An International Comparison of Achievement Inequality in Within- and Between-School Tracking Systems

Anna K. Chmielewski

Michigan State University

Updated Corresponding Address:

Ontario Institute for Studies in Education

University of Toronto

252 Bloor Street West

Toronto, ON M5S 1V6

Email: ak.chmielewski@utoronto.ca

Pre-print version—March 27, 2014

Post-print version available at: http://www.jstor.org/stable/10.1086/675529

Full citation:

Chmielewski, Anna K. 2014. "An International Comparison of Achievement Inequality in Within- and Between-School Tracking Systems." *American Journal of Education* 120(3): 293-324.

An International Comparison of Achievement Inequality in Within- and Between-School Tracking Systems

Anna K. Chmielewski

Michigan State University

Abstract

Secondary school tracking is organized in some countries on a course-by-course basis within schools and in other countries as explicit academic and vocational streaming, often in separate school buildings. This article is the first to compare these two forms of tracking, using student-level tracking data across the United States and 19 other developed countries. Results indicate that course-by-course tracking is less segregated by socioeconomic status (SES) than is academic/vocational streaming. Yet both forms of tracking have comparable achievement gaps between tracks. Among students in the same track, SES disparities in achievement are larger in course-by-course tracking than in academic/vocational streaming.

Keywords

tracking, ability grouping, course-taking, international comparison, achievement

Tracking, the practice of grouping students by achievement levels into differentiated curricula, is commonplace in the majority of American high schools (Loveless 2013). Although theoretically, tracking allows schools to tailor the curriculum to students' needs, a large literature has found that tracking contributes to educational inequality (e.g., Gamoran and Mare 1989; Lucas 1999; Oakes 1985). Some research even identifies tracking as a contributor to low and unequal US scores on international assessments (Schmidt et al. 1999; Schmidt et al. 2001; Westbury 1992).

But something called "tracking" also exists in many other countries. In some of these countries (including other Anglo countries and some Nordic countries), tracking is similar to that in the United States and consists of offering courses at varying levels of difficulty in one or more subjects within a school. Throughout this article, this practice is referred to as *course-by-course tracking*. But in many other countries (including continental European countries and some Asian countries), tracking appears much more extreme. It involves allocating students into overarching programs—often located in separate school buildings—with curricula that prepare students either for university or for trades. This article calls such practices *academic and vocational streaming.* Like US research, international research has also examined the role of tracking in shaping inequality in achievement (e.g., Brunello and Checchi 2007; Marks 2005; Schütz, Ursprung and Wößmann 2008). However, most of this research defines tracking as academic/vocational streaming and, surprisingly, often categorizes the United States and other countries practicing course-by-course tracking as countries with "no tracking" (e.g., Hanushek and Wößmann 2006).

Despite the very different appearances of course-by-course tracking and academic/vocational streaming, many researchers draw a parallel between the two and call for more comparative study of these practices (e.g., Buchmann and Park 2009; Van de Werfhorst and Mijs 2010). Both types of tracking have the same purpose of differentiating curricula by students' achievement levels, and both types often have the consequence of segregating students by socioeconomic status (SES) and ethnicity, which can exacerbate achievement disparities between these groups (Gamoran and Mare 1989; Maaz et al. 2008; Oakes 1985; Schnabel et al. 2002). Exactly how similar these achievement disparities are across the two types of tracking remains unknown because very little empirical research has compared achievement and segregation in course-by-course tracking and academic/vocational streaming.

The neglect of course-by-course tracking in international research is increasingly problematic since this form appears to be growing more common worldwide, for example, in the United States (Lucas 1999), the United Kingdom (Feinstein and Symons 1999), and France (Duru-Bellat 1996), while academic/vocational streaming declines, for example, in Sweden (Heidenheimer 1974), the United Kingdom (Manning and Pischke 2006), and recently Poland (OECD 2004). Yet across the developed world in 2003, about 45% of 15-year-old students still identified themselves as belonging to an academic or vocational stream (author's own calculation from PISA [Program for International Student Assessment] 2003 data). Thus, at present, both types of tracking are common ways that secondary school students experience curricular differentiation, making comparisons between them an important area for research. This article is the first fine-grained comparison of achievement outcomes using student-level tracking data for both types of tracking across a wide range of countries. It uses PISA data for 20 developed countries.

Differences between the Two Systems

The most visible differences between the two types of tracking are organizational and institutional. First, course-by-course tracking always occurs within schools, while in most countries, academic and vocational streams are located in physically separate buildings (although in some countries, such as Belgium, streams are located side by side in the same building) (Gamoran 2010). Second, in course-by-course tracking, it is courses that are tracked (i.e., students may take different levels of courses in different subjects), whereas in academic/vocational streaming it is students who are tracked (i.e., streams are overarching programs that determine all coursework). Third, in course-by-course tracking, theoretically all students are eligible for university, regardless of the track levels of their courses. In academic/vocational streaming, stream placement often dictates university eligibility. The important postsecondary consequences of students' track placement in academic/vocational streaming make it necessary for tracks to have consistent definitions or "charters" that transcend individual schools and extend across the entire country (Meyer 1977).

Tracking systems can also be compared in terms of rigidity, including the level of mobility between tracks, the age at which tracking begins, and the degree to which track placement is determined by more "objective" measures, such as achievement and course grades, versus by parent preferences (Broaded 1997). There is some variation across academic/vocational streaming countries in the rigidity of tracking (e.g., tracking begins as early as age 10 in Germany and Austria but as late as age 15 in France; parents have a strong influence on track placement in Germany but entrance exams determine placement in Japan) (LeTendre, Hofer and Shimizu 2003; Maaz et al. 2008). However, in general, research characterizes course-by-course tracking as less rigid than academic/vocational streaming: yearly course selections allow higher mobility between tracks, tracking does not begin until around the transition from lower to upper secondary school, and track placement is based on prior achievement and student choice (LeTendre, Hofer and Shimizu 2003). In practice, many of these assumptions about course-by-course tracking may be violated—which is addressed in the next section below—but this type of tracking is not explicitly defined in the rigid way that academic/vocational streaming is.

Similarities between the Two Systems

A great deal of research on US tracking calls into question the assumption that course-by-course tracking is a more fluid system than academic/vocational streaming. Although there is theoretically higher mobility between tracks in course-by-course tracking, course sequences that build upon prerequisite content can make transfer difficult, especially in math (Frank et al. 2008; Stevenson, Schiller and Schneider 1994). Additionally, students may not be tracked separately for each school subject, as the levels of students' courses tend to be associated across subjects (Lucas and Berends 2002). The theoretically weak link between track level and postsecondary opportunities may also not be a reality, as advanced math coursework is a strong predictor of college attendance (Schneider, Swanson and Riegle-Crumb 1998) and the strongest predictor of college completion (Adelman 1999). Perhaps most important, the supposed greater role of student choice in track placement in course-by-course tracking does not appear to produce less bias in assignment, as many scholars have described social class segregation between tracks as a systematic method of class reproduction (Bowles and Gintis 1976; Oakes 1985) or a covert means of within-school "second generation" racial segregation in nominally desegregated schools (Lucas and Berends 2002; Oakes 2008; Wells and Serna 1996). In sum, course-by-course tracking may not be as explicit as academic/vocational streaming, but it can still produce a great deal of inequality through informal processes.

Current Policy Landscape

Historically, most developed countries practiced academic/vocational streaming, either between or within schools. During the 1960s and '70s, many countries reformed their systems by delaying the point of selection to a later age or doing away with academic/vocational streams and creating comprehensive schools, including the United States (Lucas 1999), the United Kingdom (Manning and Pischke 2006), Sweden (Heidenheimer 1974), Norway, and Finland (Ariga et al. 2005). Others reformed more recently in the 1980s and '90s, including France, Spain (Ariga et al. 2005), and Poland (OECD 2004). In some countries, both academic/vocational streaming and comprehensive schools coexist, but generally one is predominant. For example, the United Kingdom still has some academic-track grammar schools, but 90% of students attend comprehensive schools (Manning and Pischke 2006). Germany has some comprehensive schools (*Gesamtschulen*), but they enroll only 10% of students (author's own calculation from PISA 2003 data).

In many countries, course-by-course tracking has increased with the decline of academic/vocational streaming. For example, when US schools dismantled overarching academic and vocational streams, they were replaced with course-by-course tracking (Lucas 1999). In the 1970s when England was midway through its transition from between-school streaming to comprehensive schools, comprehensive schools had higher rates of within-school math and English tracking than did grammar and secondary modern schools (Feinstein and Symons 1999). In France, after the age of selection was pushed back to age 15, tracking increased in lower secondary schools, largely as a result of parental pressures (Duru-Bellat 1996). In German comprehensive schools, within-school course tracking is mandated by law (Trautwein, Köller and Kämmerer 2002). Currently, on the basis of principal reports, the countries with the highest rates of course-by-course tracking are Anglo countries (the United States, Australia, Canada, the United Kingdom, and New Zealand), the countries with moderate rates are Nordic and other comprehensive systems (Iceland, Norway, Sweden, Spain, and Poland), and the countries with low rates are Denmark and Finland, as well countries practicing primarily academic/vocational streaming (Austria, Germany, Greece, Japan, etc.) (author's own calculation from PISA 2003 data). The tendency of course-by-course tracking to arise to replace academic/vocational streaming raises the question of whether countries are replacing formally stratified systems with informally stratified ones that create similar levels of inequality.

Effects of the Two Types of Tracking on SES Disparities in Achievement

The central concern of this article is whether or not the two types of tracking have comparable effects on SES disparities in achievement. The literature presented above suggests two alternatives: On the one hand, academic/vocational streaming may produce greater inequality of outcomes than course-by-course tracking because its institutional structure is more rigid. On the other hand, course-by-course tracking may produce a similar level of inequality through informal processes. As this article is interested in student SES and achievement, the two forms of tracking are compared on three major dimensions: first, the level of SES segregation between tracks; second, the size of achievement gaps between tracks; and third, the strength of the association between SES and achievement.

As the review of literature below will demonstrate, prior international tracking literature has examined these three areas but has largely ignored within-school course-by-course tracking. Some authors categorize countries such as the United States as countries with "no tracking," along with systems such as Finland's, which has very little differentiation at all. This may in fact be incorrect, as US literature has identified tracking as a major source of inequality. Other authors do include course-by-course tracking countries but measure tracking in ways that that cannot capture within-school tracking, such as calculating achievement gaps between schools rather than between tracks. Since course-by-course tracking and academic/vocational streaming occur in different locations, they can only be directly compared using student-level track placement data. This article makes two improvements upon prior literature: First, rather than simply examining academic/vocational streaming, it also incorporates course-by-course tracking practices. Second, it uses student-level data on both types of tracking across a large number of countries for the first time. If this study finds that the two types of tracking do indeed have comparable levels of SES segregation between tracks and achievement gaps between tracks, this implies that future comparative research on tracking should no longer ignore the growing practice of course-by-course tracking.

Which Type of Tracking is More Socioeconomically Segregated?

Prior research suggests that tracks are more segregated in systems of academic/vocational streaming than in course-by-course tracking. The primary explanation for this is that the age of selection into tracks is generally younger in academic/vocational streaming, and parental background exerts a stronger influence over educational transitions at younger ages—this is known as the “life course hypothesis” (Blossfeld and Shavit 1993; Mare 1980). A number of comparative studies provide some support for this idea but no direct evidence since none uses student-level track placement data for both academic/vocational streaming and course-by-course tracking. One group of studies measures segregation between schools instead of between tracks, finding greater SES segregation to be associated with earlier or more rigid tracking (Demeuse and Baye 2008; Jenkins, Micklewright and Schnepf 2006; Willms 2010) (with the exception of

Gorard and Smith 2004). However, studies of between-school segregation cannot separate the effects of tracking from other causes of sorting between schools, such as residential segregation, nor can they capture within-school segregation between classrooms.

Another group of studies that uses student-level tracking data but only for academic/vocational streaming countries shows that SES is highly predictive of track placement (Buchmann and Park 2009; Horn et al. 2006; Maaz et al. 2008; Mateju and Straková 2005). But again none of these studies calculates SES segregation between tracks in course-by-course tracking countries. The one exception is a study comparing the United States and Germany, which shows that after controlling for achievement, SES is more predictive of German school type placement than of US math or English track placement (Schnabel et al. 2002). But no study has compared a larger set of course-by-course tracking and academic/vocational streaming countries using student-level tracking data.

The lack of comparative evidence on course-by-course tracking leaves some doubt on the prediction that SES segregation is lower in these systems. This is particularly the case because research in individual course-by-course tracking countries shows that tracks are quite segregated by SES, both in the United States (Gamoran and Mare 1989; Kelly 2004; Lee and Bryk 1988; Lucas 2001) and in Australia (Lamb and Fullarton 2002). Although in course-by-course tracking, SES probably exerts a smaller direct effect through parental involvement in track placement than in academic/vocational streaming (Baker and Stevenson 1986; Kelly 2004), the indirect effect via prior SES differences in achievement could still be substantial compared to systems with more standardized and equitable primary level schooling (Broaded 1997).

Which Type of Tracking Has Larger Achievement Gaps between Tracks?

The evidence on achievement gaps between tracks is even less definitive than for segregation. Much recent comparative research argues that achievement should diverge more between students in academic and vocational streams than in course-taking tracks because they begin earlier and are kept more rigidly separate. Using as their outcome the total variability of achievement in each country (e.g., standard deviation), many of these authors find higher variability of scores in countries with earlier or more rigid academic/vocational streaming (Dupriez, Dumay and Vause 2008; Hanushek and Wößmann 2006; Montt 2011). However, without student-level tracking data, one cannot know the size of achievement gaps between tracks because high-achieving students are not necessarily in the high track, or low-achieving students in the low track. One study addresses this issue by using student-level tracking data for academic/vocational streaming countries, and in contrast to the results above, it does not find evidence of large track effects (Marks 2006). Yet this study does not use student-level tracking data for course-by-course tracking countries. Another study includes some course-by-course tracking countries by examining between-classroom variance in achievement, and finds that achievement diverges more in countries with homogeneous classrooms (Huang 2009). But this study cannot disentangle classroom and school segregation because it uses data from the Trends in International Mathematics and Science Study (TIMSS), which samples only one classroom per school in most countries.

Single country studies of course-by-course tracking also find that achievement differs widely between tracks. US tracking research shows that this is at least because of selection (Figlio and Page 2002; Slavin 1990) but probably also because of diverging achievement between tracks (Gamoran and Mare 1989; Leow et al. 2004; Schneider, Swanson and Riegle-Crumb 1998). In other course-by-course tracking countries, similar or even larger achievement differences between tracks have been found, for example in England (Kerckhoff 1986) and Australia (Lamb and Fullarton 2002; Schmidt et al. 1999).

In both course-by-course tracking and academic/vocational streaming, the evidence from single-country studies suggests that achievement gaps between tracks are driven primarily by instructional and curricular differences (Gamoran 2010). The connection between curriculum and achievement on standardized tests is often conceptualized as opportunity to learn (OTL), or students' exposure to the content of the assessment. But while organizational and institutional differences between the two types of tracking are easy to observe, it is difficult to predict the instructional implications of "academic" and "vocational" curricula versus subject-specific advanced or remedial curricula. Although OTL research originated with the earliest large-scale cross-national assessments over 40 years ago (McDonnell 1995), no recent cross-national assessment studies tie curricular exposure directly to student track placement in multiple countries across both types of tracking. For example, Schmidt and colleagues' work with TIMSS documents that OTL varies more within the United States than in many other countries and that this variation corresponds to eighth grade math course tracks (Schmidt et al. 2001). However, since the authors do not use student tracking data from any academic/vocational streaming countries, it is unknown how variation in OTL corresponds to tracks in these countries.

On the basis of existing research evidence, it remains unclear in which type of tracking OTL and achievement gaps between tracks should be larger. Although OTL measures are not available in the PISA data used in this article, it is the first study to directly compare achievement gaps between tracks in course-by-course tracking and academic/vocational streaming.

Which Type of Tracking Has Larger SES Disparities in Achievement within Tracks?

In contrast with the lack of comparative research on achievement gaps between tracks, a great deal of research has studied the relationship between tracking and inequality in achievement. Achievement inequality is often measured as the SES “gradient” (the coefficient of SES in a linear regression predicting achievement), and many authors find that gradients are steeper in academic/vocational streaming countries (Ammermüller 2005; Brunello and Checchi 2007; Duru-Bellat and Suchaut 2005; Marks 2005; Schütz, Ursprung and Wößmann 2008). However, a closer examination of these results reveals that many Anglo countries with course-by-course tracking have similar gradients to academic/vocational streaming countries, and it is Nordic countries with moderate or no course-by-course tracking that have very weak gradients. While this variation could be explained by other factors such as income inequality, more convincing evidence regarding the role of tracking would need to examine student-level track positions rather than aggregate country-level measures. Otherwise, the steeper SES gradients in academic/vocational streaming countries could well be due to inequality within tracks. Two studies cast further doubt on a causal relationship between tracking and SES gradients, by showing that the stronger association between family background and achievement in academic/vocational streaming countries is also evident in fourth grade before tracking begins (Jakubowski 2010; Waldinger 2006). Still, neither of these studies uses student-level track placement data. Therefore, it still remains to be seen whether SES gradients are steeper within tracks in course-by-course tracking than in academic/vocational streaming.

If the SES gradient is indeed steeper within tracks in course-by-course tracking, this could occur for two reasons. First, the meaning of tracks may be less consistent across schools in course-by-course tracking than in academic/vocational streaming. There is evidence that in the

U.S., not all “high tracks” are created equal. Depending on the SES profile of their schools, similar students take differing amounts of advanced math and science coursework (Crosnoe 2009). Among courses with the same title, curriculum and textbooks vary widely across different schools, and are related to school demographics (Cogan, Schmidt and Wiley 2001). Thus, it is important to take into account variation across schools, which will be done in this article using hierarchical models. The second reason the SES gradient may be steeper within tracks in course-by-course tracking is that there could be greater inequality even among students in both the same school and track. Prior research suggests that in academic/vocational streaming, individual SES effects occur primarily at the single transition point when track selection occurs, and within those tracks, SES disparities are small. In contrast, in course-by-course tracking, individual-level factors such as parental involvement may continue to have strong effects, even for students in the same track (Buchmann and Dalton 2002; Crosnoe and Schneider 2010; Oswald, Baker and Stevenson 1988; Rubinson 1986; Turner 1960).

Data

The data set used is the Program for International Student Assessment (PISA) 2003, conducted by the Organization for Economic Cooperation and Development (OECD). PISA tests nationally representative samples of 15-year-olds in each participating country, regardless of grade level, in math, science, and reading. PISA 2003 was selected as the data set for this analysis because it is the only cycle of PISA to collect detailed international data on math course tracking. This is vital to the goal of the study to examine course-by-course tracking internationally because math is arguably the most important subject of course-by-course tracking in the United States and many other course-by-course tracking countries. It is the subject that is most commonly tracked, the cumulative nature of the curriculum makes course sequences more rigid than other subjects, and math track placement influences placement in other subjects (Frank et al. 2008; Ireson and Hallam 2009; Lucas 1999; Stevenson, Schiller and Schneider 1994).

Country PISA samples are nationally representative of all 15-year-olds who are enrolled in public and private schools. In 2003, about 276,000 students in 41 countries participated, including all 30 OECD countries at the time and 11 non-OECD countries, which tend to be lower income. In the analyses to follow, the sample is limited to OECD countries excluding Mexico and Turkey, in order to obtain a more homogeneous sample of wealthy countries.[1] Also excluded are the nine countries that either do not practice any tracking for 15-year-olds or did not collect student tracking data (listed in the next section below). The Flemish and French communities of Belgium are counted as two "countries" because they have two separate school systems and, by extension, two separate tracking systems. After exclusions, the sample consists of about 99,000 students in about 3400 schools in 20 countries.

Limitations of the Dataset

One limitation of this dataset is that although it represents all OECD countries practicing academic/vocational streaming, it excludes several course-by-course tracking countries because they did not collect tracking data. Among the countries without tracking data are three that practice extensive course-by-course tracking (Canada, New Zealand, and the United Kingdom), four with moderate levels of course-by-course tracking (Norway, Poland, Spain, and Sweden), and two with very little course-by-course tracking (Denmark and Finland). Therefore, the results cannot be seen as representative of all course-by-course tracking across the OECD. However, the three available course-by-course tracking countries do represent a range of course-by-course tracking practices, as they include two Anglo countries with extensive course-by-course tracking (the United States and Australia) and one Nordic country with a moderate level of course-by-course tracking (Iceland).

A further limitation of this and any PISA dataset is that since it is cross-sectional, it does not include prior measures of achievement, meaning that we cannot observe the effects of tracking on students over time, only the existing differences between tracks at age 15. Since students in different tracks likely had different prior levels of achievement—indeed this is the purpose of tracking—SES segregation between tracks is due to a combination of SES differences in prior achievement and other factors such as school personnel bias in assignment or differences in parental involvement in assignment. Similarly, any observed achievement differences between tracks should not be attributed exclusively to tracking effects but rather to a combination of sorting and tracking effects. Although we cannot judge the effects of tracking, SES and achievement differences between tracks are still substantively meaningful, as they represent the way in which inequality in achievement is structured in a particular country, and they allow us to answer the question, "How much do we know about a student if we know his or her track?"

Finally, because PISA does not sample classrooms or administer teacher questionnaires, it does not include extensive data on curriculum and OTL. The trade-off is that because PISA randomly samples 35 15-year-olds from across the entire school, we can observe within-school variation.

Variables

Math Achievement.—The PISA math assessment is designed to measure students' mathematical "literacy," or their ability to apply their knowledge and skills in real-world situations, and includes the areas of quantity, space and shape, change and relationships, and uncertainty (OECD 2004). PISA conceptualizes the achievement of country populations and subpopulations as latent variables measured with uncertainty. Thus, student math scores consist of five plausible values for each student rather than a single score; methods for plausible values are described below. The mean score in OECD countries is 500 and the standard deviation is 100. To ease interpretation of the results, student achievement is standardized to have a mean of 0 and a standard deviation of 1 across the sample of 20 countries.

Socioeconomic Status.—The OECD-created Index of Social and Cultural Status is the most commonly used measure of SES in research using PISA data. It combines highest level of parental educational attainment, highest level of parental occupational status; and an index of household possessions, including books, a computer, and the student's own bedroom.[2] All components of SES are student reported. To ease interpretation of the results, student achievement is standardized to have a mean of 0 and a standard deviation of 1 across the sample of 20 countries.

Track.—Since the aim of this paper is primarily to compare the older and more typically-studied academic/vocational streaming to the newer and less studied course-by-course tracking, rather than to develop a new classification of tracking practices, it makes the conservative decision to classify a country's type of tracking simply by the track variable used for its students, either math class or academic/vocational program. In the three countries with course tracking data, students were asked "What type of mathematics class are you taking?" In the United States, students responded in six categories: Pre-algebra or general mathematics, Algebra I, Geometry, Algebra II, Pre-Calculus or Calculus, or Other. Since students in American high schools typically progress through these classes year by year in a particular sequence (Algebra I in ninth grade, Geometry in tenth grade, etc.), US tracking research often codes math track based on a combination of student math course and student grade level (e.g., Lucas 1999; Stevenson, Schiller and Schneider 1994). Thus, math class is recoded as “middle math track” if the student is at grade level within the course sequence, “high math track” if the student is above grade level, and “low math track” if the student is below grade level. In Australia and Iceland, students responded in only three categories: “high level”, “medium level”, or “basic level” math classes (see online appendix for more information about coding math course tracking).

In countries with academic/vocational streaming, students were asked “Which of the following programs are you in?” The country-specific responses vary in number and were recoded into three (high, middle, low) categories where possible, except in cases where there are only two tracks. The categories generally correspond to academic, high vocational, or low vocational. The *International Encyclopedia of National Systems of Education* (Postlethwaite 1995), Eurydice.org (Eurydice 2011), and any descriptions of national education systems in the empirical literature reviewed above were used as resources. There are some countries where both the math class variable and the academic/vocational program variable were collected, but the analyses below are limited to the primary form of tracking in each country, as identified by prior research. Table 1 presents a description of the tracking data.

(Table 1 about here)

Note that in several countries, a substantial portion of students are untracked, indicating that they report attending comprehensive schools (in most cases, lower secondary schools). In the analyses to follow, untracked students are dropped. This is because the unique contribution of this study is to compare tracked students in the two forms of tracking, rather than to compare tracked students to untracked students, which has already been studied in prior literature.

Generally, course-by-course tracking countries have fewer untracked students than academic/vocational streaming countries. Note also that the distribution of students across tracks differs widely across countries. For example, the percentage of students in the high track ranges from 14% in the Czech Republic to 86% in Portugal. All of the course-by-course tracking countries have three tracks and generally have a more even distribution across tracks than many of the academic/vocational streaming countries. Logically, we would expect the high track to be more elite when it includes only a small portion of students. As both the proportion of students who are tracked and the distribution of students across tracks could influence results, in order to check robustness, some of the analyses below are limited to countries with “widespread tracking” (no more than 50 percent of students untracked and no more than 50 percent of students in the high track). This leaves all three countries with course-by-course tracking but only eight of the 17 countries with academic/vocational streaming (marked with an “x” in Table 1).

Grade.—Grade is student-reported grade level in school, which ranges from seventh to twelfth grade, with a median of tenth grade.

Gender.—Student-reported gender is recoded with female equal to 1 and male equal to 0.

Ethnic/Language Minority Status.—Minority status is a dummy variable recoded from five student-reported items. It equals 1 for any student born in another country or at least one of whose parents was born in another country, for any student whose family spoke a language at home other than the national language(s), or for any nonwhite student in the United States (student race was reported only in the United States).

Methods

The objectives of this study are to compare course-by-course tracking and academic/vocational streaming in terms of the level of socioeconomic segregation between tracks, the size of achievement gaps between tracks, and the strength of the SES gradient within tracks. First, separate models are estimated for each country to describe variation across the sample, and then all countries are pooled to test for significant differences between the two types of tracking, after adding control variables.

The individual country models predicting selection into tracks are a series of multinomial logistic regressions predicting high and low track enrollment, with middle track as the reference category. In countries with only two tracks, logistic regressions are used to predict high track enrollment. The country models describing achievement gaps between tracks are ordinary least squares (OLS) regressions predicting math achievement from track enrollment. Track is entered as dummy variables for high and low track, with middle track as the reference category or as only a dummy variable for high track in countries with two tracks. The individual country models do not account for clustering of students within schools because track is a school-level variable in countries practicing between-school academic/vocational streaming. In those countries, track would be perfectly predicted by, or collinear with, students' school locations. Thus, measures of segregation and achievement gaps that are comparable across both types of tracking must ignore schools. (However, some pooled models below do incorporate the school level.)

All pooled models are hierarchical models that account for clustering of students within countries. The model testing the first research question—whether course-by-course tracking is less segregated than academic/vocational streaming—models track placement using a two-level hierarchical generalized linear model with a multinomial logit link function, where for each category *m* = 1 (low), 3 (high), relative to the reference category 2 (middle), we have

$$\eta_{mij} = \log\left(\frac{\varphi_{mij}}{\varphi_{2ij}}\right)$$

where

$$\Pr(H = m|\beta) = \varphi_{mij}$$

and η_{mij} is the log odds that student *i* in country *j* is in category *m* of track (*H*), which is estimated as follows:

$$\hat{n}_{mij} = \gamma_{00(m)} + \gamma_{10(m)}S_{ij} + \gamma_{01(m)}T_j + \gamma_{11(m)}S_{ij(m)}T_j + \boldsymbol{X}_{ij(m)}\boldsymbol{\beta} + u_{j(m)} + u_{S(m)},$$

$$u_{j(m)} \sim N(0, \tau_{00}); \quad u_{S(m)} \sim N(0, \tau_{10}),$$

[1]

where γ_{10} and γ_{01} are the main effects on student SES S_{ij} and country type of tracking T_j, γ_{11} is the coefficient for the interaction between student SES and country type of tracking $S_{ij}T_j$, $\boldsymbol{X}_{ij}$ is a vector of student-level controls (grade level, gender, and ethnic/language minority status), u_j is a country-level random intercept, and u_S is a country-level random slope on student SES. All student-level variables are country-mean centered. If course-by-course tracking is less segregated than academic/vocational streaming, then the course-by-course tracking × SES interaction should be negative when predicting high track and positive when predicting low track. As with the individual country models above, the pooled track selection model does not incorporate the school level since school exactly coincides with track in between-school tracking countries.

For the second research question comparing achievement gaps between tracks, achievement is modeled using a two-level hierarchical linear model of the form

$$\hat{A}_{ij} = \gamma_{00} + \gamma_{10}H_{ij} + \gamma_{20}S + \gamma_{01}T_j + \gamma_{11}H_{ij}T_j + \gamma_{11}S_{ij}T_j + \boldsymbol{X}_{ij}\boldsymbol{\beta} + u_j + u_H + u_S + r_{ij},$$
$$u_j \sim N(0, \tau_{00}); \quad u_H \sim N(0, \tau_{10}); \quad u_S \sim N(0, \tau_{20}); \quad r_{ij} \sim N(0, \sigma^2), \qquad [2]$$

where A_{ij} (math achievement) is the outcome of student i in country j, γ_{10}, γ_{20}, and γ_{01} are the main effects on student track H_{ij}, student SES S_{ij}, and country type of tracking T_j, γ_{11} and γ_{11} are the coefficients for the interactions between student track and country type of tracking $H_{ij}T_j$, and between student SES and country type of tracking $S_{ij}T_j$, $\boldsymbol{X}_{ij}$ is a vector of student-level controls (grade level, gender, and ethnic/language minority status), u_j is a country-level random intercept, u_H and u_S are country-level random slopes on student track and student SES, and r_{ij} is a student-level random effect. All student-level variables are country-mean centered. If achievement gaps between tracks are smaller in course-by-course tracking than in academic/vocational streaming, then the course-by-course tracking × high track interaction should be negative. This model does not include the school level so that achievement gaps in between- and within-school tracking can be compared in the same model.

The final achievement model tests the third research question of whether SES gradients are steeper within tracks in course-by-course tracking than in academic/vocational streaming. Thus, it is a three-level model that incorporates the clustering of students within schools as well as countries. It adds school-level controls for aggregated student-level variables (i.e., school mean grade level, proportion female, proportion minority, and mean SES) as well as school-level random slopes on student SES and student track. All student-level variables are school-mean centered except for track which remains country-mean centered, as it is a school-level variable in between-school tracking systems and a student-level variable in within-school tracking systems.

All school-level variables are country-mean centered. If SES gradients are steeper within tracks in course-by-course tracking than in academic/vocational streaming, then the course-by-course tracking × SES interaction should be positive.

PISA uses two-stage sampling of first schools and then students (OECD 2005), resulting in a complex sample design. Following Kreuter and Valliant (2007), the individual country OLS and multinomial logit models are estimated with Stata using provided student replicate weights and the balanced repeated replication technique, resulting in standard errors that correctly take sample design into account (OECD 2005). The hierarchical models are estimated with HLM 7.0 software using the provided final student and school weights. Following Rabe-Hesketh and Skrondal (2006), student weights are transformed to conditional within-school weights for the three-level models and conditional within-country weights for the two-level models. The country weight portion of the final student weight is applied at the country level.

All variables except achievement are derived from student questionnaires and have low levels of missingness, ranging from 0.05% for grade level to 5.46% for track. All missing data are imputed with Stata, using multiple imputation by iterative chained equations and creating five imputed data sets. For the individual country OLS and multinomial logit models, each model is estimated five times and averaged, and all standard errors are calculated to reflect uncertainty due to imputation. For the hierarchical models, HLM's multiple imputation function is used to implement this procedure automatically. Similarly, the five plausible values of achievement in PISA can also be understood in a multiple imputation framework. The proper technique for dealing with plausible values uses the same standard error formula to account for imputation variance (OECD 2005). Therefore, the same procedures described above for computing average estimates and standard errors are also used to correctly deal with plausible values of achievement.

Results

Which Type of Tracking is More Socioeconomically Segregated?

The first two columns of table 2 present descriptive results for SES segregation across the 20 countries. The SES coefficients predicting high track are positive and significant in nearly every country, and the SES coefficients predicting low track are negative and significant in every country. As predicted, the coefficients tend to be smaller in magnitude for course-by-course tracking than for academic/vocational streaming. For course-by-course tracking, the median log odds of being in the high versus the middle track are 0.50, and the median log odds of being in the low versus the middle track are -0.42. The corresponding estimates for academic/vocational streaming are 0.93 and -0.82. Translated into odds, this means that in course-by-course tracking, a one standard deviation increase in SES corresponds to a student being about 1.7 times more likely to be in the high track and about 0.7 times as likely to be in the low track. In academic/vocational streaming, this same unit increase in SES corresponds to a student being 2.5 times more likely to be in the high track and about 0.4 times as likely to be in the low track. These results suggest that course-by-course tracking systems tend to have lower SES segregation between tracks than academic/vocational streaming systems.

(Table 2 about here)

Table 3 presents results of the hierarchical multinomial logit models predicting track selection, which allow a test of whether the difference observed above between types of tracking is significant, while also controlling for other student characteristics. If the results of the pooled models are consistent with the country models, then the course-by-course tracking × SES interaction is expected to be negative and significant when predicting high track and positive and significant when predicting low track. This is in fact what the models show. In model 1, where the sample of countries is restricted to only those with widespread tracking, after controlling for student grade level, gender, and ethnicity, the coefficient on course-by-course tracking × SES predicting high track is negative and significant ($p < .001$), and the coefficient on course-by-course tracking × SES predicting low track is positive and significant ($p = .02$). In model 2, which includes the full sample of countries, the coefficient on course-by-course tracking × SES predicting high track is also negative and significant ($p < .001$). Thus, after accounting for other demographic characteristics associated with student track, course-by-course tracking systems have significantly lower levels of SES segregation between tracks than academic/vocational streaming systems, both in the sample on the whole and in the smaller sample of countries with widespread tracking.

(Table 3 about here)

Which Type of Tracking Has Larger Achievement Gaps between Tracks?

The last two columns of table 2 give a descriptive picture of achievement gaps between tracks in the two types of tracking. High track coefficients predicting math achievement are positive and significant in nearly every country, and low track coefficients are negative and significant in every country.[3] Unlike the results for SES segregation, there appears to be no noticeable difference between the two types of tracking in the size of achievement gaps between tracks. The median gap between estimated high and low track achievement for course-by-course tracking countries is about 1.3 standard deviations and for academic/vocational streaming countries is about 1.2 standard deviations.

Table 4 presents results of the hierarchical linear models predicting math achievement, which allow a test of whether the small difference in achievement gaps between types of tracking observed above is significant, while controlling for differences in the demographic compositions of tracks. Model 3 is restricted to the sample of countries with widespread tracking. If results of this model are consistent with those from the country models above, both the course-by-course tracking × high track interaction and the course-by-course tracking × low track interaction are expected to be close to zero and not significant. In fact, the findings in model 3 only partly support this prediction. While the course-by-course tracking × high track interaction is indeed close to zero and not significant (p =.83), the course-by-course tracking × low track interaction is positive and significant ($p < .001$). This indicates that in countries with widespread tracking, after controlling for student grade level, gender, ethnicity, and SES, the achievement gap between the high and middle tracks is similar in course-by-course tracking to academic/vocational streaming, but the achievement gap between low and middle tracks is smaller in course-by-course tracking than in academic/vocational streaming. When model 3 is expanded to the full set of countries (not presented here), only the high track coefficient and interaction can be estimated, and this interaction is not significant, indicating similar achievement gaps between the high track and the combined middle and low tracks across the entire sample. Therefore, the size of achievement gaps between tracks appears somewhat similar across the two types of tracking; gaps are similar between the high and middle tracks but different between the low and middle tracks.

(Table 4 about here)

Which Type of Tracking Has Larger SES Disparities in Achievement within Tracks?

The course-by-course tracking × SES interaction in model 3 above provides a first look at whether the relationship between SES and achievement is stronger within tracks for students in course-by-course tracking than in academic/vocational streaming. Controlling for student track, the course-by-course tracking × SES interaction is positive and highly significant ($p < .001$). However, there are two potential sources of steeper SES gradients within tracks in course-by-course tracking countries. The first source is between-school differences, which could arise if the average SES and achievement levels of high and low tracks are less consistent across schools in course-by-course tracking than in academic/vocational streaming. The second source of steeper gradients within tracks in course-by-course tracking is if there are larger SES disparities in achievement even among students in both the same school and track. Model 4 adjudicates between these two possibilities by adding the school level to the model. This allows us to observe whether students in the same school and the same track continue to have a significantly stronger relationship between SES and achievement in course-by-course tracking than in academic/vocational streaming. If so, the course-by-course tracking × SES interaction is expected to remain positive and significant. In fact, the interaction in model 4 does remain positive and significant, indicating that even among students in the same school and same track, SES is more predictive of achievement in course-by-course tracking than in academic/vocational streaming. The interaction is reduced in significance (from $p < .001$ to $p = .04$) and slightly reduced in magnitude, which suggests that the contrast between the two types of tracking is partially explained by differences across schools (i.e., all “high tracks” not being created equal in course-by-course tracking). Finally, model 5 tests the robustness of these findings in the full sample of countries. Here too, the course-by-course cracking × SES interaction is positive and significant (*p* =.047), consistent with the results from model 4. Thus, while prior literature has shown that overall country SES gradients are similar in the two types of tracking (and additional analyses not reported here confirm this for the current dataset), it appears that more of the gradient occurs within tracks and schools in course-by-course tracking, and more of the gradient occurs between tracks in academic/vocational streaming.

Figure 1 plots math achievement by SES in course-by-course tracking and academic/vocational streaming (limited to widespread tracking countries), and fits lines estimating the association between SES and achievement in each of the three tracks. A number of the results observed above are illustrated in this figure. First, in course-by-course tracking, there are more high-track dots at low values of SES and more low-track dots at high values of SES compared to academic/vocational streaming. This reflects the finding that course-by-course tracking has less SES segregation between tracks than academic/vocational streaming. Second, the distance between the fit lines is similar on average, reflecting the finding that achievement gaps between tracks do not differ dramatically between course-by-course tracking and academic/vocational streaming. Third, the slopes of the fit lines are steeper in course-by-course tracking. That is, SES remains highly predictive of achievement in course-by-course tracking systems, even for students in the same track. It can be calculated from model 3 that in course-by-course tracking, a low-SES student (SES = -1) in the high track and a high-SES student (SES = 1) in the high track have an estimated difference in achievement of over half a standard deviation, while in academic/vocational streaming, two similarly low-SES and high-SES students both in the high track have an estimated difference in achievement of less than one third of a standard deviation. Therefore, even though the results from the segregation models above show that a low-SES student has a lower probability of being in the high track in academic/vocational streaming than in course-by-course tracking, once that student is in the high track, she is predicted to achieve nearly as high as her high-SES classmates. In course-by-course tracking, a low-SES student in the high track is still predicted to achieve substantially lower than her high-SES classmates.

(Figure 1 about here)

Discussion

Overall, the findings of this article suggest that while course-by-course tracking and academic/vocational streaming differ in some important ways, in terms of the key outcome—math achievement gaps between tracks—they do not differ dramatically. This study has three main findings. First, as predicted by prior literature, SES segregation between tracks is higher in academic/vocational streaming than in course-by-course tracking. This finding supports the life course perspective, which predicts that academic/vocational streaming will be more segregated, as it typically begins at earlier ages when parental background holds a strong influence over educational transitions (Blossfeld and Shavit 1993; Mare 1980). Alternatively, it could be that SES is already more strongly associated with achievement prior to track selection in academic/vocational streaming countries, and the higher level of segregation is due to the indirect effect of SES on achievement. However, despite some limited evidence (Jakubowski 2010; Waldinger 2006), there is not a strong reason to believe that the untracked primary school systems of academic/vocational streaming countries produce greater SES disparities in achievement than in course-by-course tracking countries.

Second, after controlling for these differences in the demographic composition of tracks, the achievement gap between the low and middle tracks is larger for academic/vocational streaming than for course-by-course tracking, but the gap between the high and middle tracks is very similar in both types of tracking. This finding is the first evidence directly comparing achievement gaps between tracks in course-by-course tracking and academic/vocational streaming, and supports the notion that course-by-course tracking and academic/vocational streaming expose students to similarly differentiated OTL. Alternatively (or more probably, in addition), the different allocation mechanisms employed in the two tracking systems may end up sorting students similarly by achievement. Since no prior measures of achievement are available in PISA, we cannot adjudicate between these two explanations; existing achievement gaps between tracks at age 15 most likely result from some combination of track effects and selection effects.

The larger achievement gap between the low and middle track in academic/vocational streaming may be a result of tracks in these countries having more universal countrywide meanings that transcend individual schools, in contrast with greater variability in course-by-course tracking countries. Across the sample of course-by-course tracking countries, about 14% of students attend schools where no student reports being in the low math track. This may simply be due to sampling variation (although only 6% of students attend schools where no student reports being in the high track). But if a sizable number of schools truly have only two math tracks, then the lowest-achieving students in those schools would be in the middle track, while the lowest-achieving students in academic/vocational streaming countries would nearly always have a low vocational school available in their locality.

Third, although SES disparities in achievement are similar overall in both types of tracking, student track placement explains more of these SES disparities in academic/vocational streaming than in course-by-course tracking. This means that in course-by-course tracking, sizable SES disparities in achievement exist among students in the same track. These remaining SES disparities are partly explained by differences across schools, which suggests that “not all high tracks are created equal,” (i.e., that there is variation across schools in the meaning of tracks) (Cogan, Schmidt and Wiley 2001; Crosnoe 2009). But most of those SES disparities remain after accounting for the school level; that is, they are disparities between students in both the same school and track. This is consistent with the idea that there is more SES inequality at the individual or family level in course-by-course tracking countries, and more at the institutional or track level in academic/vocational streaming countries (Buchmann and Dalton 2002; Crosnoe and Schneider 2010; Oswald, Baker and Stevenson 1988; Rubinson 1986; Turner 1960).

While there are important institutional differences between course-by-course tracking and academic/vocational streaming, the results of this study suggest that the two types of tracking differ more in degree than in kind. Although SES segregation between tracks is higher in academic/vocational streaming, it is still high in course-by-course tracking. And although the achievement gap between the low and middle tracks is larger in academic/vocational streaming, the gap between the high and middle tracks is equally large in both systems. This reinforces the concern that course-by-course tracking may substitute an implicitly unequal system for an explicitly unequal one. Lucas’ (2001) theory of “effectively maintained inequality” states that as access to a level of education becomes universal, qualitative differentiation occurs within the level. The results of this study suggest that such processes occur in many different international contexts with different institutional tracking structures.

Furthermore, the similarities found between the two forms of tracking validate the practice of comparing them, as this study has done for the first time, and also strongly warrant future research in this area. As more countries reform their academic/vocational streaming systems and replace them with course-by-course tracking, international comparative research cannot continue to ignore the newer, less explicit form of tracking. Since this study was the first comparison of achievement in course-by-course tracking and academic/vocational streaming, it made some simplifying assumptions, such as examining only the predominant type of tracking in each country and grouping together all course-by-course tracking countries and all academic/vocational streaming countries. Future research should examine variation in tracking practices in more detail and should collect curricular data to examine OTL gaps between tracks and longitudinal data to observe changing achievement over time.

Although this study supports the idea that the different institutional structures of the two types of tracking do not correspond to large differences in the instructional effects of tracking, they may nevertheless correspond to dramatic differences in the social and institutional effects of tracking. These could have serious consequences for students and important implications for educational policy and practice. In previous research, we have shown that different types of tracking have different effects on students' academic self-concept (Chmielewski, Dumont and Trautwein 2013). Additionally, the less explicit structure of course-by-course tracking may leave students and parents less aware of how course choices influence postsecondary opportunities. Lucas (1999) portrays tracking in US high schools as an exceedingly complex and "hidden" system in which middle class students have a distinct advantage because of their families' greater levels of information. Schneider and Stevenson (1999) find that many students with high educational aspirations are unaware that their course work choices have not prepared them for college. A large body of literature has revealed a lack of institutional or chartering effects of tracking in the United States (Dreeben and Barr 1988; Gamoran 1986; Pallas et al. 1994), in contrast with strong chartering effects in countries with academic/vocational streaming (Buchmann and Dalton 2002; Buchmann and Park 2009). Many authors argue that this results in American students having less "realistic" educational aspirations than students in academic/vocational streaming (Buchmann and Park 2009; Kerckhoff 1977; Mateju et al. 2007; Turner 1960). Indeed, additional analyses not reported here indicate that gaps between tracks in university expectations are substantially larger in academic/vocational streaming countries than in course-by-course tracking countries.[4] As more countries replace explicit academic/vocational streaming with less explicit course-by-course tracking, they may give more low-track students the expectation of college without giving them greater preparation to succeed in college.[5]

References

Adelman, Clifford. 1999. "Answers in the Tool Box: Academic Intensity, Attendance Patterns, and Bachelor's Degree Attainment." Washington, DC: Government Printing Office: U.S. Department of Education, Office of Educational Research and Improvement.

Ammermüller, Andreas. 2005. *Educational Opportunities and the Role of Institutions*. Mannheim: ZEW Discussion Paper No. 05-44, Center for European Economic Research.

Ariga, Kenn, Giorgio Brunello, Roki Iwahashi, and Lorenzo Rocco. 2005. "Why Is the Timing of School Tracking So Heterogeneous?". Bonn, Germany: IZA Discussion Paper No. 1854.

Baker, David P., and David L. Stevenson. 1986. "Mothers' strategies for children's school achievement: Managing the transition to high school." *Sociology of Education* 59(3):156-66.

Blossfeld, Hans-Peter, and Yossi Shavit. 1993. "Persisting Barriers: Changes in Educational Opportunities in Thirteen Countries." in *Persistent Inequality: Changing Educational Attainment in Thirteen Countries,* edited by Yossi Shavit and Hans-Peter Blossfeld. Boulder, CO: Westview Press.

Bowles, Samuel, and Herbert Gintis. 1976. *Schooling in Capitalist America*. New York: Basic Books.

Broaded, C. Montgomery. 1997. "The Limits and Possibilities of Tracking: Some Evidence from Taiwan." *Sociology of Education* 70(1):36-53.

Brunello, Giorgio, and Daniele Checchi. 2007. "Does school tracking affect equality of opportunity? New international evidence." *Economic Policy* 22(52):781-861.

Buchmann, Claudia, and Ben Dalton. 2002. "Interpersonal Influences and Educational Aspirations in 12 Countries: The Importance of Institutional Context." *Sociology of Education* 75(2):99-122.

Buchmann, Claudia, and Hyunjoon Park. 2009. "Stratification and the formation of expectations in highly differentiated educational systems." *Research in Social Stratification and Mobility* 27(4):245-67.

Chmielewski, Anna K., Hanna Dumont, and Ulrich Trautwein. 2013. "Tracking Effects Depend on Tracking Type: An International Comparison of Students' Mathematics Self-Concept." *American Educational Research Journal* 50(5):925-57.

Cogan, Leland S., William H. Schmidt, and David E. Wiley. 2001. "Who Takes What Math and in Which Track? Using TIMSS to Characterize US Students' Eighth-Grade Mathematics Learning Opportunities." *Educational Evaluation and Policy Analysis* 23(4):323-41.

Crosnoe, Robert. 2009. "Low-Income Students and the Socioeconomic Composition of Public High Schools." *American Sociological Review* 74(5):709-30.

Crosnoe, Robert, and Barbara Schneider. 2010. "Social Capital, Information, and Socioeconomic Disparities in Math Coursework." *American journal of education (Chicago, Ill.)* 117(1):79-107.

Demeuse, Marc, and Ariane Baye. 2008. "Measuring and Comparing the Equity of Education Systems in Europe." Pp. 85-106 in *Governance and Performance of Education Systems*, edited by N. C. Soguel and P. Jaccard. Dordrecht: Springer.

Dreeben, Robert, and Rebecca Barr. 1988. "Classroom composition and the design of instruction." *Sociology of Education* 61(3):129-42.

Dupriez, Vincent, Xavier Dumay, and Anne Vause. 2008. "How Do School Systems Manage Pupils' Heterogeneity?" *Comparative Education Review* 52(2):245-73.

Duru-Bellat, Marie. 1996. "Social Inequalities in French Secondary Schools: From Figures to Theories." *British Journal of Sociology of Education* 17(3):341-50.

Duru-Bellat, Marie, and Bruno Suchaut. 2005. "Organisation and Context, Efficiency and Equity of Educational Systems: What PISA Tells Us." *European Educational Research Journal* 4(3):181-94.

Eurydice. 2011. "Eurydice: The Information Network on Education in Europe." Brussels: Education Audiovisual & Culture Executive Agency.

Feinstein, Leon, and James Symons. 1999. "Attainment in secondary school." *Oxford Economic Papers* 51(2):300-21.

Figlio, David N., and Marianne E. Page. 2002. "School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality?" *Journal of Urban Economics* 51(3):497-514.

Frank, Kenneth A., Chandra Muller, Kathryn S. Schiller, Catherine Riegle-Crumb, Anna Strassmann Mueller, Robert Crosnoe, and Jennifer Pearson. 2008. "The Social Dynamics of Mathematics Coursetaking in High School." *American Journal of Sociology* 113(6):1645-96.

Gamoran, Adam. 1986. "Instructional and Institutional Effects of Ability Grouping." *Sociology of Education* 59(4):185-98.

—. 2010. "Tracking and Inequality: New Directions for Research and Practice." Pp. 213-28 in *The Routledge International Handbook of the Sociology of Education*, edited by M. Apple, S.J. Ball, and L.A. Gandin. London: Routledge.

Gamoran, Adam, and Robert D. Mare. 1989. "Secondary School Tracking and Educational Inequality: Compensation, Reinforcement, or Neutrality?" *American Journal of Sociology* 94(5):1146-83.

Gorard, Stephen, and Emma Smith. 2004. "An International Comparison of Equity in Education Systems." *Comparative Education* 40(1):15-28.

Hanushek, Eric, and Ludger Wößmann. 2006. "Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence Across Countries." *The Economic Journal* 116(March):63-76.

Heidenheimer, Arnold J. 1974. "The Politics of Educational Reform: Explaining Different Outcomes of School Comprehensivization Attempts in Sweden and West Germany." *Comparative Education Review* 18(3):388-410.

Horn, Dániel, Ildikó Balázsi, Szabolcs Takács, and Yanhong Zhang. 2006. "Tracking and inequality of learning outcomes in Hungarian secondary schools." *Prospects* 36(4):433-46.

Huang, Min-Hsiung. 2009. "Classroom Homogeneity and the Distribution of Student Math Performance: A Country-Level Fixed-Effects Analysis." *Social Science Research* 38(4):781-91.

Ireson, Judith, and Susan Hallam. 2009. "Academic self-concepts in adolescence: Relations with achievement and ability grouping in schools." *Learning and Instruction* 19(3):201-13.

Jakubowski, Maciej. 2010. "Institutional Tracking and Achievement Growth: Exploring Difference-in-Differences Approach to PIRLS, TIMSS, and PISA Data." Pp. 41-81 in *Quality and Inequality of Education*, edited by J. Dronkers. Dordrecht: Springer Science+Business Media B.V.

Jenkins, Stephen P., John Micklewright, and Sylke V. Schnepf. 2006. "Social Segregation in Secondary Schools: How Does England Compare with Other Countries?". Bonn, Germany: IZA Discussion Paper No. 1959.

Kelly, Sean. 2004. "Do Increased Levels of Parental Involvement Account for Social Class Differences in Track Placement?" *Social Science Research* 33(4):626-59.

Kerckhoff, Alan C. 1977. "The Realism of Educational Ambitions in England and the United States." *American Sociological Review* 42(4):563-71.

—. 1986. "Effects of Ability Grouping in British Secondary Schools." *American Sociological Review* 51(6):842-58.

Kreuter, Frauke, and Richard Valliant. 2007. "A survey on survey statistics: What is done and can be done in Stata." *The Stata Journal* 7(1):1-21.

Lamb, Stephen, and Sue Fullarton. 2002. "Classroom and School Factors Affecting Mathematics Achievement: A Comparative Study of Australia and the United States Using TIMSS." *Australian Journal of education* 46(2):154-72.

Lee, Valerie E., and Anthony S. Bryk. 1988. "Curriculum Tracking as Mediating the Social Distribution of High School Achievement." *Sociology of Education* 61(2):78-94.

Leow, Christine, Sue Marcus, Elaine Zanutto, and Robert Boruch. 2004. "Effects of Advanced Course-taking on Math and Science Achievement: Addressing Selection Bias Using Propensity Scores." *American Journal of Evaluation* 25(4):461-78.

LeTendre, Gerald K., Barbara K. Hofer, and Hidetada Shimizu. 2003. "What is Tracking? Cultural Expectations in the United States, Germany, and Japan." *American Educational Research Journal* 40(1):43-89.

Loveless, Tom. 2013. "The 2013 Brown Center Report on American Education: How Well are American Students Learning?". Washington, DC: The Brookings Institution.

Lucas, Samuel R. 1999. *Tracking Inequality: Stratification and Mobility in American High Schools*. New York: Teachers College Press.

—. 2001. "Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects." *American Journal of Sociology* 106(6):1642-90.

Lucas, Samuel R., and Mark Berends. 2002. "Sociodemographic Diversity, Correlated Achievement, and De Facto Tracking." *Sociology of Education* 75(4):328-48.

Maaz, Kai, Ulrich Trautwein, Oliver Lüdtke, and Jürgen Baumert. 2008. "Educational Transitions and Differential Learning Environments: How Explicit Between-School Tracking Contributes to Social Inequality in Educational Outcomes." *Child Development Perspectives* 2(2):99-106.

Manning, Alan, and Jörn-Steffen Steve Pischke. 2006. "Comprehensive Versus Selective Schooling in England in Wales: What Do We know?". Bonn, Germany: IZA Discussion Paper No. 2072.

Mare, Robert D. 1980. "Social Background and School Continuation Decisions." *Journal of the American Statistical Association* 75(370):295-305.

Marks, Gary N. 2005. "Cross-National Differences and Accounting for Social Class Inequalities in Education." *International Sociology* 20(4):483-505.

—. 2006. "Are Between- and Within-School Differences in Student Performance Largely Due to Socio-Economic Background? Evidence from 30 Countries." *Educational Research* 48(1):21-40.

Mateju, Petr, Michael L. Smith, Petr Soukup, and Josef Basl. 2007. "Determination of College Expectations in OECD countries: The Role of Individual and Structural Factors." *Czech Sociological Review* 43(6):1121-48.

Mateju, Petr, and Jana Straková. 2005. "The Role of the Family and the School in the Reproduction of Educational Inequalities in the Post-Communist Czech Republic." *British Journal of Sociology of Education* 26(1):17-40.

McDonnell, Lorraine M. 1995. "Opportunity to Learn as a Research Concept and a Policy Instrument." *Educational Evaluation and Policy Analysis* 17(3):305-22.

Meyer, John W. 1977. "The Effects of Education as an Institution." *American Journal of Sociology* 83(1):55-77.

Montt, Guillermo. 2011. "Cross-National Differences in Educational Achievement Inequality." *Sociology of Education* 84(1):49-68.

Oakes, Jeannie. 1985. *Keeping Track: How Schools Structure Inequality*. New Haven, CT: Yale University Press.

—. 2008. "Keeping Track: Structuring Equality and Inequality in an Era of Accountability." *The Teachers College Record* 110(3):700-12.

OECD. 2004. "Learning for Tomorrow's World: First Results from PISA 2003." Paris: OECD.

—. 2005. "PISA 2003 Technical Report." Paris: OECD.

—. 2011. "Education at a Glance 2011: OECD Indicators." Paris: OECD.

Oswald, Hans, David P. Baker, and David L. Stevenson. 1988. "School Charter and Parental Management in West Germany." *Sociology of Education* 61(4):255-65.

Pallas, Aaron, Doris R. Entwisle, Karl L. Alexander, and M. Grancis Stluka. 1994. "Ability Group Effects: Instructional, Social, or Institutional?" *Sociology of Education* 67(1):27-46.

Postlethwaite, T. Neville. 1995. *International Encyclopedia of National Systems of Education*. Oxford: Pergamon.

Rabe-Hesketh, Sophia, and Anders Skrondal. 2006. "Multilevel Modeling of Complex Survey Data." *Journal of the Royal Statistical Society* 169(4):805-27.

Rubinson, Richard. 1986. "Class Formation, Politics, and Institutions: Schooling in the United States." *American Journal of Sociology* 92(3):519-48.

Schmidt, William H., Curtis C. McKnight, Leland S. Cogan, Pamela M. Jakwerth, and Richard T. Houang. 1999. *Facing the Consequences: Using TIMSS for a Closer Look at U.S. Mathematics and Science Education*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Schmidt, William H., Curtis C. McKnight, Richard T. Houang, HsingChi Wang, David E. Wiley, Leland S. Cogan, and Richard G. Wolfe. 2001. *Why Schools Matter: A Cross-National Comparison of Curriculum and Learning*. New York: John Wiley & Sons.

Schnabel, Kai U., Corinne Alfeld, Jacquelynne S. Eccles, Olaf Köller, and Jürgen Baumert. 2002. "Parental Influence on Students' Educational Choices in the United States and Germany: Different Ramifications--Same Effect?" *Journal of Vocational Behavior* 60(2):178-98.

Schneider, Barbara, and David L. Stevenson. 1999. *The Ambitious Generation: America's Teenagers: Motivated but Directionless*. New Haven: Yale University Press.

Schneider, Barbara, Christopher B. Swanson, and Catherine Riegle-Crumb. 1998. "Opportunities for Learning: Course Sequences and Positional Advantages." *Social Psychology of Education* 2(1):25-53.

Schütz, Gabriela, Heinrich W. Ursprung, and Ludger Wößmann. 2008. "Education Policy and Equality of Opportunity." *Kyklos* 61(2):279-308.

Slavin, Robert E. 1990. "Achievement Effects of Ability Grouping in Secondary Schools: A Best-Evidence Synthesis." *Review of Educational Research* 60(3):471-99.

Stevenson, David L., Kathryn S. Schiller, and Barbara Schneider. 1994. "Sequences of Opportunities for Learning." *Sociology of Education* 67(3):184-98.

Trautwein, U., O. Köller, and E. Kämmerer. 2002. "Effekte innerer und äußerer Leistungsdifferenzierung auf selbstbezogene Fähigkeitskognitionen, die wahrgenommene Unterrichtspartizipation und die wahrgenommene soziale Akzeptanz [The Effects of Within-Class and Between-Class Ability Grouping on Academic Self-Concepts, Reported Classroom Participation, and Perceived Social Integration]." *Psychologie in Erziehung und Unterricht* 49(4):273-86.

Turner, Ralph. 1960. "Sponsored and Contest Mobility and the School System." *American Sociological Review* 25(6):855-67.

Van de Werfhorst, Herman G., and Jonathan J.B. Mijs. 2010. "Achievement Inequality and the Institutional Structure of Educational Systems: A Comparative Perspective." *Annual Review of Sociology* 36:407-28.

Waldinger, Fabian. 2006. "Does Tracking Affect the Importance of Family Background on Students' Test Scores?": mimeo, London School of Economics.

Wells, Amy Stuart, and Irene Serna. 1996. "The Politics of Culture: Understanding Local Political Resistance to Detracking in Racially Mixed Schools." *Harvard Educational Review* 66(1):93-119.

Westbury, Ian. 1992. "Comparing American and Japanese Achievement: Is the United States Really a Low Achiever?" *Educational Researcher* 21(5):18-24.

Willms, J. Douglas. 2010. "School Composition and Contextual Effects on Student Outcomes." *The Teachers College Record* 112(4):3-4.

Notes

I would like to thank Francisco Ramirez, Sean Reardon, Barbara Schneider, John Meyer, Evan Schofer, Christopher Jencks, William Schmidt, Hanna Dumont, and the members of Stanford Comparative Sociology Workshop and Center for Education Policy Analysis seminar for comments on previous drafts. This research was supported by a grant from the American Educational Research Association which receives funds for its AERA Grants Program from the National Science Foundation under Grant #DRL-0941014. Additional funding is by the Pathways to Adulthood Postdoctoral Fellowship Program. Opinions reflect those of the author and do not necessarily reflect those of the granting agencies.

[1] This follows the country sample decisions of other comparative research (e.g., Dupriez, Dumay and Vause 2008). Results for the full sample of countries were consistent with those reported here.

[2] All analyses were also run with parent education or parent occupation only, and results were very similar.

[3] In theory, math achievement gaps between tracks in course-by-course tracking may be particularly high because the data reflect math course tracking. Indeed, the gap in reading achievement between tracks is smaller in course-by-course tracking than in academic/vocational streaming (1.09 vs. 1.24 SD, respectively). However, as in the math achievement models, only the reading gap between the low and middle tracks is significantly smaller in course-by-course

tracking than in academic/vocational streaming (*p* =.001), while the reading gap between the high and middle tracks is not significantly different in the two types of tracking (*p* =.47).

[4] The average gap in expectations between high and low tracks for the United States, Australia, and Iceland is significantly smaller (16 percentage points) than that for the countries with widespread academic/vocational streaming (49 percentage points; *p* <.001).

[5] Although there are no data available on PISA students' eventual postsecondary attainment, OECD statistics on educational attainment indicate that the proportion of students who complete university degrees (of those who initially enroll) is lower in the three course-by-course tracking countries than in the academic/vocational streaming countries in the sample. However, this is entirely the result of much higher proportions of students initially enrolling in university, as the ultimate completion rates are nevertheless higher in course-by-course tracking countries than in academic/vocational streaming countries (about 33 percent versus 25 percent, respectively) (OECD 2011). These figures pertain to International Standard Classification of Education (ISCED) 5A attainment for the cohort aged 25-34 in 2009, which is slightly older than the PISA 2003 cohort.

Table 1. Descriptions of Tracking Data by Country

Country	Sample Size (Unweighted)	Percent[a] High Track	Middle Track	Low Track	Track Missing	Untracked	Widespread Tracking[b]
Course-by-Course Tracking							
Australia	12,551	36.14	47.65	16.21	12.69	0	x
Iceland	3,350	41.60	42.20	16.20	35.79	5.16	x
United States	5,456	28.98	49.22	21.80	12.13	0	x
Academic/Vocational Streaming							
Austria	4,597	21.29	53.67	25.04	0	0	x
Belgium-Flanders	5,059	47.13	28.98	23.89	0.74	0	x
Belgium-French	3,737	47.60	27.66	24.74	0.04	0	x
Czech Republic	6,320	14.13	69.25	16.62	0	0	x
France	4,300	84.61	0	15.39	0	38.32	
Germany	4,660	33.37	36.94	29.68	0.22	10.03	x
Greece	4,627	78.12	0	21.88	0	8.85	
Hungary	4,765	37.64	41.45	20.92	0	6.08	x
Ireland	3,880	78.91	17.97	3.12	0	63.68	
Italy	11,639	40.13	34.62	25.26	0	1.57	x
Japan	4,707	75.47	0	24.53	0	0	
Korea	5,444	72.91	0	27.09	0	1.57	
Luxembourg	3,923	80.86	10.37	8.77	0	0	
Netherlands	3,992	19.94	18.69	61.37	0.13	0	x
Portugal	4,608	86.39	0	13.61	0	35.09	
Slovak Republic	7,346	53.56	23.70	22.73	0	0	
Switzerland	8,420	36.59	12.08	51.33	1.42	80.25	

[a] "Missing" track data indicates that the student did not report his or her track information and are imputed using multiple imputation. "Untracked" indicates that the student attended a comprehensive lower secondary school or other comprehensive school with no tracking. Untracked students are omitted from the analyses. High, middle, and low track percentages are computed from imputed data and exclude untracked students. Thus, row totals may not equal 100.
[b] Countries with "widespread tracking" had no more than 50 percent of students in the high track and no more than 50 percent of students in untracked comprehensive schools.

Table 2. Raw Coefficients Predicting Track Selection and Math Achievement, No Controls

	SES Coefficients Predicting Track		Track Coefficients Predicting Achievement	
Country	High Track	Low Track	High Track	Low Track
Course-by-Course Tracking				
Australia	0.43 ***	-0.40 ***	0.80 ***	-0.61 ***
	(0.04)	(0.05)	(0.03)	(0.04)
Iceland	0.54 ***	-0.48 ***	0.74 ***	-0.60 ***
	(0.06)	(0.07)	(0.04)	(0.05)
United States	0.50 ***	-0.42 ***	0.66 ***	-0.47 ***
	(0.05)	(0.06)	(0.04)	(0.05)
Academic/Vocational Streaming				
Austria	1.20 ***	-0.73 ***	0.61 ***	-0.68 ***
	(0.11)	(0.08)	(0.07)	(0.06)
Belgium-Flanders	0.98 ***	-0.97 ***	0.74 ***	-1.17 ***
	(0.07)	(0.07)	(0.03)	(0.04)
Belgium-French	0.74 ***	-0.87 ***	0.73 ***	-1.06 ***
	(0.12)	(0.11)	(0.06)	(0.07)
Czech Republic	1.22 ***	-1.03 ***	1.08 ***	-0.76 ***
	(0.08)	(0.11)	(0.05)	(0.08)
France	0.93 ***		1.10 ***	
	(0.08)		(0.09)	
Germany	1.24 ***	-0.61 ***	0.87 ***	-1.02 ***
	(0.07)	(0.09)	(0.05)	(0.08)
Greece	0.97 ***		1.01 ***	
	(0.07)		(0.06)	
Hungary	1.13 ***	-1.41 ***	0.61 ***	-0.86 ***
	(0.09)	(0.13)	(0.07)	(0.08)
Ireland	0.53 ***	-0.88 ***	0.21 **	-1.03 ***
	(0.11)	(0.21)	(0.07)	(0.19)
Italy	0.82 ***	-0.40 ***	0.33 ***	-0.63 ***
	(0.06)	(0.06)	(0.08)	(0.06)
Japan	0.75 ***		0.46 ***	
	(0.08)		(0.07)	
Korea	0.95 ***		0.99 ***	
	(0.07)		(0.06)	
Luxembourg	0.13 **	-0.61 ***	-0.15 ***	-1.41 ***
	(0.05)	(0.06)	(0.04)	(0.05)
Netherlands	0.67 ***	-0.78 ***	0.45 ***	-1.06 ***
	(0.10)	(0.09)	(0.03)	(0.06)
Portugal	0.60 ***		0.33 ***	
	(0.04)		(0.06)	
Slovak Republic	0.04	-0.98 ***	-0.03	-0.79 ***
	(0.08)	(0.08)	(0.08)	(0.09)
Switzerland	1.00 ***	-0.31 *	0.66 ***	-0.40 *
	(0.12)	(0.15)	(0.18)	(0.20)

Notes: Track selection model is based on a multinomial logistic regression predicting high, middle, and low track from socioeconomic status, with no other controls. Math achievement model is based on an ordinary least squares linear regression model predicting math achievement from dummy variables for high track and low track.
* $p<0.05$. ** $p<0.01$. *** $p<0.001$ (two-tailed tests).

Table 3. Coefficients from Hierarchical Generalized Linear Models Predicting Selection into High and Low Tracks

	Widespread Tracking Model 1		Full Sample Model 2
	High Track	Low Track	High Track
Student Level			
Grade	0.19 ***	0.17 ***	0.11 ***
	(0.02)	(0.03)	(0.02)
Female	0.21 ***	-0.22 ***	0.25 ***
	(0.03)	(0.02)	(0.02)
Ethnic/Language Minority	-0.12 ***	0.09 ***	-0.12 ***
	(0.02)	(0.03)	(0.02)
SES	0.93 ***	-0.61 ***	0.99 ***
	(0.06)	(0.06)	(0.06)
Country Level			
Course-by-Course Tracking	-0.30	-0.44	-1.40 **
	(0.20)	(0.24)	(0.44)
Course-by-Course Tracking × SES	-0.50 ***	0.22 *	-0.44 ***
	(0.07)	(0.08)	(0.09)
Intercept	-0.30	-0.48 *	0.45
	(0.15)	(0.18)	(0.28)
N (Students)	64,902		98,578
N (Countries)	11		20

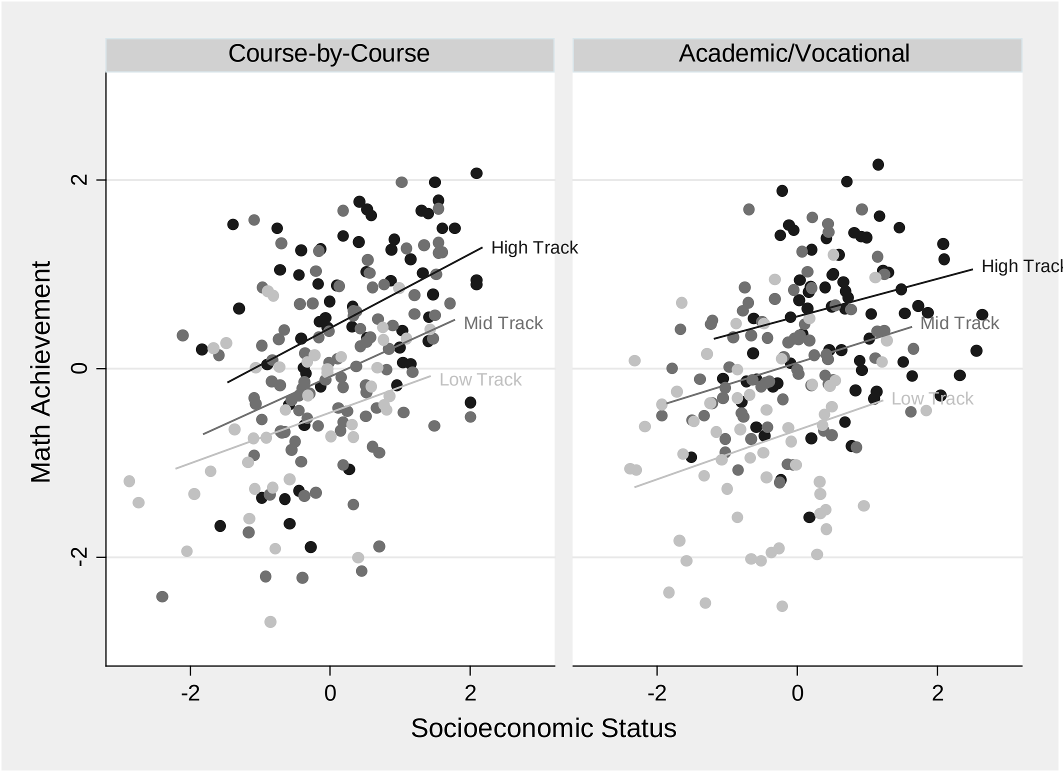
Notes: "Widespread Tracking" model (1) is based on a two-level hierarchical multinomial logistic model of students nested in countries. "Full Sample" model (2) is based on a two-level hierarchical logistic model of students nested in countries.
* p<0.05. ** p<0.01. *** p<0.001 (two-tailed tests).

Table 4. Coefficients from Hierarchical Linear Models Predicting Math Achievement

	Widespread Tracking - No School	Widespread Tracking - With School	Full Sample - With School
	Model 3	Model 4	Model 5
Student Level			
Grade	0.36 ***	0.38 ***	0.37 ***
	(0.01)	(0.02)	(0.02)
Female	-0.18 ***	-0.18 ***	-0.18 ***
	(0.01)	(0.02)	(0.02)
Ethnic/Language Minority	-0.36 ***	-0.16 ***	-0.18 ***
	(0.01)	(0.03)	(0.03)
High Track	0.50 ***	0.19	0.35 ***
	(0.07)	(0.10)	(0.08)
Low Track	-0.72 ***	-0.53 ***	
	(0.04)	(0.05)	
SES	0.16 ***	0.08 ***	0.09 ***
	(0.01)	(0.01)	(0.02)
School Level			
Mean Grade		0.30 ***	0.23 **
		(0.07)	(0.08)
Proportion Female		0.15	0.13
		(0.18)	(0.14)
Proportion Minority		-0.55 ***	-0.50 ***
		(0.13)	(0.15)
Mean SES		0.51 ***	0.67 ***
		(0.07)	(0.06)
Country Level			
Course-by-Course Tracking	-0.22	-0.17	-0.31
	(0.11)	(0.19)	(0.22)
Course-by-Course Tracking × High Track	0.02	0.18	0.12
	(0.09)	(0.19)	(0.19)
Course-by-Course Tracking × Low Track	0.30 ***	0.13	
	(0.06)	(0.10)	
Course-by-Course Tracking × SES	0.10 ***	0.09 *	0.09 *
	(0.02)	(0.04)	(0.04)
Intercept	0.05	-0.05	0.10
	(0.09)	(0.09)	(0.08)
Variance Explained (Students)	0.40	0.26	0.17
Variance Explained (Schools)		0.82	0.73
Variance Explained (Countries)	0.09	0.13	0.47
N (Students)	64,902	64,902	98,578
N (Schools)		2,300	3,426
N (Countries)	11	11	20

Notes: "No School" model (3) is based on a two-level hierarchical linear model of students nested in countries. "With School" models (4 and 5) are based on three-level hierarchical linear models of students nested in schools nested in countries.
* p<0.05. ** p<0.01. *** p<0.001 (two-tailed tests).

Figure 1. Achievement by SES and Track in Course-by-Course Tracking and Academic/Vocational Streaming, Widespread Tracking Countries



Notes: No controls. Math achievement and SES are country-mean centered to approximate random effects. Data points in figure are 200 randomly drawn (sample weighted) observations for each type of tracking.

Online Appendix

An International Comparison of Achievement Inequality in Within- and Between-School Tracking Systems

ANNA K. CHMIELEWSKI
Michigan State University

Online Appendix

Appendix A: More Information about Coding Math Course Tracking

There was some concern that the course sequence of Pre-algebra, Algebra I, Geometry, Algebra II, Pre-calculus may not be universal in American high schools. While the order of Pre-Algebra, Algebra I, and Pre-Calculus is clear, Geometry and Algebra II could logically be switched in the sequence. However, a review of recent literature on high school math course-taking suggests that Geometry preceding Algebra II is the far more common sequence (Bozick, Ingels and Owings 2008; McFarland 2006; Riegle-Crumb 2006; Schiller et al. 2010). Moreover, Geometry preceding Algebra II also appears to be the most common sequence in the American PISA sample. A close examination of course enrollment patterns by grade level in the 274 schools in the US sample revealed 101 schools where it is very clear that Geometry preceded Algebra II, two schools where it is very clear that Algebra II preceded Geometry, and 171 schools where the pattern is more ambiguous. If tracks are recoded most conservatively based on an Algebra II, Geometry sequence both in the two schools with the switched pattern and in the 171 ambiguous schools, then mean achievement and SES do not differ significantly between any tracks except for high track math achievement, which is somewhat lower under the more conservative scheme. The final decision was to recode the two schools with clear Algebra II, Geometry patterns as such and to leave the rest of the schools coded as Geometry, Algebra II, based on the assumption that ambiguity of enrollment patterns was attributable to sampling variation.

Five academic/vocational streaming countries also collected the math class variable, and one course-by-course tracking country (Australia) also collected the academic/vocational program variable. Additional analyses not reported here confirm that these secondary forms of tracking (math class in the academic/vocational streaming countries that collected this variable and academic/vocational streaming in Australia) appear less important empirically than the primary forms. Achievement gaps between math course tracks in academic/vocational streaming countries are much smaller than those between math course tracks in the three countries categorized as course-by-course tracking countries in this paper. This makes sense, as math course tracks within academic or vocational streams are located within an already narrowed achievement distribution. In Australia, less than 9% of students report being in the vocational stream, which is always located in the same school as the academic stream, and both academic and vocational students can be in either the high, middle, or low math track.

Data from course-by-course tracking systems were also examined to attempt to identify any possibly untracked schools. Principals reported whether their school groups students by ability for math classes that differed in content and/or difficulty for all, some, or no classes. Only a small proportion of principals report no math course tracking in the course-by-course tracking countries (Australia, Iceland, and the United States). As an additional check, student track reports were examined, and they were found to vary within most of these schools, suggesting that they are in fact tracked. In the United States and Australia, there are no schools where both student and principal reports are consistent with untracked math classes. In the United States, 2.4% of students attend schools where all students at the same grade level report taking the same math class, but in all of these schools, principals report at least some ability grouping for math classes. In 5.2% of schools in Iceland, principals report no tracking and all students report the same track or all are missing. These were recoded to “untracked.”

References

Bozick, Robert, Steven J. Ingels, and Jeffrey A. Owings. 2008. "Mathematics Coursetaking and Achievement at the End of High School: Evidence from the Education Longitudinal Study of 2002 (ELS:2002)." Washington, D.C.: US Department of Education.

McFarland, Daniel A. 2006. "Curricular Flows: Trajectories, Turning Points, and Assignment Criteria in High School Math Careers." *Sociology of Education* 79(3):177-205.

Riegle-Crumb, Catherine. 2006. "The Path through Math: Course Sequences and Academic Performance at the Intersection of Race-Ethnicity and Gender." *American Journal of Education* 113(1):101-22.

Schiller, Kathryn S., William H. Schmidt, Chandra Muller, and Richard T. Houang. 2010. "Hidden Disparities: How Courses and Curricula Shape Opportunities in Mathematics During High School." *Equity & Excellence in Education* 43(4):414-33.