

The Missing Manual: Using National Student Clearinghouse Data to Track Postsecondary Outcomes

Susan M. Dynarski

University of Michigan

Steven W. Hemelt

University of North Carolina at Chapel Hill

Joshua M. Hyman

University of Connecticut

This article explores the promises and pitfalls of using National Student Clearinghouse (NSC) data to measure a variety of postsecondary outcomes. We first describe the history of the NSC, the basic structure of its data, and recent research interest in using NSC data. Second, using information from the Integrated Postsecondary Education Data System (IPEDS), we calculate enrollment coverage rates for NSC data over time, by state, institution type, and demographic student subgroups. We find that coverage is highest among public institutions and lowest (but growing) among for-profit colleges. Across students, enrollment coverage is lower for minorities but similar for males and females. We also explore two potentially less salient sources of non-coverage: suppressed student records due to privacy laws and matching errors due to typographic inaccuracies in student names. To illustrate how this collection of measurement errors may affect estimates of the levels and gaps in postsecondary attendance and persistence, we perform several case-study analyses using administrative transcript data from Michigan public colleges. We close with a discussion of practical issues for program evaluators using NSC data.

Keywords: *econometric analysis, experimental research, research methodology, postsecondary education, quasi-experimental analysis*

Introduction

By many measures, those who attend and complete college are better off. They have higher earnings (Oreopoulos & Petronijevic, 2013), better health (Hout, 2011; Lleras-Muney, 2005), more robust networks, and higher job satisfaction (Oreopoulos & Salvanes, 2011). College attendance, persistence, and completion are therefore key outcomes by which researchers and policymakers evaluate the effectiveness of elementary and secondary schooling. Tracking students into and through college is critical for such evaluations. Yet measures of these outcomes have rarely been available at the levels of granularity at which

elementary and secondary education policies function: the school, the district, and the state.

The National Center for Education Statistics (NCES) proposed to address this data gap by replacing information in the Integrated Postsecondary Education Data System (IPEDS) with a unit-record system that would track individual students over time (Cunningham & Milam, 2005). In response to concerns about student privacy, and colleges' fears that the system would be used for accountability,¹ Congress effectively banned the creation of any national unit-record database in the 2008 reauthorization of the Higher Education Opportunity Act.

In the wake of the failure of the federal government to produce a dataset that tracks students into and through college, states have attempted to construct their own unit-record systems. Every state now has a longitudinal data system that tracks individual students through elementary and secondary schools (hereafter, referred to as “K–12”). Some state systems also follow students into in-state public tertiary institutions (e.g., Florida, Ohio); a few have also integrated in-state private institutions into their systems (e.g., Texas). Students who enroll in out-of-state colleges are generally lost to these systems.²

The National Student Clearinghouse (NSC), founded in 1993, has filled the void left by Congress’ ban and the gaps in states’ systems. NSC was originally tied to the loan industry and called the National Student Loan Clearinghouse. NSC gathered enrollment data from participating colleges; these data were used to confirm that a borrower was enrolled and therefore exempt from paying back (or accruing interest on) student loans. This history explains many of the strengths and weaknesses of NSC data. The NSC primarily tracks enrollment (but not credits or major), because that is what determines a student’s ability to defer payment. Although NSC has long recorded that a student had graduated from a given institution (often along with the type of degree earned), only recently (as of 2012) did it start to use Classification of Instructional Programs (CIP) codes³ to more systematically record a student’s major.⁴

The federal government has stimulated the market for NSC’s products by both imposing reporting requirements and funding states’ efforts to track their students into college. States receiving federal State Fiscal Stabilization Funds (SFSF) are required to report statistics on college enrollment and persistence rates.⁵ The Statewide Longitudinal Data Systems (SLDS) grant program (which operates within the Institute of Education Sciences of the U.S. Department of Education) funds the efforts of states to develop their data systems. SLDS has held competitions since 2005, with 47 states receiving a total of US\$612 million. Early rounds encouraged the linking of K–12 data to college outcomes, whereas the later rounds required it (Garcia & L’Orange, 2012). Many states have turned to the NSC for data to fill the gaps in their data systems.

NSC writes contracts on a state-by-state basis with local officials, with pricing a function of student enrollment.⁶

NSC data are relatively new to academic researchers and policymakers. A growing number of papers make use of NSC data for research purposes (e.g., Bettinger, Long, Oreopoulos, & Sanbonmatsu, 2012; Chingos & Peterson, 2012; Deming, Hastings, Kane, & Staiger, 2014; Dynarski, Hyman, & Schanzenbach, 2013; Goldrick-Rab, Harris, Kelchen, & Benson, 2012; Hemelt, Roth, & Eaton, 2013; Hyman, 2013; Kane, 2003; Richburg-Hayes *et al.*, 2009).⁷ These papers use NSC data to explore the effects of various programs or policies on postsecondary attendance, persistence, and attainment. A recent interchange between academics over the methods and data used to draw conclusions about the postsecondary impacts of randomly offering vouchers to high school students in New York City (Chingos & Peterson, 2012; Goldrick-Rab, 2012; Lederman, 2012) highlights the need for a more comprehensive and relevant source of evidence about potential measurement error in NSC data. As research using NSC data becomes even more common, researchers need to be aware of the benefits and challenges of working with these data.

This article examines the promises and pitfalls of NSC data for a variety of research and evaluation purposes. We first describe the basic process of obtaining data from the NSC and the structure of the data. Within the national context, we then quantify several sources of potential measurement error relevant to using NSC data to construct measures of postsecondary enrollment, persistence, and completion. To illustrate some ways in which these sources of measurement error may affect the work of researchers and policy analysts, we use administrative data from Michigan and perform several case-study analyses. Finally, we close with a discussion of salient issues for analysts using NSC data.

NSC Data Basics

The NSC was founded in 1993 with roots in the student loan industry. To this day, one of its primary services is to report on students receiving financial aid at American postsecondary institutions to both the education finance industry and the U.S. Department of Education. The

Clearinghouse supplies these basic reporting services to colleges and universities at no charge. Employers and high schools also use NSC's services to verify students' degrees and to track the postsecondary success of students after high school graduation. Academic researchers and policy analysts have also started using NSC data for a variety of purposes. The NSC now has a "Research Center" that was created in 2010 and seeks to better understand how NSC data can be used to support postsecondary research.

When researchers or policymakers obtain data from the NSC, they generally use the service called "StudentTracker." This service uses a proprietary algorithm to match a list of students supplied by the researcher to the NSC's detailed enrollment and degree information. This algorithm matches primarily on student name and date of birth. For those students who are found in NSC's database, a wealth of information is returned. Colleges submit enrollment data to the NSC several times each academic year, reporting the beginning and ending date that students are enrolled during each term, and at what intensity they are enrolled (i.e., part-time [PT] or full-time [FT]). Every college that reports enrollment information to the NSC is also supposed to report whether or not a student has earned a degree and the date the degree was earned.⁸ A subset of NSC colleges also participates in another NSC service called "DegreeVerify."⁹ For these colleges, NSC data also include the college major (e.g., English) and type (e.g., Bachelor of Arts) of all degrees earned by their students. Overall, the NSC provides a rich source of information on when and where students enroll, at what intensity, whether and when they earn a degree, and in many cases the major and type of degree earned.

Quantifying Postsecondary Outcomes in a Pre-NSC World

Prior to the NSC and state longitudinal data systems, it was very challenging for researchers and policymakers to follow students through K–12 education and into the postsecondary realm. Researchers mainly relied on nationally representative surveys that asked respondents to report information on educational attainment. Such survey data necessarily have a limited capacity to follow particular groups of students

over time. This restricts the ability of researchers and policy analysts to meaningfully evaluate state programs or policies of interest that affect a particular group or cohort of students in time, track long-term outcomes for children, and responsively steer state-, district-, or school-level policymaking. Indeed, recent work has employed NSC data to overcome such challenges to estimate the impacts of financial aid programs on student success in college (Bettinger et al., 2012; Kane, 2003), the relationships between school choice, vouchers, and postsecondary outcomes (Chingos & Peterson, 2012; Deming et al., 2014), as well as the long-run impacts of childhood educational programs or policies (Dynarski et al., 2013; Hemelt et al., 2013).

As a way to illustrate many of the challenges presented by large survey data, we perform an exercise in which we calculate college attendance rates for recent cohorts of Michigan high school students using such data sources. The sources for such an analysis are indeed limited, as it requires an annual survey with large and representative samples at the state level in which a respondent's state of high school graduation and current education level are recorded. The only surveys with sufficiently large state-level samples are the Current Population Survey (CPS) and American Community Survey (ACS). Prior to NSC, the NCES primarily used data from the CPS to calculate national college enrollment rates for high school graduates.¹⁰ Yet, neither is suitable for this type of state-level analysis, for several reasons.

First, neither the CPS nor the ACS records the state in which a person graduated from high school. In the case of the ACS, state of birth and state of residence are provided. In the CPS, only state of residence is provided.¹¹ State of residence is an imperfect measure of state of high school graduation, as young people migrate across state lines for both work and school. And, depending on whether the migration is for work or for school, CPS will record the state of residence as either the sending or receiving state. For a student who has migrated for college, the CPS considers the state of residence to be the location of the parental home. For a person who has migrated for work, the new state is the state of residence. Second, neither the CPS nor the ACS consistently records the year in which a person

earned a high school degree. The October CPS asks the question of recent high school graduates only, which does not allow the tracking of a given high school graduation cohort over time. Still, we can use some rules of thumb to try to identify a sample of high school students who graduated from a given state in a given year.

CPS

We begin with the October education supplement of the CPS from 2008 through 2010. In each wave, we are able to identify whether respondents graduated from high school in the current year or in a prior year. We use this information, along with respondents' ages, to proxy for different time periods over which these classes of high school graduates could have enrolled in college (i.e., by the first or second fall after high school graduation). So, to approximate "immediate college enrollment," we limit our sample to individuals ages 18 to 19 who identify Michigan as their state of residence and who are "current graduates" (Table 1, Panel A). As a way to extend the time period over which each class of graduates could have entered college, we use data from the subsequent CPS surveys, keep graduates from "prior years," and restrict the sample to a slightly older group of students (Table 1, Panel B). For example, to arrive at the estimates of college enrollment by the second fall after high school graduation (Panel B) for the class of 2008, we use the CPS 2009 survey wave and restrict our sample to those individuals who are "prior year (or earlier)" grads and who are 19 to 20 years old. Table 1 presents pooled estimates of college-going for the Michigan high school classes of 2008 and 2009.

ACS

The ACS is a yearly survey that was developed by the U.S. Census Bureau as a replacement for the long form on the decennial census. Thanks to the ACS, researchers and policymakers can now obtain information on social, economic, and housing characteristics of different populations on a yearly basis (instead of only every decade) at the county level. The ACS is sent to about 3 million people annually and, depending on the size of the county, 1-year, 3-year, and/or 5-year estimates of

the survey variables are available. The ACS produced test data from 2000 to 2002 and was fully implemented for the first time in 2005.¹²

The ACS allows us to increase our sample sizes representing the Michigan high school classes of 2008 and 2009. We again begin with ACS survey waves from 2008 through 2010 and restrict to individuals who identify Michigan as their state of residence.¹³ In the ACS data, we cannot identify whether a respondent is a "current year" graduate like we can identify in the CPS. Therefore, we identify high school graduates and use age to proxy for high school graduation year. We then use a combination of rising age restrictions and subsequent waves of ACS data, just as we did with the CPS, to operationalize different time periods over which respondents could have enrolled in higher education. The right half of Table 1 presents college-going estimates for recent Michigan high school graduates based on ACS data.

Estimates of College-Going Based on CPS and ACS Data: Michigan High School Classes of 2008 and 2009

On average, about 80% of the classes of 2008 and 2009 enroll in college by the second fall after graduating from high school (according to both the CPS and ACS). Although this aggregate estimate is quite steady across the two surveys, when we examine college enrollment among various demographic subgroups of interest, these estimates are based on quite limited numbers of respondents, making questions about gaps in college-going between groups difficult to answer. For example, only two Black students in Michigan responded to the 2008 and 2009 CPS surveys and no students of Hispanic ethnicity responded to the 2008 wave. Given such imprecision, any researcher or policymaker interested in estimating gaps between such groups in college enrollment rates is left with substantial uncertainty about their degree of accuracy and representativeness.

Given the larger sample sizes afforded by the ACS, estimates of gaps in college enrollment across student subgroups look quite different than those based on CPS data. This is not surprising, given the exceptionally small subgroup sample sizes we observe in the CPS data (at the state level). For example, according to the CPS, the

TABLE 1
College Attendance Rate of Michigan High School Classes of 2008 and 2009

	CPS			ACS		
	Proportion	<i>n</i>	<i>SE</i>	Proportion	<i>n</i>	<i>SE</i>
	(1)	(2)	(3)	(4)	(5)	(6)
A. Immediate enrollment						
Overall	.74	54	.06	.79	10,904	.004
Sex						
Female	.68	28	.12	.83	5,615	.01
Male	.81	26	.08	.75	5,289	.01
Race and ethnicity						
White	.80	44	.17	.80	8,974	.01
Black	.25	4	.23	.69	1,074	.01
Other	.67	6	.21	.78	856	.01
Hispanic	1.00	2	—	.62	471	.02
B. Enrollment by second fall after HS graduation						
Overall	.80	121	.04	.79	12,805	.004
Sex						
Female	.86	57	.07	.82	6,544	.01
Male	.75	64	.05	.75	6,261	.01
Race and ethnicity						
White	.82	96	.10	.80	10,571	.01
Black	.65	17	.12	.70	1,271	.01
Other	.86	7	.14	.80	963	.01
Hispanic	.67	9	.16	.67	526	.02

Note. The sample for columns (1) to (3) and (4) to (6) represents respondents from the CPS (October supplement) and ACS, respectively, whose state of residence at the time of the survey is Michigan and who we estimate to have graduated high school (HS) in 2008 or 2009. Year of HS graduation in the CPS is imputed using age, survey year, and approximate graduation year (current year vs. prior to survey). For example, Panel A, columns (1) to (3) use the 2008 and 2009 waves of the CPS and restrict to those individuals who are 18 to 19 years old and graduated HS during the same year as the survey. Panel B uses 2009 and 2010 CPS and restricts to 19- to 20-year-olds who graduated HS prior to the year of the survey. The ACS does not collect year of HS graduation data, so in columns (4) to (6) we identify HS graduates using only age and survey year: Panel A restricts to HS graduates in the 2008 and 2009 ACS who are 18 to 19 years old. ACS uses 3-year estimates. CPS = Current Population Survey; ACS = American Community Survey; HS = high school.

proportion of female students from the Michigan classes of 2008 and 2009 who attend college by the second fall after graduating from high school is 11 percentage points higher than the proportion of male students who do so. When using ACS data, this female–male gap shrinks to 7 percentage points. Similarly, the White–Black gap using CPS data is 17 percentage points, but only 10 percentage points when using ACS information.

This exercise makes clear the challenges faced by researchers and policy analysts interested in using frequent, detailed, and precise measures of postsecondary outcomes generated in close intervals to evaluate policies and make

meaningful, responsive contributions to policy-making. The use of NSC data has the potential to overcome such challenges. Yet, no data source is perfect. To help guide the growing use of NSC data in research and policy analysis, we turn to an exploration of potential sources of measurement error in this new and detailed source of comprehensive postsecondary information.

Potential Sources of Measurement Error in NSC Data

There are several main sources of measurement error in NSC data relevant to estimating

postsecondary outcomes for students. Each of these sources of error results in students who are enrolled in college to not appear in the NSC. Students who do not appear in the NSC data due to these reasons are indistinguishable from students who do not enroll in college. We describe the main sources of error, their magnitudes, and how their magnitudes vary across time and space, and by student and college characteristics.

Coverage Rates

Enrollment Coverage. The first and perhaps most salient way in which measurement error appears when using NSC data is that not all colleges report enrollment information to the NSC. Participation in NSC by colleges is voluntary, and has increased markedly over time. As of fall 2011, the NSC reports that they cover 93% of postsecondary enrollment. We calculate and report NSC coverage rates over time, by Census region, state, college type, and demographic subgroups of students. We explore the sensitivity of our calculations to different definitions of college enrollment.

We calculate the NSC coverage rate by dividing postsecondary enrollment at NSC-participating colleges by total enrollment at all U.S. colleges.¹⁴ In February 2012, we acquired a list from NSC of all participating colleges along with the date each institution started submitting data to the NSC.¹⁵ We use enrollment counts from the IPEDS, a federally generated database that lists every college, university, and technical or vocational school that participates in the federal financial aid programs (more than 6,000 institutions nationwide; NCES, 2010). We use fall enrollment counts from IPEDS and include a college as in the NSC for the fall of a given year if it joins by October 15th of that year.¹⁶

Figure 1 shows NSC coverage over time for all U.S. Title IV institutions.¹⁷ The solid line weights the coverage rate by total undergraduate enrollment, and shows that NSC coverage increased drastically in the mid-1990s through early 2000s from less than 20% to more than 80%. Since then, the rate of coverage growth has slowed, reaching 91.6% in 2011. Weighting the coverage rates by first-time (and first-year) full-time-equivalent (FTE) undergraduate enrollment (the dashed line) slightly lowers the coverage

rates.¹⁸ This suggests that coverage is lower at colleges in which there are a higher proportion of such students. For the remainder of this section, we focus on coverage rates calculated using total non-FTE-weighted undergraduate enrollment, as undergraduates (regardless of whether first time and first year) are the population of interest for the majority of analyses (and as weighting by FTE is inconsequential).

We first explore how NSC enrollment coverage varies across space within the United States. It is important for researchers and policymakers to understand how coverage varies across states and Census regions as analyses of particular programs or policies often focus on students from a certain state or region, and as students generally tend to enroll in college close to home.

In Figure 2, we illustrate how NSC enrollment coverage differs across Census regions. Since the early 2000s, coverage has been similar across all regions but the South. Enrollment coverage in the South continues to be a few percentage points lower than in other regions of the country. When we look across states, there is substantially more variation in coverage. Table 2 reports coverage rates over time by state for all Title IV institutions. Wyoming and Virginia are currently the states with the highest coverage rates at 99.6% and 97.4%, respectively. West Virginia and Louisiana bottom out the list, at 64.5% and 81.9%, respectively. Different states have experienced shifts in coverage over this time period: For example, between 1997 and 2011, New Mexico's coverage rate catapulted from 10% to 90%. Yet during the same time period, North Dakota saw a 3 percentage point decline in its overall coverage rate. For a complete set of coverage rates by college type (2-year/4-year; public/private; non-profit/for-profit) and state for all years 1995–2011, see the online set of appendix tables (Tables A1–A6; see the Online Appendix, available at <http://epa.sagepub.com/supplemental>).

NSC data also vary in the degree to which they capture enrollment at different types of colleges. In Figure 3, we observe that coverage rates are quite high for public institutions and private, non-profit (4-year) institutions, but much lower among for-profit institutions of any level. Coverage is more than 99% in 2011 among public 4-year institutions, but only 48% at for-profit schools. Although enrollment in for-profit institutions represents a

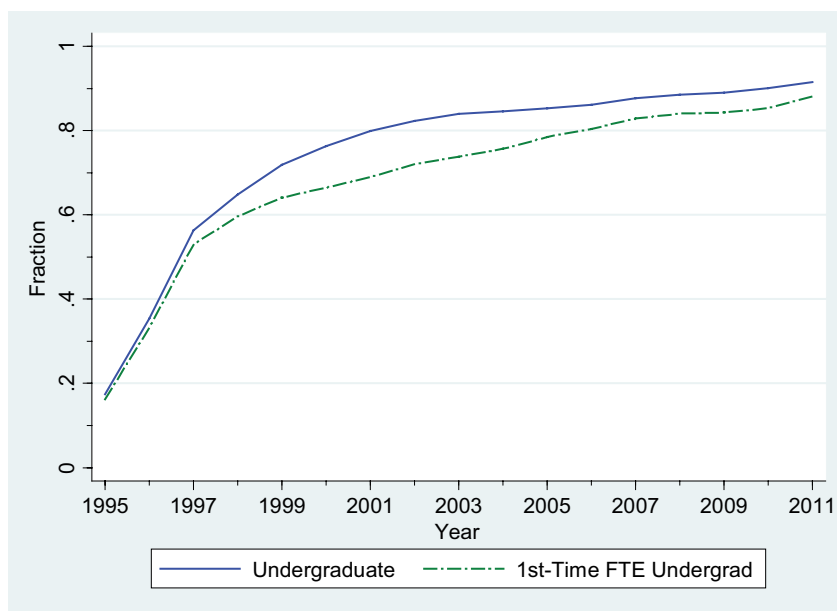


FIGURE 1. NSC enrollment coverage by enrollment measure.

Note. The underlying sample is all U.S. Title IV postsecondary institutions. Unless otherwise noted, trends are weighted by total undergraduate enrollment. Year corresponds to the fall semester. Enrollment and student characteristics are obtained from IPEDS. An institution is counted as a NSC participant if it joins by October 15 of a given year. Information on institution selectivity comes from 2004 Barron's ranking data: An institution is counted as a "selective" institution if it received a Barron's ranking of "very competitive," "highly competitive," or "most competitive." Collectively, this group of institutions admits about half or less of their applicants. IPEDS = Integrated Postsecondary Education Data System; NSC = National Student Clearinghouse; FTE = full-time equivalent.

relatively small share of total postsecondary enrollment, it has grown drastically over the last decade—from 3% to 9% of U.S. undergraduate enrollment (Ackerman, Cronin, Turner, & Bershadker, 2011). NSC coverage among for-profits has more than doubled since 2000, but improving coverage in this sector remains an ongoing challenge for the NSC. Figure 4 illustrates a persistent gap in coverage by institutional selectivity: Coverage is consistently higher among selective institutions (i.e., those admitting half or less of applicants).¹⁹ This finding largely mirrors the fact that coverage is highest at 4-year institutions of all types and lowest among the largely non-selective for-profit sectors.

To get a sense of how NSC-participating institutions, and their enrollees, compare with postsecondary institutions in the United States more broadly, we merge IPEDS data with the NSC participant list and present information on institutional and student body composition across the higher education landscape in Table 3. This exercise is similar in spirit to Figure 3, except that

here we also present the student-weighted fraction of total U.S. enrollment represented by each type of college (column 1). We confirm that although NSC coverage is low at for-profit institutions, enrollment at such institutions accounts for a relatively small share of all undergraduate enrollment (i.e., around 10%). Coverage at private non-profit 2-year institutions is also low, but these institutions account for virtually 0% of all undergraduate postsecondary enrollments (and we omit this sector from the table).

Table 3 also allows us to explore, on average, how NSC enrollment coverage varies by student characteristics. Coverage among males is about 1.5 percentage points higher than among females. Figure 5 shows that this small sex gap in coverage has existed since the early 2000s. At present, postsecondary enrollment among Black students is less well covered than enrollment among White students. To a lesser extent, this is also true for Hispanic students. The coverage rate among Black students at all Title IV institutions is 88% compared with 93% among White students.

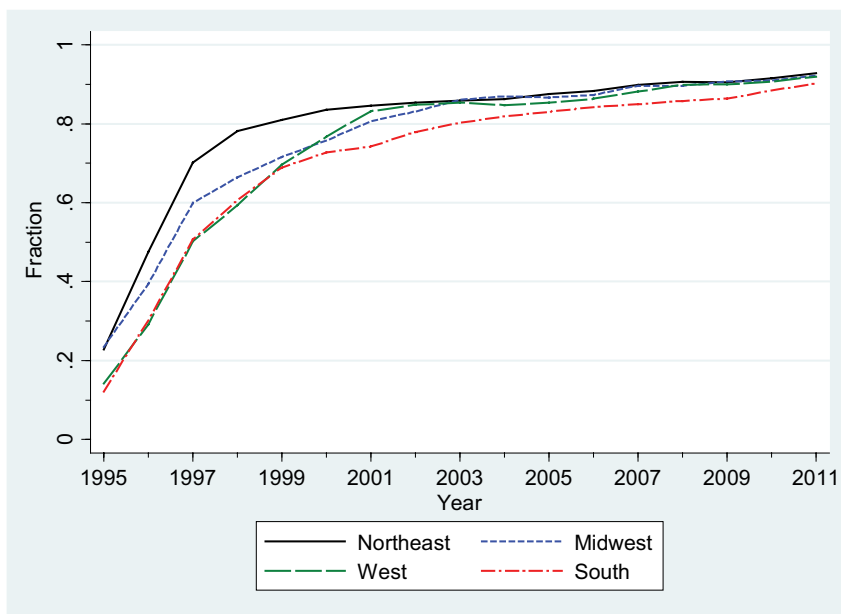


FIGURE 2. NSC enrollment coverage by geographic region.

Note. The underlying sample is all U.S. Title IV postsecondary institutions. Unless otherwise noted, trends are weighted by total undergraduate enrollment. Year corresponds to the fall semester. Enrollment and student characteristics are obtained from IPEDS. An institution is counted as a NSC participant if it joins by October 15 of a given year. IPEDS = Integrated Postsecondary Education Data System; NSC = National Student Clearinghouse.

Figure 6 shows that coverage has historically been similar for Whites and students grouped into an “other race” category,²⁰ but lower for minority students. Since the early 2000s, this White-minority coverage gap has remained steady at between 4 and 6 percentage points. Very recently, coverage has increased slightly among Hispanic enrollees. Since the early 2000s, the race gap in coverage has been more pronounced than the sex gap, with White males and females having similarly higher coverage than minority males and females.

We now take a more detailed look at gaps in enrollment coverage based on the most recent year in our data (2011–2012) presented in Table 4. We report shares of White non-Hispanic, Black non-Hispanic, and Hispanic students by region who attend a particular type of postsecondary institution, alongside subgroup-region-specific coverage rates for each institution type. These estimates allow us to better contextualize the gaps in coverage we present in Table 3, as college enrollees with different demographic characteristics are not equally distributed across geographic regions and types of postsecondary institutions.

As an illustration, we find in Table 3 an overall White–Black gap in coverage of 5.4 percentage points. We observe a White–Black gap of a similar magnitude (5.8 percentage points) within the private, non-profit 4-year sector. Yet, we also observe that college enrollment of Black students accounts for a much higher share of total undergraduate enrollment in for-profit institutions (across regions) than it does for any other sector: The top rows of Table 4 show that Blacks comprise 14% of undergraduates at any college (column 2), but 23% of students at for-profits (column 14). So, the relative concentration of Black students in poorly covered for-profit schools (especially in the South) and the relatively poor coverage of the small share of Black students at 4-year private institutions appear to account for the overall White–Black gap in coverage.

Degree Coverage. Colleges that participate in the NSC are supposed to report whether a student graduates and their date of graduation for every student on whom they report enrollment information. We have performed degree coverage rate calculations analogous to our enrollment calculations above, simply weighting by

TABLE 2

NSC Postsecondary Enrollment Coverage Rates: All U.S. Title IV Institutions by State and Year

State	1995	1997	1999	2001	2003	2005	2007	2009	2011
AL	.015	.525	.701	.716	.722	.789	.868	.884	.894
AK	.000	.000	.930	.929	.953	.949	.936	.915	.910
AZ	.000	.119	.217	.383	.450	.550	.705	.867	.892
AR	.000	.444	.720	.773	.778	.860	.852	.921	.921
CA	.181	.498	.755	.885	.909	.898	.914	.909	.932
CO	.044	.752	.834	.801	.825	.836	.791	.798	.823
CT	.569	.743	.806	.801	.835	.823	.840	.859	.844
DE	.000	.010	.542	.614	.641	.640	.618	.952	.938
DC	.055	.352	.553	.814	.826	.835	.926	.962	.964
FL	.003	.206	.570	.701	.779	.823	.825	.807	.838
GA	.246	.533	.643	.658	.872	.899	.909	.902	.948
HI	.008	.028	.073	.791	.823	.820	.864	.879	.888
ID	.151	.628	.620	.933	.965	.963	.965	.939	.954
IL	.252	.696	.778	.910	.941	.892	.919	.907	.926
IN	.000	.149	.307	.372	.569	.594	.822	.918	.938
IA	.000	.517	.612	.872	.887	.863	.833	.764	.838
KS	.000	.273	.379	.656	.809	.834	.842	.855	.863
KY	.004	.823	.817	.814	.828	.907	.948	.934	.943
LA	.294	.584	.737	.701	.754	.795	.800	.765	.819
ME	.000	.712	.808	.826	.838	.916	.931	.938	.944
MD	.027	.422	.729	.757	.900	.896	.904	.913	.936
MA	.116	.579	.732	.827	.836	.870	.933	.942	.951
MI	.093	.324	.559	.701	.764	.830	.844	.943	.950
MN	.127	.930	.942	.933	.945	.952	.952	.946	.950
MS	.079	.507	.794	.805	.823	.846	.929	.937	.953
MO	.514	.695	.808	.860	.906	.909	.922	.903	.901
MT	.000	.823	.896	.908	.914	.915	.919	.917	.924
NE	.356	.761	.846	.935	.958	.959	.959	.954	.968
NV	.000	.315	.692	.877	.968	.944	.933	.913	.915
NH	.459	.655	.709	.680	.705	.931	.937	.946	.962
NJ	.017	.509	.724	.807	.840	.852	.862	.878	.926
NM	.000	.103	.387	.620	.626	.634	.834	.913	.902
NY	.250	.784	.851	.865	.862	.869	.892	.902	.933
NC	.189	.825	.902	.912	.919	.923	.960	.964	.958
ND	.000	.956	.947	.957	.936	.919	.915	.902	.924
OH	.421	.695	.835	.856	.884	.896	.919	.906	.919
OK	.081	.331	.554	.646	.639	.640	.646	.742	.838
OR	.071	.603	.688	.787	.843	.897	.896	.901	.953
PA	.345	.792	.874	.880	.895	.902	.907	.903	.920
RI	.000	.302	.600	.804	.905	.889	.950	.954	.955
SC	.124	.803	.886	.929	.950	.947	.955	.955	.956
SD	.000	.807	.830	.826	.809	.820	.830	.826	.862
TN	.281	.828	.844	.878	.873	.858	.850	.850	.858
TX	.053	.411	.543	.620	.711	.748	.765	.810	.921
UT	.504	.757	.785	.938	.935	.928	.913	.919	.907

(continued)

TABLE 2 (CONTINUED)

State	1995	1997	1999	2001	2003	2005	2007	2009	2011
VA	.000	.866	.893	.897	.900	.897	.969	.970	.974
VT	.481	.741	.891	.913	.926	.933	.953	.950	.959
WA	.112	.727	.824	.942	.949	.950	.965	.955	.964
WV	.000	.288	.599	.688	.710	.697	.693	.722	.645
WI	.457	.862	.887	.947	.967	.973	.967	.964	.965
WY	.000	.622	.958	.942	.931	.916	.951	.958	.996

Note. NSC = National Student Clearinghouse.

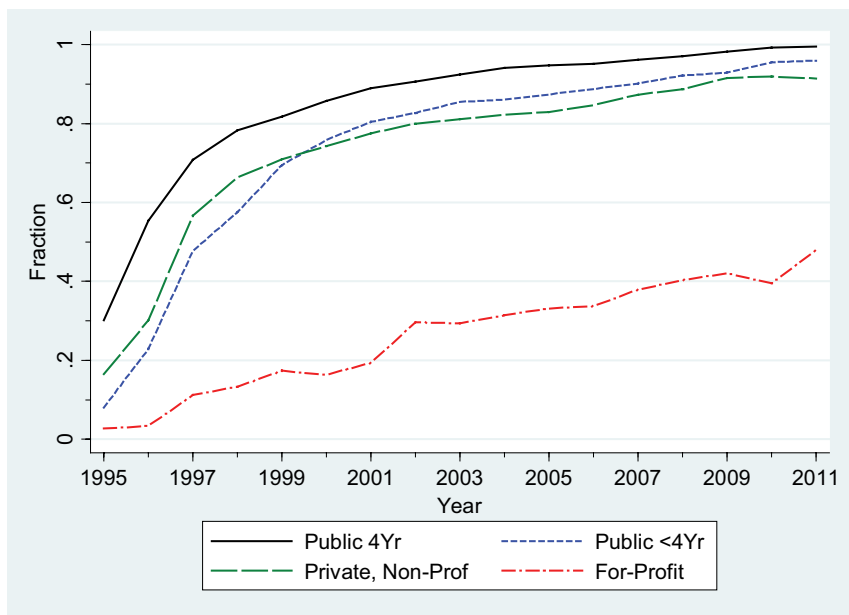


FIGURE 3. NSC enrollment coverage by institutional sector.

Note. The underlying sample is all U.S. Title IV postsecondary institutions. Unless otherwise noted, trends are weighted by total undergraduate enrollment. Year corresponds to the fall semester. Enrollment and student characteristics are obtained from IPEDS. An institution is counted as a NSC participant if it joins by October 15 of a given year. IPEDS = Integrated Postsecondary Education Data System; NSC = National Student Clearinghouse.

the number of degrees awarded instead of undergraduate enrollment counts. However, unlike with enrollment, it is not necessarily reasonable to assume that NSC-participating institutions report degree information for all of their students.²¹ Thus, these calculations may provide an overestimate of the true rates of degree coverage in NSC data. For this reason, we use state administrative data in Michigan to more carefully speak to rates of degree coverage in NSC data.

In Table 5, we descriptively explore this concern by counting the number of undergraduate

degrees (i.e., associate's and bachelor's) awarded in our NSC data to Michigan college-goers during the 2011–2012 academic year (column 2) and compare these counts with the number of degrees awarded among Michigan postsecondary institutions as reported in the IPEDS in the same year (column 1). We focus on undergraduate degrees: We exclude degrees where the title indicates the degree is at the master's or doctoral level as well as degrees with a missing title earned by students who previously earned a degree from a 4-year school.²² We also exclude degrees with a title that indicates a certificate or diploma. Yet, in our NSC

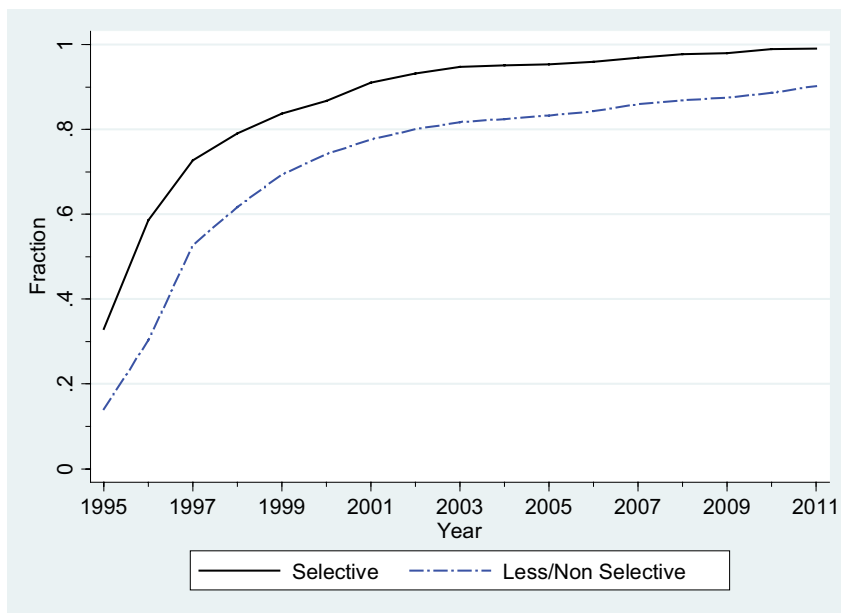


FIGURE 4. NSC enrollment coverage by institutional selectivity.

Note. The underlying sample is all U.S. Title IV postsecondary institutions. Unless otherwise noted, trends are weighted by total undergraduate enrollment. Year corresponds to the fall semester. Enrollment and student characteristics are obtained from IPEDS. An institution is counted as a NSC participant if it joins by October 15 of a given year. Information on institution selectivity comes from 2004 Barron's ranking data: An institution is counted as a "selective" institution if it received a Barron's ranking of "very competitive," "highly competitive," or "most competitive." Collectively, this group of institutions admits about half or less of their applicants. IPEDS = Integrated Postsecondary Education Data System; NSC = National Student Clearinghouse.

data, 16% of degrees have a missing degree title. We exclude (include) all such degrees to form the lower (upper) bound on NSC degrees. To arrive at an informed middle ground, we calculate the fraction of total degrees that are certificates by sector in Michigan using the IPEDS data, assume that this is the share of the degrees with missing degree titles in our NSC data by sector that are certificates or diplomas, and exclude only that number of the degrees with missing titles.²³ Columns 3, 4, and 5 report the raw degree coverage rates for our "best approximation," lower, and upper bound degree counts.²⁴

As IPEDS includes degrees earned in Michigan regardless of where students attended high school, and our sample of NSC students only includes students who attended public high school in Michigan, the degree coverage rate will be artificially low: Students who earn a degree in Michigan but attended a private high school in Michigan or a high school outside of Michigan will be counted in the IPEDS data but not in our NSC data.²⁵ Therefore, we create a weighted degree coverage rate—weighting the raw coverage rate by the

proportion of total undergraduate enrollment in Michigan that we observe in our NSC data. The resulting weighted degree coverage rates appear in columns 8–10 and can be interpreted as the rate of degree coverage conditional on being observed as enrolling in the NSC data.²⁶

On average, we observe roughly 81% of all degrees earned conditional on enrollment in a NSC-participating institution. Degree coverage is highest at public colleges (82%) and lowest at private, non-profit institutions (62%). Given the number of moving parts in this exercise, readers should interpret these degree coverage rates with care. At minimum, these estimates suggest that NSC-participating institutions in Michigan may not always report degree completion for all undergraduates. Analysts using NSC data in other states and regions to examine degree receipt should be aware of this potential drawback.

Matching Errors

A second way in which measurement error may appear when using NSC data to estimate

TABLE 3
How Representative Are NSC-Participating Colleges?

	U.S. colleges			U.S. colleges in NSC			Student-weighted coverage rate
	Proportion of students	<i>n</i> (colleges)	<i>n</i> (students)	Proportion of students	<i>n</i> (colleges)	<i>n</i> (students)	
	(1)	(2)	(3)	(4)	(5)	(6)	
All colleges	1.00	6,311	18,440,536	1.00	3,121	16,887,096	.916
Public							
2-year	.39	1,267	7,161,172	.41	925	6,871,474	.960
4-year	.36	666	6,626,325	.39	647	6,592,742	.995
Private, non-profit							
4-year	.15	1,303	2,675,633	.15	976	2,490,514	.931
Private, for-profit	.10	2,828	1,911,972	.05	531	915,839	.479
Student demographics							
Race/ethnicity							
White, non-Hispanic	.55		10,103,260	.56		9,404,816	.931
Black, non-Hispanic	.14		2,588,472	.13		2,269,436	.877
Hispanic	.14		2,607,518	.14		2,337,389	.896
Sex							
Female	.57		10,491,909	.56		9,539,568	.909
Male	.43		7,948,628	.44		7,347,529	.924

Note. Year corresponds to the fall semester. Coverage rates are calculated by dividing the total undergraduate enrollment in postsecondary institutions participating in the NSC by the total undergraduate enrollment in all postsecondary institutions. Undergraduate enrollment counts are obtained from IPEDS. A school is counted as in the NSC for that year if it joins by October 15th. See appendix tables for rates for all years 1995–2011 by state and college type. NSC = National Student Clearinghouse; IPEDS = Integrated Postsecondary Education Data System.

postsecondary outcomes for students is simply due to the matching process used by the NSC. The NSC considers their matching algorithm proprietary; but, from communication with NSC staff, we know that this process tends to err on the side of “false negatives.” We also know that the primary information used to match students to their postsecondary records is name and date of birth. In our Michigan case-study below, we test the robustness of this matching algorithm to a variety of typos or inconsistencies in student names.

Suppressed Student-Level Information

A final way in which NSC data may not fully capture information on the postsecondary experiences of a desired sample of students has to do

with the Family Educational Rights and Privacy Act (FERPA). FERPA is a federal law that protects the privacy of students’ educational records. The law applies to all schools that receive funds from the U.S. Department of Education. Under FERPA, both students and schools can block their enrollment and degree information.²⁷ So, NSC cannot release student-level information if the record is “FERPA-blocked.”

The NSC has published a report that provides state-specific statistics on the overall number of blocked records processed by NSC in 2006–2007, 2008–2009, and 2010–2011 (NSC, 2012). Nationally, the NSC finds FERPA-blocking to be most prevalent among students at 2-year institutions (both public and private). Within this 2-year group, student records at

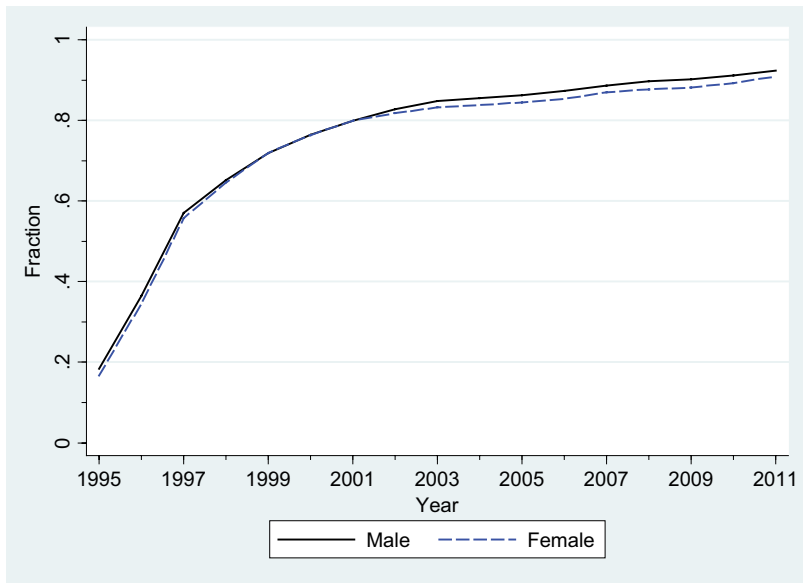


FIGURE 5. *NSC enrollment coverage by student sex.*

Note. The underlying sample is all U.S. Title IV postsecondary institutions. Unless otherwise noted, trends are weighted by total undergraduate enrollment. Year corresponds to the fall semester. Enrollment and student characteristics are obtained from IPEDS. An institution is counted as a NSC participant if it joins by October 15 of a given year. IPEDS = Integrated Postsecondary Education Data System; NSC = National Student Clearinghouse.

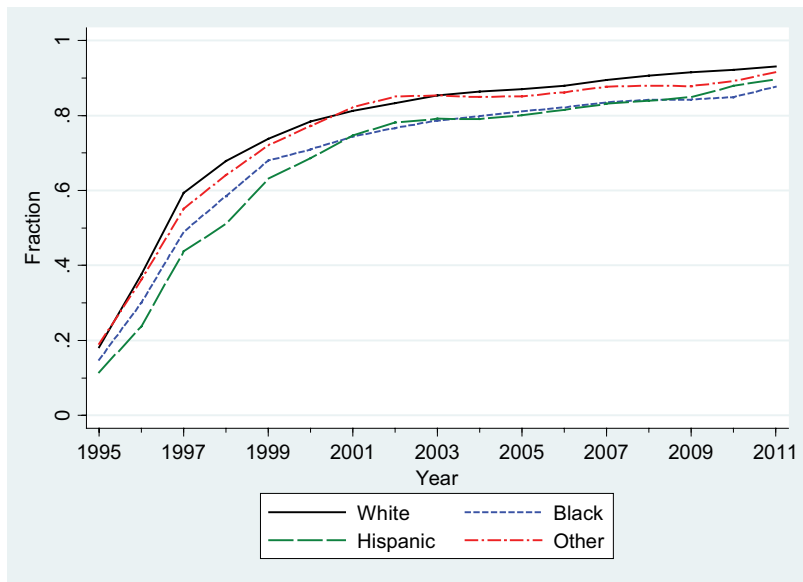


FIGURE 6. *NSC enrollment coverage by student race and ethnicity.*

Note. The underlying sample is all U.S. Title IV postsecondary institutions. Unless otherwise noted, trends are weighted by total undergraduate enrollment. Year corresponds to the fall semester. Enrollment and student characteristics are obtained from IPEDS. An institution is counted as a NSC participant if it joins by October 15 of a given year. IPEDS = Integrated Postsecondary Education Data System; NSC = National Student Clearinghouse.

TABLE 4
Distribution of College Enrollment and NSC Coverage: Where and for Whom Is Coverage Best?

Institutional sector	Any college			Public, 2-year			Public, 4-year			Private, non-profit, 4-year			For-profit		
	White, NH	Black, NH	Hispanic	White, NH	Black, NH	Hispanic	White, NH	Black, NH	Hispanic	White, NH	Black, NH	Hispanic	White, NH	Black, NH	Hispanic
Student subgroup	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Entire United States															
Proportion of students	.548	.140	.141	.521	.146	.178	.599	.116	.122	.604	.119	.079	.395	.230	.158
Coverage rate	.931	.877	.896	.959	.955	.954	.996	.994	.996	.938	.880	.918	.459	.507	.399
Northeast															
Proportion of students	.572	.122	.115	.546	.147	.150	.616	.104	.101	.592	.084	.080	.413	.240	.181
Coverage rate	.939	.880	.901	.991	.992	.997	.992	.996	.994	.957	.973	.962	.305	.330	.316
South															
Proportion of students	.527	.216	.137	.532	.229	.137	.557	.182	.144	.535	.221	.081	.349	.323	.166
Coverage rate	.921	.875	.880	.932	.932	.929	.995	.993	.998	.893	.813	.853	.339	.450	.248
Midwest															
Proportion of students	.687	.122	.058	.673	.139	.075	.751	.079	.040	.702	.096	.054	.487	.255	.066
Coverage rate	.939	.875	.921	.973	.983	.985	1.000	1.000	1.000	.949	.928	.923	.431	.502	.497
West															
Proportion of students	.439	.067	.237	.401	.067	.298	.507	.037	.181	.572	.045	.129	.370	.141	.199
Coverage rate	.928	.885	.902	.963	.963	.952	.994	.993	.992	.948	.895	.925	.624	.711	.500

Note. Enrollment shares and coverage rates are presented for 2011–2012. Race categories not shown: Asian, American Indian, multi-racial, and unknown. NH = non-Hispanic; NSC = National Student Clearinghouse.

TABLE 5
NSC Postsecondary Degree Coverage in Michigan

	Number of undergraduate degrees				Coverage	Coverage (lower bound)	Coverage (upper bound)	Weighted coverage	Weighted coverage (lower bound)	Weighted coverage (upper bound)
	IPEDS	NSC	NSC (lower bound)	NSC (upper bound)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
All colleges	94,459	36,937	32,013	38,299	.391	.339	.405	.811	.703	.841
Public	73,205	32,016	29,823	32,411	.437	.407	.443	.821	.765	.832
2-year	25,886	8,956	7,578	9,628	.346	.293	.372	.686	.581	.738
4-year	47,319	22,777	22,245	22,783	.481	.470	.481	.744	.727	.744
Private, non-profit	19,419	4,887	2,189	5,090	.252	.113	.262	.624	.280	.650
4-year	19,419	4,887	2,189	5,090	.252	.113	.262	.622	.279	.648
For-profit	1,835	111	1	798	.060	.001	.435	.877	.008	6.319

Note. Data are presented for 2011–2012. Degrees in IPEDS are associate’s and bachelor’s degrees awarded at all institutions in Michigan. Degrees in NSC are all associate’s and bachelor’s degrees earned by students who attended (public) high school and college in Michigan. See the text for discussion of raw and weighted degree coverage rates. NSC = National Student Clearinghouse; IPEDS = Integrated Postsecondary Education Data System.

for-profits experienced the lowest block rate, while students at 2-year publics blocked the most (NSC, 2012; Figure 1).

Records of students with different characteristics are blocked at different rates. More than 7% of Asian/Pacific Islanders’ records, 5% of Hispanic students’ records, and only about 3% of White and Black students’ records were blocked (NSC, 2012; Figure 2). By their nature, FERPA blocks are unobservable to researchers. Upon receipt of the NSC data, there is no way to know the attributes of students who blocked their records. The blocking patterns detailed by the NSC have different implications for states and regions, depending on the composition of the postsecondary system as well as the relative size of each student subpopulation. We explore the implications of FERPA-blocking in Michigan in our case-study analysis below.

Comprehensive Coverage Rates

Given the three sources of measurement error, we find two to be the most crucial: coverage and FERPA-blocking. We tie these two sources of measurement error together and create a “com-

prehensive coverage rate.” To arrive at this rate, we take our initial enrollment coverage rate and multiply it by 1 minus the FERPA block rate.²⁸ This procedure down weights the original coverage rate by the likely share of blocked records. Thus, this figure represents the fraction of enrollment at NSC-participating colleges observable by researchers at the detailed student level. This is the relevant statistic for researchers and policymakers who want to know what fraction of undergraduate college enrollment is captured by their NSC data.

The comprehensive coverage rate increases from 82.1% in 2006–2007 to 86.1% in 2010–2011. This comprehensive coverage rate varies by sector, with for-profit institutions experiencing the lowest coverage. Although private, non-profit 2-year institutions also have low comprehensive coverage, recall that such institutions account for virtually 0% of total undergraduate enrollment (see Table 3). We provide comprehensive coverage rates by state and sector in 2006–2007, 2008–2009, and 2010–2011 in Online Appendix Table A7 (see the Online Appendix, available at <http://epa.sagepub.com/supplemental>).

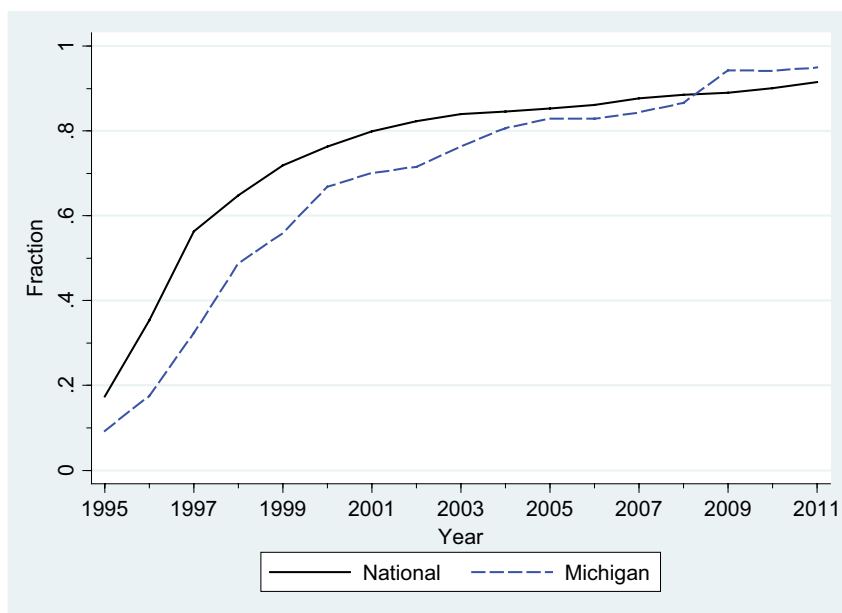


FIGURE 7. *NSC enrollment coverage in Michigan.*

Note. The underlying sample is all U.S. Title IV postsecondary institutions. Unless otherwise noted, trends are weighted by total undergraduate enrollment. Year corresponds to the fall semester. Enrollment and student characteristics are obtained from IPEDS. An institution is counted as a NSC participant if it joins by October 15 of a given year. IPEDS = Integrated Postsecondary Education Data System; NSC = National Student Clearinghouse.

Michigan Case Study: Estimates of Postsecondary Outcomes Using NSC Data

As a way to illustrate how analysts and researchers can use the estimates provided in this article in their own work, and how these potential sources of measurement error affect estimates of key postsecondary parameters, we turn to state administrative data on all public colleges and universities in Michigan.

Coverage Rates in Michigan

As a first-order exploration, we present enrollment coverage rates overall and by institutional sector for Michigan in Figures 7 and 8. In Figure 7, we observe that enrollment coverage in Michigan follows a similar pattern to the national trend, rising 4 percentage points above the national average by 2011. As is the case nationally, for-profits have the poorest coverage in Michigan (Figure 8). Prior to 2009, a few key 4-year private colleges in Michigan were also driving down the aggregate coverage rate, but joined the NSC in 2009. These colleges explain the sharp uptick in coverage between the fall of 2008 and the fall of 2009.

NSC's Matching Algorithm and Michigan Data

Even though the NSC's matching algorithm is proprietary, it is important for researchers to have a sense of how well this process functions. As a way to test the robustness of the NSC's matching algorithm to typos or inconsistencies in student names, we introduce slight variations into the names of a subset of students ($n = 10,000$) that we previously observed matching to the NSC database. We then resubmit these name variants to the NSC and examine which variants match and which do not.²⁹ The name variants we chose were either typographical errors likely to occur in administrative data, such as adding or mistyping a character; or inconsistencies likely to occur when entering a student's name at different times in different settings, such as using a nickname instead of a full name, and using (or excluding) a suffix.

From this process, we learned a few things that may improve match rates for others using NSC data: (a) Adding a letter to a first name is much more harmful to the overall match rate than deleting a letter; (b) the NSC matching algorithm seems to be very good at handling odd

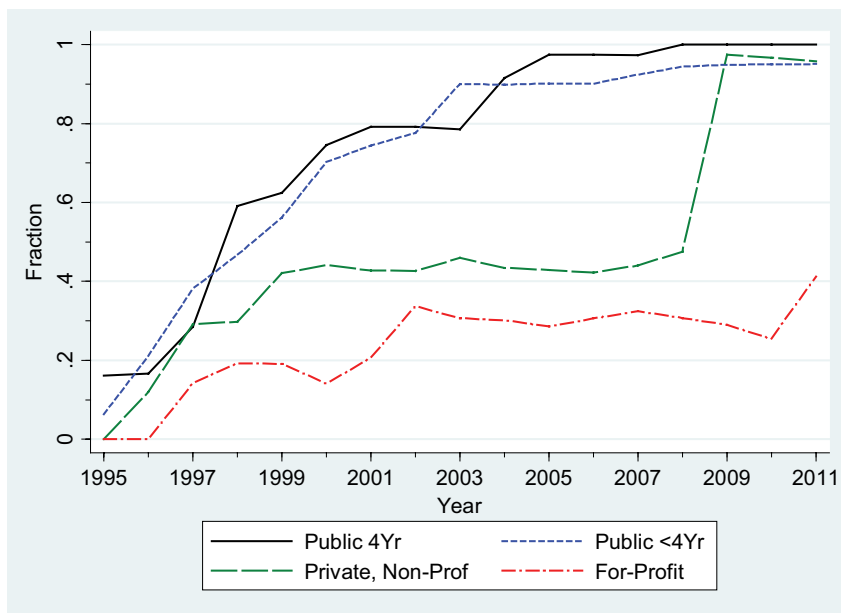


FIGURE 8. *NSC enrollment coverage in Michigan by institutional sector.*

Note. The underlying sample is all U.S. Title IV postsecondary institutions. Unless otherwise noted, trends are weighted by total undergraduate enrollment. Year corresponds to the fall semester. Enrollment and student characteristics are obtained from IPEDS. An institution is counted as a NSC participant if it joins by October 15 of a given year. NSC = National Student Clearinghouse; IPEDS = Integrated Postsecondary Education Data System.

non-alpha-numeric characters in names (e.g., “%”). Appending these types of characters to the end of first names does little to the overall match rate; (c) resubmitting names using popular first-name variants for students who did not initially match (e.g., “Chris” for “Christopher”) is likely to increase the overall match rate by about 1 percentage point; and (d) for students who have a suffix that goes along with their name, it is crucial that this suffix be placed in a field separate from the student’s first and last names. At least in our state data, suffixes often appeared as characters appended to last names—and we found it important that these suffixes be removed and placed in the separate field that NSC has for suffixes. Yet, students with suffixes represented a very small portion of our random sample (0.3%); therefore, the improvement in the overall match rate was minimal.

FERPA-Blocking in Michigan

To illustrate the consequences of the fact that FERPA-blocking rates are time-sensitive, we use NSC data from Michigan over the course of several different submissions to the NSC to estimate

rates of college enrollment. These submission time points span roughly 3 years. In Table 6, we estimate the rates of college enrollment for the Michigan high school class of 2009 by a constant point in time (i.e., the fall of 2010) separately based on three different pulls of NSC data. Therefore, as the underlying group of individuals submitted to the NSC did not change across time points and there were no changes in the stock of NSC-participating institutions, any shifts in the estimated enrollment rates can be attributed to changes in FERPA-blocking behavior.

Generally, our estimates of college enrollment for the class of 2009 are quite stable and trend downward only slightly (mostly between 0 and 3 percentage points) over time. As a way to capture all relevant student-level information, we then combine our NSC data across pulls. This allows us to observe student-level data for those students who may have attended a school that subsequently FERPA-blocked their information (or for students who blocked their information during an earlier pull only to unblock it later). Naturally, this also allows us to capture updated enrollment information for all students. In the final column of Table 6, we present pooled

TABLE 6

How Does Time-Sensitive FERPA-Blocking Affect College Enrollment Estimates?

	College enrollment rate based on submission to NSC in . . .			
	September 2010	August 2011	August 2012	Union of submissions
	(1)	(2)	(3)	(4)
Overall	.66	.64	.64	.67
Institution type				
Public, 2-year	.36	.34	.34	.37
Public, 4-year	.29	.29	.28	.29
Private, 4-year	.08	.08	.07	.08
Student sex				
Female	.69	.67	.67	.70
Male	.62	.61	.61	.64
Student race/ethnicity				
White, non-Hispanic	.68	.66	.66	.68
Black, non-Hispanic	.62	.60	.60	.63
Hispanic	.48	.44	.43	.49

Note. The underlying sample submitted to the NSC is always the Michigan high school class of 2009. Each column shows the fraction enrolled in college by September 2010 (second fall after high school graduation). Private 4-year colleges include non-profit and for-profit institutions. FERPA = Family Educational Rights and Privacy Act; NSC = National Student Clearinghouse.

estimates of the same parameters—by institution type and by student subgroup. These pooled estimates are of the same magnitude or larger compared with their corresponding parameters from any individual NSC pull. This suggests that there is a value added in submitting a group of students multiple times to the NSC.³⁰

In our Michigan NSC data, we also observe institution-specific patterns of FERPA-blocking over these three time points that suggest (at least in some cases) changes in FERPA-blocking are largely a school-level, rather than a student-level, phenomenon.³¹ We observe exceptionally large increases in the FERPA-blocking rate at particular institutions in our sample. For example, at one large public research institution, the FERPA-blocking rate jumped from 8.5% in the fall of 2010 to 24.2% about a year later. Yet, at another large public research institution, the FERPA-blocking rate barely budged over the same time period, from 0.02% to 0.09%. We observe similar examples in the 2-year sector: One community college in Michigan observed its FERPA-blocking rate increase from essentially 0 (0.01%) to nearly 31% during the same period.³²

These two institutions with large increases are outliers, but highlight the importance of multiple submissions to identify time periods where FERPA-blocking may be of relatively more concern. Researchers should be cognizant of the degree to which blocked information for the state and time period under study may affect the estimates of key college-going parameters.

Exploring Estimates of Postsecondary Enrollment and Persistence in Michigan

With an improved understanding of how these three main sources of measurement error are likely to show up in our Michigan NSC data, we now exploit the existence of an additional rich data source in Michigan as a way to illustrate implications of this collection of measurement error concerns for estimating a range of postsecondary outcomes. We first describe the student-level administrative transcript data, and then conduct analyses in which we compare estimates of college enrollment and persistence for a recent graduating class of Michigan high school students based on each data source.

Michigan's Student Transcript and Academic Record Repository (STARR). STARR consists of electronic college transcripts for all students who attend public institutions of higher education in Michigan, including community colleges. The database contains information about each course a student takes, including course title, credits earned, and grade. It also contains information on declared majors and minors, degrees or certificates a student earns, and demographic information as collected by the participating colleges.

The collection of these data was mandated as a requirement for Michigan to receive SFSF dollars under the American Recovery and Reinvestment Act (ARRA). The state has just concluded its second collection of these data. The first collection included complete transcripts for every student who was enrolled in a Michigan public college or university at any point between June 2009 and May 2011. The second collection includes the same set of information for any student who attended a Michigan public college between June 2011 and May 2012. STARR data will be collected annually going forward. STARR covers all students enrolled in Michigan public institutions, whether they attended high school in Michigan or elsewhere, are a graduate or undergraduate, or are seeking a degree or not. As these data are a collection of transcripts, linking to a student once captures their full history of attachment to an institution.

Enrollment. We focus on Michigan public high school graduates in the class of 2011. We use this class of graduates because the 2011–2012 wave of public postsecondary transcript data (i.e., STARR) is the cleanest (i.e., many bugs in the data collection process were worked out between the first and second waves). In Table 7, we estimate the proportion of this class that enrolls in postsecondary schooling within 1 academic year after high school graduation. We calculate the same statistics for a variety of sociodemographic subgroups of students.

We observe that nearly 70% enroll in college (anywhere) within the first year following high school graduation and 55% enroll in a Michigan public college during that time frame (based on NSC data alone). If we instead examine the share of these high school graduates who enroll in a

Michigan public college based on state administrative data, we arrive at the same estimate (i.e., 55%). But, if we combine the NSC and state administrative data sources, this figure climbs to 59%. This suggests that the two data sources are capturing slightly different students.

Students who we observe attending a Michigan public college in one data source but not the other tend to cluster in a few institutions. For example, nearly 30% of students we observe attending college in the STARR data but not in the NSC data are at community colleges we know recently joined the NSC. Similarly, about 25% of students we observe attending college according to NSC data but not STARR data are at a community college that had some trouble uploading complete transcript data for its students. No data source is perfect. These examples illustrate holes in coverage and how states' own postsecondary data systems can improve and be improved by NSC data. At least in Michigan, our estimates in Table 7 suggest small-to-moderate differences in the calculated college-going rate based on NSC versus state administrative data.

Persistence and Progress Through College. Given the incentives and demands created by SFSF funding, one of the most salient measures of persistence that state departments of education now calculate is the proportion of students from recent high school graduating classes who go on to obtain a year's worth of college credit.

The NSC records students' college-going experience in terms of enrollment spells (i.e., particular dates of enrollment) rather than semesters. The federal government supplied states with a formula to use with NSC data that calculates whether a student "earned" a year of credit based on the total number of weeks a student was enrolled in college. This formula considers students who meet any of the following criteria as having earned 1 year of college credit (24 credits): enrolled FT for 28 weeks, enrolled half-time for 56 weeks, enrolled less-than-half-time for 112 weeks, or any equivalent combination. This calculation requires information on the intensity at which students were enrolled during each enrollment spell. Unfortunately, a sizable share of intensity information is missing in NSC data: For the Michigan high school class of 2011, about 25% of enrollment spells are missing intensity information.³³

TABLE 7

Initial College Enrollment for Michigan's Class of 2011: NSC Versus Administrative Data

Data source	NSC		STARR	Union of NSC and STARR	
	Any college	Michigan public college	Michigan public college	Any college	Michigan public college
	(1)	(2)	(3)	(4)	(5)
Proportion enrolling in					
Any college	.69	.55	.55	.72	.59
4-year college	.39	.27	.29	.41	.29
2-year college	.29	.28	.27	.32	.31
Student sex					
Male	.64	.52	.52	.67	.56
Female	.73	.58	.58	.76	.62
Student race/ethnicity					
White, non-Hispanic	.65	.51	.45	.68	.54
Black, non-Hispanic	.70	.56	.58	.73	.60
Hispanic	.52	.42	.44	.57	.48
Student FARM status					
FARM	.58	.46	.45	.61	.50
Non-FARM	.74	.60	.61	.78	.64

Note. The sample includes all 2011 graduates from Michigan public high schools ($n = 105,750$). College enrollment is defined by the first college in which a student enrolls, during the first year following high school graduation (i.e., summer 2011, fall 2011, spring 2012, summer 2012). NSC = National Student Clearinghouse; STARR = Student Transcript and Academic Record Repository; FARM = eligibility for free or reduced-price meals.

As a first-order question, we explore the adequacy of this federal formula for translating enrollment spells into credits earned. To investigate how well this formula approximates actual credit accumulation among students, we limit our analytic sample to college-goers from Michigan's high school class of 2011 who attended a public college in Michigan and have no missing intensity information. We then compare estimates of the number of credits earned and the share of this subsample earning at least 1 year of credit based on NSC data to the same statistics calculated using the state's transcript data instead (i.e., the truth).

The results in Table 8 illustrate that the federal formula generally overestimates the share of students earning at least 1 year of college credit by about a year after high school graduation. Furthermore, this overestimate is not equal for all students. The formula substantially overestimates credit accumulation for traditionally disadvantaged groups (e.g., those eligible for free or reduced-price meals [FARM] and minority

students). For example, estimating the share of Black undergraduates who earn at least 1 year of college credit based on NSC data yields a figure of 47%. Counting up actual credits earned by Black students at Michigan public colleges in our state administrative data instead yields a figure of only 29%, for a difference of 18 percentage points.

About a quarter of the college-going sample in Table 8 has at least one missing enrollment intensity status associated with an enrollment spell in our NSC data.³⁴ How states choose to handle these missing values also has implications for estimates of postsecondary attainment. In Table 9, we impute missing intensity values in the NSC data as either less-than-part-time (LTPT), PT, or FT; calculate the corresponding share of students earning at least 1 year of college for each subgroup; and compare this figure with the "truth" we observe in our state administrative data (i.e., actual credits earned). This exercise illustrates that for FARM, Black, and Hispanic students, imputing missing intensity

TABLE 8

Lost in Translation? Using the Federal NSC Formula to Approximate Credit Accumulation at Public Colleges

	STARR		NSC	
	Number of credits	Share with >1 year of credit	Estimated number of credits	Estimated share with >1 year of credit
	(1)	(2)	(3)	(4)
All students				
All credits earned	22.25	.55	23.34	.63
Earned at 4-year	15.17		14.69	
Earned at 2-year	6.77		8.64	
White, non-Hispanic				
All credits earned	23.43	.59	23.86	.66
Earned at 4-year	16.01		15.25	
Earned at 2-year	7.08		8.60	
Black, non-Hispanic				
All credits earned	15.82	.29	20.73	.47
Earned at 4-year	10.34		11.72	
Earned at 2-year	5.31		9.01	
Hispanic				
All credits earned	18.68	.40	21.13	.51
Earned at 4-year	10.26		10.52	
Earned at 2-year	8.18		10.62	
FARM				
All credits earned	17.65	.36	21.21	.50
Earned at 4-year	9.70		10.47	
Earned at 2-year	7.80		10.74	
Non-FARM				
All credits earned	24.02	.62	24.15	.68
Earned at 4-year	17.27		16.32	
Earned at 2-year	6.37		7.84	

Note. The sample is restricted to college-goers from the Michigan high school class of 2011 who attended public colleges in Michigan and had no missing enrollment intensity information ($n = 38,996$). College enrollment is defined over the same time horizon as in Table 7. NSC = National Student Clearinghouse; STARR = Student Transcript and Academic Record Repository; FARM = eligibility for free or reduced-price meals.

values as LTPT gets one closer to the “truth,” whereas for White and non-FARM students, the best imputation choice lies between PT and FT. Still, in all cases, imputing all missing intensity values as FT only serves to exacerbate the degree to which the federal formula overstates credit accumulation.

Taken together, our findings suggest that this formula is not reliable for the very groups whose attainment is often of concern. It may be more conservative to simply report what we know: the number of semesters enrolled, and be agnostic about equivalent credits.

Special Issues for Policy Analysts and Program Evaluators

In this article, we describe and explore the most salient sources of measurement error in NSC data. We examine the implications of such measurement error for estimating postsecondary outcomes of interest to policymakers and researchers. It is our hope that these analyses will serve as a useful reference for policy analysts and researchers using NSC data to study the impacts of various programs and policies on postsecondary attendance, persistence, and attainment.

TABLE 9

Improving the Translation: Effects on Credit Accumulation of Intensity-Based Imputation

	STARR	Share of students with missing enrollment intensity	NSC: Impute missing values as . . .		
			LTPT	PT	FT
	(1)	(2)	(3)	(4)	(5)
All students		.25			
Number of credits earned	21.46		19.36	20.90	24.00
Share with >1-year college credit	.53		.48	.48	.67
White, non-Hispanic		.25			
Number of credits earned	22.57		19.75	21.34	24.52
Share with >1-year college credit	.57		.50	.51	.70
Black, non-Hispanic		.22			
Number of credits earned	15.09		17.34	18.66	21.28
Share with >1-year college credit	.27		.36	.36	.51
Hispanic		.25			
Number of credits earned	18.04		17.26	18.76	21.76
Share with >1-year college credit	.38		.38	.39	.56
FARM		.25			
Number of credits earned	16.97		17.25	18.83	21.98
Share with >1-year college credit	.34		.37	.37	.56
Non-FARM		.24			
Number of credits earned	23.23		20.19	21.72	24.80
Share with >1-year college credit	.60		.52	.53	.71

Note. The sample is restricted to college-goers from the Michigan high school class of 2011 who attended public colleges in Michigan ($n = 54,170$). College enrollment is defined over the same time horizon as in Table 7. STARR = Student Transcript and Academic Record Repository; NSC = National Student Clearinghouse; LTPT = less-than-part-time; PT = part-time; FT = full-time; FARM = eligibility for free- or reduced-price meals.

Our findings illustrate that NSC data miss a shrinking but non-trivial portion of undergraduate enrollment in the United States: Coverage is highest among public institutions (99% for 4-year; 96% for 2-year) and lowest (but growing) among the for-profit sector (48%). There is a substantial variation in coverage across states, institutional sectors, and over time. Such variation underscores the need for researchers using NSC data to evaluate programs or policies to understand the level of coverage NSC data provide for the state(s) and time period under study. Given relatively low coverage among for-profits (at this time), research should avoid using NSC data for projects that focus predominately on such institutions. Furthermore, researchers should be aware that NSC data do not comprehensively report major information and do not include student transcript information, which could limit the use of these data for certain studies.

We find that the algorithm the NSC uses to match students to their postsecondary enrollment records is quite robust to common computerized perturbations of names. Such errors are unlikely to introduce appreciable measurement error (at least based on our Michigan data). Finally, our work illustrates the value of submitting the same population of students to the NSC multiple times, as FERPA-blocking rates are time-sensitive and sometimes institution-specific. By doing so, researchers stand to capture enrollment information prior to the initial submission about students who removed their FERPA block between submission dates.

In this concluding section, we highlight a few issues relevant for program evaluators interested in estimating the impact of some treatment on college outcomes. We focus on a hypothetical research question in which the outcome of interest is enrollment in college (i.e., binary). In general,

the types of measurement errors we discuss in this article will lead researchers to incorrectly categorize some college-attenders as not attending college. This sort of misclassification of a dichotomous dependent variable, if uncorrelated with treatment status, will attenuate any treatment effect toward 0 (Bound, Brown, & Mathiowetz, 2001). The basic intuition behind this conclusion is that for a binary outcome, there can never be a positive “measurement error” for someone who did indeed attend college (outcome = 1). That is, for someone who did attend college (outcome = 1), there is no way to “positively” misclassify that person. The only feasible misclassification is to instead say that they did not attend college (outcome = 0).³⁵ Therefore, this implies that if non-coverage in NSC data is unrelated to treatment status (as in the case of a randomized experiment), then a simple adjustment using our reported “comprehensive” coverage rate will scale the treatment effect (measured in percentage points) to its true magnitude.

States typically choose the sample of students to submit to the NSC based on the grade in which they are enrolled. In fact, states most often submit classes of graduating public high school students to the NSC. Assume one is interested in the effect of a ninth-grade curriculum change on college attendance. The change is introduced in a particular year, t , such that those in eighth grade before time t are not exposed, but those in eighth grade in year t and forward are exposed. If the state only submits the analogous graduating classes of students, the analyst will not observe college outcomes for dropouts and will observe college outcomes sooner for those students (from the treatment and control cohorts) who progressed through high school on time. The curriculum could have affected dropout and the rate at which students progress through school. If so, such a submission plan can lead to measurement error that is correlated with the treatment of interest. The lesson here is to submit intact cohorts of students (e.g., first-time ninth graders by year) and to be patient about those who are progressing slowly or redefine your outcomes (e.g., within 1 year of expected on-time high school graduation).

In non-experimental and quasi-experimental settings, there may be reason to suspect that NSC coverage could be correlated with treatment

status.³⁶ For evaluations that exploit temporal variation in exposure to a treatment and compare college-going rates over time, it is important to consider whether there were changes in the number and type of institutions participating in the NSC during the outcome time period under study. A spike in NSC participation that occurs during the time when treated students would be enrolling in college (but not control students) would confound increases in college enrollment due to the treatment of interest with increases due to changes in NSC coverage. One way to address this concern is to only count attendance at colleges that have joined NSC by the beginning of the outcome period of interest: that is, code students who enroll in colleges that join the NSC later as non-enrollees. This allows the analyst to assess the effect on college enrollment among a common underlying population of colleges.

Finally, variation in NSC coverage by institutional sector points to a potential interpretation concern: If a treatment or policy of interest induced students currently attending institutions with poorer coverage rates (e.g., for-profits) to instead attend colleges that are comprehensively covered by the NSC (e.g., publics), then researchers may incorrectly characterize what is really a “choice effect” (i.e., moving from one type of college to another as a result of the treatment) as an “enrollment/attendance” effect.

Of course, all research questions are unique. We seek to provide the evidence, framework, and key guiding principles necessary to help researchers and policy analysts interested in higher education more carefully exploit the richness of NSC data—as well as explore the potential impacts of the limitations inherent in using these data on their work.

Authors' Note

The opinions expressed are those of the authors and do not represent views of the Institute of Education Sciences or the U.S. Department of Education.

Acknowledgments

We are grateful for the help of our partners at the Michigan Education Department: Tom Howell, Venessa Keesler, and Joseph Martineau, as well as for exceedingly helpful support from state analysts, including Rod Bernosky, Melissa Bisson, and Karen Conroy. We had several very helpful conversations

with Jason DeWitt and Travis Maciejewski at the National Student Clearinghouse. We are grateful to Sara Goldrick-Rab and three anonymous reviewers for helpful comments. We also thank Hassan Enayati, Joanna Frye, Emily House, and Chris Zbrozek for providing excellent research assistance.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grants R305B110001 and R305E100008 to the University of Michigan, as well as through Grant R305C110011-11A to the Teachers College, Columbia University.

Notes

1. For example, see letters and policy statements from the National Association of Independent Colleges and Universities (NAICU) directed at key policymakers that argue against the creation of such a system: http://www.naicu.edu/news_room/detail/national-association-of-independent-colleges-and-universities-views-spellings-commission-final-report-as-improved-but-still-problematic

2. A recent report from the Data Quality Campaign (DQC; 2012) finds that although 43 states in 2012 have been able to make some sort of link between their K–12 data and postsecondary information, “meaningful and useful P–20W data sharing is still lacking” (p. 9). DQC pointed, in particular, to the lack of data on students who enroll outside their home state.

3. Classification of Instructional Programs (CIP): <http://nces.ed.gov/ipeds/cipcode/>

4. Employer demand for a centralized way to verify that college graduates applying for jobs actually earned the degrees they claimed to have earned is what drove National Student Clearinghouse (NSC) to collect more information about degrees (NSC, personal communication, August 2, 2012).

5. For more information on the State Fiscal Stabilization Funds (SFSF), see <http://www2.ed.gov/policy/gen/leg/recovery/factsheet/stabilization-fund.html>; and <http://www2.ed.gov/programs/statestabilization/guidance.pdf>

6. Although the federal government has augmented NSC’s business through the SFSF program, the recent shift to federal direct lending may have dampened this

increase by reducing the extent to which NSC can rely on revenue from student loan companies.

7. In addition, corporate research (e.g., The College Board, ACT Inc.) makes use of NSC data as well (see, for example, Radunzel & Noble, 2012).

8. In practice, it appears that some colleges do not report information on degree receipt.

9. The list of participating colleges, along with the date they began reporting degree information to the NSC, can be found at <http://www.studentclearinghouse.org/verifiers/degree/schools.php>

10. As an example, see Table 170 in the National Center for Education Statistics (NCES; 1990): “Enrollment rates of 18- to 24-year-olds in institutions of higher education, by race/ethnicity” (p. 181). Access: <http://nces.ed.gov/pubs91/91660.pdf>

11. The migration survey supplement to the Current Population Survey (CPS) is administered every few years in March and captures the state of residence 5 years earlier.

12. For more information on the American Community Survey (ACS), see <http://www.census.gov/acs/www/>

13. We use the ACS 3-year estimates.

14. This assumes that institutions participating in the NSC report all of their enrollment data to the NSC. Based on our conversations with the NSC, this seems to be an entirely reasonable assumption. To the extent that this is untrue, our coverage rate estimates may slightly overstate coverage. In Michigan, we find no evidence that NSC misses enrollment from participating colleges, except for a small share of dually enrolled students (i.e., those students enrolled in college while still in high school).

15. We use the master list maintained by NSC staff. A slightly modified list is publically available at http://www.studentclearinghouse.org/colleges/student-tracker_for_outreach/participating_schools.php

16. Our choice of October 15th follows the method used in-house by NSC to calculate their coverage rate. The rates we report are virtually identical when we use a more conservative cutoff date of August 31st.

17. In NSC-reported coverage calculations, they restrict the postsecondary institutions to those that are degree-granting. Degree-granting institutions, as defined in Integrated Postsecondary Education Data System (IPEDS), are those that award associate’s degrees or higher. So this definition excludes institutions such as trade schools that offer only certificates and similar awards or diplomas. We include these institutions because we believe that most users of NSC aim to capture any and all postsecondary enrollment. Coverage rates are about 2 percentage points higher if we limit to degree-granting institutions. Analogous calculations that restrict to degree-granting institutions are available from the authors upon request.

18. Full-time-equivalent (FTE) enrollment is calculated by adding full-time enrollment counts to down-weighted part-time enrollment counts, down weighting by college-type specific fractions provided by IPEDS.

19. An institution is counted as a “selective” institution if it received a Barron’s ranking of “very competitive,” “highly competitive,” or “most competitive.” We thank Michael Bastedo and Ozan Jaquette for sharing their compilation of Barron’s ranking data.

20. Other race includes Asian students, multi-racial students, and students whose race is not known by the postsecondary institution.

21. We learned this based on our past research using degree completion information in NSC data, and from conversations with research staff at the NSC.

22. This method will mistakenly include graduate degrees that have a missing degree title and are earned by students who received their undergraduate degree in another state or at a college not in the NSC.

23. The number of degrees in column 2 by sector does not sum to the total number of degrees due to the variation across sectors in this share.

24. Note that the bounds are more informative for the overall and public institution analyses because there are fewer missing degree titles for students in public institutions. Private non-profits, and particularly private for-profit institutions in our data, have higher rates of missing degree titles, and thus the bounds are less informative.

25. As a reminder, researchers must submit names and dates of birth of students from a particular sample, in our case, Michigan public high school students; researchers cannot simply ask NSC for information on a certain population of students, such as all students who attend college in Michigan.

26. The weighted coverage rate is the unweighted rate multiplied by the inverse of the fraction of Michigan undergraduate enrollment that we observe in our NSC data. This “enrollment fraction” is calculated as follows: the denominator is the 12-month unduplicated undergraduate head count as reported by IPEDS for a year and sector. The numerator is the number of undergraduate NSC students in the same year and sector enrolled at a Michigan college (where we proxy for undergraduate by including only students who have not previously earned a degree from a 4-year college). For example, assume we observe 600 degrees in the NSC data and 1,000 degrees in the IPEDS data, yielding an unweighted coverage rate of 0.6. We observe 3,000 undergraduates in the NSC data and 4,000 in IPEDS for that year and sector for a weight of 0.75. Thus, the weighted coverage rate is $0.6 \times (0.75)^{-1} = 0.8$.

27. The overwhelming majority of “FERPA-blocked” (Family Educational Rights and Privacy

Act-blocked) records (nationally) are due to student-level blocking.

28. These state- and sector-specific rates are only available from the NSC for 3 years: 2006–2007, 2008–2009, and 2010–2011 (NSC, 2012).

29. This entire process was completed all within the same month; therefore, any changes in NSC’s interpretation of FERPA laws, algorithm processes, and so on are not likely to confound our name variant analysis—nor are changes in students’ choices about whether to block their detailed information under the FERPA.

30. Goldrick-Rab and Harris (2010) make a similar point in their informal brief to colleagues about NSC data and highlight the (potentially) sizable costs of multiple submissions.

31. NSC provides back to researchers the total number of blocked records by institution for a given submission.

32. Perhaps this is suggestive of different institutions changing the way in which a FERPA-blocking option is presented to undergraduate students (i.e., the default).

33. These missing intensity values tend to be fairly evenly distributed between 4-year and 2-year schools, but relatively more concentrated among public institutions compared with privates.

34. The share of students with missing enrollment intensity information does not vary much across demographic subgroups.

35. Therefore, this conclusion differs importantly from a case in which the outcome is continuous and the measurement error of that outcome is not associated with treatment. In such cases, the measurement error will simply cause the treatment effect to be estimated with less precision (i.e., larger standard errors), but the coefficient itself remains unbiased (Wooldridge, 2003).

36. Even in cases where the treatment has been randomly assigned to a group of students, researchers would want to ensure that the quality of the administrative information used to match students to NSC data (i.e., name and birth date) does not systematically vary by treatment status.

References

- Ackerman, D., Cronin, J. A., Turner, N., & Bershadker, A. (2011, November). *Coordinating the American Opportunity Tax Credit and the federal Pell Grant*. Paper presented at National Tax Association Conference.
- Bettinger, E. P., Long, B. T., Oreopoulos, P., & Sanbonmatsu, L. (2012). The role of simplification and information in college decisions: Results from the H&R Block FAFSA experiment. *Quarterly Journal of Economics*, 127, 1205–1242.

- Bound, J., Brown, C., & Mathiowetz, N. (2001). Measurement error in survey data. In E. Leamer & J. J. Heckman (Eds.), *Handbook of econometrics* (Vol. 5, pp. 3705–3843). Amsterdam, The Netherlands: Elsevier Science.
- Chingos, M. M., & Peterson, P. E. (2012). *The effect of school vouchers on college enrollment: Experimental evidence from New York City*. Washington, DC: Brown Center on Education Policy, Brookings Institution.
- Cunningham, A. F., & Milam, J. (2005). *Feasibility of a student unit record system within the Integrated Postsecondary Education Data System* (NCES 2005-160). Washington, DC: U.S. Government Printing Office, National Center for Education Statistics, U.S. Department of Education.
- Data Quality Campaign. (2012). *Data for action 2012: Focus on people to change data culture*. Washington, DC. Retrieved from <http://dataqualitycampaign.org/resources/details/1631>
- Deming, D. J., Hastings, J. S., Kane, T. J., & Staiger, D. O. (2014). School choice, school quality, and postsecondary attainment. *American Economic Review*, 104, 991-1014.
- Dynarski, S., Hyman, J., & Schanzenbach, D. W. (2013). Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion. *Journal of Policy Analysis and Management*, 32, 692–717.
- Garcia, T., & L’Orange, H. P. (2012). *Strong foundations: The state of state postsecondary data systems: 2012 update on data sharing with K-12 and labor*. Boulder, CO: State Higher Education Officers Association.
- Goldrick-Rab, S. (2012). *Review of “the effects of school vouchers on college enrollment.”* National Education Policy Center. Retrieved from <http://nepc.colorado.edu/files/ttr-voucherscollege.pdf>
- Goldrick-Rab, S., & Harris, D. N. (2010). *Letter to colleagues: Observations on the use of NSC data for research purposes*. Retrieved from <http://www.finaidstudy.org/documents/NSC%20Dear%20colleagues%20letter.pdf>
- Goldrick-Rab, S., Harris, D. N., Kelchen, R., & Benson, J. (2012). *Need-based financial aid and college persistence experimental evidence from Wisconsin*. Retrieved from <http://ssrn.com/abstract=1887826>
- Hemelt, S. W., Roth, K. B., & Eaton, W. W. (2013). Elementary school interventions: Experimental evidence on postsecondary outcomes. *Educational Evaluation and Policy Analysis*, 35, 413–436.
- Higher Education Opportunity Act of 2008, Pub. L. No. 110–315 (2008).
- Hout, M. (2011). Social and economic returns to college in the United States. *Annual Review of Sociology*, 38, 179–400.
- Hyman, J. (2013, December). *ACT for all: The effect of mandatory college entrance exams on postsecondary attainment and choice* (Working paper). Ann Arbor: University of Michigan.
- Kane, T. J. (2003, May). *A quasi-experimental estimate of the impact of financial aid on college-going* (NBER Working Paper Series No. 9703). Cambridge, MA: National Bureau of Economic Research.
- Lederman, D. (2012, September 13). Higher ed scholars’ voucher war. *Inside Higher Ed*. Retrieved from <http://www.insidehighered.com/news/2012/09/13/researchers-argue-over-school-vouchers-impact-college-going>
- Lleras-Muney, A. (2005). The relationship between education and adult mortality in the United States. *Review of Economic Studies*, 72, 189–221.
- National Center for Education Statistics. (1990). *Digest of education statistics*. Washington, DC: U.S. Department of Education.
- National Center for Education Statistics. (2010). *Integrated Postsecondary Education Data System (IPEDS)*. Washington, DC: U.S. Department of Education.
- National Student Clearinghouse. (2012). *Impact of directory information blocks on student tracker results*. Retrieved from http://nscresearchcenter.org/wp-content/uploads/NSC_Directory_Block_Rates.pdf
- Oreopoulos, P., & Petronijevic, U. (2013). *Making college worth it: A review of research on the returns to higher education* (NBER Working Paper Series No. 19053). Cambridge, MA: National Bureau of Economic Research.
- Oreopoulos, P., & Salvanes, K. G. (2011). Priceless: The non-pecuniary benefits of schooling. *Journal of Economic Perspectives*, 25(1), 159–184.
- Radunzel, J., & Noble, J. (2012). *Tracking 2003 ACT-tested high school graduates: College readiness, enrollment, and long-term success* (ACT Research Report Series, No. 2). Retrieved from http://www.act.org/research/researchers/reports/pdf/ACT_RR2012-2.pdf
- Richburg-Hayes, L., Brock, T., LeBlanc, A., Paxson, C., Rouse, C. E., & Barrow, L. (2009). *Rewarding persistence: Effects of a performance-based scholarship program for low-income parents*. Retrieved from <http://www.mdrc.org/sites/default/files/Rewarding%20Persistence%20ES.pdf>
- Wooldridge, J. M. (2003). *Introductory econometrics* (2nd ed.). Mason, OH: South-Western College.

Authors

SUSAN M. DYNARSKI is a professor of public policy, education, and economics at the Gerald R. Ford School of Public Policy, School of Education, and Department of Economics, University of Michigan, and a research associate at the National Bureau of Economic Research. Her research focuses on financial aid for college, inequality in postsecondary outcomes, charter schools, and labor and public economics.

STEVEN W. HEMELT is an assistant professor of public policy at the University of North Carolina at Chapel Hill. His primary fields of interest are the economics of education, program evaluation, and education policy. His current research focuses on transitions

from secondary to postsecondary education, college affordability, and differentiated accountability in K–12 schooling.

JOSHUA M. HYMAN is an assistant professor in the Department of Public Policy at the University of Connecticut, with a joint appointment in the Department of Educational Leadership in the Neag School of Education. His research focuses on labor economics, public finance, and the economics of education.

Manuscript received October 11, 2013

Revision received June 12, 2014

Accepted January 23, 2015