

Tracking is Making a Bad Situation Worse: Inefficiency and Inequality in Ann Arbor

1. INTRODUCTION:

Despite Ann Arbor's reputation as one of the most educated and liveable cities in America (Manzullo 2107), inequality in the area is among the worst in the country. Washtenaw County is the eighth most economically segregated metropolitan area in the country and ranks 80 out of 83 for income inequality in Michigan (Kay 2018). Furthermore, in third grade there is a shocking 35 point gap on reading test scores between African-American and white students (Kay 2018). If Ann Arbor is such an educated city with some of the best ranked public schools in the state (US News 2018), why are so many disadvantaged and minority students low achievers?

A public education theoretically provides equal opportunity for all students, yet many governments justify systematically decreasing this opportunity for disadvantaged and minority kids increasing it for the already advantaged. The practice I'm referring to is known as 'educational tracking' and Ann Arbor Public Schools employs it in many ways. In theory, tracking makes sense. Pro-tracking policists argue that assigning students to classes based on ability should allow teachers to cater their lessons to the ability of their students thus increasing overall academic achievement of both high and low ability students (LeTendre, Hofer, and Shimizu 2003). Also, grouping similar peers together should increase overall achievement without increasing inequality. Although some positive peer effects have been confirmed through studies (LeTendre, Hofer, and Shimizu 2003), an overwhelming amount of studies done about tracking show that its effect is not positive. Quite the contrary, studies show that the result of tracking is typically that students in lower tracks perform worse than those without tracking

(Burriss and Welner 2005). Additionally, minority and low socioeconomic students are overrepresented in lower tracks (Thomas 2018). This is a recipe for increasing inequality.

Of course, there are many kinds of tracking. While countries like Germany, Austria, Hungary and the Slovak Republic strictly track students into different schools, other countries like Canada, Japan, Norway, and the United States *typically* keep all primary and secondary school students in the same building. In Ann Arbor, students are tracked in a variety of ways. In elementary school students are tracked by (perceived) ability into reading groups. Seventh grade math teachers help track students into either Algebra I (high-track) or Algebra A (low-track); this track determines the path of math courses they will take through the end of secondary school. In high school, students are tracked into accelerated (AC) English, science, language, and math classes. The existence of A.P. (Pioneer and Skyline) and IB (Huron) courses is perhaps the most significant form of tracking in Ann Arbor Public Schools. A study done in 2003 shows that A.P. tests are a strong indicator of bachelor's degree attainment (Camara 2003). Maybe these classes would be beneficial if the students that take them accurately represented the population of public schools, but this is not the case. AP Cohort Data Reports show that in 2011, African-Americans made up 14.6 percent of American high school graduates but represented less than 4 percent of the A.P. students with a score of 3 or better on an exam (Newman 2011). It also shows that only about one in ten students with low-income families will take A.P. courses, while one in four students from middle to high-income homes do (Thomas 2018). These statistics suggest that high track classes might not actually represent kids with high ability, instead they typically represent kids with more privilege.

Detracking is the increasingly popular act of either delaying tracking to later in a student's education or getting rid of it all together, keeping all students in the same

classroom(Burris and Welner 2005). In the past few decades, a growing collection of studies from around the world has been providing evidence that detracking leads to a significant decrease in educational inequality gaps in years of attainment (Burris and Welner 2005). More specifically, the effect of detracking usually closes this gap by increasing the average years of education attained by “low-achievers”(Burris and Welner 2005). Since the studies focus on such a wide variety of tracking methods from across the globe, I acknowledge that some statistics are not directly applicable to AAPS policy. However, I hypothesize that the consistency in results that show decrease in educational attainment inequality would remain consistent in Ann Arbor. With Ann Arbor Public School’s current method of tracking, there is a disproportionate number of minorities children from low-income families represented in low-tracks. I believe that the result is increased inequality. I argue that delaying or abolishing tracking methods in AAPS would allow these low-achievers a better chance at true equal opportunity. Given Ann Arbor’s current unacceptable state of inequality, AAPS should look toward detracking to begin to alleviate the gap in attainment.

2. LITERATURE REVIEW:

Literature about tracking has two major themes. The first asks how tracking affect overall average educational attainment. The second asks how does it affect the inequality in educational attainment. Statistics usually show no effect or a slightly positive effect on overall average educational attainment. They consistently show significant positive effects on inequality.

There are countless ways to employ tracking in public schools. Moreover, there are countless countries with wildly varying cultures and societies that use tracking in their school systems. As a result, literature on tracking is diverse. This is both a good and a bad thing. It’s

good because the more countries and methods included, the more complete the conclusions on tracking. The difficulty is that it is hard to compare the effects of tracking between countries since there are so many (sometimes immeasurable) cultural and social variables that affect inequality and educational outcomes. I find studies performed within one country to be the most trustworthy because it means there are minimal social and cultural differences between schools and students (or at least the differences are uniform). Perhaps the most referenced country in the study of tracking is Germany which has an extremely strong tracking structure. This section provides a summary of literature that aims to show how tracking directly affects education and inequality in Germany as well as literature that explains its effects around the globe.

2.1 Effects of Tracking in Germany

The most compelling research on the effects of tracking on educational attainment and inequality was done in Germany by Simon Lange and Marten von Werder (2016). Historically, all children in Germany are tracked into one of three tracks after fifth grade (around age 10). Teachers sort students based on (perceived) ability and parents have the option to override teachers' recommendations but typically only do when the recommendation is different from their own educational track (Lange and von Werder 2016). They will either follow the academic track called gymnasium (high ability), the higher level vocational track (medium ability), or the lower level vocational track (low ability). However, in Lower Saxony (one of Germany's sixteen states), a unique educational reform occurred in the 1970s which delayed this tracking by two years (Lange and von Werder 2016). Research on Lower Saxony uses data collected before and after the reform using Lower Saxony as the treatment group and several German states whose policy remained consistent with pre-reform Lower Saxony as the control (Lange and von Werder 2016). This research aims to answer the question: how does this two year delay in tracking

affect overall average years of educational attainment and educational inequality between kids whose parents are highly educated and those whose parents have a low level of education?

Through use of difference-in-difference estimation, results show that the reform had no effect on overall average educational attainment but significantly decreased the inequality (Lange and von Werder 2016). Specifically, the results show that the gap in years of education decreased by 1–1.5 years and average educational attainment saw no significant change after the reform (Lange and von Werder 2016). In this case, delaying tracking from after grade five to after grade seven significantly decreases inequality of educational attainment with respect to parents' educational attainment. Krause and Schüller (2014) also conclude that family background is a strong indicator of track recommendation “with ethnic inequalities actually emerging from socio-economic inequalities” in the German education system .

The situation in Lower Saxony is reliable in identifying effects of tracking because it examines a unique situation in which a group was isolated by state policy and the age that tracking began was delayed with practically all other aspects of education held constant. This perfectly sets up the difference-in-difference estimation. Lange and von Werder explain that the states used as the control groups were uniform with each other and Lower Saxony *other* than aspects changed by the detracking reform. With that said, another change made by the reform was that it made it easier for parents to override recommendations made by the school system and send them to the track of their choice. Most commonly, a parent's choice in their child's track will reflect their own track. This aspect of the reform would work to increase inequality in attainment unlike the tracking delay which decreases the gap. The authors emphasize that all results should reflect a combination of both aspects of the reform which does not diminish the accuracy of the findings since they work in opposite ways. The way recommendations are made

by (dominantly male) teachers in Germany tend to be discriminatory against minorities, girls, and low-income students (Lange and von Werder 2016) In their difference-in-difference estimation they only used variables that remain constant over time. They controlled for variables that might affect educational attainment such as date of birth, migrant status, first language, city of education etc. and performed one estimation for each gender. Given that this is an experiment held on a national stage over many years, they have to account for many policy changes and regional differences. They address dozens of the most significant with results from other studies and projecting those findings effects into to their study.

Results in this literature suggest that the delay in tracking decreased the educational inequality is not simply because tracking is “bad”, but rather that the age of ten is too early. This is ultimately *why* this paper is important and *why* these studies are done; in theory educational tracking would be efficient yet when it is implemented, its effects are often overall insignificant or even negative. Economists and policymakers are in a struggle to find out whether tracking is in fact “good”, and if so how its positive effects can be maximized and its negative effects (such as inequality) be minimized. Statistics Lower Saxony suggest tracking too early is inefficient and harmful. First, ability is much easier to recognize as a child gets older. Second, it suggests that the peer effect of grouping low-achieving students early on can resonate their low-achievement even though they may have high potential.

2.2 Cross-National Studies

Countries such as Austria and Germany track kids by separating them into different schools by (perceived) ability at a young age. Conversely, countries such as the United States, Japan, and Sweden keep all children in the same schools but might track them in classes or

groups for differing ability levels later on. Hanushek and Wössman (2015) compare educational outcome due to tracking between eighteen countries rather than within one country. Their research compares educational outcomes in countries that track students into separate schools by age ten with countries who keep students in comprehensive schools. They aim to find how early tracking (by age 10) affects educational inequality and mean performance. It is important to note that Hanushek and Wössman measured both mean achievement and inequality using international test scores (the IEA) rather than years of attainment. Using difference-in-difference estimation to compare countries with and without early tracking, they found that the presence of early tracking is a strong indicator of increased inequality of achievement. From the 18 countries used in the data, the four countries with the highest increase in relative inequality were all countries with early tracking (Hanushek and Wössman 2015). Specifically, “Germany with an increase of 0.71, Greece 0.30, Czech Republic 0.25, and Italy 0.22” (Hanushek and Wössman 2015). The same study found no significant effect on mean educational performance.

Complexities in studying cross-national patterns arise: “macro issues of institutional structure are extraordinarily difficult to evaluate within individual countries largely because the variations in structure that exist within countries are almost certainly related to the characteristics of the families and schools choosing to follow an anomalous pattern” (Hanushek and Wössman 2015). I acknowledge that using cross-national studies is not the most accurate way to analyze the American school system. However, I believe that showing that early tracking consistently increases educational inequality in many countries (as in Hanushek and Wössman 2015) makes it reliable to apply research from Germany and other countries to the American school system.

2.3 Detracking in America

Luckily, there are already documented cases of school districts detracking in the United States. Argys, Brewer, and Rees (1996) noted that many schools across the country are ending the practice of tracking in response to research showing its negative effects on minorities and children from low-income families. Their research focuses on schools that track 10th, 11th, and 12th grade students into high-achievement classes (like A.P. or “Honors”) and schools that no longer practice that kind of tracking. They designed a regression of test score data from the National Center for Education Statistics controlling student, classroom, and teacher characteristics. Using this regression Argys, Brewer and Rees compare between schools that detracked and schools that did not change their tracking methods. This research shows 3 important trends in data. First, socioeconomic status is a strong indicator of track placement with “students from better backgrounds more likely to be placed in upper-level tracks” (Argys, Brewer, and Rees 1996). Second, detracking has a significant positive effect on achievement for students who were formerly in lower tracks (Argys, Brewer, and Rees 1996). Third, detracking American high school student has a negative effect on high-track students. Detracking proponents tend to argue that detracking in America would improve achievement for low-track students with no significant effect on other students (Argys, Brewer, and Rees 1996). This evidence suggests that this assumption is “overly optimistic”; although detracking does benefit low track students and decreases inequality in achievement, it seems to be at the expense of high track students (Argys, Brewer, and Rees 1996). Again, it is important to note that this research aimed to find the effect of tracking on student achievement through test scores rather than students’ educational attainment. Inequality is likewise defined as the gap in achievement rather than gap in attainment.

Research done in New York state which focuses on attainment rather than achievement provides more optimistic results. When high schools in a suburban New York City district were detracked, all students saw benefits. Research was done using the attainment of a high school diploma as the educational outcome of interest. Using binary logistic regression analyses before and after the detracking of the district, Burriss and Welner (2005) show these gains in attainment. “When all students were taught the high-track curriculum, achievement rose for all groups of students--majority, minority, special education, low-SES, and high-SES” (Burriss and Welner 2005).

3. Synthesis of Literature

Literature on the effects of tracking is consistent in one important way: detracking results in decrease in inequality. In all research presented in this paper (from Germany, international analysis, and within the US), detracking leads to a smaller gap in both attainment inequality and achievement inequality. Furthermore, each case presented in the literature showed that detracking tends to have a significantly positive effect on educational achievement and attainment for students from lower tracks. Most discrepancies in results are about how detracking affects students from high-tracks. For example, regarding high schoolers in the US, Argys, Brewer, and Rees (1996) suggest that detracking has a significantly negative effect on high track students while Burriss and Welner (2005) show that detracking positively affected high track students in New York. However there is one important distinction to make between these results. Argys, Brewer, and Rees used achievement (national test scores) as their outcome while Burriss and Welner used educational attainment (attainment of a high school diploma) as their outcome. In the literature presented in this paper, these results are consistent. Although

detracking tends to have negative effects on high-track students' achievement, it tends to have either no significant effect or a positive effect on their average educational attainment. In other words, the only negative effect of detracking in the evidence presented is the harm to high-track students' achievement. So is attainment or achievement more important? I argue that decreasing the attainment gap is more important because years of education is a significant indicator of lifetime income (Burris and Welner 2005) whereas test scores like the National Center for Education Statistics tests are simply measurements for educational achievement.

Research also provides evidence that privilege is a strong indicator of track where being a non-minority, high-SES student indicates a higher track (Krause and Schüller 2014). Fundamentally, high tracks are supposed to represent high ability rather than high privilege. This begs the question: how much do we want to help high track students compared to low track students? I argue that since minority and low-SES students are overrepresented in lower tracks (Thomas 2018), improving attainment and achievement in lower tracks should be prioritized over students higher tracks in order to decrease inequality.

Due to the diversity in literature, applying results to Ann Arbor Public Schools requires trust that observed decreases in inequality from detracking are truly caused by detracking. I argue that the consistency in results that show decreased inequality in attainment (and achievement) would carry over to Ann Arbor Public Schools.

4. Policy Recommendations

It is clear from research that detracking schools consistently results in decreased inequality. Given Ann Arbor's current state of educational inequality, I recommend that AAPS implement detracking.

International studies such as the German and cross-country studies show strong evidence that tracking before age 10 increases inequality. This shows that tracking elementary age students is harmful (Simon Lange and Marten von Werder 2016). Research also shows that teachers who sort students by ability at the young age of ten (or younger) tend to be inaccurate in their judgement (Krause and Schüller 2014). Studies on American schools presented in this paper provide strong evidence that detracking high school age students leads to decreased inequality (Argys, Brewer, and Rees 1996).

Given that AAPS tracks students in almost every stage of their education, my recommendations to decrease inequality through the detracking of schools are as follows:

- Detrack elementary schools by abolishing reading groups and other learning groups in which students are sorted by ability.
- Stop sorting eighth grade students into either high track (Algebra I) or low track (Algebra A) math classes.
- In high school, rather than separating students by ability into accelerated (AC) english, science, language, and math classes, keep them in comprehensive classes, and get rid of AC classes.
- Detracking students in A.P. and I.B. classes would likely be the policy change with the most resistance as taking these classes (and their tests) are strong indicators of graduating from college, and students and parents know this (Camara 2003). The pragmatic approach to detracking A.P. and I.B. courses would be to keep them and allow and encourage students of all achievement levels to take them. I believe the approach that would lead to the largest decrease in inequality would be to rid of A.P. and I.B. courses.

Lastly, I recommend that these policy changes are implemented in the form of an experiment. In other words, I recommend that only certain elementary, middle, and high schools are detracked (a test group) while the others continue tracking (control group). This will allow AAPS to observe data of detracking schools and compare it to data from schools in their own district in order to show that detracking does indeed decrease educational inequality.

WORKS CITED

- Argys, Laura M., Daniel I. Rees, and Dominic J. Brewer. "Detracking America's Schools: Equity at Zero Cost?" *Journal of Policy Analysis and Management* 15, no. 4 (1996): 623-45.
- Burris, Carol Corbett, and Kevin G. Welner. "Closing the Achievement Gap by Detracking." *The Phi Delta Kappan* 86, no. 8 (2005): 594-98.
- Camara, Wayne J.. "College Persistence, Graduation, and Remediation." *The College Board Office of Research and Development*. (2003): 1-5.
- Hanushek, Erik A. and Ludger Wössman. "Does Educational Tracking Affect Performance and Inequality? Differences in Differences Evidence Across Countries" *National Bureau of Economic Research* 61, (2015): C63-C76.
- Kay, Grace. "New Initiative Aims to Tackle Inequality in Ann Arbor, Washtenaw County." *The Michigan Daily*, The Michigan Daily, 1 Feb. 2018.
- Krause-Pilatus, Annabelle & Schüller, Simone, 2014. "Evidence and Persistence of Education Inequality in an Early-Tracking System: The German Case," IZA Discussion Papers 8545, Institute for the Study of Labor (IZA).

- Lange, Simon and Marten von Werder. "Tracking and the Intergenerational Transmission of Education: Evidence from a Natural Experiment." *Economics of Education Review* 61, (12, 2017): 59-78.
- LeTendre, Gerald K., Barbara K. Hofer, and Hidetada Shimizu. "What Is Tracking? Cultural Expectations in the United States, Germany, and Japan." *American Educational Research Journal* 40, no. 1 (2003): 43-89. <http://www.jstor.org/stable/3699425>.
- Logsdon, Kullen. "Ann Arbor, Washtenaw County Launch 'One Community' Equity Initiative." *MLive.com*, MLive.com, 25 Jan. 2018, www.mlive.com/news/ann-arbor/index.ssf/2018/01/washtenaw_county_leaders_launc.html.
- Manzullo, Brian. "Ann Arbor the Most Educated City in America, Study Finds." *Detroit Free Press*, Detroit Free Press, 25 July 2017.
- Thomas, Jacqueline Rabe. "Access to AP Courses Often Elusive for Low-Income Students." *The CT Mirror*, The CT Mirror, 15 Dec. 2018, ctmirror.org/2018/05/14/advanced-placement-debate-open-closed-gate/.
- Tugend, Alina. "Who Benefits From the Expansion of A.P. Classes?" *The New York Times*, The New York Times, 7 Sept. 2017,
- "The Best High Schools in Michigan, Ranked." *U.S. News & World Report*, U.S. News & World Report, www.usnews.com/education/best-high-schools/michigan.
- "Understand the Class of 2017 Results | AP Results: Class of 2017." *AP Program Results: Class of 2016*, AP, 26 Oct. 2018, reports.collegeboard.org/ap-program-results.