individually. So, if one school with two seats give the following priority list for three students: s: i_1, i_2, i_3, this means that this school prefers the class (i_1, i_2) to any other class, and (i_1, i_3) is preferred to (i_2, i_3), so we have, by consequence, $\wp_s : (i_1, i_2) \succ (i_1, i_3) \succ (i_2, i_3)$.[16]

2.3.1 Another Simple Example

Let's use the previous example (subsection 2.2.1) to made a new allocation with the GS algorithm. Remember the preferences of students and schools stated in the example and let's reverse the preferences of student i_2 (all the other preferences are the same):

$$i_2 : s_3,\ s_2,\ s_1$$

[15] The change of an algorithm *one-to-one* to a *many-to-one* slightly alternates its operation and interpretation. Imagine the group of students (I) in place of the group of men (M), and the group of schools (S) in place of the group of women (W). The difference is that every school has an upper limit of vacancies to accept students. We may see this like a polyandrous marriage. In order to extend the reasoning of the GS *one-to-one* to *many-to-one* we need to "copy" the school q times, where q is the school number of seats, each copy have the same priority over the students. The students are indifferent to any place in the same school, but any seat in a most preferred school is preferred than any seat of other least preferred school. So we need to expand school by their seats.

[16] We are simply assuming that the school preferences are only just over individuals, not over individuals as groups. In future works we are plan to consider the preferences of groups of individuals, because by this way school have more freedom to judge what is the best composition to form a class. These compositions can take into account a mix of *skills* and *backgrounds* of students.

Suppose now that school s_2 have only **one** seat, the others schools still with two seats each and have the same priorities. With this new preferences and vacancies, let's see all the steps of student-optimal GS algorithm:

1. The students i_1, i_2 and i_3 try to enter in school s_3. Student i_4 propose to school s_2. As school s_3 have received three offers (i_1, i_2, i_3) it must to pick the two priority students (i_2, i_3) and reject i_1. s_2 hold the unique offer of i_4.
2. Student i_1 rejected in the first step, need to offer to his/her second option, which is s_2. School s_2 now have two offers, the former from i_4 and the new from i_1. As now s_2 have only one seat, the school have to take the student with biggest priority, which, between i_1 and i_4 is i_4. So, i_1 is rejected again.
3. Student i_1 rejected in the second step, need to offer to his/her third option, which is s_1. As the school s_1 had not yet received any proposal, it accepts the proposal of i_1 and the algorithm stops (by coincidence i_1 is the highest priority in school s_1, but this could have any other order).

The final allocation of this slightly modified example is presented in the matrix A_2 bellow:

$$A_2 = \begin{array}{c} \\ i_1 \\ i_2 \\ i_3 \\ i_4 \end{array} \begin{array}{c} \begin{array}{ccc} s_1 & s_2 & s_3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \end{array}$$

The reader should note that the TTC will produce the same allocation (with also 3 steps) in this new example, that is just a coincidence because the example have a small dimension, we will see examples in the next section where this equivalence not always holds. The examples of Abdulkadiroğlu & Sönmez (2003) [2] also show this difference for the same set of preferences of students and schools.

3 Models and Simulations

In this section, we exemplify our simulation work with one simple example that consider the distance and the proficiency and other simulation that presents the rules used in the model for the more general situation of BH's enrollment mechanism.

First, we construct a hypothetical model that simulates a very simple city (*The Grid City Example*). This model provides a more complete understanding of the entire simulation and hence we expect the reader to be able to judge its usefulness and procedures. This hypothetical model also splits between two parts, *one* illustrating the situation before the current mechanism (the queues mechanism), approximated by the TTC algorithm, and *another* showing the simulated allocation obtained with the GS algorithm.

In the second procedure we apply this generic model to the observed data of BH (Simulation for Observed Data). For the simulation procedures we selected a

sample of 16,354 students and 296 schools.[17] The logic of this model is the same developed in the Grid City, but with characteristics (distance and vacancy in the schools) approaching to those observed in the city of BH.

3.1 The Grid City Example

Suppose a city where the streets are all orthogonal forming a grid and that each block has a distance of one kilometer. The figure 3 shows the layout of the place of residence of students (the red circles) and the location of schools (represented by small black squares). There are four students and four schools. By simplification, each school have just one seat and the student's preferences are formed like in equation 1 (Lexicographical preferences). Suppose also that schools s_2 and s_3 have the highest level of proficiency, $X = 3$, and schools s_1 and s_4 have $X = 2$, (the school classification are presented in table 1).

The matrix D represented below brings the distance of the school to each student. Students can't pass the blocks on the diagonals. This is a taxicab Geometry city, where the distance between two points is the sum of the *cathetus*.

$$D = \begin{array}{c} \\ i_1 \\ i_2 \\ i_3 \\ i_4 \end{array} \begin{array}{c} \begin{array}{cccc} s_1 & s_2 & s_3 & s_4 \end{array} \\ \begin{bmatrix} 2 & 3 & 4 & 9 \\ 1 & 6 & 7 & 8 \\ 4 & 3 & 2 & 3 \\ 5 & 6 & 7 & 12 \end{bmatrix} \end{array} \tag{3}$$

Table 1 Example of the list of schools and the average proficiency

School	Level of Proficiency (X)
s_1	2
s_2	3
s_3	3
s_4	2

Considering equation 1, the distance matrix presented in matrix of equation 3, and the school proficiencies of table 1 we have the following preferences $\wp_i$ for the

[17] The universe of the students of the 5th grade of BH is greater than 22k and the city have a total of 382 schools. In georeferencing procedure, 16k students could be precisely identified (location of their prospective residence with a point on the map of BH) and there was 296 schools in the final sample, we pulled out 86 schools in which there was no georeferenced students in the sample.

students. The higher the level of X and the smaller the distance more preferred is the school:

$$\wp_i : \begin{cases} i_1 : s_2 \succ s_3 \succ s_1 \succ s_4 \\ i_2 : s_2 \succ s_3 \succ s_1 \succ s_4 \\ i_3 : s_3 \succ s_2 \succ s_4 \succ s_1 \\ i_4 : s_2 \succ s_3 \succ s_1 \succ s_4 \end{cases} \tag{4}$$

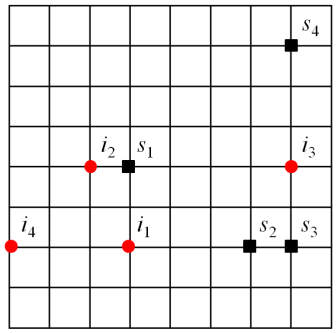

Fig. 3 Example of the Grid city. Each quarter (square) have one kilometer in all sides. There are four students and four schools. The students are represented by the red circles and schools by black small squares. Each school have just one seat and the student's preferences are formed like in equation 1. Schools s_2 and s_3 have the highest level of proficiency, $X = 3$, remember equation 2. School s_1 and s_4 have $X = 2$, meaning they have the lowest level of proficiency.

3.1.1 Simulating the Baseline

With $\wp_i$ given by 4, consider that the queue system gives the following order for priorities (which we also represent by $\wp$ since the mathematical meaning is interchangeable, but now the subscript s of schools, $\wp_s$):

$$\wp_s : \begin{cases} s_1 : i_3 \succ i_1 \succ i_4 \succ i_2 \\ s_2 : i_2 \succ i_4 \succ i_3 \succ i_1 \\ s_3 : i_3 \succ i_2 \succ i_4 \succ i_1 \\ s_4 : i_1 \succ i_3 \succ i_2 \succ i_4 \end{cases} \tag{5}$$

Considering $\wp_i$ and $\wp_s$ and doing the TTC algorithm we obtain a particular matching, let's call it μ, associating each student with theirs respective number in school set. In the matrix form we have 1's in the diagonal:

$$\mu = \begin{array}{c} \\ i_1 \\ i_2 \\ i_3 \\ i_4 \end{array} \begin{array}{c} \begin{array}{cccc} s_1 & s_2 & s_3 & s_4 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

The μ allocation is one of $4! = 24$ possible allocations and gives a particular total distance travelled by the students of 22 km (the trace of Matrix D). The average distance traveled is 5.5 km. For now, suppose that all this students have a proficiency of 222 in expectancy, $E(y_i) = 222$, which is the mean of proficiency for BH 5$^{\text{th}}$ grade students (controlled by other factors). Controlled by other effects, suppose also that students loose, on average, 2 points of proficiency for every kilometer that they travel. The result for each student is obtained generically by equation 6:[18]

$$y_{i,s} = \upsilon_i + \gamma \cdot d_{i,s} \tag{6}$$

This is a very simple equation to describe the student i performance at school s. The υ_i is a latent proficiency for this student, what she can do by herself without considering what school she is enrolled in. So, the γ generically represents what the student could gain or lose studying at one school which stays at distance $d_{i,s}$, in this paper we are only supposing that $\gamma < 0$, giving some prior suggested by BH school-student data. Since we are interested only in general aspects of the mechanism, and not only in the particular case of one of the agents, we assume a latent individual trait as being equal for all agents ($\upsilon_i = \upsilon, \forall i \in I$).[19] So, considering the values discussed in the previous paragraph, we could obtain the expectancy for the proficiency as:

$$E(y_{i,s}) = \overline{y} = \upsilon - \gamma \cdot \overline{d} \tag{7}$$

Or:

$$\overline{y} = 222 - 2 \cdot \overline{d} \tag{8}$$

The average distance depends on the allocation mechanism adopted by city government. In particular, we could rewrite equation 8 given a specific allocation (μ_n), or any of $4! = 24$ possible allocations, $\{1, \cdots, n, \cdots, 24\}$. Table 2 give all the possible values of $\overline{y}$ and $\overline{d}$ and figure 4 shows it in a graphic way:

$$\overline{y}_{\mu_n} = 222 - 2 \cdot \overline{d}_{\mu_n} \tag{9}$$

In table 2 note that μ_7 have the same mean distance of the allocation obtained with the TTC algorithm, in fact, we confirm that this is indeed the allocation of TTC, along with three other allocations that return the same mean distance and therefore the same mean proficiency. In the next subsection we will explore which allocation is obtained from the GS algorithm.

[18] The parameters used for the simulations in this study were obtained from regressions with data of PROEB and *Censo Escolar* MEC/INEP. The PROEB contains the proficiency of each student and also some personal characteristics like age, race, years of schooling of the parents, items of permanent consumption at home, and many others, totaling 32 variables. The *Censo Escolar* have other 26 variables for students and we used also an administrative databank of BH municipality which have other 9 exclusive variables for schools. For scope concerns, these regressions are not presented in this paper, those interested in these aspects should contact the first author for more details.

[19] We discuss the realism of these assumptions on the Conclusion, section 5.

Table 2 Relation between all allocations of the grid city example and distance and proficiency

Allocation	Mean Distance	Mean Proficiency	Allocation	Mean Distance	Mean Proficiency
μ_1	6.5	209	μ_{13}	5.0	212
μ_2	6.5	209	μ_{14}	5.0	212
μ_3	6.5	209	μ_{15}	4.5	213
μ_4	6.5	209	μ_{16}	4.5	213
μ_5	6.0	210	μ_{17}	4.5	213
μ_6	6.0	210	μ_{18}	4.5	213
μ_7	5.5	211	μ_{19}	4.5	213
μ_8	5.5	211	μ_{20}	4.5	213
μ_9	5.5	211	μ_{21}	4.5	213
μ_{10}	5.5	211	μ_{22}	4.5	213
μ_{11}	5.0	212	μ_{23}	3.5	215
μ_{12}	5.0	212	μ_{24}	3.5	215

Source: Simulated data from grid city example.

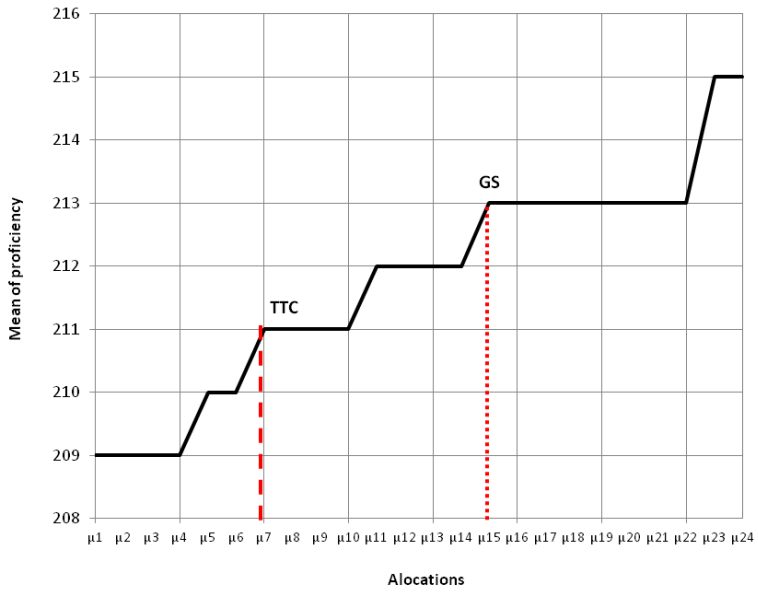

Fig. 4 Relation between all allocations of the grid city and the students' average proficiency.

3.1.2 Obtaining the allocation by the GS algorithm

In this simulation the students preferences ($\wp_i$) still the same of the previous subsection, but to obtain a simulation of the GS algorithm similar to the current BH

educational allocation system we have to change the preferences of the schools (or priorities), now the schools have priorities over the students accordingly to their distance (and only the distance), let's call this new preferences profile as $\wp_s^{'}$:

$$\wp_s^{'}: \begin{cases} s_1 : i_2 \succ i_1 \succ i_3 \succ i_4 \\ s_2 : i_1 \succ i_3 \succ i_2 \succ i_4 \\ s_3 : i_3 \succ i_1 \succ i_2 \succ i_4 \\ s_4 : i_3 \succ i_2 \succ i_1 \succ i_4 \end{cases} \tag{10}$$

Given $\wp_i$ and $\wp^{'}s$, we could also obtain the allocation of GS.To build the preferences $\wp_s^{'}$ we used the D matrix in equation 3 and we had to make a tiebreaker criteria for the distances. We place always one step ahead the students with smallest index. However, in this example, any other criterion would produce the same results in terms of final allocation. We should note that TTC and GS using $(\wp_i, \wp_s)$ and $(\wp_i, \wp^{'}s)$ produce equivalent allocation in each example μ and $\mu^{'}$, but this is not necessarily true for the general case. Suppose:

$$\wp_i^{''}: \begin{cases} i_1 : s_1 \succ s_2 \succ s_3 \succ s_4 \\ i_2 : s_1 \succ s_4 \succ s_3 \succ s_2 \\ i_3 : s_2 \succ s_1 \succ s_3 \succ s_4 \\ i_4 : s_4 \succ s_2 \succ s_3 \succ s_1 \end{cases} \quad \wp_s^{''}: \begin{cases} s_1 : i_4 \succ i_3 \succ i_1 \succ i_2 \\ s_2 : i_2 \succ i_4 \succ i_1 \succ i_3 \\ s_3 : i_4 \succ i_1 \succ i_2 \succ i_3 \\ s_4 : i_3 \succ i_2 \succ i_1 \succ i_4 \end{cases}$$

Each school with just one seat. These two give different allocation in GS and TTC. Following the GS algorithm of page 11 we obtain a new allocation μ' (represented by μ_{15} in table 2). In this example the GS algorithm managed to shorten the distances from students to school, the average distance of μ' is 4.5 kilometers:

$$\mu' = \begin{array}{c} \\ i_1 \\ i_2 \\ i_3 \\ i_4 \end{array} \begin{array}{c} \begin{array}{cccc} s_1 & s_2 & s_3 & s_4 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

From the students point of view we see that i_3 and i_4 remains with the same contentment as before in TTC, but i_1 have a increased a lot of her welfare (from her third option to the first), in the other hand, i_2 lost your first option, getting the third. By making the priorities of schools closer to the distance criterion (one of the variables also considered by parents and students) the BHM approached at least one common interest of these two sides of the mechanism, but that's not guaranteed that this was an improvement for all participants in the mechanism (just that it is stable).

From table 2 and figure 4 is possible to identify two allocations that minimizes the distance (We can identify these allocations computationally), $\mu_{23} = \{(i_1,s_2);(i_2,s_1);(i_3,s_4);(i_4,s_3)\}$ and also $\mu_{24} = \{(i_1,s_3);(i_2,s_1);(i_3,s_4);(i_4,s_2)\}$. But accordingly to the preferences $\wp_i$ and $\wp_s^{'}$, both of these allocation have (i_3,s_3) as a blocking pair, therefore, the allocation that minimizes the distance is not stable.

In this subsection we presented the general rules that we use to simulate the BHM. The TTC algorithm simulates a previous situation to the current mechanism, in which parents and students look to enroll in schools according to their preferences and could find a place in queue by chance. The TTC simulates our baseline since no previous proficiency study was done before the intervention of the school enrolment mechanism and the GS algorithm proposed (with schools giving priority to students who live closer) will serve to make a comparison with the current observed allocation.

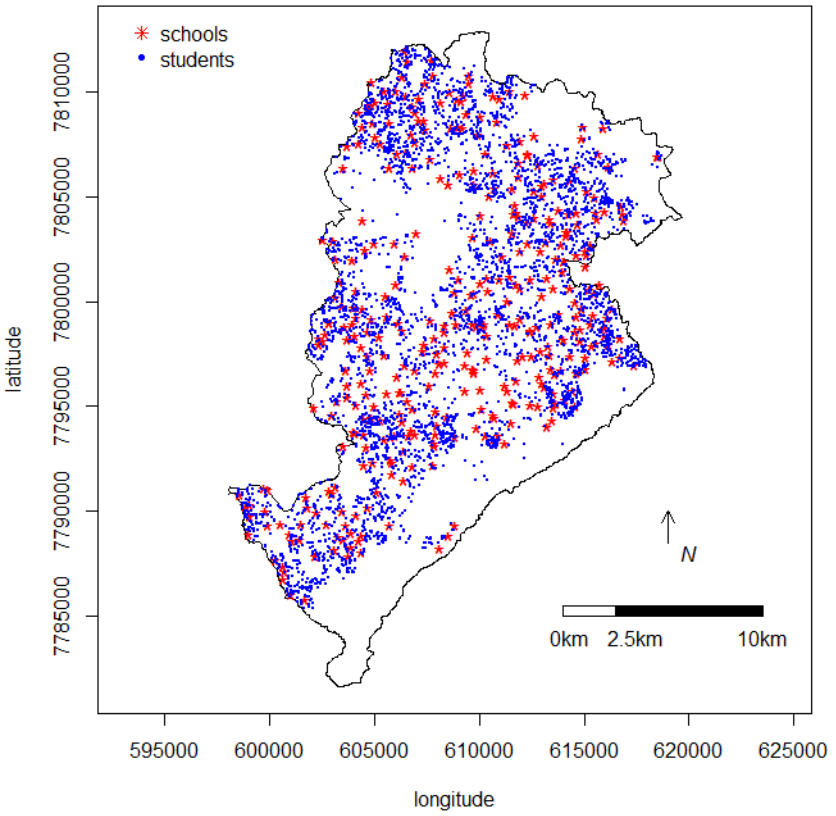

Fig. 5 Students and schools geocoded in Belo Horizonte municipality. The schools are all represented by red asterisks and a sample of students is represented by blue bullets.

3.2 Simulation for observed Data of BH

Knowing the procedures performed in subsection 3.1 we can now extrapolate this exercise to the real data of Belo Horizonte city. To accomplish this we need a geocoded database of students and schools in the city. Fortunately BH have a geocoding system quite exemplary, all public schools present in the city are georeferenced (see [7]). Students were geocoded from the available address in PROEB database and using the method applied by [8] consisting of geocoding from given addresses.

The initial 5th grade PROEB database possessed 27,195 students, after cleaning the data to merge students between PROEB and educational census we obtained a lower initial number of 22,964 students. It was also necessary to geocode all these students, after the parsing procedure we obtained 16,354 students accurately localed. Figure 5 shows a map of BH with all geocoded schools of the database and one sample with one third of geocoded students.[20]

The idea here is that from geocoded database with students' grades, compare three allocations: μ_{TTC}, μ_{GS} and μ_{obs}. The first allocation is our *baseline*, trying to simulate the situation prior to the mechanism, the second one is the GS algorithm, to verify if the observed allocation is in the core, the latter allocation is the one observed that was held in 2010 for students of 5th grade.

4 Results

Here we present the simulation results obtained with the sample of 16,354 students and 296 schools of BH. The main simulations presented here are based on lexicographical preferences described by the equation 1 and 2.[21] To simulate how the student's average score will behave we use an equation similar to equation 9. In fact, regressions estimated for these simulations were based on the linear-log models, we took the logarithm of average distance to obtain equation 11, but after that we put the parameters linearly for greater clarity of exposition.

$$\bar{y}_{\mu} = 229 - 3.7 \cdot (\bar{d}_{\mu}) \tag{11}$$

The latent proficiency is now $\upsilon = 229$ and $\gamma = -3.7$, this parameter was obtained after estimations controlled by several variables of student's characteristics and some of his/her family. So, for each kilometer added to the distance traveled by the student he/she loses 3.7 proficiency points.

[20] For better visualization we put a sample of students, otherwise the excess of blue spots could hide the schools.

[21] As previously argued on page **??**, many others simulations are presented in the original work [**?**], in all of them the only thing that change are the preferences of parents/students. The Schools' priorities is kept fixed, with the exception of the TTC simulation, where the priorities are random, as explained in subsection 2.2.

As our main results we have that the difference in proficiency given by the difference between TTC and BH allocation is very important. By the simulation we show that the BH system aggregates 8.8 points in average proficiency. Since proficiency standard deviation is around 45 points, this is a small change of less than 20% of standard deviation, but considering that students gain only one fourth of a standard deviation from changing one grade to another, this result looks promising, in the situation prior to the current BH's allocation mechanism, would be as if students were still in the 4th grade only because of the allocation.

Table 3 Allocations simulated with $\wp_i$ and $\wp_s^{'}$ to obtain average distance and proficiency.

Type	Average Distance	Simulated Average Proficiency	Difference in Proficiency
TTC (*baseline*)	6.6	204.58	-8.88
BH (*observed*)	4.2	213.46	-
GS	3.9	214.57	+1.11

Other result is that by linear approximation of proficiency, we see that the current mechanism is close to the allocation produced by the allocation produced by GS algorithm, this is because the average distance is relatively close, a little bit smaller in the GS algorithm, but we could conclude that the actual mechanism is very close to one with parents choosing the nearest school among those of best quality (level of proficiency).

There are some blocking pairs in current allocation (*observed*) but we can say that the current system is relatively stable when it comes to real purposes. In general, less than 10% of parents of enrolled students come to ask for a review of the system or fill a formal declaration. Otherwise it would be difficult to restrain population protests over 20 years of operation of this mechanism.

Another simulation was performed assuming the creation of ten new seats in the 30 best performance schools (that is close to 14% of the simulated seats). Thus schools with higher proficiency level now have more seats, supporting more students.

Figure 6 presents this results, the school demand was constructed with the new demand over previous demand ($D = \frac{n'_s}{n_s}$), where n'_s is the new number of students at school s and n_s is the former number of students. If $D < 1$ the school have excess of capacity, for $D = 1$ the use of this school is the same, and $D > 1$ is for over demanded schools.

In Figure 6 we see this by the red line, 40 of the old schools would be completely empty in this new simulation, 14 are underused, all the 30 best performance schools have their vacancies filled. This simulation shows that the creation of new seats in the best performing schools have to be preferred, however, our schooling experience

may recommend caution with this recommendation because the school may not be ready to operate with the same quality in a new scale.

Finally we must point out that since minimize the distances is an important element of our simulation, there are other allocations that would minimize further distances and this could get better results in proficiency, however, such allocations may be more unstable (with a higher proportion of blocking pairs) and we would need more developments in future works to verify more hypotheses.

5 Conclusion

If our assumption that students of 5^{th} are penalized when studying away from home is right, we can conclude that the system of Belo Horizonte must have helped to improve student performance on average about 5% (8.8 points in mean proficiency). To build the BH's allocation model, we had to start from very general assumptions which we believe that are feasible (see subsection 2.1). Parents prefers schools with good average proficiency scores (maybe small differences are negligible, but not differences of the proficiency grade level in our model). Between schools with the same score level, parents and students prefers the most closest school. So, we have supposed lexicographical preferences for parents/students and employed the TTC algorithm to simulate the situation previous the actual mechanism (since no

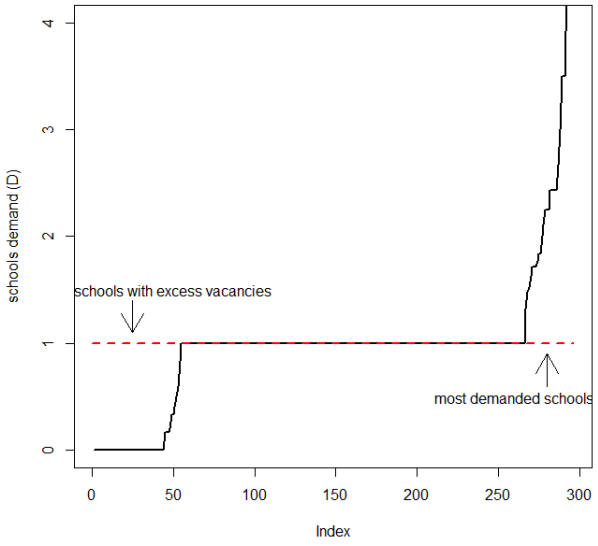

Fig. 6 School occupancy with more new seats

monitoring was done previously the mechanism implementation). And, finally, we consider that the actual real mechanism is very similar to the GS algorithm.

In fact, as we saw in section 4, the actual allocation in BH is not stable, so it is not really given by a GS algorithm, the real mechanism was not built based on the GS algorithm, it does not ask the parents' preferences and only provides some rules for schools priorities, but not a list of the preferences.

So, the model adopted here have some simplifying assumptions that can be relaxed in subsequent works, specially, we could investigate how parents classifies the schools, and how they differ. Also, perhaps an ideal education system should consider class compositions, a school may want that their students came form different parts of the city or with various levels of income, or proficiency, this class composition priorities was not considered in our present work.

Indeed, the model presented in this paper can be added with endless possibilities, other variables related to the allocation, not only those related to the distance, can be added. Geographical obstacles can be incorporated to fit the model to the reality of school districts, and not only considering rectilinear (subsection 3.1) or euclidian (subsection 3.2) distances.

We consider that it is important that we have sought for a practical application of matching which not only seeks to verify its stability and Pareto optimality, but also seeks a way to relate the BH's mechanism with the results of students in proficiency. We hope that this work can therefore contribute to understanding the mechanisms of allocating students operating in Brazil and in the worldwide.

References

1. Atila Abdulkadiroğlu, Parag A Pathak, and Alvin E Roth. The new york city high school match. *American Economic Review*, pages 364–367, 2005.
2. Atila Abdulkadiroglu and Tayfun Sönmez. School choice: A mechanism design approach. *The American Economic Review*, 93(3):729–747, 2003.
3. Luís Carlos Martins Abreu. Mecanismos de seleção de gale-shapley dinâmicos em universidades brasilieiras; sisu, sisu (alpha) e sisu (beta). 2013.
4. Clodoveu DAVIS Jr and SA de Belo Horizonte. Gis: dos conceitos básicos ao estado da arte. *Espaço BH*, (1), 1997.
5. Clodoveu A Davis Jr and Frederico T Fonseca. Assessing the certainty of locations produced by an address geocoding system. *Geoinformatica*, 11(1):103–129, 2007.
6. Gustavo Andrey Fernandes. O sistema de matrícula escolar de são paulo: uma abordagem à luz da teoria dos jogos, 2007.
7. Frederico T Fonseca, Max J Egenhofer, Clodoveu A Davis Jr, and Karla AV Borges. Ontologies and knowledge sharing in urban gis. *Computers, Environment and Urban Systems*, 24(3):251–272, 2000.
8. Frederico Torres Fonseca and Carlos André Zuppo. School pre-registration and student allocation. *URISA Journal*, 1:30–40, 1994.
9. David Gale and Lloyd S Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, pages 9–15, 1962.
10. Marcus Vinícius Pinto. Cadastramento escolar: democratização do acesso à escola pública. *Informática Pública*, 1(2):139–156, 1999.

11. Alvin E Roth and Marilda A Oliveira Sotomayor. *Two-sided matching: A study in game-theoretic modeling and analysis*. Number 18. Cambridge University Press, 1992.
12. Lloyd Shapley and Herbert Scarf. On cores and indivisibility. *Journal of mathematical economics*, 1(1):23–37, 1974.
13. Marilda Sotomayor. Mecanismos de admissão de candidatos às instituições. modelagem e análise à luz da teoria dos jogos. *Brazilian Review of Econometrics*, 16(1):25–63, 1996.
14. Chung-Piaw Teo, Jay Sethuraman, and Wee-Peng Tan. Gale-shapley stable marriage problem revisited: Strategic issues and applications. *Management Science*, 47(9):1252–1267, 2001.