1. Title: Students in flight

Reporters: MaryJo Webster, Beena Raghavendran, Anthony Lonetree, Beatrice Dupuy; Data visualizations by Jeff Hargarten and Billy Steve Clayton

1. Dates published: September 17, September 20, September 25
2. Topic and synopsis of story: The number of students enrolling in schools outside their home district is four times larger than it was in 2000, and black and Asian students are more likely than those from other racial groups to take advantage of Minnesota’s open enrollment and charter school options. This is becoming problematic for the Minneapolis and St. Paul school districts, two of the state’s largest, because both have reached the point that nearly one-third of kids who live within their boundaries don’t attend their schools. In the suburbs, some districts are filling classroom seats that would otherwise be empty due to an inflow of students, others are struggling to keep students and some have essentially become revolving doors.
3. How the story got started: In 2015, we wrote a series of data-driven stories showing how the Minneapolis and St. Paul public school districts had become more segregated in the past 15 years. Through our reporting, we kept hearing that open enrollment – and specifically charter schools – were pulling kids away from these two large districts and making it difficult for them to remain integrated. So we decided to put in a data request for open enrollment data, broken down by race, to see if we could prove that.
4. Was your work in any way based on another news organization’s previously published or aired report? No.
5. Major types of documents used and if FOI requests were needed. Did you have difficulties obtaining any electronic information? How did you resolve it? Did you use FOI for data under state or federal law? We relied heavily on databases for this. More about that below.
6. Major types of human sources used: parents, officials in the districts that we focused on, education experts.
7. Results:
8. Follow-up:
9. Advice to other journalists planning a similar project:

The heart of our reporting was measuring the share of students living in each district not choosing their home district. We chose to break it down by race, but you wouldn’t necessarily need to do that in order to get an interesting story. We did not measure students attending private schools or being home-schooled because we couldn’t get those figures broken down by race. So essentially we focused on the population of families who are choosing public schools and then measured whether they are choosing their home district, another district or a charter school. Important note: Minnesota’s legal definition of “open enrollment” doesn’t include children attending charter schools. So they gave us two datasets – one counting up the kids going to charter schools; and another counting the kids using open enrollment to attend another traditional school district elsewhere. We combined the two for our analysis.

The numbers you would need to get are: the total number of students enrolled in the district; the number of enrollees who come in via open enrollment and live somewhere else; the number of students who live in the district but either open enroll elsewhere or attend a charter school. For each district, we estimated the number of “residents” (regardless of where they go to school) by first subtracting the incoming open enrollees off the total enrollment number, then adding back in the students who open enroll elsewhere or attend charter schools. This became our denominator. Then we could take the number who “leave” and divide by the resident number to come up with a percentage. We followed the same steps for looking at individual racial groups.

FINDING PARENTS:

1. Difficulty, uniqueness of effort, or other special circumstances related to this subject:

This was the first time anyone in Minnesota had analyzed this data. Academics and our predecessors at the Star Tribune had looked at open enrollment before, but not broken down by race and it did not include the children going to charter schools (attending a charter school is not considered “open enrollment” under Minnesota’s legal definition). Our newspaper had reported on school choice numerous times in the past 30 years and our editors didn’t want to repeat what had been done in the past. Breaking it down by race proved to be among the most interesting of our findings because most people assume school choice is only used by well-to-do, white kids. Our data showed that was definitely not the case, especially in Minneapolis.

1. Length of time taken to report, write and edit the story: This was done piece-meal across the better part of a year. None of the reporters were freed up for their beat for any significant amount of time, so the work was done in between other responsibilities. As a result, it’s hard to come up with the number of hours involved. We obtained the first batch of data in December 2016 (just a few years’ worth of data) and once we saw that we had something notable, we requested data going back as far as possible. Then we had to wait a couple months for that additional data. MaryJo Webster completed the analysis in April 2017, but we decided to hold off on the reporting and aim for a September (back-to-school) publication date.
2. Data analysis questions….
   1. List any websites that were invaluable: The Minnesota Department of Education publishes most of its data to this data portal, <http://education.state.mn.us/MDE/Data/>, where we obtained test scores, high-level open enrollment data and overall enrollment data going back in time.
   2. Did you obtain or build any electronic databases?

The story relied heavily on databases that we obtained under the Minnesota Data Practices Act from the Minnesota Department of Education. The agency was immensely accommodating, generating these data files without charging us and in a prompt manner. We were able to work directly with their data analyst and get our questions answered without hassle.

The data we used included some that MDE publishes on its data portal (such as overall enrollment figures for each school year in each district; and standardized test results), but the primary data source was data we requested showing the number of students who don’t attend their home district, broken down by race, what district they live in and the district or charter district they attended.

The data is summarized because Minnesota law restricts release of student-level data (even if it doesn’t contain any identifiers). MDE also redacted any data where less than 3 students of a racial group were leaving one district to go to another. Later in the reporting process, we went back to MDE and asked for the same data, but instead of breaking down by race, to break it down by whether the students are on the free/reduced lunch program or not (a proxy for poverty).

* 1. Any difficulties? Initially it was a little daunting to try to figure out what to do with this data. The initial structure of the data, with multiple records for each district, was not the right structure for the analysis we wanted to do. It took some thought to figure out what re-organizing needed to be done. The main thing was figuring out the denominator in order to be able to say what percentage of students were leaving each district. We looked at American Community Survey tables on school enrollment to see if those numbers were complete enough to give us an accurate “school-age children” count, regardless of where they attend school.

However, it quickly became apparent that the margins of error were too high, especially when we were looking at racial groups. The Department of Education had totals of children being homeschooled or going to private school, so we could get a rough idea of how many students we were not counting, however the state didn’t have race information on those students. We settled on using only public school children as the denominator – figures we could calculate ourselves using the various datasets we had requested from the state education department.

We were also a little hampered by the fact that this was all summary data at the district level – we aren’t legally able to get student-level data in Minnesota – and that the state had redacted some of the numbers that affected the racial breakdowns. We couldn’t see which schools these kids were not attending. And we knew that our results were conservative, so we spent a good deal of time cross-checking against other data to ensure the redactions hadn’t caused massive gaps.

* 1. (e. and f.) Analysis: Yes, there was significant analysis done by MaryJo Webster. She primarily used MySQL and toward the end relied heavily on Microsoft Excel to answer small questions and to share the data with reporters. The analysis involved calculating the number of “residents” in each district (regardless of where they go to school), the number of students who enroll in the district but live elsewhere and the number of students who live in the district but go to a charter school or open enroll in another district. We did this at the district level for all students, then also ran the same calculations for each racial/ethnic group (white, black, Asian, American Indian and Hispanic).

These calculations required combining the open enrollment data we had obtained through our request, and the regular enrollment data that MDE publishes on its data portal. We used data from the 1999-00 school year through the 2016-17 school year, generating these calculations for each year.

Once these calculations were in place, then it was possible to look at things like: change over time in the share of students leaving the district; which racial groups had the highest rate of leaving districts; the ratio of students leaving to the students coming in (this allowed us to say that Minnetonka gained 17 students for each 1 it lost and Minneapolis lost 22 students for each 1 it gained). We looked at each district, then also collapsed the figures up to bigger groups (all the district in Minneapolis and St Paul; all the district in the suburbs; all the districts in the rest of the state) to see broader geographic patterns.

We published two data visualizations. The first, highlighting our key findings, was used as an “introduction” to the series and it was published online a few days before the first story published. The second was an interactive, created in D3, that allowed readers to see the open enrollment trends in their school district. This allowed us to expand the breadth of our piece, since it included hundreds of school districts that weren’t featured in our stories.