

“Anything for my child”: Incentives in the San Francisco Elementary School Matching Process¹

Ambika Acharya, Tara Balakrishnan, Sarah Rosston

Introduction

Making sure their children get a high-quality education is a major priority for parents, and because they are so driven to see their children succeed, they will invest a lot of energy and resources to maximize their child’s chance of attending a good school. In this project, we look at the San Francisco elementary school matching process, made by the San Francisco Unified School District (SFUSD). In this process, students are assigned an elementary school based on their preferences and with regards to a variety of different factors including: school proximity to their home and census tract information about whether the student lives in an area with low test scores. We will analyze whether this system is strategy-proof, the impacts it has on school diversity and also do case studies to see how parents respond to and interact with the matching process.

We specifically evaluate the effects of this system on the diversity of schools. More than 25% of elementary schools in the SFUSD are “racially isolated,” meaning at least 60% of students are the same race. Ensuring diversity in schools is not only important for equity, segregation in schools also creates legal issues for school districts; SFUSD was sued by the NAACP in 1983 because of *de facto* segregation in schools, so the school district is concerned about legal implications of diversity. Although the addition of census tract was meant to serve as a proxy for race, since using race as an assignment criteria was barred from the matching process by a court case, we believe that elementary schools in San Francisco are relatively segregated because this system prefers students who are close to the school district, and those who rank more schools--a group that tends to be wealthier and English-speaking.²

Current System

Matching Algorithm

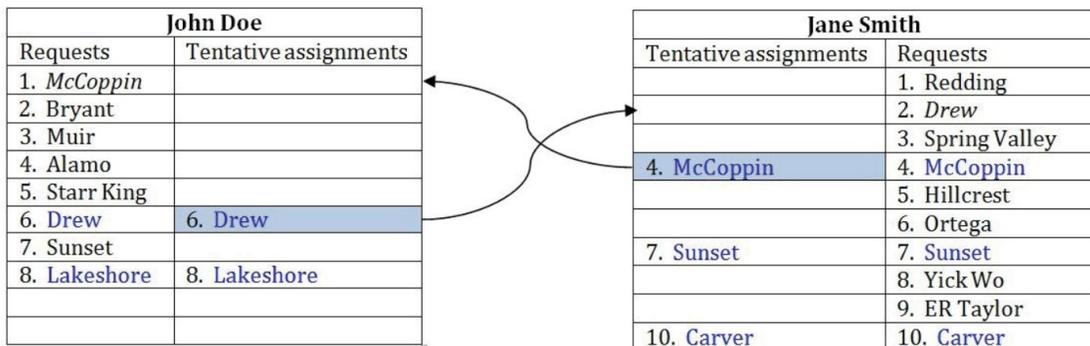
The algorithm as outlined by SFUSD:

- Students rank the schools in the order in which they truly like them. They are encouraged to only rank the schools that they would actually attend and leave the remainder of schools off their list.
- Then the algorithm determines the preference level that each student has for a particular seat, in an order determined by a random lottery number assignment.

¹ Our video: <https://youtu.be/-uCcspAOhp8>

² <http://sfpublicpress.org/news/2015-02/as-parents-get-more-choice-sf-schools-resegregate>

- The highest preference is given to students who attended the associated preschool or have an older sibling at that school. These factors however are less interesting to consider as it's impossible to strategize around.
- Preference is then given to students who live in attendance areas(neighborhoods) which are categorized as CTIP-1.
 - CTIP is the Census Tract Integration Preference, where CTIP-1 refers to neighborhoods with low average test scores.
 - The motivation behind this to give students who live in traditionally underperforming areas the chance to attend a better school. CTIP-1 low-income areas house are around 60% African-American or Latino, so this also progresses the de-segregation mission of the SFUSD.
- The 3rd highest preference is given to local students who live in the same attendance area as the school they are trying to match with. (This is known as their “attendance area school.”)
- There are a few other living situations which give a student preference, but fewer students benefit from them.
- The algorithm then assigns students a temporary seat at various schools(one to many matching) based on their rankings and their preference level. At this stage, student may have multiple tentative offers.
- Lastly it swaps students' seats at school such that all students involved in the transfer are better off, and until no more swaps can be performed where everyone involved is better off, as in the example below where John Doe and Jane Smith trade tentative assignments.



3

Properties of the Algorithm

The algorithm is both pareto optimal and mostly strategy proof. If students would not accept certain schools, either by moving out of San Francisco or attending a private school, they still benefit from ranking all schools.

³ <http://www.sfusd.edu/en/assets/sfusd-staff/enroll/files/2012-13/Assignment-Transfer%20process%20explained.pdf>

Proof: mechanism is pareto optimal

Suppose the mechanism produces an assignment where student A could improve her assignment without making another student worse off. Then every other student either keeps the same school assignment or switches to a school he or she prefers. Since a spot at a school must open for A to improve her assignment, there must be some non-zero number of students who would all be better off by switching assignments, but this is a contradiction because then there would be a cycle in the graph and matching ends when there are no remaining cycles.

Proof: mechanism is (mostly) strategy proof

If a student ranks School 1 and School 2 in the reverse order of her true preference, she could be reassigned from her preferred school to her less preferred school, but if she ranks them in order of preference, she can only move from the less preferred school to the more preferred school. Since a student can get multiple tentative assignments which are not dependent on rankings, ranking a school higher cannot improve the chances of a cycle. This goes against advice from several parenting blogs⁴ to rank an attendance area school higher to increase the likelihood of being in a cycle; because receiving an offer is only contingent on a school being somewhere in a student's ranking, a student does not increase her likelihood of a tentative offer at her attendance area school by ranking it 2nd instead of 8th.

While some parenting blogs give incorrect advice, they correctly advise submitting a list of preferences beyond just the schools they would be willing to send their child to. Students cannot improve chances of a good assignment by mis-ordering their choices, but they can improve their outcome by ranking schools that they consider less preferable than private schools or moving to another school district. Consider the following example where the capacity of all schools is 1 and the only public school Sheila Hill is willing to attend is Lilienthal.

	Anna Yao (attendance area = Drew)	Sheila Hill (attendance area = Clarendon)	José Garcia (attendance area = Lilienthal)	Daniel (attendance area = Malcolm X)
1.	Clarendon	Lilienthal	Drew	Lilienthal
2.	Drew	<i>none</i>	Lilienthal	Clarendon
3.	Lilienthal	<i>none</i>	Clarendon	Drew
4.	Malcolm X	<i>none</i>	Malcolm X	Malcolm X

Table 1: Student Preferences

⁴ <http://www.sfkfiles.com/2012/10/the-inner-workings-of-sfusd-lottery.html>

In the first round, Anna receives a tentative offer at Drew, José at Lilienthal, Daniel at Clarendon, and Sheila gets a spot at Malcolm X. Since Anna, Jose and David form a preference cycle as shown in Figure 1, they trade offers so Anna gets assigned to Clarendon, Jose to Drew, Daniel to Lilienthal and Sheila to Malcolm X, an assignment without preference cycles as shown in Figure 2.

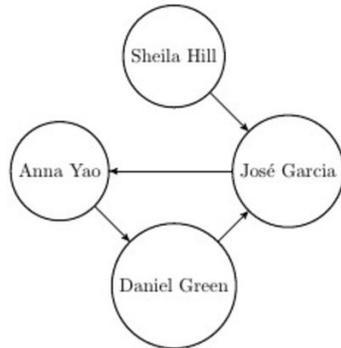


Figure 1: Tentative Offers

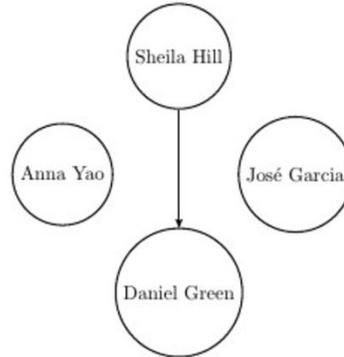


Figure 2: Result after removing cycles

If Sheila had added attendance area school in her rankings, however, she would have been assigned to Lilienthal, her first choice school. If she ranked all schools, Anna would receive a tentative offer at Drew, Sheila at Clarendon, José at Lilienthal, and Daniel at Malcolm X, forming the preference cycle in figure 4 (this is based on each of their respective attendance areas which is one of the tie-breakers).

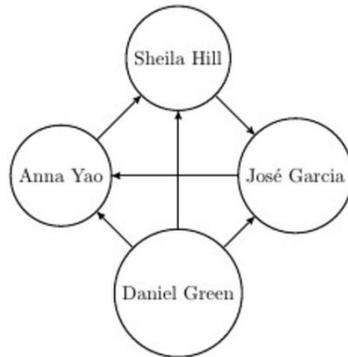


Figure 3: Tentative Offers

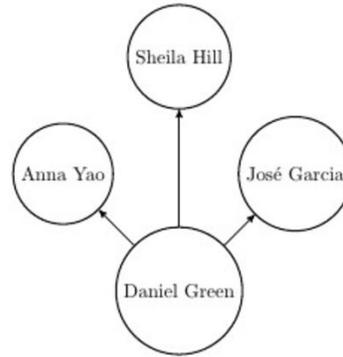


Figure 4: Result after removing cycles

When the preference cycle is removed, Anna would receive a spot at Clarendon, Sheila at Lilienthal, José at Drew, and Daniel at Malcolm X, an assignment without cycles as shown in Figure 4. In this case, Sheila is better off because she ranked a school she would not be willing to attend.

Simulation

Data

Because the San Francisco school board does not release any data on students' rankings for the system, we generated student objects by piecing together information from multiple datasets published by SFUSD after the 2016 matching process. The data generation process took a bulk of our time because of the lack of centralized information about this process including student rankings, demographics of attendance areas CTIP tracts, and mappings between attendance area and census tracts. Interestingly, even Al Roth, a Stanford Economics professor and Nobel Laureate was not allowed to help implement the matching system when he was working on a project to help improve the system because of concerns about student data.⁵

Methods for Generating Students and Schools

The list of schools we use is the list of elementary public schools in San Francisco. Since some schools have language immersion programs, we count the immersion program as a separate school to make the matching process more intuitive because immersion programs do not have attendance areas and have a set of seats to allocate separate from the rest of the school. For example Parker Elementary has a general program for kindergarteners and a Chinese language immersion program. In our school list we add both 'Parker' and 'ParkerCHN.' Our list contains 107 schools and each has a name and capacity for their kindergarten class(how many students⁶ they will accept from the matching process.)

Now, we describe the schema for a student object and how we went about constructing each attribute:

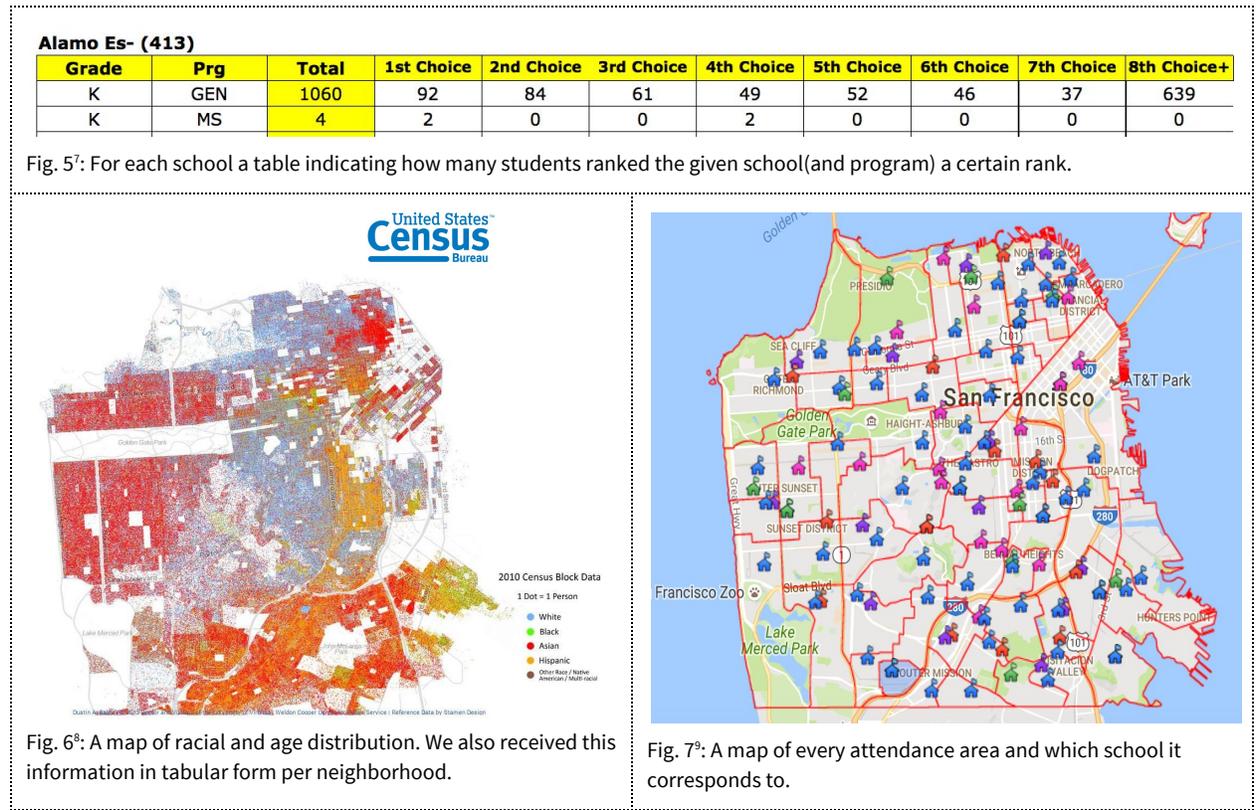
<p><u>Student</u></p> <p>race = {hispanic, white, black, american indian, asian, native hawaiian, multiracial}</p> <p>attendance area school = {school from list of schools}</p> <p>CTIP = {1,2}</p> <p>rankings = {ordered list of schools}</p> <p>private_school_cutoff = {number at which they stop ranking schools because they would rather send their child to private school}</p>

A student's attendance area school is defined as the school they have preference for due to proximity based on their neighborhood. A student's neighborhood is labeled by its Census Tract Integration Preference(CTIP), which is either CTIP-1(for low-income) or CTIP-2.

⁵ <https://marketdesigner.blogspot.com/2010/09/san-francisco-school-choice-goes-in.html>

⁶ [http://web.sfusd.edu/Services/research_public/rpa_student_enrollment/SFUSD%20School%20Site%20List%20and%20Summary-%20Student%20Enrollment%20\[Most%20Current\].pdf](http://web.sfusd.edu/Services/research_public/rpa_student_enrollment/SFUSD%20School%20Site%20List%20and%20Summary-%20Student%20Enrollment%20[Most%20Current].pdf)

To generate students we combined the the following datasets: a cumulative summary of the rankings for every school from the 2016 matching process[Fig. 5], census tract(neighborhood) race and age breakdown taken from the United States census[Fig. 6], and a map of the attendance areas across San Francisco[Fig. 7].



We manually overlaid the data from maps [Fig. 2] and [Fig. 3] block by block to calculate for every census tract (neighborhood) the CTIP, the number of kindergarten-aged students, the number of kindergarten-aged students per race and the attendance area school. Because we only had data on the racial distribution across all ages and on the number of 5-9 year-olds, we assume that 20% of 5-9 year-olds are kindergartners and that the racial distribution is constant across age groups, an assumption that probably underestimates the number of minority students.¹⁰ We then went through each census tract and generated a student object for the number of kindergarten-aged students for every race, assigned each student the CTIP of that census tract and assigned the attendance area school of that census tract.

Since we do not know how individual students rank schools, we aggregate data about how many times a school was ranked in each position from [Fig. 1] to create a probability distribution for

⁷ http://www.sfsud.edu/en/assets/sfsud-staff/enroll/files/2016-17/Report_Requests.pdf
⁸ <http://www.nmilholland.com/blog/projects/kidsdotmap/>
⁹ <http://enrollinschool.org/lookup/area.php>
¹⁰ <https://www.brookings.edu/blog/up-front/2011/08/26/america-reaches-its-demographic-tipping-point/>

ranking a school. We defined the probability of a school being next on a student's ranked list as the score of a school as a fraction of the total score of remaining schools where score is defined as below.

$$score(school) = \left(\sum_{i=1}^{8+} (100 - 10i) * ranks_i(school) \right) + 1000I\{school = attendance\ area\}$$

In the score, $ranks_i(school)$ is the number of times a school was ranked in position i . Because we believe students generally prefer shorter commutes, we increase the score for an attendance area school. We also assume students care whether a school is an immersion school, so we multiply the score of immersion schools by a "immersion preference" drawn from a normal distribution centered at one and bound by zero. This method produces similar distributions of rankings for each school without every student having the same rankings.

While these numbers reflect a distribution of choices similar to the actual choices of students, the school district says that preferences are actually substantially different by race. Because we do not have information about the strength and impact of this effect, we do not take this into consideration in our model.

Additionally, we presume that only parents who are wealthy can afford to not rank every school and instead have a cutoff after which they would rather send their child to a private school than to a school they haven't ranked. Since the number of rankings a student puts down affects decision making in the matching algorithm, we determine a cutoff for a student from a CTIP-2 neighborhood by selecting a random cutoff between 9 and 107 (we assume all students rank 8 schools.)

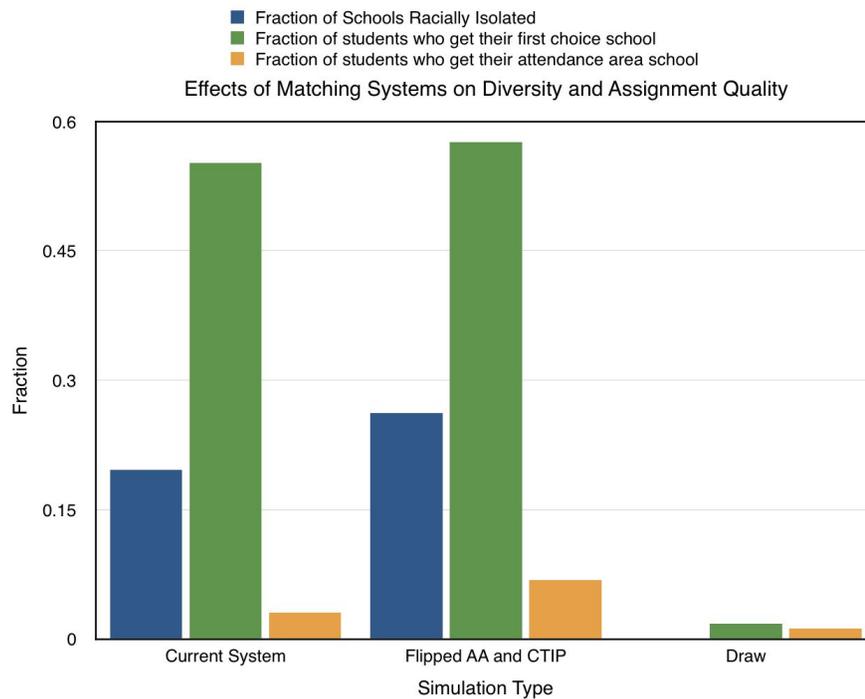
Simulation Results

Using the students and schools we created, we wrote a simulation to assign these students to schools using the currently SFUSD matching system. If we assume all students rank all possible schools, the assignment problem becomes too computationally intensive for us to simulate since it requires repeatedly removing cycles on a graph with hundreds of thousands of edges. The SFUSD algorithm description is not clear about whether it only removes 2-person preference cycle or it removes multi-person cycles. Only considering 2 person cycles which would be computationally simpler, but not pareto-optimal so we assume the school district considers all preference cycles. True student preferences might also lead to a computationally simpler problem if most students can be assigned to their top choice school immediately.

Because of these considerations, we ran the simulation we wrote on 10% of the students, and set the capacity of every school to 10% of its capacity. We believe that running different matching systems on this subset will still yield similar results to running the same algorithms on the entire student set, with many fewer computations. We ran three simulations of matching systems and compared their effects of three metrics to better understand why the SFUSD has settled on the current system.

The three systems we used are: the current SFUSD system as described in the previous section; the current SFUSD system with preference for Attendance Area school higher than CTIP; the draw, a lottery system based off of the Stanford Draw.

Since one of our major goals was to identify the effects different matching systems have on diversity in schools, we calculated the fraction of schools in each system that were racially isolated, meaning they have at least 60% of a single race. We also compared the three systems on the percentage of students who received their first choice school since the satisfaction of the families is a big indicator for a good system. Lastly, we compared the systems based on students who received their attendance area school, which indicates how many students were able to go a school within walking distance. We recognize that many students might not want to go to their attendance area school and may not rank is highly meaning it might not be the best indicator of a successful system, however we still take note of this metric. Here is a summary of results from our simulation:



We see that in terms of parent satisfaction, the current system and the modified system do well, giving over half the students their first choice. The draw, however, can barely give any students their first choice since if many students have similar first choices they will fill up quickly and leave other students disappointed. However, due to the randomness of the draw, it creates no racially isolated schools, while both the current system and modified current system have around 20-30 racially isolated schools. Because the SFUSD matching system is heavily driven by diversity and racial desegregation, using the current system where students in CTIP1 neighborhoods (which are essentially a proxy for race as described by the SFUSD), is more effective in creating racially diverse schools than the modified system. All three matching systems have similar low fractions of students who receive their attendance area school and since this metric is not directly correlated to a student's preference, it is not a significant indicator of a good matching system. After seeing the comparisons against other systems, we get a better understanding that the SFUSD chose the

current system to give low-income students better opportunities and to increase racial diversity in schools, all while keeping parents' satisfaction high.

Another option would be to assign all students to their neighborhood school. Historically this strategy was attempted, but at the time housing segregation was particularly prevalent. So today, this option would decrease racial isolation, leaving only 12 schools racially isolated according to the data we collected, but would lead to many schools whose students are almost all white and asian or almost all black and latino. While racial isolation would decrease, only 17%¹¹ of kindergartners rank their attendance area school as their first choice, so students would be less happy with their assignment and higher income students could potentially leave the school system at higher rates or congregate more around high-performing schools. Additionally, these estimates include the 30% of students who attend private schools, a group that does not reflect the diversity of the city, so the system could lead to more than 12 isolated schools. Because so few parents want to send their children to their attendance area school, moving away from a choice system does not seem like a reasonable way to address school diversity.

Anecdotal Evidence

In addition to analyzing the SFUSD matching system from a high level simulation, we also scoured parenting blogs to see what real parents are saying about the system. While the majority of the blogs feature parents complaining about how many hours they spent researching schools and stressing out about the results¹², we also found many parents who post strategies they think will work and horror stories of their experiences.

One common strategy many parents talk about is leveraging the amount of weight the matching system puts on CTIP. The use of CTIP as a proxy for race is problematic because even though they are small, census tracts are not necessarily homogenous and because parents with the means to do so can choose to move to a CTIP-1 area during the school assignment process to gain an advantage in the system. There is anecdotal evidence that the label can make an address more attractive; a blogger on SF K Files, a popular blog about Kindergarten in SFUSD advises "If you would truly consider moving OUT of the city altogether if you can't get into a public school you can live with... Consider moving out of your non-CTIP address in an attendance area with a highly requested school and move to a CTIP area or an attendance area for a less requested school."¹³ This advice makes sense given that CTIP has priority over attendance area at every school, but also means that kindergartners in areas designated as CTIP-1 might not be a good proxy for race given that parents have some power to choose their CTIP designation.

While some parents may believe that they should rank an attendance area school higher up on their list, the advice on blogs seems to be moving away from the idea that mis-ranking schools could

¹¹ <http://www.sfusd.edu/en/assets/sfusd-staff/enroll/files/2015-16/4th-annual-report-april-8-2015.pdf>

¹² <https://ww2.kqed.org/news/2015/03/17/anxious-parents-try-to-game-system-in-san-francisco-school-lottery/>

¹³ <http://www.sfiles.com/2013/04/how-to-play-san-francisco-school.html>

lead to a better outcome.¹⁴ The current advice on parenting blogs is to list lots of schools, but does not advise listing every school, possibly because ranking is time-consuming, especially the writers on the blog who try to visit every school they rank, and possibly because unpopular schools are not likely to bring a student into a preference cycle. This indicates parents who are read blogs like SF K Files probably believe the only way to “game” the system is moving to a CTIP-1 address or ranking choices they consider worse than choosing a private school are move, and thus believe the system is mostly strategy-proof. Parents who spend less time focusing on school assignment may have less information. Evidence suggests poorer families rank fewer schools and receive worse assignments, so even though they have an advantage by living in a CTIP-1 area, lack of information or commute difficulties might prevent students in CTIP-1 areas from attending better schools.¹⁵

Low-income families, who are predominantly minorities may not have the resources to understand the school matching process or research and tour schools. Ranking more schools and submitting an application before the deadline both increase the odds a student is assigned to his or her top-choice school. Inequities in information have been a problem in the district, and the students who do not submit rankings are mostly minorities.¹⁶ The San Francisco Chronicle interview of students at Cleveland elementary school found that parents of students at the 83%-Latino school had not attended the district’s enrollment fair and had not toured schools, while most parents at nearby Clarendon, which is high-performing, racially diverse, wealthy, and the most requested school in the city, many parents had toured and ranked schools.¹⁷

Conclusion

To truly achieve desegregation and parent satisfaction, it is likely that the SFUSD should invest resources in areas outside of improving the matching system. Our simulation has shown that the matching system is the most effective it has ever been and the only way to really game the system is to rank more schools than you wish to send your child to, something that puts non-English speaking and low-income families at a disadvantage. Yet the current system doesn’t educate parents with low income and education levels to realize that getting into a good school can help their child succeed and it doesn’t help those parents find the money and resources necessary to send their child to a good school across the city. To get at the heart of the problem, we need to provide better support for parents entering the matching journey and also find ways to bring all schools to a common high standard. This will help ensure that children of all races and socioeconomic statuses will enjoy the benefits of the public school system for decades to come.

Appendix

Our video can be found here: <https://youtu.be/-uCcspAOhp8>

Our code repository can be found here: https://github.com/aacharya14/school_incentives

¹⁴ <http://www.sfkfiles.com/>

¹⁵ <http://sfpublicpress.org/news/2015-02/as-parents-get-more-choice-sf-schools-resegregate>

¹⁶ <http://sfpublicpress.org/news/2015-02/as-parents-get-more-choice-sf-schools-resegregate>

¹⁷ <http://www.sfchronicle.com/schools-desegregation/>

References

Address Lookup Tool. *SFUSD*. <http://enrollinschool.org/lookup/area.php>

Brooks, John. "Anxious Parents Try to Game System in San Francisco School Lottery. 17 March 2015.
<https://ww2.kqed.org/news/2015/03/17/anxious-parents-try-to-game-system-in-san-francisco-school-lottery/>

"Fact Sheet." *SFUSD*. 26 March 2012
<http://www.sfusd.edu/en/assets/sfusd-staff/enroll/files/2012-13/Assignment-Transfer%20process%20explained.pdf>

Featherstone, Clayton, et al. "School Choice in the San Francisco Unified School District (SFUSD)."
<http://isites.harvard.edu/fs/docs/icb.topic781759.files/10.2%20School%20choice%20in%20San%20Francisco.%20Clayton%20Featherstone.pdf>

Frey, William. "America Reaches its Demographic Tipping Point." 26 August 2011.
<https://www.brookings.edu/blog/up-front/2011/08/26/america-reaches-its-demographic-tipping-point/>

"The Inner Workings of the SFUSD Lottery." *The SF K Files*. 29 October 2012.
<http://www.sfkfiles.com/2012/10/the-inner-workings-of-sfusd-lottery.html>

Knight, Heather. "Is desegregation dead?" *San Francisco Chronicle*
<http://www.sfchronicle.com/schools-desegregation/>

Milholland, Nancy. "Children 5-17 Years Old by Race in San Francisco as a Dot Density Map" 8 November 2014.
<http://www.nmilholland.com/blog/projects/kidsdotmap/>

Roth, Al. "San Francisco School Choice Goes In-House." *Market Designer*. 23 September 2010.
<https://marketdesigner.blogspot.com/2010/09/san-francisco-school-choice-goes-in.html>

"School Site List and Summary" *SFUSD*. 7 November 2016.
[http://web.sfusd.edu/Services/research_public/rpa_student_enrollment/SFUSD%20School%20Site%20List%20and%20Summary-%20Student%20Enrollment%20\[Most%20Current\].pdf](http://web.sfusd.edu/Services/research_public/rpa_student_enrollment/SFUSD%20School%20Site%20List%20and%20Summary-%20Student%20Enrollment%20[Most%20Current].pdf)

Smith, Jeremy Adam. "As Parents Get More Choice, S.F. Schools Resegregate." *San Francisco Public Press*. 2 February 2015.

<http://sfpublicpress.org/news/2015-02/as-parents-get-more-choice-sf-schools-resegregate>

"SY 2016-2017 Total Requests by School/Grade/Program with Choice Ranking." *SFUSD Educational Placement Center*. 16 March 2016.

http://www.sfusd.edu/en/assets/sfusd-staff/enroll/files/2016-17/Report_Requests.pdf

US Census (2010). US Census Bureau. www.census.gov.

"Student Assignment: 4th Annual Report: 2014-15 School Year." *SFUSD*. 18 April 2015.

<http://www.sfusd.edu/en/assets/sfusd-staff/enroll/files/2015-16/4th-annual-report-april-8-2015.pdf>

"SFUSD Lottery - Lessons Learned" *The SF K Files* 6 April 2013.

<http://www.sfkfiles.com/2013/04/how-to-play-san-francisco-school.html>