Rationalizing mistakes in school choice mechanisms

Autor: Izquierdo Nofuentes, Alex (Graduado en Economía).
Público: Comunidad Educativa. **Materia:** Diseño de Mercado. **Idioma:** Inglés.

Title: Rationalizing mistakes in school choice mechanisms.

Abstract
Transactions take place in a variety of situations where money cannot be used as the tool to match supply and demand, such as when assigning hospital beds to patients or students to school/college places. In these situations, different mechanisms must be devised so that the most efficient allocation of these very scarce resources could be made. This paper focuses on school choice mechanisms, and carries out an econometric study about the specific mechanism used in Sabadell. Furthermore, it tries to explain the reasons that cause people to make mistakes when submitting their list of school preferences for their three-year-old children.
Keywords: Economics, Market Design, Matching, School Choice, School Allocation, Statistics, Econometrics

Título: Racionalizando errores en mecanismos de asignación escolar.

Resumen
Las transacciones ocurren en muchas situaciones donde el dinero no se puede utilizar como herramienta para emparejar oferta y demanda, como asignar camas de hospital a pacientes o estudiantes a plazas escolares. En estas situaciones, se deben idear diferentes mecanismos para que se pueda hallar la asignación más eficiente de estos escasos recursos. Este artículo se centra en los mecanismos de asignación escolar, y lleva a cabo un estudio econométrico sobre el mecanismo específico usado en Sabadell. Además, trata de explicar las diferentes razones que causan a la gente cometer errores cuando solicitan escuelas para sus hijos de 3 años.
Palabras clave: Economía, Diseño de Mercado, Emparejamiento, Asignación Escolar, Estadística, Econometría.

Recibido 2017-07-18; Aceptado 2017-08-01; Publicado 2017-08-25; Código PD: 086041

1. INTRODUCTION

School choice mechanism consists in the process through which every year new students are assigned to the available seats in the schools at their disposal. Abdulkadiroğlu and Sönmez provide this definition:

> In a school choice problem there are a number of students, each of whom should be assigned a seat at one of a number of schools. Each school has a maximum capacity but there is no shortage of the total seats. Each student has strict preferences over all schools, and each school has a strict priority ordering of all students. Here, priorities do not represent school preferences but they are imposed by state or local laws (2003: 733).

This problem is very similar to the *college admission problem* introduced by Gale and Shapley (1962), which has been extensively studied, with the difference that "in school choice, schools are objects to be 'consumed' by the students, whereas in college admissions, schools themselves are agents who have preferences over students" (Abdulkadiroğlu and Sönmez, 2003: 733).

To solve this problem, different mechanisms have been studied. The decision of which one to apply must be carefully pondered, since the implications will determine the specific final outcome, and therefore the environment in which the children will be raised. As Abdulkadiroğlu and Sönmez put it, "the outcome of a school choice problem is an assignment of schools to students such that each student is assigned one school and no school is assigned to more students than its capacity. We refer each such outcome as a matching" (2003: 733).

However, it is not always possible to assign to all students their top choice, and for this reason literature tries to find the most efficient possible outcome. In this context, "[a] matching is Pareto efficient if there is no other matching which assigns each student a weakly better school and at least one student a strictly better school" (Abdulkadiroğlu and Sönmez, 2003: 733).

In this paper, I explain the Boston Mechanism (which is the one used in Sabadell), comparing it with other mechanisms, and I analyze the data obtained, where we can see consistency with the previous empirical evidence. Finally, I try to rationalize people mistakes in a specific scenario, looking for their cause and explaining it.

2. LITERATURE REVIEW

School choice has been widely studied, and different mechanisms have been proposed in order to be used to allocate students to school places. I will briefly explain the main ones. Then, I will point out the different criticisms directed to the Boston mechanism (BM) and why it should be changed, and under which circumstances this mechanism works more efficiently.

2.1 School choice mechanisms

2.1.1 Boston Student Assignment mechanism

The so-called Boston mechanism has been used in the city of Boston (and some others) since July, 1999, and it is the one used in Sabadell too, as I will explain later. According to Abdulkadiroğlu and Sönmez (2003), the mechanism works as follows:

1. Each student submits a preference ranking of the schools.
2. For each school a priority ordering is determined. Students in the same priority group are ordered based on a previously announced lottery.
3. The final phase is the student assignment based on preferences and priorities:

Round 1: In Round 1 only the first choices of the students are considered. For each school, consider only the students who have listed it as their first choice and assign seats to these students following their priority order until either there are no seats left or there is no student left who has listed it as her first choice.

Round 2: Consider the remaining students. In Round 2 only the second choices of these students are considered. For each school with still available seats, consider only the students who have listed it as their second choice and assign the remaining seats to these students one at a time following their priority order until either there are no seats left or there is no student left who has listed it as her second choice.

In general, at

Round k: Consider the remaining students. In Round k only the kth choices of these students are considered. For each school with still available seats, consider the students who have listed it as their k^{th} choice and assign the remaining seats to these students one at a time following their priority order until either there are no seats left or there is no student left who has listed it as her k^{th} choice.

The mechanism finishes when each student is assigned to a seat.

2.1.2 Student-Proposing Deferred Acceptance mechanism

As aforementioned, the school choice problem is closely related to the college admissions problem. Thus, school priorities can be interpreted as college preferences and a version of the *Gale-Shapley deferred acceptance algorithm* (Gale and Shapley, 1962) can be applied. The mechanism works as follows (Abdulkadiroğlu and Sönmez, 2003: 735):

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.

In general, at

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 process finishes when no student is rejected and each student is assigned to one seat.

2.1.3. Top Trading Cycles mechanism

The mechanism works as follows (Abdulkadiroğlu and Sönmez, 2003: 735):

Step 1: Assign a *counter* for each school which keeps track of how many seats are still available at the school. Initially set the counters equal to the capacities of the schools. Each student points to her favorite school under her announced preferences. Each school points to the student who has the highest priority for the school. Since the number of students and schools are finite, there is at least one cycle. (A *cycle* is an ordered list of distinct schools and distinct students (s_1, i_1, s_2, ..., s_k, i_k) where s_1 points to i_1, i_1 points to s_2, ..., s_k points to i_k, i_k points to s_1). Moreover, each school can be part of at most one cycle. Similarly, each student can be part of at most one cycle. Every student in a cycle is assigned a seat at the school she points to and is removed. The counter of each school in a cycle is reduced by one and if it reduces to zero, the school is also removed. Counters of all other schools stay put.

In general, at

Step k: Each remaining student points to her favorite school among the remaining schools and each remaining school points to the student with [the] highest priority among the remaining students. There is at least one cycle. Every student in a cycle is assigned a seat at the school that she points to and is removed. The counter of each school in a cycle is reduced by one and if it reduces to zero the school is also removed. Counters of all other schools stay put.

The mechanism finishes when all students are assigned to a seat in one school.

2.2 Critiques to the Boston Mechanism

BM has been widely criticized due to not being strategy-proof. This means that sometimes the best strategy for a student is not to reveal their real preferences, (i.e. not listing the schools in the order they prefer), but to strategize by making a list that combines their preferences for schools and their probability of being accepted in them. In other words: "if a student does not gain admission to his first choice school, it may be that his second choice is already filled to capacity with students who listed it as their first choice. That is, a student may fail to get a place in his second choice school that would have been available had he listed that school as his first choice" (Abdulkadiroğlu, Pathak, Roth and Sönmez, 2006: 2).

In order to be able to maximize their probability of entering in a particular school, first the student must learn how the mechanism works. Pathak and Sönmez stated that students who have learned this are called "sophisticated students" (2008: 1639), while the ones that do not know it (and therefore list the schools according to their true preferences) are called "sincere students" (2008: 1639). "The strategy space of each sincere student is a singleton under the Boston game. Each sophisticated student, on the other hand, recognizes the strategic aspects of the student assignment process, and the support of her strategy space is all strict preferences over the set of schools, plus remaining unassigned" (Pathak and Sönmez, 2008: 1639). Therefore, sophisticated students are the only ones who play the game, deciding the final outcome for everyone. Since becoming a sophisticated student implies a cost that not everyone is capable of assuming, the mechanism has been accused of being unfair, and this has been the main reason of it being replaced in some cities. "Boston Public Schools stated that one of their main rationales for changing their student assignment system is that it levels the playing field. They identified a fairness rationale for a strategy-proof system. In this paper, we examined this intuitive notion and showed that the Boston mechanism favors sophisticated parents at Pareto-dominant Nash equilibrium" (Pathak and Sönmez, 2008: 1639).

2.2.1. Empirical evidence

Thanks to a study carried out in the city of Boston, we can see that under BM, most students are accepted into the schools marked as first choice, so some people could state that the mechanism is performing adequately and therefore satisfies students and families. "However, given the incentives of the Boston mechanism, treating stated choices as true choices does not give an accurate depiction of the performance of the mechanism. It would be a mistake, for instance, to conclude that 80% of students in Boston are satisfied with their assignment based on numbers that might not reflect the true preferences" (Abdulkadiroğlu et al, 2006: 13).

This happens because a majority of students decides to apply for their guaranteed option, i.e. the school they have already a seat in, as their first choice. "At the elementary school level, about 16% of students are assigned to their

guaranteed choice. This fraction increases to 29% and 52% at the middle and high school level" (Abdulkadiroğlu et al, 2006: 13).

Under BM, it makes no sense to rank an over-demanded school as their second choice or lower ("[f]or a given year, define a school to be [over-demanded] if the number of students who rank that school as their first choice is greater than the number of seats at the school") (Abdulkadiroğlu et al, 2006: 14). Evidence shows that students understand this issue, with only a few of them ranking over-demanded schools as their second or third choice. Abdulkadiroğlu et al (2006) conclude that, in elementary school, only a 36% of them rank an over-demanded school as their second choice, and only a 28% as their third choice.

This percentage of people could be understood as the ones who are being "sincere students" and, therefore, revealing their true preferences and not strategizing the game. These students are the ones who will be greatly benefited from changing the mechanism used to one that was strategy-proof.

2.3. Support to the Boston Mechanism

Criticisms against BM are based on its lack of efficiency and fairness, and that is the reason why the Deferred Acceptance mechanism (DA) has been suggested to substitute it. However, the analysis in which DA shows better properties than BM is based on a very strong assumption: schools have strict priorities over students, and thus there is no uncertainty about the final outcome (there are no ties that must be broken randomly).

In reality, schools do not have strict priorities but "weak priorities" (Miralles, 2008: 3). There are only a few categories, and consequently this fact produces several ties that must be broken randomly, rendering the assumption of full information (on which the efficiency of DA is based) impossible. According to Abdulkadiroğlu, Che and Yasuda, "[t]his makes the assumption of full information particularly problematic. Not only is it unlikely for students to know others' preferences, but it is simply impossible for them to know others' - even their own - priorities at schools if they are chosen randomly after students submit their rankings" (2009: 3).

Moreover, taking into account that schools present weak priorities and students have similar ordinal preferences (since families value schools based on similar qualities, such as reputation or the neighborhood they are in), concepts such as Pareto efficiency lose relevance. As Miralles put it: "DA performs very poorly if students' ordinal preferences are perfectly correlated. Precisely because DA is strategy-proof, it cannot make any distinction among students if all of them share identical ordinal preferences" (2008: 4).

While DA resolves ties using random lotteries, without taking into account their "cardinal preferences" (Abdulkadiroğlu et al, 2009; Miralles, 2008), i.e., the intensity with which every student prefers each school to the others, BM makes its strategizing component useful to solve ties efficiently. We can see a very useful example to understand this in Abdulkadiroğlu et al (2009: 5). Hence, obtaining strategy-proofness involves a loss in the total welfare, and the other way around, making it difficult to choose between either mechanism.

As far as fairness is concerned, naïve students benefit from BM in the sense that they have more possibilities to enter top-popular schools, since sophisticated students will not rank them in their lists. Furthermore, to help the most disadvantaged and not lose BM's efficiency, a correction could be made: "[c]orrected reported rankings would remove schools with no remaining slots to last positions. When all students are sophisticated, this correction is innocuous. Simulations show that this device works fine for naïve students while largely preserving overall efficiency" (Miralles, 2008: 5).

3. THE DATA

3.1. Data analysis

In order to construct the model, I have used data provided by the "Departament d'Ensenyament de la Generalitat de Catalunya", concerning the admission process for seats in P3 (first year of elementary school for 3-year-old children) for the academic year 2016-2017 in the schools of Sabadell. In this data, we can observe the list of school preferences that parents have submitted, as well as their final allocation, the school priorities related to every student, and the offer and demand for every school for the last 3 years.

With this data, I was able to detect different mistakes which people committed when submitting their list of preferences:

- **Mistake 1**: when people rank an over-demanded school (i.e. a school that had more applications than available seats in the k^{th} round) as their k^{th}+n choice. To detect this mistake, first I created a table with schools and rounds in order to check in which round the available seats in the school were 0 (i.e. in which round the school became over-demanded); afterwards, I created an algorithm (which was replicated) that showed which parents had committed this mistake. This is the mistake committed by most parents.

- **Mistake 2**: when people list very few schools, taking great risk unnecessarily. The detection of this mistake has been difficult, in the sense that people who committed this mistake could obtain two different outcomes:
    - They could obtain a seat in one of the schools ranked, so no observable mistake would be detected.
    - They could not get any seat in any of the schools listed, which entails that they are assigned in one of the schools with available seats when the rest of students have been allocated. This was easy to detect with the data available, and, thus, this outcome has been the one taken into account.

- **Mistake 3**: when people list a school in which they have less priority than they need to enter. This mistake also has a strategic component, since, after the main process, they are put into the waiting list and the claiming process begins, in which some irregularities can be detected, whereby the people in the waiting list could be assigned a seat. Two cases were observed in which someone obtained a seat with less priority that someone who did not get it. For this reason, these two cases were not considered, as it is not clear if it could be marked as a mistake or not.
- **General Mistake**: when parents have committed any or several of the aforementioned mistakes.

As mentioned before, the main mistake committed by parents is the mistake 1, although it has consequences (i.e. people could have obtained a better seat if they had not committed this mistake) in very few cases. Tables 1 and 2 illustrate some information about the data.

Table 1: Number of people who committed the mistake

Sample size: n=1923	People who committed the mistake	People who were affected by the mistake
Mistake 1	664	52
Percentage	34.45	2.70
Percentage conditional to mistake		7.83
Mistake 2	58	58
Percentage	3.01	3.01
Percentage conditional to mistake		100
Mistake 3	11	9
Percentage	0.57	0.47
Percentage conditional to mistake		81.82
General Mistake	688	80

Percentage	35.74	4.16
Percentage conditional to mistake		11.63

Source: Own elaboration from the data provided by the "Departament d'Ensenyament".

Table 2: Number of people who entered each option

Sample size: n=1923	Number of people who entered in each option	%
1st Option	1803	93.76
2nd Option	36	1.87
3rd Option	14	0.73
4th Option	9	0.47
5th Option	2	0.1
6th Option	1	0.05
Assigned to default option	58	3.02

Source: Own elaboration from the data provided by the "Departament d'Ensenyament".

Similarities can be observed with regard to the results published in previous empirical studies (Abdulkadiroğlu et al, 2006), in which the number of people who ranked an over-demanded school was 36%, similar to the 34.45% found in this study. In the same sense, it could be understood that the majority of families are satisfied with the outcome of the mechanism, since almost 94% of students are admitted into their top choice (however, taking into account how the mechanism works, it cannot be concluded that their top choice is their most preferred one).

3.2. The Boston mechanism in Sabadell

In Sabadell, students are assigned seats at public schools through a centralized student assignment mechanism, very similar to the so-called Boston mechanism. In the spring of each school year, the following groups of students (or their families) are required to submit a preference ranking of schools:

- Students who seek a spot in 1st year of primary school (students who turn 6 years old during that scholar year).
- Students who seek a spot in 1st year of high school (students who turn 12 years old during that scholar year).

A preference ranking of schools may also be submitted by:

- Students who seek a spot in the 1st year of elementary school (students who turn 3 years old during that scholar year).
- Students who seek a spot in kindergarten.

Students in the remaining non-transition courses are not required to submit a list of schools unless they request a transfer.

In Sabadell, students are allowed to rank a maximum of 10 schools. Taking into account the preference that the mechanism gives to first choices, it makes sense that only 1.2% of students listed up to 10 schools (and none of them were given their 10th option).

For the 1st year of elementary school (which has been the one studied in this paper), students are assigned a random number, which will be used to break ties. When the number of applications is greater than the number of seats in the school, students are ordered on the basis of how many points are collected (this procedure is related to the first school listed, and is kept for the rest of schools). Points are earned as follows[97]:

- General criteria:
    - When the student has a sibling studying in the same school, or a parent works in that school: **40 points**.
    - When the address is near the school, or the place where the father/mother/legal tutor works is near the school:
        - If the address is in the influence area of the school: **30 points**.
        - If the place where the father/mother/legal tutor works is near the influence area of the school: **20 points**.
        - If the address is in the same town but not in the area of influence of the school: **10 points**.
    - When parents or tutors receive an economic help *(renda mínima d'inserció*): **10 points**.
    - When the student, their parents or their siblings have a disability ≥ 33%: **10 points**.
- Complementary criteria: These criteria are taken into account when there are ties in the general criteria.
    - Large family or single parent family: **15 points**.
    - The student suffers from a chronic disease which affects their digestive, endocrine or metabolic systems: **10 points**.
    - The father, mother, siblings or tutors have studied in the school: **5 points**.
- When there are ties after the complementary criteria have been applied, a random lottery number is used to break it.

After the process has finished, the claiming process begins, and minor adjustments may be made. This process is conducted personally for parents who request it, and thus it is not realized through a centralized mechanism.

3.3. The model

With the data provided by the "Departament d'Ensenyament de la Generalitat de Catalunya", different variables to explain the mistakes committed by families when submitting their preferences lists were created. To create the variable "Wealth", additional data provided by the "Secció del Cadastre de l'Ajuntament de Sabadell" was used. The variables are detailed as follows:

- Dist_worst[98]: distance from the address to the worst-case scenario (i.e. the school listed by the student as their last option), in minutes. This is used as a proxy of how fond the parents are of the school. The closer to the school, the more they like it. The square of this variable is also included in the model.
- Wealth: average price in €/m^2 of the cadastral island where the address is located. This is used as a proxy of the family sophistication, how much resources they have to learn correctly how to strategize.

[97] Source: "Generalitat de Catalunya. Departament d'Ensenyament." http://queestudiar.gencat.cat/ca/preinscripcio/estudis/obligatoris/documentacio/index.html

[98] We assume that the worst-case scenario for student *i* is the school listed as their last option, because we imagine that the student knows in some way that their last listed option is the worst that could happen to them if they were to be unlucky.

- Pop1415_1: average popularity between the years 2014-2015 for the school listed as their first choice.
- Pop1415_worst: average popularity between the years 2014-2015 for the school listed as their worst-case scenario.
- Points: total points earned by the student in the general criteria. The square of this variable is also included in the model.

3.3.1. Main statistics

Table 3: Main statistics

Main statistics, using observations 1 - 1923

(absent values were not taken into account)

Variables	Average	Standard deviation
Dist_worst	18,4	41,7
Wealth	517,	96,7
Pop1415_worst	-0,0475	0,226
Points	-0,0863	0,396
sq_Points	0,165	0,225

Source: Own elaboration from the data provided by the "Departament d'Ensenyament" and "Secció del Cadastre".

3.3.2. Probit

In order to study the data, a Probit model was used. However, some problems arose.

First of all, I have no access to all the data for the variable "wealth". Most of the sample analyzed comprised people living in the city of Sabadell. As mentioned before, to create the variable "wealth" for these people, data provided by the government was utilized. For people living in other cities, 3 different groups have been made:

1. For people living in the city of Barcelona, data related to the average cadastral price of the neighborhood was found in the webpage of the city hall.
2. For people living in cities similar to Sabadell, an approximation was created by using the available public data.
3. For people living in other cities/towns, no data was available. Therefore, these observations were lost.

Secondly, the variables Pop that I used to measure the quality of the schools were also an indicator of how difficult it is to be accepted into them. Thus, the possibility of making a mistake is greater when the school is a very popular one (Pop is bigger), but also decreases if the quality of your worst-case scenario is greater (Pop is bigger). I have no access to better indicators for the quality of schools (such as Pisa results by schools); thus, I utilized a Pobit model with Instrumental Variables[99] by using the variable Pop1415_1 as an instrument for Pop1415_worst, which was the best indicator I could obtain. This variable is not the best indicator for the quality of schools, so further improvements of the model need to be made.

[99] See the Appendix for more information.

Lastly, since half of the people only listed one school, the variables "Pop1415_1" and "Pop1415_worst" are strongly correlated (in fact, they are the same for half of the sample). However, I deem the other half to be important, since it consists of completely different options which are significant variables which need to be taken into account.

Table 4: Correlation matrix

Correlation coefficients, using observations 1 - 1923

(absent values were not taken into account)

Critical value at 5% (two tails) = 0.0447 for n = 1923

Pop1415_worst	Pop1415_1	
1.0000	0.5558	Pop1415_worst
	1.0000	Pop1415_1

Source: Own elaboration from the data provided by the "Departament d'Ensenyament".

With the Probit model with instrumental variables, I try to determine the reasons that caused people to rank over-demanded schools.

Several attempts have been made before finding the final outcome, provided as follows. This outcome is thought to be more accurate. However, I hope to make further improvements in the future, as more information becomes available and I develop new ideas.

3.3.3. Results

Table 4: Probit model with endogenous regressors

```
Probit model with endogenous regressors
ML, using observations 1-1884
Dependent Variable: GeneralMistake
Instrumented: Pop1415_worst
Instruments: const, Dist_worst, sq_Dist_worst, Wealth, Points, sq_Points, Pop1415_1
Parameter covariance matrix: OPG

                 Coeficiente    Desv. Típica      z      valor p
  ------------------------------------------------------------------
  const           -0.306395      0.167875      -1.825    0.0680  *
  Dist_worst      -0.00156112    0.00215299    -0.7251   0.4684
  sq_Dist_worst    2.46121e-06   8.84631e-06    0.2782   0.7808
  Wealth          -4.72352e-05   0.000305820   -0.1545   0.8773
  Points          -0.616842      0.199432      -3.093    0.0020  ***
  sq_Points        0.309088      0.153214       2.017    0.0437  **
  Pop1415_worst    1.82284       0.581883       3.133    0.0017  ***

Log-likelihood        -659.1800  Akaike criterion    1350.3599
Schwarz criterion     1439.0184  Hannan-Quinn        1383.0119
Conditional ll      -1217.819610 Cragg-Donald stat.    101.392

Overall test (Wald) = 17.1268 (6 df, p-value = 0.0088)
Endogeneity test (Wald) = 14.8633 (1 df, p-value = 0.0001)
```

Source: Own elaboration from the data provided by the "Departament d'Ensenyament" and "Secció del Cadastre".

In analyzing the results, we can discard the null hypothesis that people's mistakes are not affected by the independent variables proposed (overall test Wald), and we can observe which variables affect the possibilities of committing a mistake.

Surprisingly, with the results found, it would seem that the wealth of the family (as proxy of the family's sophistication) and the distance to the school (as proxy of how much they like the school) do not affect how well students "play" the matching game. These variables were included in the model because they were expected to be significant, as other studies have proposed (Abdulkadiroğlu et al, 2006). However, with the data available and the model constructed, they do not affect the mistakes committed by students.

The results show that the priority which students have in their top choice affects positively their possibilities of committing a mistake. In other words, the more points they have, the more likely it is for them to commit the mistake. This makes sense, since reflecting upon the strategy to play well has a cost, and paying this cost is not needed if the student has a very high priority in their top choice.

The negative sign in the square points variable indicates that this effect diminishes as the student obtains more points (if they acquire enough points to know it is almost sure they will obtain their top choice, it is not very important if they have even more than that).

The positive sign in the Pop1415_worst variable indicates that the more popular (better quality) the worst school in which the student can obtain a seat is, the more likely it is to commit the mistake. This makes sense too, since if the worst that could happen if they commit the mistake is to go to a good school, it is preferred to commit the mistake than to pay the cost of strategizing.

A second model using Pop1415_1 as an exogenous variable and CAEP_worst as an instrument for Pop1415_worst was constructed, with very similar results. In that model, Pop1415_1 is shown as not significant.

A third model was constructed, with another specification of the distance variables, by using the following dummies, as in Calsamiglia, Fu and Güell (2016).

- Dist_worse>5: dummy variable: it takes value 1 if the distance from the address to the worst-case scenario is greater than 5 minutes.
- Dist_worse>10: dummy variable: it takes value 1 if the distance from the address to the worst-case scenario is greater than 10 minutes.

With this model, we found all the distance variables not to be significant, possibly because of the interaction between the variables.

4. CONCLUSIONS

With the information provided by the model constructed, we can obtain some ideas, even though these ideas could be developed in the future, and new and better variables could be added to the model.

Students will obtain a determined level of utility, depending on the school in which they obtain a seat. This level of utility will be the highest possible if they obtain a seat in their top choice school, the lowest if they obtain a seat in their last choice school (since we assume students understand that the worst that could happen to them if they commit the mistake is going to their last listed school), and every level in-between.

When the lowest level of utility that could be attained increases, so does the probability of committing a mistake, because the potential utility gain is reduced, and, therefore, thinking of a better strategy which prevented the mistake from being committed (and increasing the probability of obtaining the maximum utility level) would not be cost-effective. It is for this reason that, if the quality of the worst-case scenario school increases, so does the probability of committing the mistake. This idea is consistent with what was found in the model.

On the other hand, if the distance to the worst-case scenario school increases (i.e. the school is less desirable), the probability of committing the mistake decreases.

However, the distance to the school was not significant at all in the model, in contradiction with the anecdotal evidence that other papers may suggest (Abdulkadiroğlu et al, 2006).

Furthermore, we included some control variables that measured the cost of strategizing (submitting the best possible list) for each student (inversely related to the sophistication of the student). The more sophisticated the student is, the

more incentives they have in order to submit the best possible list, and hence the probability of committing a mistake is reduced.

The sophistication of the student was measured by using an indicator of the family wealth, and in the model constructed it did not affect the probability of committing the mistake, unlike may be suggested by the anecdotal evidence present in other studies (Abdulkadiroğlu et al, 2006). This could happen because the measure of wealth utilized is not a proper measure of the sophistication needed for this "game".

Finally, the points measure the uncertainty of the utility obtained by playing a particular strategy. When the points increase, the information needed to know whether the student will enter in their top choice is reduced, and hence the probability of entering the top choice increases (the risk of not entering the top choice diminishes). Moreover, in the limit, once a certain number of points has been achieved, it may be established that the student will not make it to the second round. Thus, it does not matter what schools the student list as their $(1+n)^{th}$ choice (if these schools are over-demanded or not), increasing the probability of committing the mistake. This idea is also consistent with the results found in the model, where the square points variable was added too, and through this variable it is shown that the effect of the points in the probability of committing a mistake is smooth: when the points increase, the probability of committing the mistake increases, but the increase is smaller as points reach a high level).

APPENDIX

Endogenous Variable

When constructing the model, an endogeneity problem with the Pop1415_worst variable was found. That is the reason why a Probit model with instrumental variables was used. When constructing a Probit binary model, the signs of the variables' coefficient were not the expected (and not significant). The model without instrumental variables is provided as follows:

Table 5: Binary Probit model

Modelo 1: Probit, usando las observaciones 1-1923 (n = 1886)
Se han eliminado las observaciones ausentes o incompletas: 37
Variable dependiente: GeneralMistake
Desviaciones típicas basadas en el Hessiano

	Coeficiente	*Desv. Típica*	*z*	*Pendiente*[*]
const	-0.400328	0.165623	-2.4171	
Dist_worst	-0.00174258	0.00149829	-1.1630	-0.0006504
sq_Dist_worst	2.8157e-06	2.27331e-06	1.2386	1.05093e-06
Wealth	2.41435e-05	0.000310377	0.0778	9.01132e-06
Points	0.0350159	0.103616	0.3379	0.0130693
sq_Points	0.222274	0.15721	1.4139	0.0829617
Pop1415_worst	-0.299195	0.159604	-1.8746	-0.111672

Media de la vble. dep.	0.357900	D.T. de la vble. dep.	0.479510
R-cuadrado de McFadden	0.004678	R-cuadrado corregido	-0.001013
Log-verosimilitud	-1224.296	Criterio de Akaike	2462.592
Criterio de Schwarz	2501.387	Crit. de Hannan-Quinn	2476.879

[*]Evaluado en la media
Número de casos 'correctamente predichos' = 1217 (64.5%)
f(beta'x) en la media de las variables independientes = 0.480
Contraste de razón de verosimilitudes: Chi-cuadrado(6) = 11.5086 [0.0739]
Contraste de normalidad de los residuos -
Hipótesis nula: el error se distribuye normalmente
Estadístico de contraste: Chi-cuadrado(2) = 16.0198
con valor p = 0.000332155

Source: Own elaboration from the data provided by the "Departament d'Ensenyament" and "Secció del Cadastre".

The complete model of the Probit model with Instrumental Variables (including the first-stage regressions) is provided as follows:

Table: 6-Complete Probit model with endogenous regressors

```
Probit model with endogenous regressors
ML, using observations 1-1884
Dependent Variable: GeneralMistake
Instrumented: Pop1415_worst
Instruments: const, Dist_worst, sq_Dist_worst, Wealth, Points, sq_Points, Pop1415_1
Parameter covariance matrix: OPG

                 Coeficiente    Desv. Típica      z      valor p
  ------------------------------------------------------------------
  const           -0.306395      0.167875       -1.825    0.0680  *
  Dist_worst      -0.00156112    0.00215299     -0.7251   0.4684
  sq_Dist_worst    2.46121e-06   8.84631e-06     0.2782   0.7808
  Wealth          -4.72352e-05   0.000305820    -0.1545   0.8773
  Points          -0.616842      0.199432       -3.093    0.0020  ***
  sq_Points        0.309088      0.153214        2.017    0.0437  **
  Pop1415_worst    1.82284       0.581883        3.133    0.0017  ***

"First-stage" regressions

                Coeficiente   Desv. Típica     z       valor p
  ------------------------------------------------------------------
  const          -0.00328193   0.0264542     -0.1241   0.9013
  Dist_worst     -0.000108744  0.000371095   -0.2930   0.7695
  sq_Dist_worst   1.30321e-07  2.83160e-06    0.04602  0.9633
  Wealth         -9.95159e-06  4.81792e-05   -0.2066   0.8364
  Points          0.167339     0.0169574      9.868    5.72e-023 ***
  sq_Points      -0.0911429    0.0187148     -4.870    1.12e-06  ***
  Pop1415_1       0.302399     0.0236832     12.77     2.46e-037 ***

Log-likelihood        -659.1800  Akaike criterion    1350.3599
Schwarz criterion     1439.0184  Hannan-Quinn        1383.0119
Conditional ll      -1217.819610  Cragg-Donald stat.    101.392

Overall test (Wald) = 17.1268 (6 df, p-value = 0.0088)
Endogeneity test (Wald) = 14.8633 (1 df, p-value = 0.0001)
```

Source: Own elaboration from the data provided by the “Departament d’Ensenyament” and “Secció del Cadastre.”

School seat application form which parents must fill out

R/N:Y0388/Y155

Generalitat de Catalunya
Departament d'Ensenyament

Sol·licitud de preinscripció a ensenyaments de segon cicle d'educació infantil, educació primària i educació secundària obligatòria en centres educatius sufragats amb fons públics. Curs 2017-2018

Dades de l'alumne/a

DNI/NIE/Passaport | Nom | Primer cognom | Segon cognom

Identificador de l'alumne/a del Registre d'alumnes (RALC)[1] | Targeta sanitària individual de l'alumne/a, TSI (si l'alumne/a en té, introdueix-ne el codi alfanumèric)

Tipus de via | Adreça | Núm. | Planta | Porta

Codi postal | Municipi | Localitat | Districte

Província de residència | País de residència

Telèfon | Data de naixement | ☐ Home ☐ Dona | Nacionalitat

País de naixement | Província de naixement | Municipi de naixement

Adreça electrònica on voleu rebre informació de la preinscripció

Llengües que entén:[2] ☐ Català ☐ Castellà ☐ Cap de les dues | Germans al mateix centre:[3] ☐

Nova incorporació al sistema educatiu (alumnes estrangers que s'incorporen al sistema educatiu espanyol o que fa menys de 2 anys que hi són): ☐

☐ Alumne/a en situació d'acolliment ☐ Escolarització per mobilitat forçosa (art. 84.7 de la LOE; cal acreditar-ho documentalment)
☐ Alumne/a de menys de 18 anys emancipat ☐ Alumne/a de més de 18 anys tutelat

Dades del pare, mare o tutor/a

DNI/NIE/Passaport | Nom | Primer cognom | Segon cognom

DNI/NIE/Passaport | Nom | Primer cognom | Segon cognom

Dades escolars de l'alumne/a

Codi del centre actual | Nom del centre

Nivell d'estudis actual
☐ Llar d'infants ☐ Segon cicle d'educació infantil ☐ Educació primària ☐ Educació secundària obligatòria

Curs | Idioma estranger que estudia al centre

Necessitats educatives específiques (només si escau)
☐ Tipus A: alumnes amb discapacitat, trastorns greus de la personalitat o de la conducta o amb malalties degeneratives greus
☐ Tipus B: alumnes en situacions socioeconòmiques o socioculturals desfavorables

Plaça sol·licitada

Cal presentar una única sol·licitud amb la llista de centres ordenats per ordre de preferència.

Ensenyament | Nivell

Voleu plaça d'ofici?[4] Sí ☐ No ☐

Codi del centre	Nom
1.	
2.	
3.	
4.	
5.	
6.	
7.	
8.	
9.	
10.	

A42-V03-17

Generalitat de Catalunya
Departament d'Ensenyament

Criteris específics a l'efecte de barem

Simultaneïtat amb ensenyaments professionals de música i dansa o amb programes d'alt rendiment esportiu[5] ☐ SI ☐ No

L'alumne/a prové d'un centre i ensenyament adscrits ☐ SI ☐ No

Criteris generals a efecte de barem

Existència de germans al centre sol·licitat en primer lloc o pares o tutors legals que hi treballin[6] ☐ SI ☐ No

Domicili al·legat a l'efecte del criteri de proximitat al centre[7] (escolliu una opció)

☐ L'habitual dins l'àrea d'influència
☐ L'habitual dins el municipi, però fora de l'àrea d'influència
☐ L'habitual dins el districte, però fora de l'àrea d'influència (només per a Barcelona)
☐ El lloc de treball dins l'àrea d'influència (especifiqueu l'adreça del lloc de treball i la raó social)

Raó social	Adreça	Municipi	Codi postal

Beneficiari/ària de la renda mínima d'inserció ☐ SI ☐ No

Discapacitat de l'alumne/a, del pare, mare o germans ☐ SI ☐ No

Criteris complementaris a l'efecte de barem

Família nombrosa o monoparental ☐ SI ☐ No

Malaltia crònica de l'alumne/a que afecti el seu sistema digestiu, endocrí o metabòlic, inclosos els celíacs ☐ SI ☐ No

Alumnes que hagin tingut el pare, la mare, els tutors o els germans escolaritzats, en ensenyaments declarats actualment gratuïts i universals,[8] al centre per al qual es presenta la sol·licitud ☐ SI ☐ No

Manifestació de caràcter voluntari sobre l'opció pels ensenyaments de religió[9]

Ensenyaments de religió ☐ SI ☐ No

En el cas d'optar pels ensenyaments de religió, tipus d'opció que es vol cursar:
☐ Catòlica ☐ Evangèlica ☐ Islàmica ☐ Jueva

Declaració del pare, mare, tutor/a o alumne/a major d'edat

Nom i cognoms	En qualitat de (Marqueu l'opció corresponent)
	☐ Pare ☐ Mare ☐ Tutor/a ☐ Alumne/a major d'edat

Declaro que són certes les dades que faig constar en aquest document i que **no he presentat cap sol·licitud en cap altre centre** per al mateix ensenyament.

Lloc i data

Signatura

D'acord amb la Llei orgànica 15/1999, de 13 de desembre, de protecció de dades de caràcter personal, les dades de la sol·licitud s'incorporaran al fitxer "Escolarització d'alumnes", del qual és responsable la Direcció General d'Atenció a la Família i Comunitat Educativa. Podeu exercir els drets d'accés, rectificació, cancel·lació i oposició mitjançant un escrit adreçat a la unitat esmentada (Via Augusta, 202-226, 08021 Barcelona).

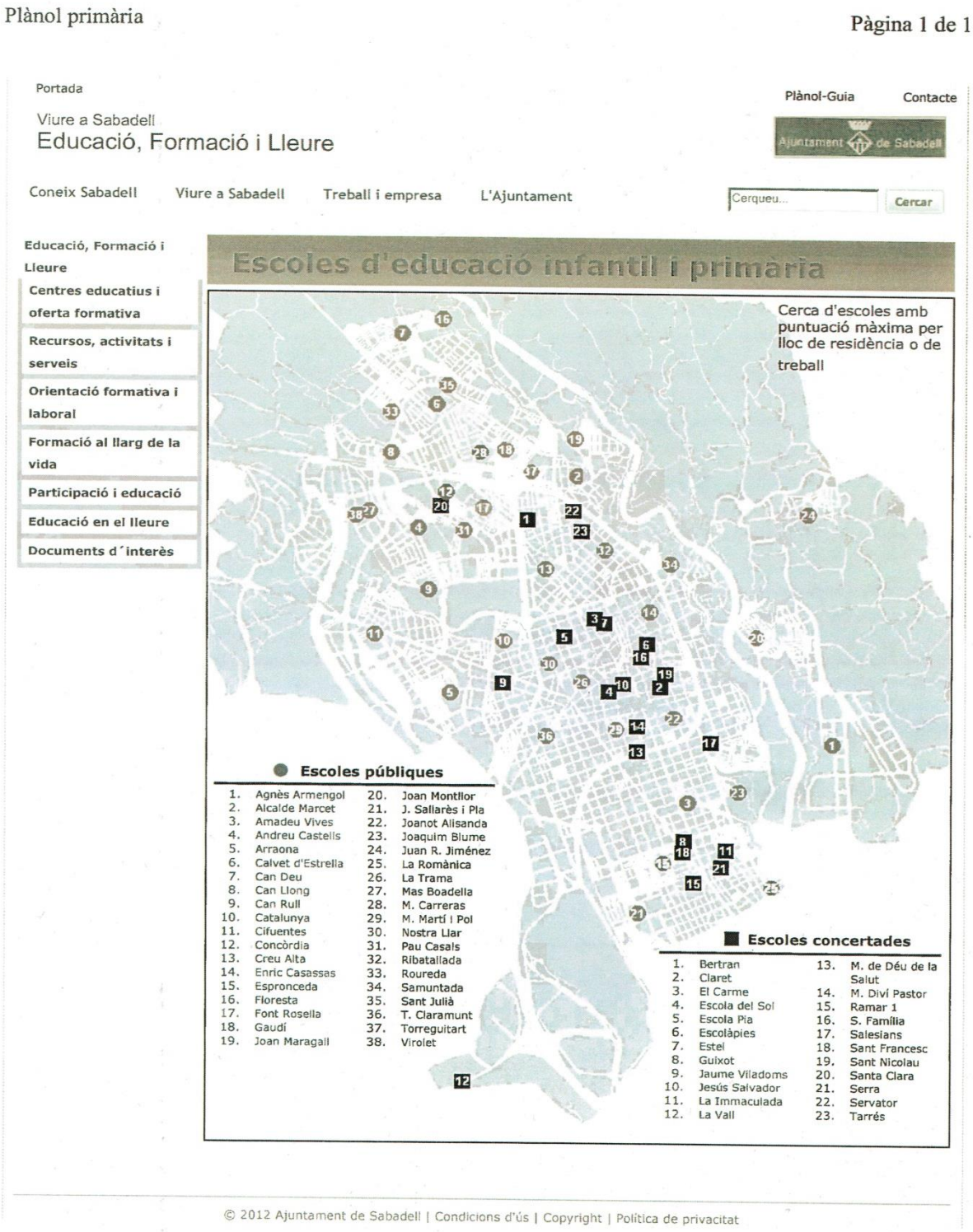
1. Tots els alumnes escolaritzats a Catalunya des del curs 2015-2016 en tenen.
2. D'acord amb l'article 21.2 de la Llei 1/1998, de 7 de gener, de política lingüística, i l'article 11.4 de la Llei 12/2009, del 10 de juliol, d'educació, en el curs escolar en què els alumnes inicilin el primer ensenyament, els pares, mares o tutors dels alumnes la llengua habitual dels quals sigui el castellà poden sol·licitar, en el moment de la matrícula, a la direcció del centre en què siguin admesos, que els fills rebin atenció lingüística individualitzada en aquesta llengua.
3. Si dos o més germans presenten la sol·licitud al mateix centre per al mateix nivell educatiu, cal marcar la casella a totes les sol·licituds.
4. En els ensenyaments de P3 i 1r d'ESO s'assigna plaça d'ofici si no s'ha obtingut plaça en cap dels centres sol·licitats. Per a la resta d'ensenyaments, si es vol l'assignació d'ofici cal indicar-ho a la sol·licitud. Si no s'indica, s'entén que no es demana.
5. El/la sol·licitant que vulgui obtenir una plaça a l'educació secundària obligatòria (ESO) i curs/i (o hagi de cursar en el proper curs 2017-2018) ensenyaments professionals de música o dansa o participi en programes d'alt rendiment esportiu, ho ha d'indicar a la sol·licitud de preinscripció.
6. D'acord amb l'article 10.1 del Decret 75/2007, aquest criteri també s'ha de marcar si es té un germà/ana escolaritzat en un centre públic de primària que té consideració de centre únic amb el centre on es demana plaça, és a dir, quan només està adscrit a aquest centre.
7. Les dades del domicili que s'al·leguen a l'efecte de barem poden ser consultades per les persones interessades que participen en aquest procés de preinscripció en el mateix centre, ensenyament i nivell.
8. Són gratuïts i universals els ensenyaments següents: el segon cicle de l'educació infantil, l'educació primària, l'educació secundària obligatòria, els programes de qualificació professional inicial i la formació professional de grau mitjà.
9. En el moment de la matrícula cal especificar l'opció d'aquesta matèria, que serà vàlida per a tota l'etapa, sens perjudici de la possible modificació, que s'haurà de comunicar al centre, per escrit, amb prou antelació a l'inici de qualsevol curs.

A42-V03-17

2/2

Source: "Generalitat de Catalunya. Departament d'Ensenyament."
http://educacio.gencat.cat/documents/FormularisModels/CentresGestioAdministrativa/A42.pdf

Map detailing the location of preschool and primary education institutions in Sabadell



Source: "Ajuntament de Sabadell."

Bibliografía

- Abdulkadiroğlu, A., Che, Y. and Yasuda, Y. (2009). "Resolving Conflicting Preferences in School Choice: The 'Boston' Mechanism Reconsidered". *SSRN Electronic Journal.*
- Abdulkadiroğlu, A., Pathak, P., Roth, A. and Sönmez, T. (2006). "Changing the Boston School Choice Mechanism". *NBER Working Paper No. 11965.*
- Abdulkadiroğlu, A. and Sönmez, T. (2003). "School Choice: A Mechanism Design Approach". *American Economic Review*, 93(3): 729-747.
- Calsamiglia, C., Fu, C. and Güell, M. (2016). *Structural Estimation of a Model of School Choices: the Boston Mechanism vs. Its Alternatives*, unpublished manuscript.
- Miralles, A. (2009). *School Choice: The Case for the Boston Mechanism*, unpublished manuscript.
- Pathak, P. and Sönmez, T. (2008). "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism". *American Economic Review*, 98(4): 1636-1652.
- Roth, A. (2016). *Who Gets What - and Why: Understand the Choices You Have Improve the Choices You Make*. London: William Collins.