Overcoming the Rural Hospital Theorem: Compulsory Social Service Allocation in Colombia

Santiago Velez-Ferro*†

November 5, 2018

The most recent version is available here.

Abstract

A long-standing policy concern in many countries is the difficulty of filling medical positions in rural areas. In Colombia, the Ministry of Health requires newly-graduated health professionals to work in a rural or marginalized urban area for a year in order to receive professional certification. The decentralized mechanism used until 2013 to allocate graduates to slots was one that health professionals could manipulate to avoid an assignment. In 2014, a single-offer centralized mechanism that cannot be manipulated to avoid an assignment, based on Gale and Shapley's deferred acceptance algorithm, was adopted. Following a revealed preference approach, I estimate health professionals' hospital preferences using the 2014 data. Using these estimates and the fact that under the decentralized mechanism health professionals were able to avoid positions that fall below their acceptance threshold, I obtain the average marginal utility a health professional would require to accept a position by simulating the outcome had the decentralized mechanism still been in use. Then, I simulate the outcome of the centralized mechanism in the absence of the requirement that students accept the assignment determined by the mechanism. I find that, given the choice, about 30% of physicians would be left unassigned, implying that it is important for the policy's success that assignments be mandatory. One feature of the centralized mechanism is that, in the case of multiple individuals having the same priority for a particular position, the tie is broken randomly. I show that breaking the ties in favor of those who listed a specific hospital as preferred can yield welfare gains of up to 12%. Finally, I show that moving from the random lottery to a merit-based tie-break, based on the results of the examination that health professionals take at the completion of their studies, can entail welfare gains under certain conditions yet raise inequality concerns under others.

*Email: velez@econ.umd.edu

†I am grateful to Paula Jaramillo and Çağatay Kayı for their encouragement, support, and helpful comments. I thank Erkut Ozbay, Andrew Sweeting, Lawrence Ausubel, and Daniel Vincent for helpful comments, discussions and insights. Thanks to Diego Restrepo and the Department of Human Talent at the Ministry of Health for providing the data and support. All errors are mine. The views expressed here are those from the author and do not reflect the views of the Colombian Ministry of Health.

1 Introduction

The quality of health care systems depends highly on the availability of health professionals. However, rural communities tend to suffer from a lower availability of these professionals. The small population and scale makes the loss of a single health professional likely to have far-reaching consequences (NRHA Policy-Brief 2012). Rural areas lack health professionals primarily becaus health professionals try to avoid them. The allocation of medical residents in the US achieved through the National Residency Matching Program (NRMP) has shown the following pattern: "A number of hospitals, particularly Rural Hospitals, fail each year to fill as many positions as they have available, and find that a high percentage of the positions they do fill are filled by foreign medical school graduates" (Roth, 1986). In the health professionals market, a *stable* allocation is one in which there does not exist a pair of hospital and health professionals not mutually assigned and that they prefer each other. The so called "*Rural Hospital Theorem*" states that for any stable allocation the unfilled positions will be the same. Hence the solution to the vacancy of rural hospital positions cannot be found within this type of allocations. Therefore Roth (1986) concludes that "this maldistribution seems unlikely to be changed by any system that does not involve some element of *compulsion*, or some change in the relative numbers of available positions".

In Colombia, the Ministry of Health has tackled this problem through a program called Compulsory Social Service[1], hereafter CSS. The health professions that participate are Medical Lab Science, Nursing, Medicine, and Dentistry. The CSS requires health professionals to work at a hospital they assign them to for one year, 70% of which are located in rural areas, but some are located in marginalized urban ones. The compulsion element comes in through the fact that professionals in these fields require to be professionally certified by the Ministry of Health; this happens only after participating in the CSS.

Since October 2014, the CSS allocation of health professionals has been determined using a single-offer centralized mechanism based on Gale and Shapley's (1962) deferred acceptance algorithm (DA). Health professionals report their five preferred positions and also rank all of the states where they might be assigned. Based on that information, the ministry creates a ranking of all

[1] The name in Spanish is *Servicio Social Obligatorio*.

hospitals for each participant, and that ranking is used in the assignment process. Additionally, the Ministry of Health assigns a (coarse) priority order to each hospital. [2] The allocation is determined using the deferred acceptance algorithm, henceforth DA, with the rankings and priorities[3] as inputs.

To study and evaluate the performance of three dimensions of the program, this paper uses empirical estimates of the preferences. The first dimension I evaluate is the compulsory element, or the fact that participation is compulsory rather than based on individual rationality. Second, I estimate the allocative efficiency of the health professionals in the centralized mechanism. Third, I measure the consequences of moving from a single random lottery to a merit-based tie-break.

Colombia's CSS mechanism has one of the strongest compulsory elements of any centralized mechanism that has been studied in the literature. I show that compared to the previous decentralized mechanism, the centralized mechanism when adopted results in significant welfare gains. The first counterfactual answers the empirical question of whether, given the welfare gains, the program still needs the compulsory characteristic to successfully allocate health professionals. To this end I refer to the fact that in October 2012, the allocation was determined by a decentralized system of lotteries that was regularly manipulated by physicians to avoid being assigned. I simulate the outcome of the centralized mechanism in the absence of the requirement that students accept the assignment determined by the mechanism. I find that if physicians were given the opportunity of not ranking all hospitals, about 30% of the them would be left unassigned, implying that for the policy to succeed, assignments must be mandatory. This contrasts with the 6% rejection rate observed in the actual allocations.

The second counterfactuals are designed to test the allocative efficiency of the implemented mechanism. Improving the allocative efficiency, besides improving the welfare of health professionals, would also reduce the rejection rate. Since this mechanism uses the DA algorithm to determine the allocations, it is known to *respect priorities* and to be *strategy proof.* The first of these characteristic establishes that whenever there is position that a health professional prefers relative to the

[2] In this program hospitals do not report preferences of health professionals. Hence, the hospitals are not strategic agents in this setting, making this allocation problem one-sided. In the literature this is known as *school choice* (Abdulkadiroglu and Sönmez, 2003).

[3] A single random tie-break is used to break ties in priorities.

one she is assigned, that position must be assigned to someone with at least the same priority. The latter property establishes that no health professional benefits from misrepresenting his preferences. These properties, although desirable, constrain the possible allocations and (possibly) entail efficiency costs. In these counterfactuals I measure the cost in efficiency of these properties. In order to measure the cost of the *respecting priorities* property, I compare the allocative efficiency of the allocations determined by the implemented mechanism with the *Top Trading Cycles.* This is due to the fact that the former is known to always yield *health professional-efficient* allocations while also being *strategy proof.* I find small welfare gains implying that *respecting priorities* does not come at a major cost. This is due to the fact that the preferences are significantly heterogeneous.

To measure the cost of the *strategy proof* property, I evaluate the allocative efficiency of mechanisms that, although not *strategy proof*, are known to be difficult to manipulate. When determining the allocation, treats equally a reported (one of the most preferred ones) and an implied preference through the ranking of states (potentially one of the least desired). I show that breaking ties in priorities in favor of reported hospitals could induce welfare gains of $8\% - 12\%$. This welfare increase results in a reduction of the rejection rates to levels of $2\% - 5\%$.

Another proposed mechanism is one which the allocation is determined in two rounds, using the reported preferences in the first and the implied preferences, i.e. those implied from state reports, on the second. This results in welfare gains of up to 13% while also having no rejections in some cases. This mechanism is neither *strategy proof* nor *respects priorities.* However it respects the priorities of the reported preferred hospitals and is strategy proof in in the relative ranking of the reported preferences.

In the aforementioned mechanism treating differently the reported preferences than the not reported ones resulted in significant welfare gains. I study next what the optimal number of hospitals to report is, under the scenario where health professionals report truthfully. I find that in some cases the number is significantly low while on others it is a high number. Most of the environments where DA is used a truncation of possible reports is also at hand. I propose changing the termination rule of the deferred acceptance algorithm in order to reduce the cost of congestion while also allocating all slots.

Overall, I show that the cost of *strategy proofness* has first order consequences on the aggregate welfare of the allocation. Therefore, in this setting the fine-tuning the algorithm has significant welfare consequences. This contrasts with other centralized mechanisms settings in which it has been observed that the fine-tuning of the algorithm is of second order (Abdulkadiroğlu et al., 2017).

The last counterfactual studies the consequences of moving to a merit-based tie-break. When designing the current program, the ministry decided not to use the results of the end of study nation's exam (known as SaberPro). They did this because the goal of the mechanism is to tackle the inequality of the allocation of health professionals; if they used the results, then the better health professionals would be allocated to the most desired positions. Conversely, not using the results would improve the odds that very good professionals would be allocated to remote positions. The allocation of physicians was not congested—that is, there were more positions than participants. In this case, the quality of allocated professionals would be a zero sum game in which the more desired positions would have a higher chance of going to good physicians. This, in turn, would generate professions that are strongly congested. By giving the professionals with lower scores a higher probability of not being assigned, a merit-based tie-break allows for an increase in the quality of health professionals. Nevertheless, because of regional priorities, a small fraction of the hospitals are made worse off.

If statements about welfare are to be accurate, preference estimates must be accurate, too. Since the allocation mechanism is based on DA, the stated preferences are treated as the true preferences. This is due to the fact that DA is known to be strategy-proof. Had the health professionals not been restricted in the number of hospitals they report directly it is dominant for them to report their true preferences (Jaramillo et al., 2016). This true even if their objective was not to be assigned at all. Since professionals report five (5) hospitals, they might misreport their true preferences in order to improve their chances of getting into a reported position. I do not observe evidence of this strategic behavior being systematic. Since their ranking will include all positions, there are many positions at which health professional can be allocated regardless of their report. In particular, were one to be strategic with the report, the last position reported is the one that results in the highest incentive to do so. This would result in having a higher probability of being assigned to the last position reported than to the second to last.

Overall, this paper evaluates a program designed to solve the rural hospital theorem problem of the health professionals market. This theorem happens to be framed in the context of rural hospitals but is applicable to all markets where stability is expected in the outcome. Therefore the lessons learned from the evaluation of the program at hand are valuable for other markets in which vacant positions are seen as market failures.

1.1 Related Literature

This paper makes six significant contributions to the literature. First, it contributes to the public policy literature that examines the inequality of the allocation of health professionals to rural and urban positions. Several current policies tackle this problem. For example, in the US, foreign health professionals are waived the two year residency requirement if they work in a Health Professional Shortage Area(J1w). In the case of Japan, Kamada and Kojima (2015) propose a centralized mechanism that uses hard regional caps to allocate a higher share of medical interns to that country's rural regions. Agarwal (2017) discusses how monetary incentives would modify the allocation of residents in the US. The approach taken by the Ministry of Health is different in that it uses a centralized mechanism and makes participation in the allocation compulsory.

Second, this paper contributes to the empirical study of preferences and allocations in the case of health professionals. To the best of my knowledge Agarwal (2015) is the only scholar who has carried out this type of analysis, which he applies to family medicine residents in the US. Agarwal's estimates are obtained from the analysis of observed matches. This paper, in contrast, uses reported preferences. The estimates of the preferences show that the health professionals' preferences are significantly heterogeneous.

Third, this paper contributes to the empirical literature on centralized mechanisms, particularly those that use DA to determine the allocation. In the context of school choice (Abdulkadiroğlu et al., 2017) estimate parents' preferences for high schools in New York and compares the allocation efficiency of centralized and decentralized mechanisms. In the context of College Admissions, Luflade (2017) studies the value of information in the performance of the sequential use of truncated

DA to allocate college positions in Tunesia. For the case of health professionals in Colombia, this paper shows how on top of using a centralized mechanism, a compulsory feature is central to successfully allocating health professionals to rural areas.

Fourth, this paper contributes to several influential theoretical results in the large literature of mechanism design. This is done by comparing several algorithms that have been proposed previously in terms of allocation efficiency. Erdil and Ergin (2008) propose an algorithm that eliminates the welfare reducing cycles that the coarse priorities may induce. I find small welfare gains from moving to that algorithm. There is a family of algorithms that can be described as adding rank priorities to the priorities determined by the Ministry of Health. In this family we can find the Immediate Acceptance (Kojima and Ünver, 2014) and the New Haven algorithms. Within this family, I compare the efficiency of giving a priority to reported hospitals above any other priority has substantial welfare gains but at a cost of a having a significant number of the original priorities not being respected. I show that breaking the ties in favor of the health professionals who report a hospital as preferred is also within this family. However, this mechanism respects priorities while having welfare gains that range from $8\% - 12\%$.

These results can be seen as a contribution to the literature that measures the costs (gauged in terms of the welfare of health professionals) of the properties of the allocations that results from DA. Roth (2008) and Abdulkadiroğlu et al. (2017) do this in the case of the high school slots allocation in New York City. In sharp contrast to their findings, I find that because the preferences are completed (due to the fact that the rankings are complete), the fine tuning of the algorithm used to determine the allocation is a first-order concern.

Fifth, this paper also contributes to the literature on merit/effort-based priorities in centralized allocations. This is exemplified by the case of the cadet branch-of-choice in the US Military Academy (Sönmez and Switzer, 2013). I show that under certain conditions merit based tie-break can entail welfare gains under certain conditions yet raise inequality concerns under others.

Finally, to provide additional evidence that the stated preferences are the true preferences, I estimate the utilities using the Attention Logit model developed by Abaluck and Adams (2017). To estimate utility and attention this model exploits the fact that (assuming that health professionals are perfectly rational utility maximizers) the cross characteristic derivative matrix of hospitals [4] should be symmetric. The attention estimates are established in order to recover the aforementioned symmetry. I observe that the estimated utility and attention are similar for all health professions. I argue that this is evidence that there is no systematic misrepresentation of the preferences. I use the random coefficients model to make the counterfactuals because allowing Attention Logit to have random coefficients, in order to allow for preference heterogeneity (to allow for preference heterogeneity) is beyond the scope of this paper.

This paper is organized as follows. In the next section I discuss the institutional details and the October 2014 allocation data. In Section 3, I detail the estimation and the results of a random coefficients model of the preferences of health professionals. Section 4 evaluates the performance of the program in the absence of the compulsory feature. In section 5, I describe the theoretical properties of centralized mechanisms and evaluate the allocative efficiency (in terms of health professional welfare) that would have occurred if other allocation mechanisms had been used. In section 6, I evaluate the consequences of moving from a random tie-break to a merit-based one. In section 7, I discuss the assumption that reported preferences are the true preferences and present evidence that support this is the case. Section 8 concludes.

2 Institutional Background

The Compulsory Social Service, hereafter CSS, of health professionals in Colombia was created to tackle the problems of inequality in the allocation of these professionals to rural areas and marginalized urban areas in Colombia. It started in 1949 for physicians, 1951 for dentists, 1971 for medical lab scientists, and 1971 for nurses. CSS lasts one year.

The inequality in the allocation of health professionals is not exclusive to Colombia and neither is the compulsory social service solution. Currently Bolivia, Costa Rica, Ecuador, Honduras, Nicaragua, Mexico[5], Peru, and Salvador also require, at least for physicians, compulsory social service. All of these programs emphasize that the *social service* is a contribution to society. The

[4] This is the discrete choice equivalent of the Slutsky matrix.

[5] In 1936, Mexico was the first country who implemented a compulsory social service for physicians. Today this program is known as *Servicio Social en Medicina*

programs differ primarily in whether the service must be undertaken before graduation (as part of their studies) or after graduation (as a first professional work experience).

The CSS in Colombia can be seen as mechanism to overcome the Rural Hospital Theorem effects of the country's health professionals market. In the matching literature the Rural Hospital Theorem predicts that under any stable allocation the set of unmatched hospital positions is the same. The name of this theorem reflects the fact that rural hospitals tend to have the highest rate of unmatched or vacant positions. When designing the centralized allocation mechanism, the Ministry of Health it was a central objective that the mechanism *minimized the number of unfilled positions.*

A centralized mechanism was needed because prior to July 2014, the allocation mechanism was a system of state-level (departamento) lotteries marred by serious incentive and efficiency flaws. Specifically, the mechanism generally did not minimize the number of unfilled positions. In the July 2012 allocation, for example, almost 600 physicians were not allocated even though 300 positions were left unfilled. The main purpose of the mechanism, to allocate health professionals to rural areas, was not being accomplished.

If a health professional is not assigned, she is exempted from the CSS and can ask to be professionally certified. The incentive flaws occurred because health professionals, and specially physicians, were gaming the mechanism to avoid being allocated. Since October 2014, the allocation in Colombia has been decided and implemented by the Ministry of Health, which uses a centralized mechanism. This mechanism, whose design was guided and implemented by economists, implemented a version of Gale and Shapley (1962) deferred acceptance algorithm. Jaramillo et al. (2016) discuss in detail the motivations needed to implement Gale and Shapley's (1962) deferred acceptance algorithm as well as the incentive flaws of the previous mechanisms.

Professionals can apply for three types of positions in the CSS. First are positions in rural areas (at least 70% of the positions). Second are positions in marginal urban areas (at most 25% of the positions). Third, are research positions (at most 5%). The first two types can only be obtained through the centralized mechanism. Investigation positions are assigned directly by the entities that conduct the investigations, and professionals do not participate in the centralized allocation.

To determine the allocation, the Ministry of Health imposes a priority order that needs to be satisfied by every hospital. Consequently, hospitals in this setting are not strategic agents. Thus, the allocation is a one-sided problem known in the literature as *school choice* (Abdulkadiroglu and Sönmez, 2003).

Health professionals are given three types of priorities that define the priority order under which they are accepted at each hospital. In the first type (which corresponds to Colombian law), a professional who is indigenous is given priority at hospitals located in indigenous regions. Similarly, a professional who is "raizal" (a native of the Archipelago of San Andres, Providencia, and Santa Catalina) is given priority at hospitals in the Archipelago. In the second type the priorities follow from Health Ministry regulations. Specifically, priorities are given to professionals who 1) are mothers with small children, 2) have impairments or disabilities, or 3) need special medical treatment. The third priority type takes into account hospital's regional preferences: hospital prefer professionals who graduated from their state or who were born in their state. Within this system Colombian law takes precedence over the Ministry's regulations and the latter take precedence over hospital preferences. Hence, the priority structure is as follows:

In summary, the priority structure is the following:

- Being indigenous,
- Being a "raizal" (native of Archipelago of San Andrés, Providencia, and Santa Catalina),
- Being pregnant,
- Being a mother or a father who has small children,
- Being a professional who has impairments or disabilities,
- Being a professional who needs special treatment,
- Graduating from a University from the same state as the hospital,
- Being born in the same state as the hospital.

Under this system the priority order would work as follows. Assuming that health professionals will qualify for no more than one of the Ministry-regulation priorities, pregnant women will have the top priority at all hospitals that do not grant priorities given by the Colombian law. In any given state, pregnant women who were born in and who graduated in that state will be placed first, while those who graduated in but were not born in that state will be placed next. Finally, those who were born in that state but did not graduate in it will be placed third in the priority order.

Health professions who participate in the CSS belong to one of four professions[6] Medical Lab Scientist[7], Medicine, Nursing, and Dentistry. To be professionally certified, professionals in these fields need a valid professional ID card that is authorized by the Ministry of Health. The social service is compulsory because to become certified, health professionals must participate in the CSS.

The Ministry allocates the positions on a quarterly basis. As mentioned earlier, professionals who are not assigned can ask to be professionally certified immediately. If a professional rejects her allocation, she is not allowed to participate in the centralized allocation during the next two allocations (i.e., she will need to wait 9 months before participating again) and because she is not certified, she cannot work in her profession during that period. A 9 month penalty is a high penalty considering that it is for a 1- month position. However, it has been observed that about $3-6\%$ of the health professionals reject the position they are assigned to.

The mechanism incorporates the *compulsory* characteristic by not allowing the professionals to report and outside option in addition to the penalty for not accepting a position. This is due to the fact that if there is an equal number of professionals and positions, allowing professionals to report an outside could result in positions being left unfilled. As mentioned earlier, one of the objectives of the mechanism is to minimize the unfilled positions. Hence, the mechanism is also compulsory in the respect that a professional cannot simply accept an outside option and pay a penalty for it.

A health professional can choose any position that currently is unassigned or that will become available before the next allocation. However, because Colombia has a large number of hospitals, professionals are required to report up to five preferred hospitals. They also are asked to rank the states in which they would like to fulfill their CSS obligation. Thus, their ranking of hospitals is constructed as follows: preferred hospitals are placed at the top; then all hospitals in each reported

[6] This health professions are all undergraduate studies.

[7] In Colombia this profession is known as *Bacteriology and Clinical Lab* and professionals in this area are usually referred to as simply *bacteriologists*.

state (excluding those reported as preferred) are randomly allocated in the professional's ranking following the state's stated order.

Ties in the ranking of Professionals (Hospitals) in the Hospital's (Professional's) rankings (priorities) are broken using a single tie-break. Using these rankings and priorities the allocation is determined using the *Deferred Acceptance*-Algorithm:

> **DA:**
>
> *Step 1:* Each professional proposes to her fist ranked choice. Using its priority order, each hospital tentatively assigns seats to its proposers one at a time. The professional is rejected if no positions are available at the time of consideration.
>
> *Step $k > 1$:* Each professional who was rejected in the previous step proposes to her next ranked hospital. Each hospital considers its new proposers together with the previously assigned proposers and assigns its positions one at a time using the hospital's priority order. The professional is rejected if no positions are available at the time of consideration.
>
> The algorithm terminates either when there are no new proposals or when all rejected students have exhausted their rankings.

The mechanism rests on the assumption that no professional is willing to incur the penalty of not being able to work in his or her profession for 9 months. This a substantial penalty if one takes into account that the assignments are due for one year. As Jaramillo et al. (2016) prove, the restriction on preferences under this assumption does not deteriorate the mechanism's strategy-proofness because the weakly dominant strategy is for each professional to report their true restricted preferences (i.e., report their true preferences with the restriction that the outside option is placed last). This is true even in the limiting case in which no positions are acceptable for a health professional. This is due to the fact that, as shown below, since the allocation is determined using DA, she needs to be rejected by all of the hospitals at some point in the algorithm.

2.1 October 2014 Allocation

In October 2014, a total of:

- 194 Medical Lab Scientists participated for 83 positions in 81 hospitals.
- 708 Dentists participated for 109 positions in 97 hospitals.
- 386 Physicians participated for 1025 positions in 544 hospitals.
- 828 Nurses participated for 194 positions in 148 hospitals.

Note that the ratio of professionals to available positions varies by profession. In the case of physicians, for example, there are more than 2 positions available per physician (a case of very low congestion), whereas in the other professions the opposite is true. Dentists, for instance, face the most congested allocation, with over 7 participants per position. The reason why there are so many positions available for physicians is the mechanism used before the centralized one was being manipulated to avoid allocation.

I obtained secure access to administrative data for all health professionals that participated in the first allocation that used the centralized mechanism. For each health professional, I have data on the hospital's reported ranking (up to five). I have information on the state where the professional was born, the state where he or she graduated, and her or his gender. Although I have no data on professionals' exact city of birth, most of Colombia's biggest cities (where 75% of the country's population resides) are state capitals, and most universities where a health professional degree can be obtained are located in these cities. Therefore, I use the capital of the state as a proxy for location. From this I construct key variables, such as the distance from the hospital to the capital of the state of graduation or origin.

The webpage interface used by the Ministry of Health forces professionals to report five positions. However, the interface also allows participants to report the same position more than once. Indeed, about 7% of participants report a preferred position more than once, presumably in an attempt to game the system. But this choice does not give participants an extra chance to receive their preference, nor does it allow them to have a full five choices. Moreover, when reporting the ranking of states, these by default were ranked in alphabetical order. Around 8% of the participants decided not to modify the displayed ranking. I observe that the different professions differ mainly in terms

	Med-Lab	Nurses	Med-Docs	Dentists
Female (%)	80	91	58	76
Reports St. (%)	92	89	94	91
Repeats Reported Pos.(%)	6	8	8	7
Diff. St. of Grad (%)	43	25	34	34
Dummy Priority (%)	6	13	10	9
Total Num. of				
Prof. with Priority	12	104	39	64
Num. Priority Raizal	0	0	2	2
Num. Priority Indigenous	2	11	2	6
Num. Priority Pregnant	7	50	25	31
Num. Priority Mother	1	19	6	9
Num. Priority Father	1	2	3	9
Num. Priority Handicap	1	12	4	9
Num. Priority Safety	0	0	0	1

Table 1: Professionals Characteristics by Profession

of the share of participants who are female. The profession with the most females is Nursing (91%), whereas Medicine has the lowest share (58%).

In the case of non-regional priorities, the one most observed is being pregnant. I observe that Nursing has the highest share of professionals who have non-regional priorities (13%), while among Medical Lab scientists this share is only 6%. Hence, the priority structure for all of the professions is quite coarse because most participants will only have a regional priority. Moreover, participants who were both born and raised in the same state are given precedence in that state over those who were not. Because of the manner in which the priority structure was developed, subjects who have any of the aforementioned priorities will compete among themselves for all of the positions. Then the subjects who do not have any priority will compete among themselves for the remaining positions, as seen in Table 1.

With regards to hospital characteristics, I observe that about 80% of positions were occupied at the time of the allocation–that is, they will become available during the next three months. In all cases, a significant number of the positions are located in state-capital cities (18%). I also observe that positions on average are located 79km from their capital. Moreover, around 25% are located

	Med-Lab	Nurses	Med-Docs	Dentists
Rural (%)	71	78	85	86
Assigned (%)	73	75	79	87
Dist. Capital (km)	89	66	70	77
Compensatorio (%)	20	20	12	14
Coca (%)	25	32	19	26
Quota	1.02	1.21	1.88	1.13

Table 2: Hospital Characteristics by Profession

in municipalities where coca was produced in 2013. Additionally, 16% of the positions offered a *compensatorio* day, or an additional paid rest day mandated by Colombian law for employees who work four consecutive weekends, as seen in 2.

Hospitals need authorization from the Ministry of Health to open a CSS position. When they do so, hospitals offer for that position a wage that is allocated by the mechanism and set in reference to the Colombian Monthly Minimum Wage. There is no negotiation over wages: the Colombian Monthly Minimum Wage is set on a yearly basis and wages are automatically updated so that there is no need to reapply for the position (an administrative process that can take several months). Moreover, when a professional applies to the CSS he or she observes a list of available positions and the wage offered for each. In other words, during each quarterly allocation, wages are exogenously determined. A histogram of the wages is shown in Figure 1. In 2014, the Colombian Minimum Monthly Wage $\$616,000COP$, approximately $\$300USD$. As is evident from the histogram, there is a significant dispersion on the wages offered. Most of the hospitals are public state institutions and as such the wages belong to the state budget.

Because professionals choose both the hospital where they will work and the county where they will live, the characteristics of the counties in which hospitals are located play an important role in hospital choice . The Municipal Panel Data CEDE [8] provides demographic information about the municipalities in which the hospitals are located. I find that both geographic location and the Unsatisfied Basic Necessities Index, hereafter UBN, which is a measure of poverty in each count, affect hospital choice.

[8] See Acevedo et al. (2014) for a description of the Municipal Panel Data CEDE.

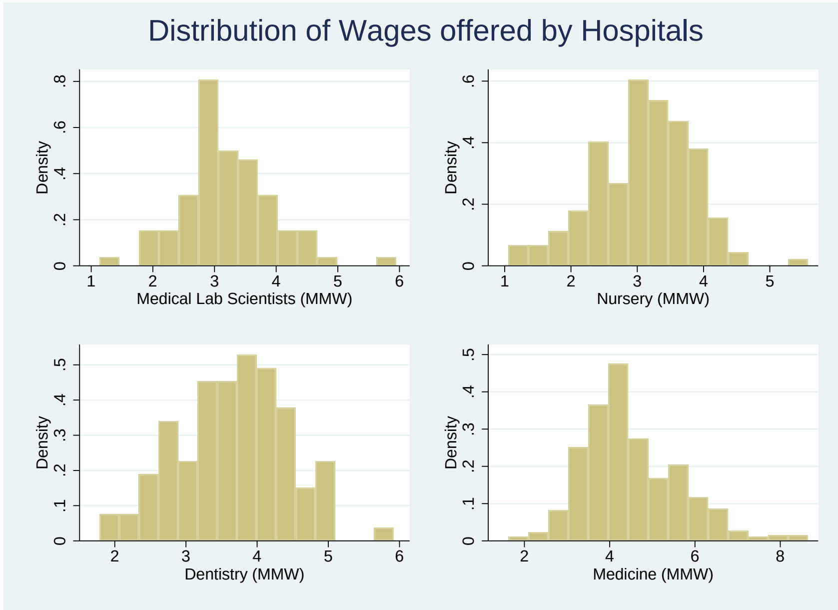


Figure 1: Histogram of wages offered by hospitals
(in Colombian Minimum Monthly Wage)

3 Estimating Health Professional's Preferences

The Ministry of Health publishes on a webpage information about available positions, and it includes both an email address and a telephone number where questions can be answered [9].

Health professionals' preferences exhibit consistent regularities. They prefer hospitals close to their current location, higher wages, and hospitals in wealthier municipalities (as measured by the UBN-index), as shown in 3. Overall more than 90% of health professionals report at a least 3 hospitals. 3 shows the regularity of health professionals' preferences.

Figure 2 shows a histogram of the distance to the first reported hospital from the capital of the state where they were born[10]. From it we can observe that despite having a high average, more than $120km$ for all health professions, the preference for distance is highly skewed and most of the health professionals prefer to be close to the capital of the state they were born. Overall we observe that the preference for location, in the sense of being either close to the capital city of the state

[9] For a detailed version of the instructions (in Spanish) go to https://tramites.minsalud.gov.co/tramitesservicios/DefaultSSO.aspx-Accessed May 2018

[10] In the Appendix a similar histogram is shown for the distance to the first reported hospital from the capital of the state of graduation

	Avg.	1st	2nd	3rd	4th	5th
Med-Lab ranking choice (%)		100	96	96	96	91
Mean Dist. Graduation (km)	446	202	224	264	299	310
Mean Dist. from St. of Origin (km)	476	181	222	245	287	302
UBN index	49	40	46	48	44	44
Wage (MMW)	3.2	3.29	3.33	3.24	3.22	3.22
Same State of Graduation (%)	4	47	39	24	18	8
Same State of Origin (%)	4	56	44	33	20	17
Nurses ranking choice (%)		100	97	90	76	55
Mean Dist. Graduation (km)	488	123	158	172	195	238
Mean Dist. from St. of Origin (km)	493	132	164	178	203	252
UBN index	42	27	30	31	32	33
Wage (MMW)	3.0	3.20	3.12	3.09	3.13	3.11
Same State of Graduation (%)	4	66	56	46	41	34
Same State of Origin (%)	4	64	53	46	39	31
Med-Docs ranking choice (%)		100	95	87	72	54
Mean Dist. Graduation (km)	405	106	106	120	146	161
Mean Dist. from St. of Origin (km)	428	128	139	153	160	188
UBN index	41	23	24	25	24	25
Wage (MMW)	4.49	4.73	4.73	4.68	4.84	4.86
Same State of Graduation (%)	4	70	67	63	55	51
Same State of Origin (%)	4	65	59	54	52	47
Dentists ranking choice (%)		100	97	96	93	72
Mean Dist. Graduation (km)	428	170	183	237	254	285
Mean Dist. from St. of Origin (km)	443	165	182	241	250	284
UBN index	44	34	35	36	37	36
Wage (MMW)	3.6	3.82	3.76	3.82	3.82	3.80
Same State of Graduation (%)	4	51	40	25	20	20
Same State of Origin (%)	4	58	44	30	24	21

Table 3: Preferences regularities of Health professionals by Ranking

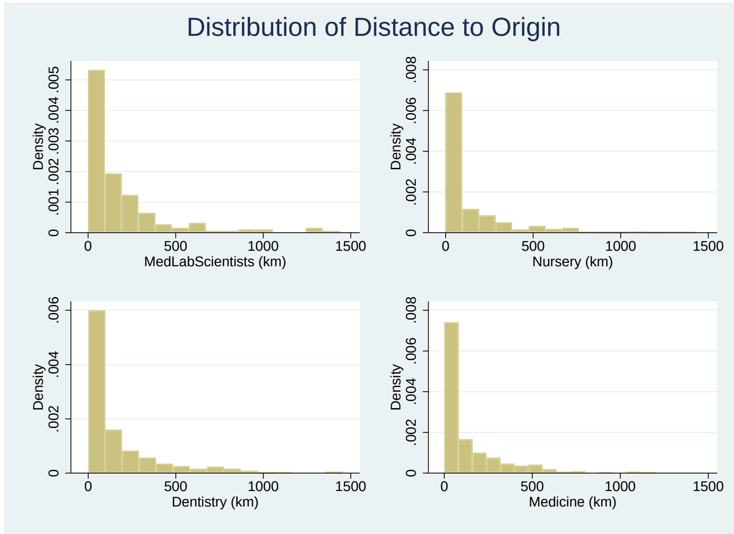


Figure 2: Travel distance from hospital to capital of state of origin city (km)

they were born or the city where they graduated from, is stronger than the preference for wage. Moreover, from Table 3 we can observe that the reported hospitals wage is not decreasing while it increases strongly in distance. Nevertheless, health professionals tend to prefer hospitals that pay wages higher than the average.

3.1 A Random Coefficients Model of Preferences

To model preferences for rural hospitals I use a standard random utility framework. Denote with U_{ij} the utility of health professional i from performing the CSS at hospital j. I assume that health professionals choose the hospitals that maximize their utility. Their utility is assumed to depend on the wage offered by the hospital W_j, the distance to their place of graduation D_{ij}, and a set of characteristics of the municipality where the hospital is located M_j, resulting in the following utility:

$$U_{ij} = W_j + \beta_i^D D_{ij} + \beta_i^M M_j + \epsilon_{ij} \tag{1}$$

Where β_i^D, and β_i^M represent the weight that health professional i places on the distance to the town of origin[11], and characteristics of the hospital's municipality, respectively. In the estimation, these variables are interacted with the gender of the health professionals. Thus, I allow different preference distributions and correlations on the weights by gender.

I use the same weight on wage for all health professionals in order to find utilities in a common numeraire, namely nominal wage. In particular, the wage coefficient is constrained to be equal to 1. Therefore, all else being equal, since all health professionals are likely to like higher wages, treating the wage as numeraire is a scale normalization. To estimate their different weights for distance to the town of origin, and hospital's town characteristics, I use the reported choices of the health professionals and a random coefficients model (McFadden and Train (2000), Train (2009)). Hence, this coefficients allowed to be random in order to allow for heterogeneity of preference in the choice behavior.

Because in our data set most health professionals report several hospitals, I have an observation for each hospital as ranked by each health professional. The first observation for each health professional indicates when the first ranked alternative was chosen among all alternatives. The second(nth) observation identities the case in which the second alternative was chosen when the first (previous n-1) chosen alternative(s) was (are) removed. Removing the previously chosen alternatives is important because it creates variation in the choice set. This variation allows a better prediction of both the agents' utilities and the ranking beyond the reported choices. This is so because when a health professional makes multiple choices that share a common attribute, I can infer that the individual has a strong preference for that attribute because independence in the additive error terms across choices makes observing such a pattern very unlikely.

I assume that ϵ_{ij} in equation (1) is a distributed i.i.d. extreme value and captures the idiosyncratic tastes for hospitals. The vector of coefficients β follows a multivariate normal mixing distribution and therefore allows for preference heterogeneity by allowing random coefficients. A dummy variable for each state is included in the municipality's characteristics. The distribution of the weights of these variables is constrained to have a zero mean. I include these variables to allow for a correlation between hospitals within the same state, which, in turn, allows for a rich variety of

[11] This distance was calculated using the *geocodeopen* command in Stata (Anderson, 2013).

distribution patterns. This is analogous to a nested logit in which states define mutually exclusive nests.

The main assumption in this estimation is that conditional on observed hospital and county characteristics, unobserved idiosyncratic tastes are independent on wages. A violation to this assumption may occur if, given the controls, health professionals prefer hospital offering higher wages. It is unlikely in this case since, as shown in Table 3, preferences regularities show that the preferences for location are stronger than those for wage. Furthermore, this assumption is plausible the given the rich set of controls.

Estimation was carried out using a hierarchical Bayes procedure. Table 4 presents the results of the random coefficients estimation of health professional preferences for rural hospitals for a selection of the variables. The distribution of the weights of the coefficients are consistent with expectations. They show that health professionals prefer higher wages, being closer to their state of graduation, and wealthier municipalities. In all cases there is significant heterogeneity among professionals. professionals. Moreover, the estimates show that female health professionals have a stronger preference for staying close to the capital of the state in which they were born. In the appendix the complete table of the means and standard deviation of the weights is shown. There it can be observed that the coefficients of the standard deviations of the dummy variables of states show that there is a significant heterogeneity in preferences for states. This indicates that the decision by the Ministry of Health to ask for a ranking of states produced a relatively good description of the rankings.

		Med-Lab	Nursing	Med-Docs	Dentists
Means					
Wage					
		1	1	1	1
Travel Distance Origin					
	F	-1.413	-1.863	-2.455	-1.444
	M	-1.192	-1.417	-1.798	-1.413
UBN-Index					
	F	-0.056	-0.114	-0.141	-0.083
	M	-0.023	-0.108	-0.118	-0.109
Coca					
	F	-1.419	-0.644	-1.095	-1.124
	M	-0.147	-2.013	-0.997	-0.338
Living Place					
	F	-0.866	-0.256	-0.577	-0.587
	M	-0.163	-0.882	-0.179	-0.541
Weekends Shift					
	F	-2.296	-0.555	-0.176	-1.231
	M	-0.001	-0.382	-0.114	-0.031
Std. Deviation					
Travel Distance Origin					
	F	1.506	1.427	2.430	1.244
	M	1.875	2.202	1.542	1.455
Travel Distance Graduation					
	F	1.265	2.068	1.862	1.131
	M	1.358	2.157	2.736	0.237
UBN-Index					
	F	0.237	0.667	0.178	0.261
	M	0.214	0.603	0.191	0.470
Coca					
	F	1.630	1.102	1.940	1.252
	M	0.568	2.291	0.875	1.012
Living Place					
	F	1.111	0.663	0.491	1.016
	M	0.585	1.254	0.418	0.942
Weekends Shift					
	F	2.276	1.028	0.414	1.297
	M	0.037	0.779	0.329	0.202

Table 4: Preference Estimates of Health Professionals

3.2 Measuring Welfare with the Estimates

The estimated model provides the posteriors of the weights for each health professional. With these posteriors I estimate the utility of each health professional at each hospital. However, these estimated utilities need not result in the same ordinal order implied by the reported preferences. The reported preferences then introduce restrictions in the unobserved terms ϵ_{ij}.

Accurately incorporating the restrictions implied by the reported preferences plays a central role in the welfare analysis that I develop below. Previous studies, such as Abdulkadiroğlu et al. (2017), reveal that higher-ranked alternatives tend to have the highest unobserved terms. In the counterfactuals I use reported preferred and not reported hospitals utilities. Therefore, it is important to incorporate the fact that reported hospitals will have a high unobserved term and that the not reported ones will have a low one. This is done in order for the welfare analysis not to be driven by the unobserved terms.

In this research I estimate unobserved terms by means of placing restrictions on the bounds of each ϵ_{ij}. The bounds are set in order to maximize the probability of observing the reported ranking. Denote with $\hat{U_{dj}}$ the estimated utility of health professional d at hospital j. The constraints imposed by the reported preferences cause a hospital ranked in the j-th position among N alternatives to satisfy:

$$\hat{U}_{dj+1} + \epsilon_{dj+1} > \hat{U}_{dj} + \epsilon_{dj} > \hat{U}_{dj-1} + \epsilon_{dj-1} \tag{2}$$

Therefore, the ranking constraints the possible values the error term ϵ_{dj} can have. Let $\underline{\epsilon}_j(\bar{\epsilon}_j)$ denote the lower (upper) bound of ϵ_j. The lower (upper) limit of ϵ_j will be bounded by the upper(lower) limit of ϵ_{j+1} (ϵ_{j-1}) as implied by the following restriction:

$$\hat{U}_{dj+1} + \underline{\epsilon}_{dj+1} = \hat{U}_{dj} + \bar{\epsilon}_{dj} \tag{3}$$

With N alternatives ranked, there are $2(N-1)$ variables because the first (last) ranked alternative has no restrictions on the upper (lower) bound. Moreover, there are $N-1$ restrictions like the one mentioned above. Thus, I the optimization problem in order to find the maximum probability of observing the ranking is:

$$\max_{(\underline{\epsilon}_{di},\bar{\epsilon}_{dj})} \log Prob(\epsilon \geq \underline{\epsilon}_1) + \sum_{k=2}^{N-1} \log Prob(\bar{\epsilon}_k \geq \epsilon \geq \underline{\epsilon}_k) + \log Prob(\epsilon < \bar{\epsilon}_N) \tag{4}$$

After having estimated the boundaries for each unobserved term (i.e., the restrictions implied by the ranking order of the alternatives), I estimate the expected value of each term and add them to the expected utilities. These are the final estimates that will be denoted $E\left[u_{ij}|r_i, H_j, M_j, \beta_i\right]$ (the expected utility of professional i at hospital j given his reports r_i), and they are the estimates that will be used to estimate the counterfactuals.

Denote with $E\left[u_{ij}|r_i, H_j, M_j, \beta_i\right]$ the expected utility of professional i at hospital j given his reports r_i, the hospital's characteristics H_j, the municipality's characteristics M_j, and the estimated coefficients of the professional β_i. Also denote with $\mu : \mathcal{P} \rightarrow \mathcal{H} \cup \mathcal{P}$ a matching, such that each professional is assigned either to a single hospital or to itself. Define the estimated average professional welfare as:

$$\bar{W}(\mu) = \frac{1}{|\mathcal{P}|} \sum_{i \in \mathcal{P}} E\left[u_{ij}|r_i, H_j, M_j, \beta_i\right]$$

.

The difference between average professional welfare between two matchings, μ and μ', is: $\bar{W}(\mu) - \bar{W}(\mu')$.

4 Individual Rationality vs Compulsion

Mechanisms based on DA are known to have good efficiency properties. Azevedo and Leshno (2016) show that under a sufficient heterogeneity of preferences, the probability of an allocation resulting from DA being efficient converges to 1 as the number of participants increases. Having the aforementioned property is called being *asymptotically efficient*. In health professionals' preferences, the ratio between the preference for wage and location (measured by the travel distance cost) gives us a measure of preference heterogeneity. Because the preference for wages is low relative to the preference for location, it is possible to conclude that health professionals have significantly heterogeneous preferences–a heterogeneity that reflects the fact that health professionals are scattered throughout the country. Given that the preferences are heterogeneous and that the allocation is determined using DA, the allocation result should have good efficient properties.

In this section I test whether the efficiency improvement is high enough to make the compulsory feature unnecessary. In other words, I test what would happen if the allocation was based on the individual rationality that the single-offer centralized allocation mechanisms are built upon. I do this by exploiting the allocation mechanism changes carried out from 2012 to 2014.

In section 2, I characterized the single-offer centralized mechanism implemented since October 2014. Prior to October 2012, a decentralized system was used to determine the allocation. Each health professional enrolled in a state, and among the enrolled health professionals, each state conducted a lottery to determine the allocation. If in a given state there were more health professionals than participants, the unassigned professionals were exempt from the CSS. In the final allocation that used this decentralized mechanism, 135 (out of 380) of the physicians enrolled in a state that had only 2 positions available, causing 133 of them to be exempt.

During the previous two allocations, others applied to the same state with the goal of not being assigned. In the April (June) 2012 allocation, a total of 74(225) physicians applied for 6(7) positions. Because this strategy was used only in the allocation of physicians, I focus on that profession. Excluding the state just mentioned, 35 states participated in the allocation, and professionals were exempt from only 3 of them. Forty-two (42) health professionals were exempted, 39 of whom came from one state: Valle del Cauca. However, Valle del Cauca had a significant number of positions available (52) and many graduates in the 2014 allocation. Thus, I assume that these graduates were not attempting to avoid the allocation.

In 2013, the Ministry of Health established enrollment limits in each state. The number of enrolled health professionals could not double the number of positions. This constraint reduced significantly the number of unfilled positions. However, the fact that there were more exempted physicians than unfilled positions motivated the ministry to design the centralized mechanism. Due to the presence of this constraint, I use the 2012 data rather than the 2013 data to find the marginal utility for physicians of accepting a position.

A similar number of physicians participated in the October 2012 and the October 2014 allocations: 382 in the former and 386 in the latter. Because I do not observe detailed data on participants in the 2012 allocation, I examine the 2014 participants. I assume that the characteristics of the two cohorts are significantly similar. To make the 2014 allocation more closely resemble the October 2012 situation, I randomly choose the participating positions in the 2014 allocation, when 213 more positions were available, in order to have the same number of positions in each state.

I simulate the outcome that would have occurred if the decentralized mechanism used in 2012 had still been in use in October 2014. I model the decision of physicians of enrolling to their preferred state (in order to work for one year) or to avoid being allocated. The last decision uses the simplifying assumption that when ever applying to the state with 2 positions, all the physicians that do not want to be allocated are successful in doing so. To this end I use information about the rankings of states submitted by physicians. For those physicians who did not modify the default ranking of states, and who, therefore, reported them in alphabetical order, I use their state of origin as the preferred state. I assume that had the decentralized mechanism was still in place and that in the lottery physicians would have enrolled in their reported preferred state. I do not find significant differences when I include the state of graduation.

I estimate the average marginal utility a physician would requires in order not to avoid being allocated. I do so by making, in equilibrium, the number of physicians that decide to avoid being allocated the same as the ones that applied to the state that had 2 positions in 2012. I estimate this marginal utility to be 1.91 Colombian Monthly Minimum Wages. Notice that I assume this marginal utility remained the same in 2012 and 2014. Given that the percentage of the PIB spent on health in Colombia between 2012 and 2014 increased from 4.47% to 4.64%, the aforementioned value is most likely to have increased.

I then simulate the outcome of the October 2014 allocation if there had been voluntary participation–that is, if the physicians had been able to apply only for positions that yielded them a utility higher than the aforementioned acceptance threshold. I find that about 115 (30%) of the physicians would have been unassigned. Notice that the number of physicians that would be left unassigned under the centralized mechanism (115) is similar to the number of physicians that strategized to avoid being allocated (133). This is so despite the fact that the average utility of the assigned physician went from 2.95 Colombian Monthly Minimum Wages to 6.02 Colombian Monthly Minimum Wages. In other words, for the policy to be successful assignments need to be mandatory.

Notice that this counterfactual has been carried out in an allocation that has low congestion, when about 3 positions were available for each physician. Allocations conducted in January and July tend to be much more congested. A higher congestion is likely to cause a lower utility per allocated physician. Hence, the compulsory feature is important for the program's success.

Another way to motivate physicians to participate in the hospital allocation is to increase wages. Figure 3 shows the number of physicians who would be allocated if the wage was increased optimally for the vacant positions. The term "optimally" refers to the fact that filling the first position comes at a lower cost than filling the second one. The results show that an average increase of 600 USD (2 Colombian Minimum Monthly Wages) would cause 35 more physicians to be allocated. This would entail having about 82 physicians not being allocated while having many options to choose from. Therefore, the alternative to making compulsory the participation of raising optimally the wages implies a very high cost.

Increasing the wage is also a competing way to motivate physicians to be allocated to the positions in the participating hospitals. 3 shows the number of physicians that would be allocated if the wage was increased optimally for the vacant positions. Optimally refers to the fact that filling the first position comes at a lower cost than filling the second one. The results show that an average increase of 600USD (2 Colombian Minimum Monthly Wages of the time) would result in 35 more physicians being allocated – resulting in an allocation rate change of 10%. This would entail having about 82 physicians not being allocated while having many options to choose from. Therefore, the alternative to making compulsory the participation of raising optimally the wages implies a very high cost.

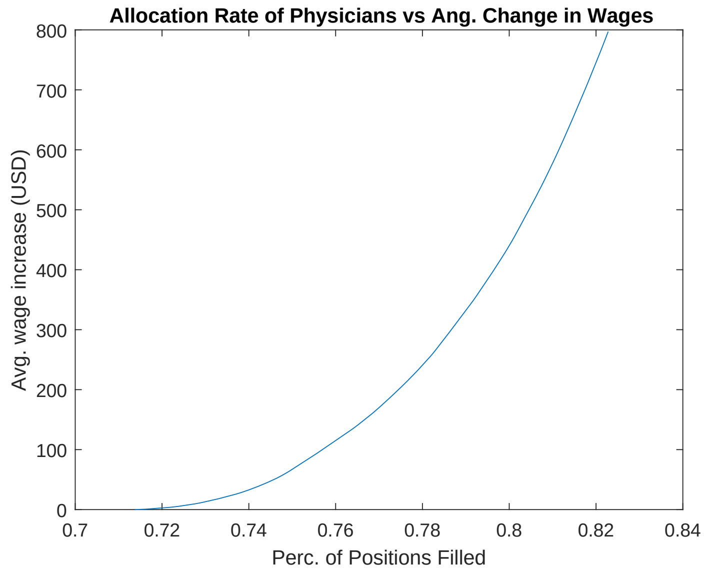


Figure 3: Increase of Physicians Allocated with Wage

5 Comparing Alternative Mechanisms

The CSS's design, as a centralized mechanism, was built on the assumption that the cost of rejecting a position was so high, that no health professional would do so. However, in practice it has been observed that, depending on the profession, about $4\% - 6\%$ of the positions are rejected. Hence, improving the welfare of the health professionals will also have the purpose of reducing the aforementioned rate of rejections. I incorporate the fact that positions are rejected on the welfare estimates by calculating the outside value of rejection in order to have, on average, the same rejection rates in each profession as observed.

In this section, I will first outline the properties and trade-offs of matchings. Then I compare the welfare in terms of efficiency for the health professionals and estimated rejection rate of different algorithms. Finally, I discuss the welfare implications of changing the total number of hospitals in the ranking of health professionals.

In the physians case, the fact that there are almost 3 positions per each professional makes the allocative analysis futile. This is due to the fact that a high share of health professionals is getting what they want. Therefore, in this section I will focus on the efficiency analysis of the other professions, i.e. the ones that were significantly congested. For those professions, in contrast to the results reported by Abdulkadiroğlu et al. (2017) and Abdulkadiroğlu et al. (2009) that conclude for their settings that any algorithm modification has a second order relevance when compared to DA, I observe that in this case there are significant welfare gains from the fine-tuning of the mechanism.

5.1 Properties and Trade-offs in Matchings

As previously defined, a *matching* is an allocation in which each health professional is assigned to at most one hospital and no hospital is assigned more health professionals than its capacity. A *mechanism* is an algorithm that takes the preferences and priorities and determines a *matching*. In centralized matching mechanisms, agents elicit rank-ordered lists. In this case health professionals report preferences over hospitals, and then a centralized authority, in this case the Ministry of Health in Colombia, determines the allocation using the preferences, coarse priorities, and a random tie-break.

The desirable properties for a matching in this problem are: *respecting priorities*, *strategy-proofness*, *efficiency*, and *minimizing unfilled positions*. The first three are common priorities of centralized mechanisms while the last is related to the compulsory nature of this problem (Jaramillo et al., 2016).

A matching *respects priority* if (a) there is no health professional-hospital pair such that they are not assigned to each other in a matching but the health professional would prefer to be matched to the hospital and the hospital has unfilled positions or (b) another health professional (who has a lower priority than the health professional at the the hospital) is matched to the hospital. This property ensures that no professional-hospital pair is willing to block the allocation or would like to match outside the centralized clearing house. Roth (2008) shows how clearing houses used to allocate medical interns in the UK that implement mechanisms with these properties (or stability, in the case of two-sided markets) have remained longer in time.

A matching is *efficient*[12] if there is no other matching such that each health professional finds it at least desirable and at least one health professional prefers it. This property means that the mechanisms maximize ordinal welfare in the respect that to make one health professional better

[12] This definition of efficiency considers only the health professionals as strategic agents.

another one would need to be made worse. However, the welfare estimates are carried out using the cardinal representation obtained with the estimated preferences.

The third property is *strategy-proofness*: no health professional should benefit from misrepresenting his preferences. This property ensures that the process of going from preferences to the elicited ranking reports is straightforward. As shown above, the type of preferences allowed in this setting is a subset of the preferences traditionally assumed in the school choice literature. Thus, if a mechanism is strategy-proof under the broad set of preferences, then it will remain so under the smaller set of allowed preferences (Jaramillo et al., 2016). This is why when I estimate the preference I assume that the reported preferences are the true preferences. I conduct robustness checks on this assumption in section 7.

The last property is *minimizing unfilled positions*: for each possible case, the mechanism assigns a matching such that there is no professional-hospital pair in which the hospital has unfilled positions but the health professional is not matched to any hospital. Given that all health professionals (hospitals) are listed in the rankings (priority orders) of hospitals (health professionals) this property is going to be satisfied.

The following example illustrates a case in which having one more professional than available positions precipitates significant welfare costs for the allocated health professionals:

5.1.1 Example

Suppose there are 4 health professionals, denoted with $hp_{i\in 1-4}$, and three hospitals with one positions each, which are denoted with $H_{i\in 1-3}$ with the following preferences and priorities:

P_{hp_1}	P_{hp_2}	P_{hp_3}	P_{hp_4}	Pr_{H_1}	Pr_{H_2}	Pr_{H_3}
H_2	H_3	H_2	H_2	hp_3	hp_2	hp_1
H_1	H_1	H_3	H_1	hp_4	hp_4	hp_4
H_3	H_2	H_1	H_3	hp_2	hp_3	hp_3
				hp_1	hp_1	hp_2

The allocation determined by using DA is: $(hp_1, H_3)(hp_2, H_2)(hp_3, H_1)$ and leaving hp_4 unassigned. However, had hp_4 not been present, the allocation would have been:

$(hp_1, H_1)(hp_2, H_3)(hp_3, H_2)$. This allocation is strictly preferred by every allocated health professional. Therefore, the costs of congestion can be substantial.

5.2 Competing Mechanisms

I compare the welfare output of different algorithms, all of which satisfy the minimizing unfilled positions property. The results derive from simulating the tie-breaks and determining the allocation with them 100 times. The implemented mechanism is based on the DA algorithm, which is known to be strategy-proof and efficient.

If a priority structure is strict, then it would Pareto-dominate any other allocation that respects priorities; however, this expectation does not necessarily hold in this case because the priority structure is significantly coarse (Erdil and Ergin, 2008). A policy maker that is willing to sacrifice some strategy-proofness for efficiency, would use the SOSM[13]. This algorithm improves the allocation upon the one resulting from DA by removing the welfare-decreasing cycles induced by use of the tie-break. I find that the SOSM yields few welfare gains for all professions. This is so because I use a single-tie break rather than multiple ones. The Appendix describes other mechanisms under both single and multiple tie-break scenarios.

The second mechanism I compare to the implemented mechanism uses the TTC algorithm to determine allocation. This approach makes each allocation efficient and maintains the strategy-proofness property. Each subject who has the highest priority at a hospital is endowed with a seat and is allowed to trade positions. The trade does not take priorities into account, and, consequently, priority violations are allowed in the resulting allocations. Indeed, we observe that violations do occur and they involve on average 13-27 professionals. Comparing these two mechanisms reveals that in this scenario with strategy-proof mechanisms, due to the aforementioned heterogeneity of preferences, there is no trade-off between efficiency and respecting priorities (Abdulkadiroglu and Sönmez, 2003). The welfare gains achieved from moving to this mechanisms are not significant. However, the rejection rate is reduced by in 1 unit.

In a unique characteristic of the implemented mechanism, preferences are completed for each

[13] SOSM stands for the *Student Optimal Stable Match*. In this setting this algorithm produces a *Health Professional Optimal Respecting Priorities Match*.

health professional using the reported ranking of states. As noted below in a discussion of welfare analysis, this fact generates significant competition between the health professionals for positions. Furthermore, imposing the strategy-proofness condition with completed preferences comes at a first order cost in welfare. The next mechanisms proposed are designed to limit competition for positions between health professionals in order to obtain significant welfare gains. They belong to a family of mechanisms that uses ranked-priorities, i.e. a priority given to a health professional because of where she ranked the hospital. These family of mechanisms if described in the Appendix. The proposed mechanisms yield higher welfare than the TTC and the first one respects priorities. We conclude that in this case the strategy-proofness property is much more costly in terms of efficiency than the respecting priorities property.

The third proposed mechanism is to break ties in favor of the professionals who report a hospital. This modifies how the mechanism handles ties in priority, and it meets the respect priorities criterion. Regarding incentives to report the truth, a professional might want to misreport his preferences in order to get an exemption, yet this action could induce a strategic behavior equivalent to the one that occurs when truncated lists are reported Haeringer and Klijn (2009). A position undesired by all professionals cannot be avoided. The welfare gains from moving to this mechanism range from $6\% - 8\%$. Notice that this mechanism respects all priorities. As of the rejection rate, it is predicted to be reduced to 2%.

The final proposed mechanism uses two rounds. In the first round, the set of allocations uses only the reported hospitals. In the second round, the remaining hospitals use the state rankings. This mechanism presents professionals with the highest incentives to manipulate: if a professional avoids being matched in the first round she is very unlikely to be matched in the second. This mechanism is a version of the *Parallel DA* mechanism used in China to determine college admissions (Chen and Kesten, 2017). The welfare gains achieved by moving to this mechanism can be as high as 55%. Again, this gain carries the cost that occurs when the priorities of many health professionals are not respected (as is the case of Nursing with around 101). Table 5 summarizes the results when these mechanisms are compared.

In contrast to the results reported by Abdulkadiroğlu et al. (2017) and Abdulkadiroğlu et al.

(2009) who conclude that for their settings any algorithm modification has a second-order relevance compared to DA, I find that when the mechanism is fine-tuned, significant welfare gains are achieved.

5.3 How many hospitals to report?

In the alternative mechanisms discussed in the previous section, allowing reported preferences to have an edge produces significant welfare gains. In this section I explore what would happen if a truncated DA was used–a condition that is consistent with the majority of school choice applications, whose reports need to be truncated [14]. In this application professionals were asked to report up to five hospitals. The number was not decided using a technical motivation.

In the next set of counterfactuals, the change in welfare is conditional on the number of hospitals that professionals are allowed to report using DA with truncated reports. In all of these cases I assume that professionals report truthfully–that is, when allowed to report N hospitals, they report their preferred ones. Therefore, when they are allowed to report five (5) or less, I use their reported hospitals. When professionals report more than the reported number, I use those that have the highest estimated welfare. Figure 4 summarizes the results. Interestingly, due to the fact that the physicians allocation was not congested, their utility increases with the number of positions. For the other professions, the number is four or five.

There is a trade-off in hand with the number of positions reported. On the one hand, the higher the number of hospitals reported the higher the share of positions allocated. On the other, the higher the competition between health professionals for these positions and hence the lower the welfare of the allocated health professionals. Due to the preference heterogeneity and the congestion of some markets, the maximum is achieved at a low number of positions for the very congested markets (Nursing and Dentistry) while achieving it at a high number for the other ones.

In most of the environments where DA is used, there is a truncation of possible reports. To reduce the cost of congestion while allocating all positions, I propose the following modification of the termination rule of the deferred acceptance algorithm: The algorithm terminates when there are

[14] Rumania and Boston are the only cases I know of in which the number of options that can be reported is unrestricted. In the former, the mechanism employed uses a serial dictatorship to make the allocation

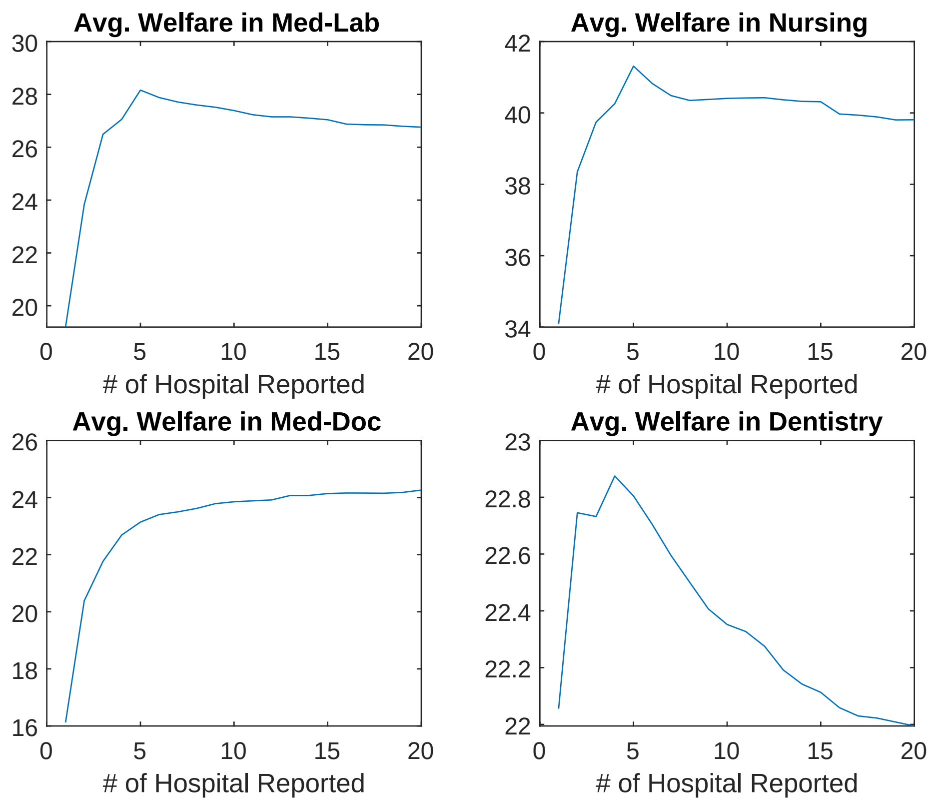


Figure 4: Average Utility by Num. of Hospitals Reported

no new proposals, when all the positions are enrolled, or when all rejected students have exhausted their ranking. This mechanism is not strategy proof: a professional may want to modify his reports in order to precipitate an early termination of the algorithm. However, it would be difficult to manipulate because the odds of inducing early termination are roughly one over the number of professionals. This modification is advantageous because under it exempt professionals would not need to "apply" to and be rejected by each hospital. That said, the mechanism does not respect priorities; it could end a health professional's exemptions before she is given the chance to a apply to hospitals in which she has a high priority. The modification of the termination rule would result in welfare increases of $8-9\%$ and a significant drop in rejection rates.

Another mechanism can also be applied to this problem: add the reported hospital priority and modify the termination rule of the DA algorithm. Under this and the previous mechanism, the same manipulation incentives would be present, yet welfare would increase $9-12\%$ and rejection rates would drop significantly.

6 Merit-based tie-break

The SaberPro is an end-of-study nationwide exam. To graduate all students in all fields of study in the country must take the test, which is used to decide admissions into Colombian graduate studies. In this section I analyze what would happen if the exam results were used to break ties within the coarse priority structure. In a procedure known as merit-based tie-breaking, this process is standard procedure in several school choice settings.

I recovered the results for 189 (out of 196) Medical Lab Scientists, 713 (out of 828) Nurses, 337(out of 386) physicians, and 661 (out of 708) Dentists. The tests usually consist of two sections: a generic section that is the same for all fields; and an advanced section specific to each field. Unfortunately, the advanced sections specific to particular fields were modified many times prior to the October 2014 allocation. Consequently, to break the ties in the priority structure, I use as the means the average in the generic portion of the test.

As in the previous sections, I run 100 simulations of the allocation using the random tie-break, and I compare it to the allocation produced by the merit-based approach. 6 shows the results.

	Welfare Gains vs DA (%)	Num. of Prof. with Not Respected Prio.	Num. of Allocated Prof. with Not Respected Prio.	% of Pos. Rejected
Med-Lab				
Implemented Mechanism	0	0	0	5
TTC	5	14	0.2	4
DA-Stopping	8	9	0.3	1
Report TB	9	0	0	3
DA Stopping+Report TB	12	10	1	0
Sequential DA	13	13	0.1	1
SOSM	0	0	0	4
Resp. Prio. Allo.	2	3	0	4
Nursing				
Implemented Mechanism	0	0	0	4
TTC	0	14	0	3
DA-Stopping	8	88	0	3
Report TB	5	0	0	4
DA Stopping+Report TB	9	89	0	3
Sequential DA	8	112	0	0
SOSM	0	0	0	4
Resp. Prio. Allo.	0	2	0	4
Med-Doc				
Implemented Mechanism	0	0	0	4
TTC	0	26	26	4
DA-Stopping	0	0	0	4
Report TB	1	0	0	4
DA Stopping+Report TB	1	0	0	4
Sequential DA	1	9	9	4
SOSM	0	0	0	4
Resp. Prio. Allo.	0	0	0	4
Dentistry				
Implemented Mechanism	0	0	0	3
TTC	2	11	0.13	2
DA-Stopping	9	64	0	1
Report TB	9	0	0	1
DA Stopping+Report TB	12	64	0	1
Sequential DA	12	65	0	0
SOSM	0	0	0	3
Resp. Prio. Allo.	2	4	0	2

Table 5: Comparison Between Different Algorithms (Complete Preferences)

	Med-Lab	Nurses	Med-Docs	Dentists
Worse	10	34	85	11
	(1.47)	(2.83)	(3.23)	(2.69)
Same	11	28	213	31
	(5,96)	(5.78)	(7.36)	(6.71)
Better	60	86	74	54
	(3.4)	(3.04)	(6.45)	(3.52)

Table 6: Merit-based vs Random Tie-break

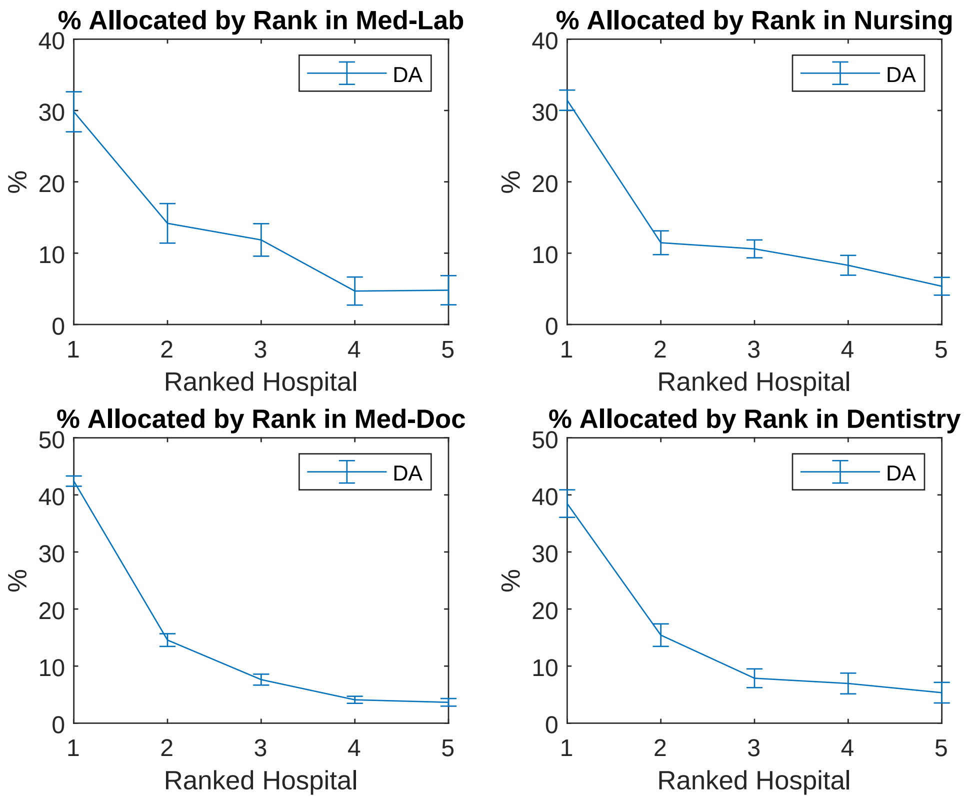
In the case of the allocation of physicians, because there are more positions than physicians, the quality of these positions is a zero sum game. Hence, the more desired positions have a higher chance of receiving good physicians, which generates inequality within the country in the quality of the allocated professionals.professionals. This outcome is contrary to the objective of the program, which is to give rural communities the chance to acquire good health professionals.

When the merit-based approach is applied to the other professions, congestion leads to a significant increase in the average quality of allocated health professionals, from which most hospitals benefit. However, in some states, regional priorities cause health professionals to be redistributed, and this negatively affects some hospitals. With regards to the welfare of health professionals, I do not observe a significant change int he average welfare when there is a merit-based tie break.

7 Robustness Check

The accuracy of the statements about welfare made in the previous sections depend naturally on the accuracy of the preference estimates. These estimates are based on the assumption that the reported preferences are the true preferences. Beyond the fact that the allocation is determined using the DA algorithm, there are several reasons to make this assumption.

As noted above, a professional cannot manipulate the allocation to achieve an exemption. A health professionals' strongest motivation to report strategically would be to over-report positions in their state of origin or graduation because doing so would give them a higher priority. This, in turn, would produce a higher probability of allocation. Two regularities show that this strategy



would not be effective. First, as reported in Table 3, states that give a higher priority experience a decrease in their ranking position. Indeed, Haeringer and Klijn (2009) show that in the truncated DA there are no incentives to misreport preferences within the reported preferences. The same result holds in the environment that I examine here. In other words, health professionals have no incentive to misreport preferences. If the last positions reported had a higher probability of being from a priority-giving state, this would constitute evidence of strategic behavior, but I have found no such evidence. In fact, I observe that the probability of being allocated to the last reported hospital is lower than the probability of being allocated to the second-to-last reported hospital, as shown in Figure **??**.

Second, the greater the congestion in a allocation, the greater the incentives to report a priority giving state. This should produce a higher weight in the travel distance coefficient. However, as table **??** shows, the biggest average weight of travel costs is observed in the least congested allocation (that of physicians).

To provide additional evidence of nonstrategic behavior, I estimate the utilities using the Attention Logit model of Abaluck and Adams (2017). In their model both utility and attention are estimated by exploiting the fact that the cross-characteristic responses [15] should be symmetric, assuming that health professionals are perfectly rational utility maximizers. However, because of imperfect consideration, this property fails. The goal of attention estimates is to recover the aforementioned symmetry.

It is possible to build an Attention Logit model that allows random coefficients. However, due to the large number of hospitals that health professionals can choose from, the estimation takes a considerable amount of time because the model ideally calculates the probability that each subset of alternatives is considered. In this study, because there are a large number of options, the estimation uses an approximation of this assumption by simulating 1500 possible subsets. Adding the random coefficients would entail a computational challenge beyond the scope of this paper. Because preference heterogeneity plays an important role in welfare analysis, I use the random coefficients model in my welfare analysis .

The Attention Logit model provides evidence that there is no systematic strategic behavior. The model estimates both the attention parameters–the probability that a given option is considered– and the utility parameters. I find similar estimates for both, indicating that the incentives that guide the search for positions are similar to the incentives that motivate participants to choose among them. Strategic behavior would break this similarity. Table 7 shows that the coefficient of priority-giving states is higher (relative to the coefficient on wages) for all professions in the Attention estimation than in the Utility estimation.

8 Conclusions

This paper evaluates the Compulsory Social Service (CSS), a program developed by Colombia's Ministry of Health to tackle inequality in the allocation of the country's health professionals. This is a long-standing policy concern in many countries because of the difficulty of filling medical

[15] This is the discrete choice analogue of the Slutsky matrix, which is symmetric and negative semidefinite for the utility-maximizing agent.

	Med-Lab	Nursery	Dentists	Physicians
Utility				
Wage	0.451	0.583	0.371	0.330
	(0.077)	(0.0269)	(0.03865)	(0.024)
Priority State	1.574	1.863	1.828	1.954
	(0.132)	(0.053)	(0.075)	(0.068)
Travel Dist.	-0.304	-0.234	-0.347	-0.348
	(0.025)	(0.0135)	(0.0191)	(0.0024)
UBN	-0.012	-0.0218	-0.012	-0.030
	(0.0024)	(0.0011)	(0.0014)	(0.0023)
Attention				
Wage	0.451	0.583	0.371	0.33
	(0.081)	(0.037)	(0.027)	(0.019)
Priority State	1.620	1.9124	2.179	2.62
	(0.372)	(0.337)	(0.159)	(0.12)
Travel Dist.	-0.28	-0.358	-0.268	-0.49
	(0.014)	(0.008)	(0.008)	(0.011)
UBN	-0.069	-0.0219	-0.0123	-0.033
	(0.0027)	(0.001)	(0.001)	(0.001)
Constant	-0.43	-0.048	-0.147	0.1122
	(0.25)	(0.14)	(0.170)	(0.10)

Table 7: Attention Logit Estimates

positions in rural areas. A salient characteristic of this program is that it is compulsory: health professionals in the areas of Medical-Lab Science, Dentistry, Nursing, and Medicine must participate in this allocation to become professionally certified. Since October 2014, the allocation has been determined by a single-offer centralized mechanism that uses the Deferred Acceptance algorithm to determine the allocation.

I use the data on the reported preferences in the first allocation, which was conducted in October 2014, to estimate the preferences of each health professional for every hospital. I estimate a random coefficients model of the preferences. These preferences allow for a correlation in the coefficients, and they are specifically modeled to allow for a correlation for hospitals within each state. I find the health professionals' preferences for hospitals to be significantly heterogeneous.

Referring to these estimates, I show that moving from the previously-used decentralized system of lotteries to the current centralized mechanism has produced significant welfare gains for physicians. From the fact that under the decentralized mechanism health professionals were able to avoid positions that fell below their acceptance threshold, and simulating the outcome had the decentralized mechanism still been in use, I obtain the average marginal utility a health professional would require to accept a position. I then simulate the outcome of the centralized mechanism in the absence of the requirement that students accept the assignment determined by the mechanism. I find that, given the choice, about 30% of physicians would reject their hospital assignment, which implies that for the policy to be successful, assignments must be mandatory.

Then I study the allocations' efficiency from the viewpoint of health professionals. Allocations determined through use of the DA algorithm are known to be strategy proof and respect priorities. However, these characteristics constrain possible allocations and, therefore, (possibly) they entail welfare costs. I show that the respect priorities component is not very costly in terms of average welfare. Thus, I show that under several mechanisms all priorities are respected. I also show that the cost of *strategy-proofness* is of the first order. Slight deviations can result in welfare gains of up to 12%. This reflects the fact that the preferences of health professionals include all possible hospitals.

Finally, I show that in case of the physicians market, using a random tie-break rather than a merit-based one was a good policy decision. The former would have allocated the best physicians to the most desirable positions, leaving those located further away with worse physicians (as measured by the results of their end-of study-exams). Doing a merit-based tie-break would benefit most of the other health professions, which face significant congestion (i.e., their ratios of health professionals to positions are very high). Behind this fact is the intuition that the less attractive health professionals would have a greater chance of not being assigned.

References

(????), "Rural healthcare workforce." https://www.ruralhealthinfo.org/topics/health-care-workforce. Accessed: 2018-09-30.

Abaluck, Jason and Abi Adams (2017), "What do consumers consider before they choose? identification from asymmetric demand responses." Technical report, National Bureau of Economic Research.

Abdulkadiroğlu, Atila, Nikhil Agarwal, and Parag A Pathak (2017), "The welfare effects of coordinated assignment: evidence from the new york city high school match." *American Economic Review*, 107, 3635–89.

Abdulkadiroğlu, Atila, Yeon-Koo Che, and Yosuke Yasuda (2011), "Resolving conflicting preferences in school choice: The "boston mechanism" reconsidered." *The American Economic Review*, 101, 399–410.

Abdulkadiroğlu, Atila, Parag A Pathak, and Alvin E Roth (2009), "Strategy-proofness versus efficiency in matching with indifferences: Redesigning the nyc high school match." *American Economic Review*, 99, 1954–78.

Abdulkadiroglu, Atila and Tayfun Sönmez (2003), "School choice: A mechanism design approach." *The American Economic Review*, 93, 729–747.

Acevedo, Karina Margarita, Iván David Bornacelly, et al. (2014), "Panel municipal del cede." Technical report, UNIVERSIDAD DE LOS ANDES-CEDE.

Agarwal, Nikhil (2015), "An empirical model of the medical match." *American Economic Review*, 105, 1939–78.

Agarwal, Nikhil (2017), "Policy analysis in matching markets." *American Economic Review*, 107, 246–50.

Anderson, Michael L. (2013), "GEOCODEOPEN: Stata module to geocode addresses using MapQuest Open Geocoding Services and Open Street Maps." Statistical Software Components, Boston College Department of Economics, URL https://ideas.repec.org/c/boc/bocode/s457733.html.

Ashlagi, Itai, Afshin Nikzad, and Assaf Romm (2015), "Assigning more students to their top choices: A tiebreaking rule comparison." Technical report.

Azevedo, Eduardo M and Jacob D Leshno (2016), "A supply and demand framework for two-sided matching markets." *Journal of Political Economy*, 124, 1235–1268.

Calsamiglia, Caterina, Chao Fu, Maia Güell, et al. (2014), "Structural estimation of a model of school choices: The boston mechanism vs. its alternatives." Technical report.

Chen, Yan and Onur Kesten (2017), "Chinese college admissions and school choice reforms: A theoretical analysis." *Journal of Political Economy*, 125, 99–139.

Erdil, Aytek and Haluk Ergin (2008), "What's the matter with tie-breaking? improving efficiency in school choice." *The American Economic Review*, 98, 669–689.

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

Haeringer, Guillaume and Flip Klijn (2009), "Constrained school choice." *Journal of Economic Theory*, 144, 1921–1947.

He, Yinghua et al. (2012), "Gaming the boston school choice mechanism in beijing." *Toulouse School of Economics Working Papers*, 12–345.

Jaramillo, Paula, Çağatay Kayı, and Santiago Velez (2016), "Matching problems with priorities and preferences:
compulsory social service allocation in colombia."

Kamada, Yuichiro and Fuhito Kojima (2015), "Efficient matching under distributional constraints: Theory and applications." *American Economic Review*, 105, 67–99, URL http://www.aeaweb.org/articles?id=10.1257/aer.20101552.

Kojima, Fuhito and M Utku Ünver (2014), "The "boston" school-choice mechanism: an axiomatic approach." *Economic Theory*, 55, 515–544.

Luflade, Margaux (2017), "The value of information in centralized school choice systems (job market paper)."

McFadden, Daniel and Kenneth Train (2000), "Mixed mnl models for discrete response." *Journal of applied Econometrics*, 447–470.

Pathak, Parag A and Jay Sethuraman (2011), "Lotteries in student assignment: An equivalence result." *Theoretical Economics*, 6, 1–17.

Roth, Alvin E (1986), "On the allocation of residents to rural hospitals: a general property of two-sided matching markets." *Econometrica: Journal of the Econometric Society*, 425–427.

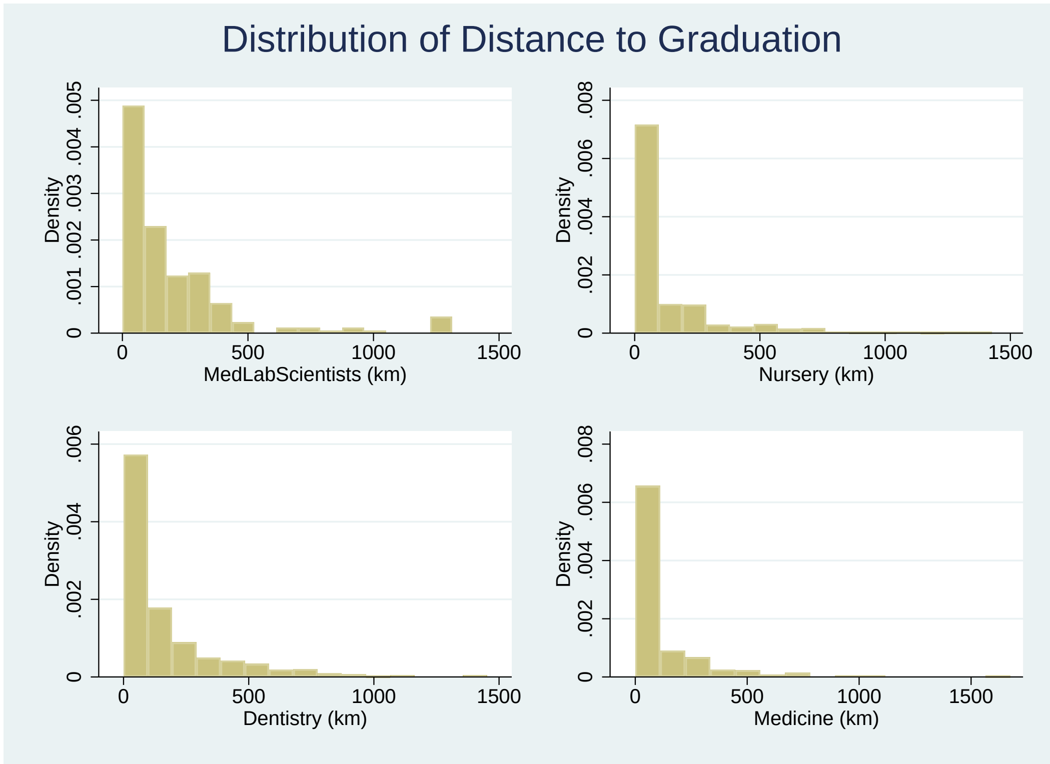
Roth, Alvin E (2008), "Deferred acceptance algorithms: History, theory, practice, and open questions." *international Journal of game Theory*, 36, 537–569.

Shapley, Lloyd and Herbert Scarf (1974), "On cores and indivisibility." *Journal of mathematical economics*, 1, 23–37.

Sönmez, Tayfun and Tobias B. Switzer (2013), "Matching with (branch-of-choice) contracts at the united states military academy." *Econometrica*, 81, 451–488, URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA10570.

Train, Kenneth E (2009), *Discrete choice methods with simulation*. Cambridge university press.

A Distribution of Distance from Preferred Hospital to Graduation City



Travel distance from

hospital to graduation (km)

B Top Trading Cycle Algorithm

TTC:

Step 1: Assign a counter for each hospital which keeps track of how many unfilled positions there are. Initially, set the counter of a hospital equal to its quota. Each professional points to his preferred hospital. Each hospital points to the professional who has the highest priority (If two professionals are indifferent in terms of priority relation, then the strict preference relation of the hospital is used for the assignment.) Since the number of professionals and hospitals are finite, there is at least one cycle. (A *cycle* is an ordered list of distinct hospitals and distinct doctors $(d_1, h_1, d_2, ..., d_k, h_k)$ where d_1 points to h_1, h_1 points to d_2,...,d_k points to h_k, and h_k points to d_1.) Moreover, each hospital can be part of a at most one cycle. Every professional in a cycle is assigned a position at the hospital he points to and is removed. The counter of each hospital in a cycle is reduced by one, and if it is reduced to zero, the hospital is removed. Counters of other hospitals stay put.

Step k: Each remaining professional points to the preferred hospital among the remaining ones, and each remaining hospital points to the professional with the highest priority among the remaining ones.(If two doctors are indifferent in terms of priority relation, then the strict preference relation of the hospital is used for the assignment.) There is a cycle. Every professional in a cycle is assigned a position at the hospital he points to and is removed. The counter of each hospital in a cycle is reduced by one, and if it is reduced to zero, the hospital is removed. Counters of other hospitals stay put. A professional that is never a part of a cycle is exempt from the compulsory social service.

C Complete Table of Estimates

		Med-Lab	Nursing	Med-Docs	Dentists
Means	Gender				
Travel Distance Origin					
	F	-1.413	-1.863	-2.455	-1.444
	M	-1.192	-1.417	-1.798	-1.413
Travel Distance Graduation					
	F	-0.532	-1.626	-1.656	-0.694
	M	-0.708	-3.024	-1.511	-0.028
UBN-Index					
	F	-0.056	-0.114	-0.141	-0.083
	M	-0.023	-0.108	-0.118	-0.109
Coca					

	F	-1.419	-0.644	-1.095	-1.124
	M	-0.147	-2.013	-0.997	-0.338
Living Place					
	F	-0.866	-0.256	-0.577	-0.587
	M	-0.163	-0.882	-0.179	-0.541
Weekends Shift					
	F	-2.296	-0.555	-0.176	-1.231
	M	-0.001	-0.382	-0.114	-0.031
Tot. Pop.¡100,000					
				-0.047	
Tot. Num. Positions					
				0.061	
Std. Deviation					
Travel Distance Origin					
	F	1.506	1.427	2.430)	1.244
	M	1.875	2.202	1.542	1.455
Travel Distance Graduation					
	F	1.265	2.068	1.862	1.131
	M	1.358	2.157	2.736	0.237
UBN-Index					
	F	0.237	0.667	0.178	0.261
	M	0.214	0.603	0.191	0.470
Coca					
	F	1.630	1.102	1.940	1.252
	M	0.568	2.291	0.875	1.012
Living Place					
	F	1.111	0.663	0.491	1.016
	M	0.585	1.254	0.418	0.942
Weekends Shift					
	F	2.276	1.028	0.414	1.297
	M	0.037	0.779	0.329	0.202
Tot. Pop. ¡100,000					
				0.183	
Tot. Num. Positions					

			0.114	
Antioquia	3.613	3.180	2.443	2.250
Atlantico	2.599	2.313	0.993	
Bogota	4.127	2.504	2.415	3.531
Bolivar	2.659	1.958	5.473	2.377
Boyaca	3.145	1.550	1.011	2.631
Caldas	2.022	2.482	1.739	2.266
Caqueta	2.383		1.571	1.995
Cauca		1.924	1.912	2.026
Cesar	2.563	1.948	1.678	1.575
Cordoba	3.510	1.875	0.568	1.949
Cundinamarca	2.543	2.005	1.300	2.041
Choco	2.569	1.937	0.833	
Huila	2.606	1.721	1.280	2.237
La Guajira	2.047	1.808	2.607	1.703
Magdalena	2.233	1.872	1.110	3.749
Meta		1.926	0.948	1.676
Narino	2.419	3.917	1.725	2.189
N. De Santander	2.705	2.160	0.775	1.707
Quindio			2.781	
Risaralda			3.013	1.904
Santander	3.890	2.537	1.253	2.292
Sucre	3.008	2.355	2.797	
Tolima	2.074	1.707	0.499	1.675
Valle del Cauca	3.758	3.163	4.418	1.965
Arauca	3.227	2.341	2.395	1.826
Casanare	2.597	1.774	0.776	1.515
Putumayo	2.346	2.579	2.415	2.369
San Andres			5.377	
Amazonas	2.527	1.939	4.450	2.610
Guainia		2.705	1.706	2.587
Guaviare		2.112	0.768	
Vaupes			0.718	

Table 8: Preference Estimates of Health Professionals

D Single vs Multiple Tie-Break

I use different algorithms proposed in the literature and compare their efficiency, as measured by average welfare, with the one resulting from the implemented mechanism. In order to make our results comparable with the results in the *school choice* literature (e.g. Abdulkadiroğlu et al. (2009), He et al. (2012), and Calsamiglia et al. (2014)) I compare the efficiency of the allocation as if the implemented algorithm used only the reported preferred hospitals. I will refer to it as the *DA* allocation.

It is important outline that the implemented algorithm with these characteristics is *strategy proof* and *respects priorities*. Also, the priority structure of the hospitals is so coarse that the mechanism has a significant randomness, i.e. it relies a lot on the single tie-breaks. Define a matching $\mu = \mu'$ iff $\mu(i) = \mu'(i)$ for every $i \in \mathcal{P}$. I conducted 1'000,000 simulations of the *DA* allocation and found all resulting matchings being different.

The first algorithm I compare *DA* with is the *Top Trading Cycles* or *TTC* proposed initially by Shapley and Scarf (1974) and in the *school choice* setting by Abdulkadiroglu and Sönmez (2003). This algorithm is *strategy proof* and *efficient*. Hence, comparing *DA* with *TTC* yields a measure of how costly it is in terms of efficiency the *respecting priorities* property. I find that the welfare actually decreases for the allocation of physicians and nurses. For the other two professions the gains from TTC are slightly higher.

The next algorithm I compare *DA* with is the *Student(Professional) Optimal Stable Match* or *SOSM* proposed by Erdil and Ergin (2008). The single tie break used can generate welfare decreasing cycles. Comparing this algorithm with *DA* gives a measure of how much welfare is lost on average due to this welfare reducing cycles. I find that there are small welfare gains when moving from *DA* to *SOSM* in every profession there are some potential gains from using this mechanism.

Another proposed modification in the literature is moving from a single tie-break to a multiple tie-break, i.e. instead of having the same tie-break for all hospitals allowing each hospital to have their own. Ashlagi et al. (2015) show that when comparing the allocations resulting from *DA* under single and multiple tie-breaking rules, the former has more subjects allocated at the firsts positions and very lasts positions while the latter tends to allocate more concentrated on the middle. Our results coincide with their results (as shown in table 9). In the particular case I are studying I find that moving from the single tie-break to the multiple one induces a welfare loss for all professions.

I compare all of the aforementioned mechanisms under the single and multiple tie-breaking rules. As shown by Pathak and Sethuraman (2011) I find no difference in the *TTC*'s welfare between the two. For the other mechanisms I consistently find that the single tie-break is superior in welfare than the multiple tie-break.

Finally, I modify the compare the *DA* with the *Augmented Choice Deferred Acceptance* (or *ACDA*) proposed by (Abdulkadiroğlu et al., 2011). This algorithm ideally should ask the professionals at which positions they would like

Avg Welfare		Med-Lab	Nursery	Med-Doc	Dentists
Single Tie-Break					
	DA STB	3.83	7.79	14.58	6.27
	MLDA STB	4.31	8.22	14.71	6.29
	SOSM STB	4.14	8.00	14.59	6.33
	TTC STB	4.16	8.00	14.50	6.48
	MLTTC STB	4.41	8.35	14.70	6.29
Multiple Tie-Break					
	DA STB	3.48	7.58	14.58	6.16
	MLDA MTB	4.01	7.96	14.71	6.18
	SOSM MTB	4.12	8.07	14.59	6.55
	TTC MTB	4.15	8.03	14.50	6.56
	MLTTC MTB	4.45	8.24	14.70	6.36

Table 9: Comparing Different Algorithms

to have an additional advantage in the tie-break. They show how this mechanism is *strategy proof.* I first calculate the probability of being assigned at each hospital for every professional under *DA*. This is done through a computation of 10,000 simulations. Then the additional advantage in the tie-break is given to each professional in their most likely hospital. I define this algorithm as *Most Likely Deferred Acceptance.* I find this algorithm produces a welfare gain of in cases relative to the Deferred Acceptance. However, the welfare gains relative to the TTC are ambiguous. Interestingly, the resulting allocation first order stochastically dominates the allocation and results in an increase of almost 2% in the allocated professionals. Interestingly, following the same principle with the *TTC* I find once again that it procedure, i.e. giving a higher to the priority in the tie-break to the most likely alternative, results in similar welfare gains and a first order stochastic improvement.