THE MISMEASUREMENT OF STEM: EVIDENCE FROM COLLEGE COURSE CLASSIFICATIONS[1]

Daniel Douglas—Trinity College

Hal Salzman—Rutgers University

Khudodod Khudododov—Rutgers University

Daniyal Rahim—University of Toronto

STEM acronymically refers to four areas of inquiry – *S*cience, *T*echnology, *E*ngineering, and *M*athematics. But as its use has become ubiquitous, STEM has taken on social and political meaning far beyond the sum of its component parts. It has come to denote the fields that, collectively, are most valued in education and the economy, whether in pre-school education policy or post-graduate program priorities, as well as in workforce training and employment.

The public is regularly told of existing or looming issues of K-12 students' STEM readiness, college students' STEM degree earning, STEM workforce shortfalls, and inequalities therein (Salzman and Douglas, 2022; Teitelbaum, 2014; Salzman, 2022). Across all these realms, the amalgamation of fields called STEM is considered more important than all other fields.

But what exactly constitutes "STEM"? Scientist, mathematician, and engineer are occupations denoted by this definition. But does it include all those in the sciences, from lab technicians to astrophysicists? If STEM includes the two-year-degree-bearing lab technologist, why does it exclude practicing physicians? Given the extent and regular use of mathematics in accounting, should actuaries be counted? Technology is an important component of modern life, so perhaps it is fitting to include HVAC technicians? Conversely, if we exclude those who are practitioners of technological applications, then do we exclude the vast majority of engineers, the largest share of whom are calculating loads, construction specs and inspecting them? Whatever your answer to these questions, they do not accord to any standard definition of STEM, for there is none.

One commonly used definition in the past half-century has been the National Science Foundation's (NSF) "S&E" category that counts those who "understand the fundamental laws of nature and are skilled in the techniques of scientific research" (Bush, 1945). Owing at least in part to the different organizational scopes of the NSF and the National Institutes

---

[1] Daniel Douglas, Trinity College Department of Sociology, daniel.douglas@trincoll.edu ; Hal Salzman, E.J. Bloustein School of Planning & Public Policy and J.J. Heldrich Center for Workforce Development, Rutgers University, HSalzman@Rutgers.edu; Khudodod Khudododov, J.J. Heldrich Center for Workforce Development, Rutgers University, khudodod.khudododov@ejb.rutgers.edu; Daniyal Rahim, Higher Education Program, University of Toronto, daniyal.rahim@gmail.com
We are grateful for the patient support of this research by the Alfred P. Sloan Foundation and our program officer Danny Goroff (Grant Nos. 2012-6-13 and G-2016-7310), with additional support from the National Science Foundation (Grant No. NSF-DGE-1561687); Sean Simone provided insightful review and information about the NCES course classifications and surveys.

of Health (NIH), S&E excludes nearly all health practitioners (such as medical doctors, nurses, and medical technicians), while including those with any science degree who work in a research laboratory – even if their job is primarily doing repetitive lab tests following rigid protocols. Using such a definition, the biology graduate who becomes a physician is considered part of the "leaky pipeline" having begun but left a STEM field. In fact, the Census finds only about one-quarter of all STEM graduates occupy STEM jobs (Landivar 2013).[2] More recently, the National Academies of Sciences, Engineering, and Medicine examined undergraduate education in STEM fields, and called for more comprehensive indicators of educational quality, including not only STEM majors, but also STEM coursework taken by students who either never enter, or eventually leave these fields of study (NASEM 2018).

For all the importance ascribed to STEM, there is a startling lack of clarity on its meaning. Does STEM denote a field of study? An occupation? Or a body of knowledge that is applied across any number of occupations? Indeed, discussions of, and policymaking about STEM employ all of these permutations.

The lack of a precise and agreed-upon definition of what counts as a STEM course, or a STEM major, is not just an esoteric technical matter. It is a problem that, we find, would alter the evidence base used in many policy, funding, and education decisions. It is also an urgent matter because the extent of mismeasurement of STEM is significant enough to require serious reconsideration of at least some findings that rely on current STEM classification schema and coding. We argue that these mismeasurement problems are consequential, much like trying to understand an astrophysics phenomenon with an imprecise constant in an equation.[3]

If, as discussed below, and in Douglas and Salzman (2020), we are undercounting math course taking by 20 percent because we overlook statistics courses taught outside of math departments, or are over-counting course taking in another field, we might draw incorrect conclusions about where postsecondary curriculum development should occur. Similarly, if we designate emerging medical practitioners as "leaving STEM", yet they are the largest single group of students enrolled in biology courses, we may be skewing policy efforts toward a STEM retention problem that may not exist and is certainly overstated due to a classification error.

Similarly, when considering that mathematics claims one of the STEM letters and Computer Science is firmly within any concept of STEM programs, what do we make of a schema that classifies a student with an interdisciplinary major in Computer Science *and* Mathematics as a non-STEM graduate? Further, if she began in mathematics and then changed to an interdisciplinary Computer Science and Mathematics major, she would be counted as one of the women "leaving STEM" during her college years. While these are mostly errors in taxonomy and the effect of administrative restructuring, and usually small even if

---

[2] Following NSF, the Census Bureau includes social scientists in their accounting of STEM degrees and occupations (Landivar, 2013). Our analysis excluding social science graduates, using this Census Bureau study, finds only a third of STEM graduates in STEM occupations (Salzman, 2015; Salzman and Benderly, 2019).

[3] For example, see the effort to refine the estimate of Hubble's Constant, a fundamental tool in determining the size and nature of the universe (Billings, 2018).

nontrivial distortions of the data, they are compounded when, for political reasons, Drama Therapy, along with Classics and Literary Reportage become classified as a STEM majors, as it was by New York University in 2020.[4]

In this paper, we take a first step in clarifying the analytic categories of STEM in education. This, we propose, is a necessary first building block for STEM analysis – to understand what constitutes STEM coursework, the constituent element of a STEM education. We first review the STEM definitional problems we have identified in the process of examining two sets of nationally-representative data, provide analysis of the extent of potential mismeasurement, and estimates of impact. We then outline an approach to resolving the mismeasurement problems in nationally-representative postsecondary student surveys.

Background

The Politics of Counting STEM

Before there was STEM, there was Science and Engineering (S&E). This collection of academic programs was first counted and classified by the National Research Council and the National Science Foundation. Congress funded the study, recognizing the role of S&E in defense, to assess the stock of science and engineering human resources that could be called upon should another Manhattan project or war effort be needed. Responsibility for the survey was given to NSF, which focused on the disciplinary areas that NSF supported, including biological and physical sciences, engineering, computer and mathematical sciences, and social sciences. Notably, this S&E classification excluded most health-related fields, which were supported by the National Institutes of Health. But this initial programmatic classification became a very important source of noise in the discussion of STEM retention when it was used for purposes outside of its intended use. As discussed above, in all official or widely used statistics, students who follow the very common trajectory from a biological sciences program to a discipline in the health professions would be, and are still classified as having left STEM. That is, in the policy world, based on these statistics, the STEM workforce is seen as being depleted by biology graduates going into health professions.

---

[4] In a rather remarkable justification, New York University appears to have classified a version of Art History and Classics majors as falling under Archeology, which is an established STEM major. The University says those majors can be classified as a field of Archeology because the former majors, similar to Archeology, also "[i]nclude instruction in…conservation and museum studies…and the study of specific selected past cultures," as long as they took several "classical archeology" courses in addition to their standard Art History or Classics courses to qualify. Individual schools assign the instructional classification code for each major (CIP) and, according to DHS, that classification is accepted by DHS with minimal review or oversight. The Department of Education's National Center for Educational Statistics, which administers the development of the CIP codes and data collection, does not review these types of classification decisions by universities; moreover, NCES uses the DHS STEM designation list for determining majors to be classified as "STEM" in NCES's datasets (e.g., BPS and B&B). "DHS Department of Homeland Security's Response to Ranking Member Charles E. Grassley's February 17, 2022 Letter" (June 2, 2022). Sen. Grassley's query to DHS: https://www.grassley.senate.gov/imo/media/doc/grassley_to_homeland_security_dept.stemoptprogram.pdf DHS's response providing the justification discussed here: https://www.grassley.senate.gov/imo/media/doc/homeland_security_depttograssleystemoptprogram.pdf. Also see NYU website: https://steinhardt.nyu.edu/degree/ma-drama-therapy

The STEM acronym is credited to biologist Judith Ramaley, who, as assistant director at NSF in 2001, created the acronym in a revision of term then in use for Science, Mathematics, Engineering, and Technology—SMET (because it lent itself to being mispronounced as "smut"; Hallinen, 2015). This expanded acronym was meant to reflect the importance of mathematics and technology as they relate to engineering and science. However, STEM quickly became politicized as denoting the fields specifically important for the economy, innovation, and national competitiveness, and to the exclusion of other fields as secondary in importance for support or as an educational or career track. A steady stream of notable policy papers, many lacking supporting data or other evidence, made dire claims of coming shortages of STEM graduates and argued for the need to produce thousands or millions more as the default solution to economic crises(e.g., National Academy of Sciences, 2007; National Science Board, 2015; PCAST 2012)[5]. Reports such as these have become the conventional wisdom, and have led to substantial funding for narrowly targeted programs and consequential legislative changes in areas ranging from early childhood education to immigration.

For classification purposes, the first general problem that emerges in the transformation from S&E to STEM is the addition of "T" – which stands for technology, but that has no specific disciplinary referent. Mathematics was always included in S&E programs, funding, and data collection, but was not given its own initial in the label perhaps because it was so small a field (about 6 percent of each S&E bachelor's cohort). Programs of study in Computer Science and Information Technology may fit most neatly under the "technology" umbrella because these are not science fields by the standard definitions, and typically do not involve courses of study similar to engineering.[6]

At the baccalaureate level, it is unclear what the "technology" category encompasses beyond Computer Science and a few small programs that fall under a generic "Information Technology" label. At the sub-baccalaureate level, in a rather remarkable expansion of the definition, STEM has been used to include HVAC[7] and automotive technicians, though not electricians or most other skilled trades[8] (Rothwell, 2013). Following the logic of these expansive STEM definitions – which, as NYU has done with bachelor degree programs, is to ferret out any occupation that has even nominal claim to activities or titles falling under a STEM category – the hidden class of STEM occupations might thus include janitors with the

[5] As examples of these reports' diagnoses, the President's Council on Jobs and Competitiveness (2011) called for 10,000 more engineers each year and President's Council of Advisors on Science and Technology (2012) called for one million more STEM graduates over the decade. Both proposals came shortly after Apple's co-founder Steve Jobs told Obama that Apple would have located its vast low-cost manufacturing operation, with 700,000 jobs, in the United States instead of China if only the company had been able to find enough U.S. engineers to support its operations (Salzman, 2013).

[6] Some computer science programs are in engineering schools. Most often, it is Computer Engineering that resides in these schools and typically requires the same core curriculum required of other engineering majors.

[7] https://www.hvcc.edu/programs/all/stem/heating-air-conditioning-refrigeration-technical-services-aos.html

[8] https://www.azwestern.edu/sites/default/files/awc/institutional-research/Automotive_Techology_STEM_Emphasis.pdf

title “environmental engineer.” Thus, at the Bachelors level, “technology” seems to have only Computer Science and some small technician majors as falling in that category and at the sub-baccalaureate level, it can be used expansively but without particular substantive rationale.

Unlike technology, mathematics is a clearly bounded field of study. In terms of annual bachelor’s degree graduates, mathematics is a fraction of the STEM fields, accounting for about 25,000 of the over 450,000 annual STEM bachelor’s degree graduates (NCES, 2021). However, as we will see below, mathematics’ broad application both within and beyond science and engineering creates a distinct set of issues for its classification.

In this paper, we focus on the current use of STEM classifications in higher education. Specifically, we detail the technical aspects of creating classifications of STEM courses, and thus, of STEM degrees. This classification system is part of a larger project examining the STEM content across the curriculum of higher education and its connection to labor market outcomes. In the next section, we discuss the theoretical boundaries that delineate STEM subject matter. We then address these conceptual definitions to our practical task of coding NCES transcript data (BPS 04/09), and specifically to its course classification index, the College Course Map. We then proceed to detail our method of reclassifying courses in these transcript data, and show how this reclassification affects estimates of STEM course completion/credit earning among bachelor’s degree completers overall and by major area. Finally, we show the significant impact of different classification schemas for estimating the extent of STEM education among college graduates and in the workforce.

What are the STEM disciplines?

Defining STEM coursework first requires delineating the boundaries of STEM. We begin by examining each discipline separately (science, engineering, mathematics); as discussed above, “technology” is not an established discipline and we discuss it separately as a field of study. Drawing on historical and empirical research, we articulate the boundaries of each discipline and then consider the course classifications and the extent to which they do and do not reflect those criteria. While some individual courses and courses of study would fall within all reasonable definitions of a discipline (e.g., introduction to biology as a science course), others can vary depending on where one draws the lines around the discipline for the purpose of the classification.

Science can be defined as the systematic pursuit of knowledge and understanding of the natural and social world (cf. British Science Council, 2006). In its broadest interpretation, science includes all natural, physical, and social sciences and courses in other disciplines that fall under this definition. At least two problems emerge when trying to determine meaningful boundaries for science courses and programs. Though the agency does not offer an explicit definition of the term, the National Science Foundation, which collects data on fields under its funding purview, includes the social or human sciences (e.g., Psychology, Political Science, and Economics). However, more recent discussions of STEM limit its scope to the natural or laboratory sciences (e.g., Biology, Chemistry, and Physics), thus excluding social sciences, and raise the question of what criteria determine the boundary of science.

Classifying Social Sciences. Colloquially, sciences are described as “hard” or “soft”, implying that some fields are more rigorous than others. Whether that “rigor” refers to methodological norms, the reliability of the findings in the field, or the subject matter, the social sciences are typically listed as “soft” and thus at the bottom of the hierarchy of sciences.[9] But empirical evidence suggests that the natural and social sciences are more similar than different in their characteristic features, and that a non-hierarchical view of the sciences is more appropriate. This is especially true at the frontiers of their ongoing research, even if the natural sciences display more agreement on foundational concepts (Fanelli, 2010).

As an example, if science draws its boundary around a method – e.g., falsification – then the social sciences are not distinct from natural sciences. It is more often their content areas that distinguish regions of science (Humphreys, 1990). Thus, social science, by the greater inherent variability of the phenomena studied, leads to less robust and often false findings (Open Science Collaboration, 2015). On the other hand, natural sciences are not without findings that are subsequently shown to be unsubstantiated or wrong (Prinz et al., 2011; Ioannidis, 2005).[10] Indeed, all scientific progress can plausibly be defined as the overturning of once firmly established knowledge (Kuhn, 1962).

However, our objective here is to provide rigor and precision to the analysis of STEM fields, education, and employment. Following popular convention, we restrict our definition to fields of study based in the physical and life sciences. However, using this “hard sciences” delineation to exclude the social sciences in their entirety elides the diversity of subject matter within social science fields. Although sociology might be generally considered outside of the “hard” sciences, a blanket exclusion of all formally defined social science disciplines would sweep out subfields like physical geography, physical anthropology, paleontology, neuropsychology and other fields whose contents and methods fall wholly within nearly all definitions of with the physical and biological sciences. For an accurate and logically consistent analysis of life and physical sciences, we need to identify and

---

[9] Perhaps ironically, it was the philosopher and often regarded “father of sociology” August Comte who suggested that the different branches of science could be regarded in a hierarchical system, ranking sociology at the bottom in its level of precision. Other sociologists followed Comte in distinguishing the hard and soft sciences according to levels of empiricism and on the degree to which they agree on theory and method (Fanelli, 2010).

[10] A 2016 survey by *Nature* found that “More than 70% of researchers have tried and failed to reproduce another scientist’s experiments, and more than half have failed to reproduce their own experiments. Those are some of the telling figures that emerged from Nature’s survey of 1,576 researchers who took a brief online questionnaire on reproducibility in research.” (Baker, 2016). Ioannidis (2005) found a large share of medical research is not supported in replication studies or in detailed assessments of the original study, often suggesting an effect of largely pharmaceutical industry funded research. Although it has been argued that it is pressures from funders to incentives for professional achievement that have led to “bad science” rather than inherent limitations of the scientific method, it does appear that, even when conducted with rigor, the scientific method in the natural sciences is far less certain or reliable than suggested by the “hard/soft” distinction.

account for the courses and programs within the social sciences that are indistinguishable from the formally defined life and physical sciences.

Health Sciences. The classification of "health fields" presents a different set of issues. First, as noted above, the health fields under the purview of the NIH were excluded in NSF's data collection. This was likely a matter or jurisdiction rather than any substantive rationale. Medical researchers are the one category that are included in science,[11] but otherwise health fields are excluded and most official and commonly used STEM classifications continue to exclude them. Most researchers are unaware of this exclusion and, thus, the "leaky pipeline" arguments often cite the loss of science graduates that counts "leaving science" as, for example, those graduating with a degree in biology and entering medical school or nursing. The "leaving science" analysis becomes even more arbitrary when classifying undergraduate majors. In many schools, a student intending to enter medical school earns a biology degree is counted as a science graduate. However, in schools with separate pre-med majors (there does not appear to be any clear differences between colleges that do and those that don't offer a distinct pre-med major), a student graduating with a pre-med or health degree is not counted as a science graduate even though the curriculum is nearly identical.[12] That is, the count of science students can exclude pre-health students if their college happens to be one that offers a distinct pre-health or pre-medicine major option.

Related to this is the classification of allied health fields such as physical therapist, dieticians, and radiology technicians. When these are bachelor's degrees, they include a science curriculum that is similar to general life sciences majors. Although some curricula in the college catalogs indicate more variation than the one or two course difference between pre-med and biology course requirements, it does not appear to be substantial enough to warrant a non-science classification; the variation involves substitution of other courses that have science content and the total science credit difference is small.

The third group of health majors and occupations are those related to health administration and healthcare policy. These are clearly not science fields and their inclusion in "health" appears to result from a classification of health that has no disciplinary foundation or definition. Since health was excluded from STEM (or S&E), and there was little policy attention to fields within non-STEM, these fields received little further scrutiny until recently. Some recent statistical analyses started to take account of the health fields and created a separate classification that would be listed alongside STEM, such as the 2013 US Census Bureau report that distinguished between STEM, STEM-related, and non-STEM occupations (Landivar, 2013).

---

[11] The original text of Public Law 507, which established the NSF, includes basic research in the medical sciences as under the purview of the foundation (National Science Foundation Act, 1950).

[12] From our examination of college catalogs, the difference is sometimes an advanced physiology course instead of a second organic chemistry class or similar variations. We find no substantive differences in the requirements between pre-med and biology major.

While noting health as an adjunct to STEM, this approach does not address the definitional and substantive problems we have just discussed. It moves solidly science-based fields such as physician and nurse into a "related" category while keeping the bachelors-level lab tech in science,[13] and includes as "STEM-related," health administrators, who would have had minimal science course requirements, alongside physicians and nurses. Including health fields in a STEM-related category addresses the problem of excluding practitioners with a life science education equal, or surpassing, many in the officially counted science occupations. But it also includes a large number of health occupations that have little-to-no formal science education while excluding those such as neuropsychologists, archeologists, and others with non-science degrees but, in fact, having life or physical science course credits comparable to those graduating with science majors. Our approach, which classifies individual courses rather than programs, is thus more suited to the task of dealing with the relationship between health fields and STEM.

Technology is perhaps the most difficult element of the acronym to define. While it could plausibly refer to academic programs such as computer science, others interpret "technology" as referring to pedagogical approaches that cut across STEM and non-STEM fields. This raises the question whether the T in STEM is meant to be silent (Path, 2016). However, there is also a body of literature that attempts to distinguish technology from science in general, and from applied science specifically (Feibleman 1961, Layton, 1974, Frey 1991). Thus, using the STEM classification rather than S&E, analysis should include technology courses exist across disciplines as distinct from courses in both Science and Engineering.

Engineering. Definitions of engineering vary far less than the other STEM acronym components; the common definition of engineering focuses on the application of mathematical and scientific principles to the design of tools and the effective use of resources. Importantly, the distinction between engineering and engineering technology is that the latter focuses more on the use of tools and techniques, while the former is concerned with the creation of those tools. Definitions of engineering generally address this distinction. The Engineers Council for Professional Development (ECPD), the predecessor to the modern Accreditation Board for Engineering and Technology (ABET)—defined the field as follows:

> "The creative application of scientific principles to design or develop structures, machines, apparatus, or manufacturing processes, or works utilizing them singly or in combination; or to construct or operate the same with full cognizance of their design; or to forecast their behavior under specific operating conditions; all as respects an intended function, economics of operation and safety to life and property." (ECPD, 1947)

[13] The criterion of "practitioner" as differentiating physician from lab tech also lacks consistency considering that x% of science graduates working in "laboratories" are processing medical tests such as blood samples and thus not engaged in work falling under the definition of science. Further, many medical practitioners working in university-affiliated hospitals often participate in medical science research as part of their work but are excluded in the counts of medical researchers.

Because engineering courses are much less likely to be taught outside of engineering departments, their classification is more clear-cut than that of other STEM courses. One logistical issue for classifying coursework is the overlap between engineering and other fields like the natural sciences (e.g., chemical engineering), or computer science (e.g., computer or software engineering, or electrical engineering). We address the issue of overlapping classification below as a practical, or classification matter, but these variations in course offerings or majors would not affect inclusion within the overall STEM classification (i.e., whether or not computer science is classified as engineering or technology, it would still be considered "STEM").

Mathematics is defined in a number of ways, and its status relative to science is debated among professionals, even as its importance to scientific practice is unquestioned. Nonetheless, existing definitions of mathematics seem to all share an emphasis on symbolic manipulation and its application to questions of quantity, change, and relation.[14] As a category of coursework in schooling, mathematics usually includes statistics and sometimes encompasses logic and quantitative reasoning.

Within postsecondary education, it would be practically easier to limit this category to courses taught in mathematics departments. However, if the substantive goal is to identify the extent of training in formal mathematics and/or statistics and/or quantitative reasoning, it is important to include the range of such courses taught outside of disciplinary mathematics departments. Moreover, because of the integral nature of mathematics to other disciplines, from economics to experimental psychology to physics, restricting an accounting of mathematical education to only courses taught in mathematics departments or by mathematicians has little substantive justification. For example, there seems to be no substantive rationale for considering an introductory statistics course taught in the mathematics department as a mathematics course, but not one taught in economics or sociology. Although some contend that a statistics course taught within disciplinary mathematics would provide more mathematics content and less applied content, this is a weak argument. On the one hand, math departments often teach statistics courses for non-mathematicians with the content adjusted accordingly. Furthermore, at the undergraduate and introductory levels, differences between these two sorts of statistics courses are unlikely to substantively differ if the goal is to account for course content falling within nearly all definitions of "mathematics."

---

[14] One definition – from the Encyclopedia of Mathematics – reads as follows: "The science of quantitative relations and spatial forms in the real world." (Encyclopedia of Mathematics, nd). The National Center for Science and Engineering Statistics also supplies a definition of mathematics - combined with computer science - that emphasizes symbolic manipulation and logical reasoning: "employs logical reasoning with the aid of symbols and is concerned with the development of methods of operation using such symbols or with the application of such methods to automated information systems." (NCSES 2021).

Thus, at the operational level, mathematics can include a continuum of content that ranges from quantitative reasoning to advanced calculus. Mathematicians, philosophers (the original mathematicians), and curriculum designers are likely to debate the boundaries of what is, and what is not properly regarded as "mathematics". However, for most purposes in policy, funding, and educational assessment, the core concern is the extent of education in quantitative thinking (also a rather imprecise characteristic) or mathematical reasoning (some regard these as largely similar), computational abilities and level of mathematical education to infer capabilities for particular occupations and types of tasks.[15]

There is no STEM classification or census that, to our knowledge, consistently or logically captures the extent of mathematics course taking based on substantive criteria. Existing classifications primarily base the decision on the department in which the course is taught, and sometimes on the major of the student taking the course. NCES, for example, has used the major of the student enrolled in some mathematics or statistics courses to determine whether the course is classified as mathematics. In our analyses, we do not offer any *a priori* criteria for a mathematics classification schema but, instead, include the widest range of courses with identifiers that allow adjustments to reflect the analytic purpose.

"STEM-X." To this point we have conceptualized STEM by referring to its constituent parts. But, as education researchers and practitioners have pointed out, some topics, skills, and themes are relevant across STEM fields.[16] While clearly designating the boundaries of each element of the STEM acronym is an important step in creating a robust classification schema, many of the actual observations we classify – here college courses – will inevitably straddle these boundaries. Courses with titles like "Biostatistics," "Medical Information Systems," and "Data Analytics" immediately suggest the overlapping nature of undergraduate coursework. Thus for classification purposes, it is necessary to have a distinct category for courses with content that crosses two or more of the acronym elements. In our classification system, we assigned such courses the label STEM-X.

Summary. STEM classification depends on the definitions of each of the component fields as well as consideration of whether "STEM" as an aggregate category includes fields that may not fit into a single disciplinary definition. As we will discuss below, an aggregate STEM classification could include "technology" courses that have no traditional disciplinary home and to mathematics in which subject matter spans disciplines or involve quantitative reasoning that falls outside most formal definitions of mathematics.

More broadly, the purpose for which researchers will use a given classification of STEM should inform the criteria used to develop the detailed schema or decision rules. Further, the schema may differ depending on whether one is classifying STEM courses, majors, the overall content of a program of study (i.e., as distinct from formal major), occupations, or industries. Depending on the purpose for a particular accounting of STEM courses, the

[15] Though the need for basic numeracy at work is apparent, the extent to which formal and advanced mathematics is a relevant work skill is a subject of debate (Douglas & Attewell 2017; Handel 2016).
[16] See, for example, the Next Generation Science Standards, which include eight "Cross-Cutting Skills" (National Research Council, 2013).

inclusion or exclusion of some categories may be warranted; however, without a transparent and consistent classification schema, it is not possible to modify or create specific categories with validity.

We now turn to the currently used classification schema for courses and programs of study, and then proceed to outline our approach to classifying coursework.

Existing Program and Course Classifications in Higher Education

Course classification depends, first, on operational definitions of each classification but also on the process by which these data are reported and compiled. This, in turn, is influenced by the organizational objectives of both the reporting institutions of higher education and the data collection and analysis organization which consists of the Department of Education's National Center for Education Statistics (NCES) and their contractors who increasingly are responsible for larger shares of the data definition, collection, compilation, and analysis activities.

Program and course classifications are typically used for statistical reporting of the supply of STEM workers based on the numbers graduating with a given major, or of courses to assess the extent of components of STEM education (e.g., the number of mathematics or biology courses taken by undergraduates). NCES is the agency responsible for collecting data from all post-secondary institutions. It is the primary source for official government higher education statistics and the most widely used educational statistics by researchers and policy analysts. Post-secondary course classification is based on the College Course Map (CCM) (Bryan & Simone, 2012). The CCM was intended to provide a classification schema with a strong empirical foundation and clear rationale that was consistent with the overall schema of disciplines and programs in postsecondary education. Thus, to understand how course classification developed, it is necessary to begin with the classification of academic programs.

The Classification of Instructional Programs (CIP) was first developed by the US Department of Education in the 1980s.[17] The CIP was intended to "assist in collecting, reporting, and interpreting data about instructional programs" (Malitz 1987, pg. 1). Specifically, CIP was meant to aid postsecondary institutions in reporting, by field, the number students studying and the number of degrees earned, to the federal government. The authors of the CIP system immediately identified two caveats for users of CIP data. First, since CIP data are compiled by institutions, rather than by a centralized government entity, one should not assume that all postsecondary institutions use the codes in exactly the same way. Second, and perhaps more importantly, institutions use the CIP system to classify instructional programs, not individual courses (Malitz 1987, pg. 3). That is, the CIP was intended to classify an instructional program as a whole rather than its constituent parts. In its entirety, an instructional program could be differentiated from other majors by

[17] CIP was the first NCES classification system that attempted to *describe and classify* higher education programs. An earlier system, the Higher Education General Information System (HEGIS) taxonomy provided a list of titles, but did not make any attempt to describe the programs (Malitz, 1981).

educational content, which would involve a range of different disciplinary courses. For example, a highly quantitative economics major might require mathematics courses that provide a level of quantitative reasoning and computational ability similar to a mathematics degree (the typical bachelor's degree in mathematics requires 40-50 course credits of mathematics) but, in its entirety, the instructional program would be differentiated by the context, application, and broader body of knowledge provided in the major.[18]

Practically, this distinction became an empirical concern for the Department of Education when college transcripts were studied in one of the first longitudinal educational surveys – the National Longitudinal Survey of the High School Class of 1972 (NLS-72). To classify the college courses taken by this cohort of students, researchers originally relied on the CIP system. But the CIP was not well-suited to classifying individual courses. This initial transcript coding effort classified up to 40 percent of courses incorrectly. Clifford Adelman, then a senior associate at the US Department of Education's Office of Research, was charged with devising a more appropriate system of course classification. Adelman was aware of the necessity of having accurate data on course taking among US undergraduates, writing that:

> "We can easily provide information on how many Associate's degrees were awarded in Nursing…[the data] doesn't tell us, for example how much science the Nurses studied or what kind of science it was, hence how well Nursing graduates are prepared to adapt to changes in biomedicine." (Adelman, 1990:1)

Thus, he sought to create and use a more careful process to code the courses in the NLS-72 transcript data that reflected the disciplinary content of courses. To begin with, Adelman revised the CIP system to align with the task of course coding. The process started by deleting redundant categories, and creating new ones when necessary, with the basic principle that a course category should only exist if there are enough occurrences of that course to be considered a distinct field of study but not so narrow that it referred to only a few courses or small segment of credits.[19] Category creation also relied on external sources, including sub-fields indicated by disciplinary journals, curriculum statements from professional associations, and consultation with experts in the fields, including focus groups to find consensus on disciplinary boundaries and content (Adelman, 1990).

Once the categories were revised, the actual coding process adhered to the notion of "primacy of subject"; a course was categorized primarily based on what was taught, and secondarily on the department of instruction. The intention was to use course classification as means of understanding the types of subject matter students have taken by the time they leave college and enter the workforce. Course coding also utilized data about the individual students – e.g., major and class year – to more accurately classify the courses they took. This was particularly important for courses with ambiguous titles – which we

---

[18] For further discussion of the dimensions of mathematics course taking, see Douglas & Salzman (2020).
[19] The specific threshold of 22 course occurrences was empirically derived by the research team (Adelman, 1990: p. 4).

discuss below. Adelman did much of the work himself – to ensure the consistency and reliability of the process – but consulted with subject experts to resolve questions related to course content.[20] Their effort produced the first College Course Map, or CCM, in 1990 (Adelman, 1990).

The CCM retained the basic CIP taxonomy—6-digit detailed codes, with 2-digit general fields. But, as detailed above, the classification is based largely—though not entirely—on course subject matter rather than the discipline of the department or instructional program in which it is taught. Periodically, the map is revised to reflect emergent fields and course subject matter. The CCM was first updated in 1995 using transcript data collected from NCES's High School and Beyond 1982 survey (Adelman, 1995). The 2003 revision corresponded with transcript analysis from the National Educational Longitudinal Survey of 1988 (Adelman 2003, 2004). After the 2003 update, Adelman left the Department of Education and subsequent revisions were done by Department of Education staff and contractors. The most recent CCM – finished in 2012 – corresponds to the transcript analysis from both the Beginning Postsecondary Students (BPS) 04-09 and Baccalaureate and Beyond (B&B) 2009 surveys (Bryan & Simone, 2012). As the transcript data from the newest NCES postsecondary survey – the BPS 2012-17 – are analyzed, we can expect a new update to the College Course Map.

The CCM coding process is, relative to CIP coding, more centralized, and more labor-intensive. In the most recent completed transcript coding (alongside the BPS 04/09), this included the development of task-specific keying and coding software, a 5-day training seminar for 71 staff members, and frequent quality control meetings and spot checks over the course of the 18 months of coding.[21] In the BPS 04/09, NCES requested and coded more than 25,000 transcripts received from over 2,600 postsecondary institutions, a total of nearly 600,000 courses.[22]

CCM course codes generally reflect the disciplinary home of a course. For example, a biology course taught in a biology department receives a general, two-digit code for biology; the detailed or four-digit classification is based on course content (e.g., human anatomy, microbiology, etc.). Coding of a specific course rather than instructional program requires a detailed analysis of course content when the title does not indicate a clear disciplinary home or detailed content specifications. In the documentation for the original CCM, Adelman (1990) offered the example of a course titled "Human Growth and Development" which was observed in Biology, Psychology, Anthropology and Education. NCES coders encountering such a course would use information in the college's course catalog to ascertain the detailed content, and information about the students' program of study to identify the proper disciplinary code for the course. This course, if offered by the

---

[20] Portions of the course coding process were sponsored by the NSF and the Mathematical Sciences Education Board. As such, classification of STEM coursework was done with some assistance (Adelman, 1990).

[21] This process is described in detail in the chapter 4 of the BPS 04/09 methodology report (Wine et al., 2011). Notably, the coding of courses is done by contractors rather than NCES staff.

[22] According to the methodology report, coding of an individual transcript typically took between 40 and 80 minutes (Wine et al., 2011).

biology department would be coded a biology course whereas the same course title for a course offered by the psychology department would be coded as a psychology course.

Some courses involve consideration of additional criteria for classification, such as a course simply titled “Statistics”, which could be taught in any number of departments. With such a course, coding staff would use the college’s course catalog and information about the student to ascertain whether this course would be coded 13.0603 (Education Statistics and Research Methods), 14.9995 (Engineering Mathematics, Engineering Statistics…), or 27.0501 (Statistics, General). Student major is also used to classify cross-listed courses to consistent with the student’s major. In the above example, a statistics course cross-listed in both mathematics and education departments would be assigned the Education Statistics code (13.0603) if the student is an education major, while it would be coded as General Statistics (27.0501) for a mathematics major in the same class. Although this procedure may be a reasonable process to infer content when the title and course description are ambiguous, such as “Human Growth and Development,” this process can create imprecision for analysis of the content of a student’s coursework if the same course is classified differently depending on the major of the student. In some cases this creates significant distortion or noise, as when the same statistics course is classified differently for different students, implying they took different courses or received different disciplinary content, as discussed below.

The CCM system described above serves as a platform for our analysis of STEM content on the basis of course taking. The well-documented and centralized process gives us more confidence than the decentralized and institution-specific means by which students degrees are assigned CIP codes. But NCES’ variable derivation process – by which CCM coded courses are aggregated into broad classifications like “STEM,” or narrower groups like “advanced mathematics” – had, based on our investigation, substantive shortcomings that could distort analyses that used this coding.

Problems with NCES definitions of STEM coursework

The data files for the BPS 04/09 PETS, the first set of transcripts we analyzed in detail, include derived variables indicating the number of STEM *courses* taken by students at various intervals (e.g., first year, first two years, etc.). When we examine the definitions and source codes of these variables, the BPS derived variable for STEM *courses* is based on the list of eligible *degree programs* for the National Science and Mathematics Access to Retain Talent (SMART) Grant Program, which existed from 2006 to 2011.[23] The variable excludes foreign language programs (which were SMART eligible), and includes additional statistics courses not taught in SMART eligible – STEM – fields. Though SMART contains a readily available list of CIP, and thus CCM, codes, this raises two issues. The first is coverage: the SMART list excludes some programs that would generally be considered STEM (e.g., 31.0505 –Kinesiology and Exercise Science), and includes others that would generally be

[23] Information on the now discontinued SMART program can be found at: https://web.archive.org/web/20111114072015/http://www.studentelligence.com/federal-grants/smart-grant/national-smart-grant.html

considered non-STEM (e.g., 42.2705 – Personality Psychology; 11.1006 – Computer Support Specialist). The second issue is the larger course/program of study distinction discussed above. Insofar as the SMART list generally considers only the four "core STEM" fields of study – Biology, Engineering, Mathematics, and Physical Sciences – it ignores many courses with substantial STEM content that are taught outside these fields.

Technical Issues with NCES Transcript Data

Our inquiry into the BPS 04/09 transcript data also revealed some technical issues with the transcript variables. These issues fall into three general categories. First, because NCES coded transcripts submitted from all institutions attended by a given student, some courses taken by transfer students were double-counted: once from the sending institution's transcript and again from the receiving institution's transcript. This problem is non-trivial given that 60 percent of students in the BPS 04/09-PETS sample attended more than one institution (Wine et al., 2011). Second, because the NCES count of transcript courses is meant to be comprehensive of all attempts, repeated courses were double-counted in cases that a passing grade was received in more than one attempt. Both of these could lead to over-counts of the unique earned credits, particularly distorting assessments of STEM course attainment. For example, a student receiving a "D" in their first enrollment in a Calculus I class and a "C" in their second enrollment would appear to have twice as many calculus credits as the student who received an "A" in their first and only enrollment in Calculus I. Repeated course taking to improve an initially poor passing grade (failed courses received no credit so would not be reflected in course credit variables) was concentrated in several STEM courses such as Calculus and Organic Chemistry, and in certain non-STEM courses such as Freshman composition and introductory history (American or European).

A final issue concerns the accuracy of the student's major listed in the record. The BPS uses the major listed in the transcript (and the school record before graduation) and the student's survey response. There were a substantial, though not large [? What % did we find problems?] number of cases where the major was missing or inconsistently reported in the administrative record and the survey report, and some cases where the listed major was inconsistent with the coursework (e.g., a biology major who did not have the required number of science courses and/or insufficient course credits reported). We developed a process to resolve most of these cases and recoded as "missing" those cases we could not resolve, which we describe in Appendix A.

The initial CIP classification achieved a reasonable classification of instructional programs that, although had some structural inconsistencies, did not appear to lead to significant distortions except in some specific fields that had small enrollments. Over time, however, the development of new interdisciplinary majors and the change in content (e.g., increased mathematics and quantitative courses in non-STEM courses) along with efforts to differentiate a group of courses as the poorly defined "STEM" agglomeration rather than specific disciplines, led to greater distortion of the CIP classification. When applying the CIP classification to courses in the development of the CCM, the initial development effort was conducted with expertise and rigor, though the method and difference between courses and instructional programs, as discussed above, led to greater noise in the CCM than in the CIP. Similarly to the CIP, the increased emphasis on differentiating course content using the STEM classification amplified the initial noise and subsequent development of the classification, and use of the DHS classification schema in particular, led to quite significant distortion. Other issues, as discussed above, combined to undermine the practical utility of the STEM classification of courses with particular consequences for some subjects such as mathematics (see Douglas and Salzman, 2020; and discussed in the next section).

Some noise in large datasets such as the BPS is to be expected. But the inconsistencies and errors we found, alongside some inherent limitations of the CIP and CCM classification process, and the lack of construct validity of the "STEM" and several other classification variables, limit the utility of readily available data for the purpose of characterizing undergraduate STEM education.[24] The strength of the BPS is the exhaustive and extensive raw data collection, the rigorous study design, data collection fieldwork, and sample size such that it was possible to develop first-level/primary classifications of many key variables, resolve conflicts, and develop classifications and variables with robust internal and construct validity.

Resolving data and variable inconsistencies and errors allowed us to develop a reliable dataset for our analysis, as discussed in the following section. The revised, robust dataset then provided the basis for developing constructs that had greater validity, such as "STEM" course credit hours that reflected the usually intended purpose of that construct, or variable. For example, when measuring mathematics education as measured by number of course credits, most analyses would not intend to measure students with repeated course taking as higher mathematics achievement than students who completed the course on their initial course enrollment. Other analyses might focus on credits taken for other purposes, and we provide a means to differentiate course taking and outcomes, as discussed in the next section.

Our development of the BPS dataset (and later the B&B dataset using similar procedures and constructs) involved first reviewing and correcting the primary data such as course credits, subject classification, student major, and other variables. We then developed variables to reflect constructs such as STEM course taking and majors that reflected the intended concept in accordance with specific definitions. Some of the concepts do not have a single or established operational definition; in those cases, we articulate one or more definitions and corresponding construct in our analysis. By clearly defining each construct and the variables, we provide the means for other researchers to use different definitions and analyses.

[24] For example, the National Academies of Science, Engineering, and Medicine (2018) suggest that the BPS (or a more frequent data collection modeled on the BPS Transcript data) could be used to measure course-to-course progress through STEM fields (pg. 168). Such an indicator would need to be grounded in a valid definition of STEM coursework.

The focus of this analysis is on STEM course taking and completion (although the dataset and our revisions provide the means for other analyses of college pathways, which we are conducting. Importantly, by correcting classification errors we provide the means for researchers to establish STEM definitions that reflect their research objectives and then develop a STEM classification that has construct validity – that is, a measure that provides a more accurate reflection of the phenomena (courses, majors) that correspond to their STEM definition.

Data and Methods

To assess the extent of STEM education we focus on the types of coursetaking of college students who complete a bachelor's degree within six years of first enrollment (this is the criterion used in the BPS dataset and completion data beyond six years is not available). The two central tasks were: (1) developing a logically consistent definition of STEM that reflected the definitions of each discipline and considering fields that could be considered "STEM" as an aggregate category outside of a specific component discipline; (2) accurately counting course credits that eliminated erroneous duplicate reporting in the administrative records and that accounted for repeated course taking of the same course.

Guided by our definitions of the components of the STEM acronym, the detailed contents of the CCM, and addressing the technical issues of double-counted courses, we undertook a comprehensive recount of the BPS 04/09 PETS data file, with the goal of more accurately counting the number of STEM courses taken by US undergraduates. Our name for this recount is The "Comprehensive Net Course Count" (CNCC). This section discusses the rationale for determining the criteria for defining STEM courses, assessing the data from the administrative and survey records, and the process for coding the variables.

Coding Approach

The first step in this process was a detailed reading and coding of the CCM. We assessed the portion of each CCM entry that outlined the usual learning goals of the course. Based on this review of course code entries, each 6-digit course code in the CCM was classified as either (N)on-STEM, or as one of the elements of the STEM acronym - (S)cience, (T)echnology, (E)ngineering, or (M)athematics.

As noted above, attempting to fit all STEM courses neatly into S, T, E, or M ignores the reality that a given course's content could appropriately place it in more than one category. We thus classified courses as STEM-(X) if they were cross disciplinary, containing learning goals that met the criteria of more than one of the acronym elements.

The science (S) classification we used in our coding of courses differentiates between social and the physical and life sciences, following the NSF's distinction.[25] It also attends to the

[25] The National Center for Science and Engineering Statistics, NSF's statistical agency, provides a single definition of Social Sciences: "*Social sciences* is concerned with an understanding of the behavior of social institutions and groups and of individuals as members of a group." Unfortunately, it does not provide a single

gaps (mentioned above) in STEM course counting within social science subfields. Similarly, when examining health, we classify practitioner courses based on their science and technology content.

Because statistics courses are methodology requirements in most science disciplines, as well as in business, engineering, education, and some health programs, we examined CCM entries to identify these courses. Further, most colleges offer quantitative reasoning courses, primarily aimed at liberal arts and humanities students, which may fall outside standard definitions of either mathematics or statistics. For our comprehensive classification of mathematics courses, we coded these as (M) as well. Since courses retain their detailed coding, different criteria could be used to create variables reflecting different definitions or analytic purposes; our purpose is to assess mathematics courses that encompass the full range of credible definitions, but researchers could use our classification schema to analyze "mathematics" as encompassing a less comprehensive set of course types. Further details on our coding are provided in Appendix C.

Refining the Transcript File

The CNCC metric assesses the number of unique course completions and excludes multiple completions of the same course at the same school, as well as duplicate credits reported for a single course as often occurred in transcripts for students who transferred schools. As noted above, this entailed removing all twice-counted courses resulting from multiple transcripts and repeated course attempts. Our approach to resolving double-counted courses is described further in Appendix B.

Data Analysis

The first analysis presented in this paper (Table 1) demonstrates the impact of our revised classification on aggregate student transcript measures as compared to the existing derived variables in the BPS 04/09 PETS file. To create a uniform frame of reference, our analysis focuses on students who completed bachelor's degrees. It also examines how the CNCC method impacts the observed average number of STEM and mathematics credits (Tables 2 and 3, respectively) earned by students who completed bachelor's degrees in different disciplines.[26] Density plots provide graphic representation of the distribution and shares of students in STEM and non-STEM majors with various levels of STEM credit earning (figures 1, 2 and 3).

Results – Comparing Methods of STEM Counting

definition of natural sciences, though the listed fields (other than mathematics) imply that natural sciences refer to: Environmental Sciences, Life Sciences, and Physical Sciences (NCSES, 2021). We use the designations of social, life, and physical sciences to differentiate broad groupings of the sciences.

[26] NCES accurately classifies major field of study for most students. But there were a number of students in the data whose major was not immediately clear. We discuss our process for classifying unidentified or inconsistent student majors in Appendix C.

Table 1. Overall Comparison of Credit Counts among Bachelor's Degree Completers, NCES and CNCC definitions

	Summary Statistics							
	Mean	Standard Deviation	Median	25th %ile	75th %ile	Range	N (weighted)	N (unweighted)
NCES Derived Variable Count								
Total credits	129.8	29.1	129.0	121.0	142.0	298.0	881,250	5,110
STEM credits	32.8	31.8	19.0	12.0	42.0	174.8	870,420	5,040
Mathematics credits	9.2	8.2	6.0	4.0	11.0	81.8	803,420	4,600
Science credits	17.9	20.7	9.0	6.0	20.0	139.3	820,660	4,730
Engineering Credits	24.3	25.2	10.0	3.0	46.0	112.5	123,750	720
Comprehensive Net Course Count (CNCC)								
Total credits	128.5	28.6	128.0	120.0	141.0	298.0	881,150	5,110
STEM credits	37.3	32.1	24.0	15.0	50.0	185.0	874,640	5,060
Mathematics credits	10.3	8.2	9.0	6.0	13.0	81.8	827,520	4,760
Science credits	17.9	20.4	9.0	6.0	20.0	132.6	822,360	4,750
Engineering Credits	24.2	25.1	10.0	3.0	46.0	112.5	123,750	720
Technology credits	10.9	14.8	5.0	3.0	11.0	101.5	474,850	2,700

Source: BPS 04/09 PETS data, Authors' calculations.
Note: N sizes coarsened.

Table 1 presents weighted descriptive statistics for the number of total, STEM, Mathematics, Science and Engineering credits among Bachelor's Degree completers in the BPS 04/09 survey sample. The top panel presents estimates based on NCES-derived variables. The bottom panel presents estimates based on our generated CNCC variables. Comparing the two panels, we note that the CNCC measures result in a slightly lower number of total earned credits (128.5 vs. 129.8) among all students, which we attribute to the reduced number of double-counted courses. But the CNCC also shows a markedly higher number of total STEM credits earned (37.3 vs. 32.8), and particularly of mathematics credits (10.3 vs. 9.2). The observed numbers of Science and Engineering courses were not substantially altered by the CNCC method. By applying clear definitions of elements of the STEM acronym, our approach thus identifies more STEM courses taken by undergraduates, particularly in Mathematics and Technology.

Table 2. STEM Credits Earned by Student Major among Bachelor's Degree Completers, NCES and CNCC Definitions

STEM credits earned	NCES		CNCC			
	Mean	SD	Mean	SD	Mean Diff	N (weighted)
STEM majors						
Agriculture/Natural Resources	57.3	26.7	64.8	26.8	7.5	15,080
Architecture/Eng Tech	47.6	37.1	86.3	28.3	38.7	10,100
Computer Science	74.3	27.1	77.5	26.3	3.3	19,650
Engineering	106.6	26.3	107.1	26.5	0.5	50,480
Life Sciences	78.4	21.5	79.7	21.7	1.3	60,250
Mathematics & Statistics	70.7	23.3	71.7	23.7	1.0	11,430
Physical Sciences	79.1	22.4	79.3	24.1	0.2	14,700
Total STEM	**81.9**	**30.6**	**85.7**	**28.4**	**3.8**	**181,930**
Non-STEM Majors						
Business	19.4	11.3	25.0	12.1	5.6	151,840
Communications	14.0	7.7	17.6	9.6	3.7	50,650
Education	19.7	12.7	20.4	13.0	0.8	67,300
Health	38.0	20.1	48.3	29.3	10.3	52,330
Humanities	15.2	11.9	19.9	15.1	4.7	132,770
Social Sciences	17.1	12.3	21.8	13.3	4.8	166,220
Total non-STEM	**19.0**	**13.9**	**23.9**	**17.0**	**4.9**	**623,830**
Other/Unknown	30.7	21.2	35.1	22.9	4.4	56,280
Total non-STEM+Other/Unknown	**19.9**	**15.0**	**24.6**	**17.8**	**4.8**	**692,710**

Source: BPS 04/09 PETS data, Authors' calculations.
Note: N sizes coarsened.

Table 2 compares CNCC and NCES estimates of earned STEM credits by student major among bachelor's degree earners. The top panel of the table presents STEM major categories, and the bottom panel presents non-STEM major categories. Among STEM majors, the number of STEM credits earned is virtually unchanged between the NCES and CNCC measures. The exceptions to this general observation are Architecture and Engineering Technology majors, whose number of STEM credits increased from 48 to 86, and Agriculture/Natural Resource majors, whose number of earned STEM credits increased from 57 to 65. Architecture was not part of the SMART program list used to create the BPS derived STEM course variable. This explains the sharp increase in STEM credit counting in the CNCC measure for graduates in this field of study. The substantial increase for Agriculture/Natural Resource majors may be a result of more careful counting of environmental science courses in the CNCC measures.

Among non-STEM graduates, the coding and classification corrections led to substantial differences in STEM credits. Nearly all non-STEM major categories saw an increase of 3.5 to 5.6 STEM credits using the CNCC method. Particularly large differences among business (+4.2) and humanities and social science (+4.7) majors may be due to specific course classification shifts in these fields. Health majors show the largest differences between the two methods because many health courses are similar to disciplinary science courses but, because "health" is classified as a non-STEM course and major, even some disciplinary science courses taken by a health major may be coded as non-STEM. The CNCC method more accurately classifies these courses which dramatically increases their observed number of earned STEM credits (+10.3). It is important to note that the "net" count offsets the increase in unique courses coded by eliminating repeat course records, a phenomenon which is more common for calculus and organic chemistry. The NCES count would include both instances of a repeated calculus I class, but would not count a statistics class taught outside of the mathematics department or, in some cases, a statistics course offered by mathematics department but taken by a non-STEM student.

Table 3. Mathematics Credits Earned by Student Major among Bachelor's Degree Completers, NCES and CNCC Definitions

	NCES		CNCC			
Math & Stats credits earning	Mean	SD	Mean	SD	Mean Diff	N (weighted)
STEM majors						
Agriculture/Natural Resources	8.7	4.3	9.2	4.5	0.5	14,590
Architecture/Eng Tech	10.8	9.2	11.3	8.1	0.4	8,590
Computer Science	16.4	9.8	17.5	9.5	1.1	19,500
Engineering	21.8	7.2	21.9	7.1	0.1	50,020
Life Sciences	9.2	4.6	10.0	4.6	0.8	58,760
Mathematics & Statistics	46.8	13.4	47.0	13.3	0.1	11,430
Physical Sciences	15.6	9.7	16.2	9.9	0.6	14,600
Total STEM	**16.5**	**12.2**	**17.0**	**11.9**	**0.5**	**178,060**
Non-STEM Majors						
Business	8.8	4.6	11.7	5.3	2.8	147,460
Communications	5.3	3.1	5.9	3.4	0.6	46,720
Education	9.4	8.4	9.5	8.3	0.1	61,870
Health	7.0	4.1	7.9	4.3	0.9	48,780
Humanities	5.8	4.2	6.3	4.4	0.6	107,880
Social Sciences	6.1	4.0	7.7	4.4	1.6	149,290
Total non-STEM	**7.1**	**5.0**	**8.5**	**5.5**	**1.4**	**584,140**
Other/Unknown	7.7	5.0	8.3	5.0	0.7	55,120
Total non-STEM+Other/Unknown	**9.2**	**8.3**	**8.4**	**5.5**	**-0.8**	**649,460**

Source: BPS 04/09 PETS data, Authors' calculations.
Note: N sizes coarsened.

Table 3 compares CNCC and NCES estimates of earned Mathematics credits by bachelor's degree earners arrayed by student major. The overall observation is similar to that in Table 2. The switch from NCES to CNCC measures does not dramatically affect the average numbers of mathematics credits earned by STEM majors. Among non-STEM majors we see more pronounced increases, specifically among business (+2.8) and social science (+1.6) majors. These observed changes are likely attributable to the more thorough counting of business and social science statistics and research methods courses within the mathematics category. The small but positive increases among nearly all majors – STEM and non-STEM – may be due to better counting of pre-calculus courses, which were not classified accurately in the NCES counts.[27] It is important to note that the aggregate counts may not reflect larger changes among specific subgroups in the population as, for example,

[27] In the BPS 04/09 transcript file, four NCES-derived variables count mathematics credits earned. One counts the number of pre-collegiate math credits earned (typically remedial courses); the second counts college-level courses below the level of Calculus I; the third counts calculus and other advanced math courses, the fourth counts applied mathematics and statistics courses. The CCM code for pre-calculus (27.9996) was not included in any of these derived variables.

lowering the mathematics course credits for lower performing students who repeated courses which offset the significant increases in mathematics courses when including statistics courses taught by faculty outside of the mathematics department and/or to non-STEM students.

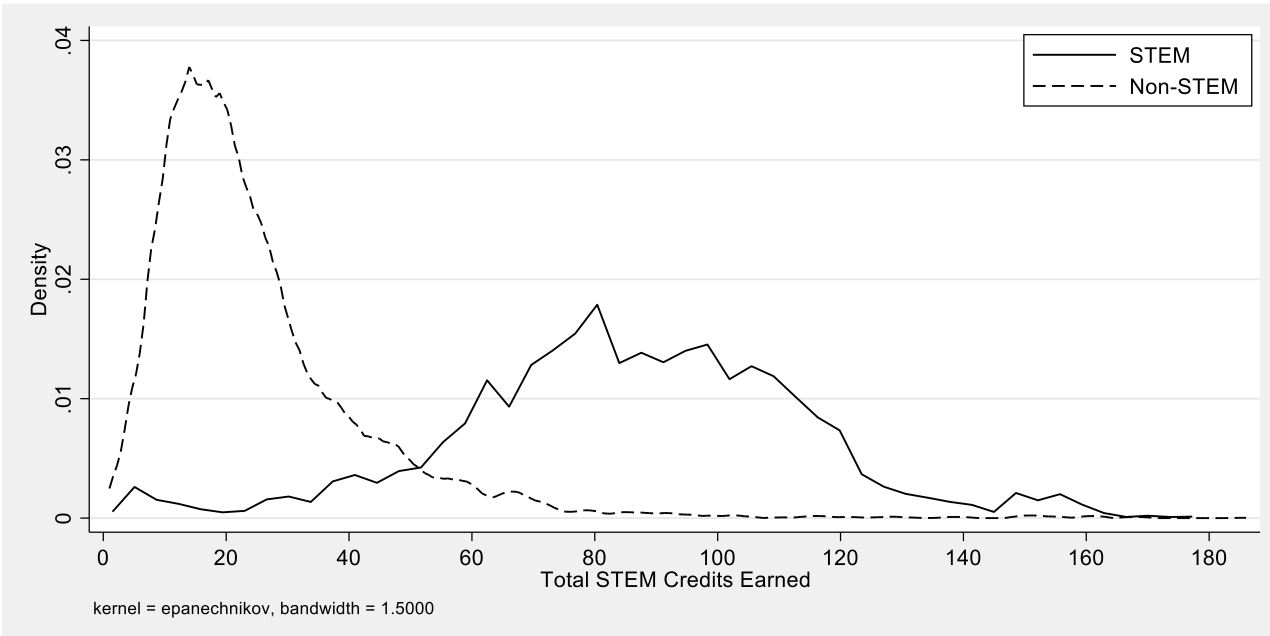

Figure 1. STEM credits earned by broad major category, CNCC definition

Source: BPS 04/09 PETS data, Authors' calculations.

To provide a fuller picture of STEM credit earning among US bachelor's degree holders beyond central tendency statistics (Tables 1 and 2), Figure 1 plots the distribution of STEM credits earned. Predictably, the peaks of these distributions do not overlap; students who earned STEM degrees earn far more STEM credits on average than those with non-STEM fields. We note two important observations from this graph. The first is that STEM degree holders do vary widely in their number of earned STEM credits. The clearest peak is around 80 credits, but the larger share of students cluster between 60 and 120 credits.

The second observation concerns the distribution of STEM credit earning among non-STEM degree holders. While this plot peaks around 20 STEM credits earned, it has a heavy positive or right-tail skew with substantial density through 60 credits, and a tail extending through 180 credits. Simply put, many non-STEM degree holders have completed substantial STEM coursework. Importantly, there are far more non-STEM bachelor's degree graduates, so the density plot for this non-STEM group represents far more students in terms of absolute numbers. That is, the density plot shows the distribution of each population but does not show the numeric differences. The STEM graduate population is approximately 21 percent of the cohort, and thus non-STEM population is 3.7 times larger. Specifically, the density plot showing .01 on the y-axis for a given credit level would equal about 1,800 STEM students and about 6,800 non-STEM students with that credit level.

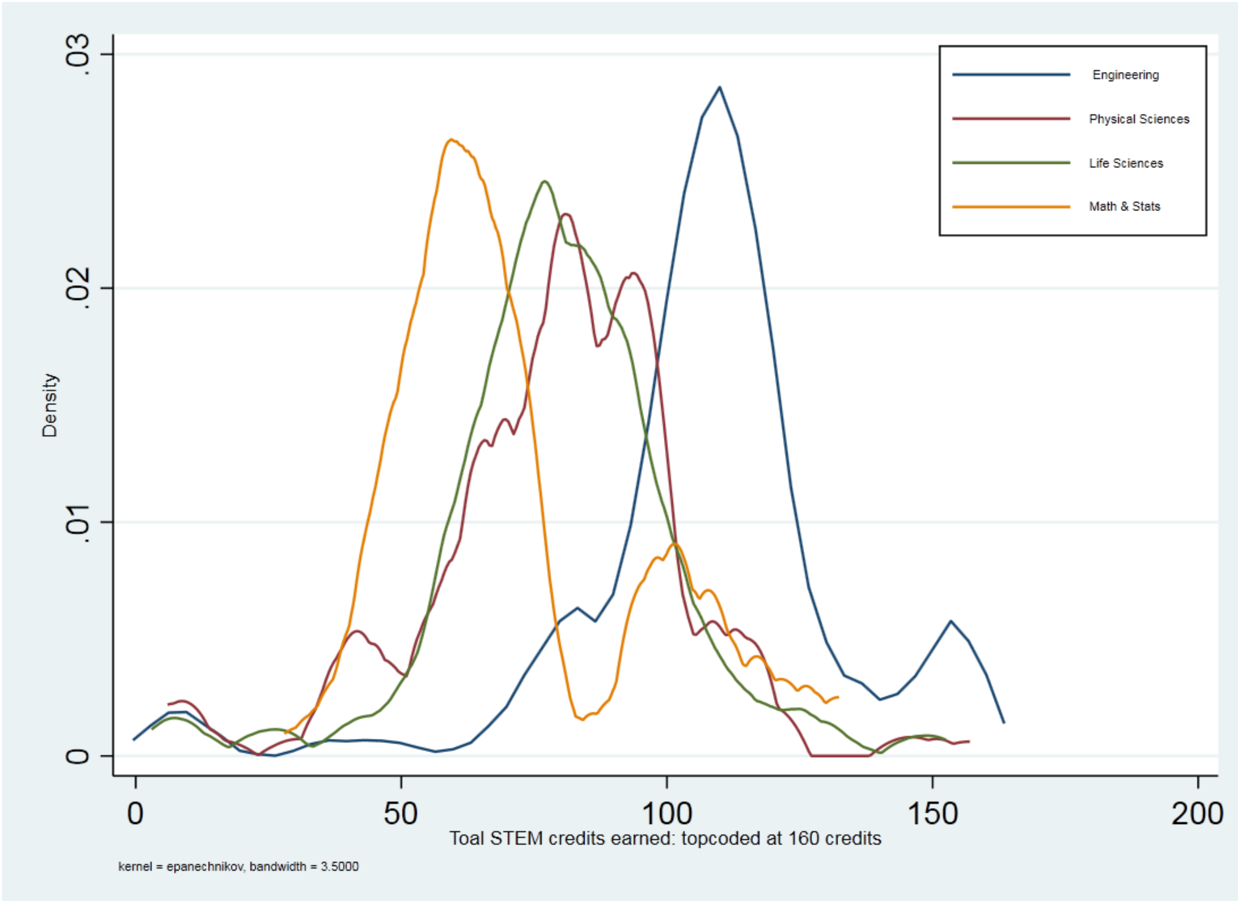

Figure 2. STEM credits earned by detailed STEM major, CNCC definition

Source: BPS 04/09 PETS data, Authors' calculations.

Figure 2 examines earned STEM credits among STEM majors to zoom in on the first point made in Figure 1- the mound between 60 and 120 credits. Similar to what we noted in Table 2, Figure 2 shows the heterogeneity among STEM majors. Mathematics (yellow) and students are clustered around 60 STEM credits, with an additional peak at roughly 100 credits. Life science (green) and physical science (red) students have similar distributions centered around 75-80 credits. The distribution for engineering students is centered at over 100 credits.

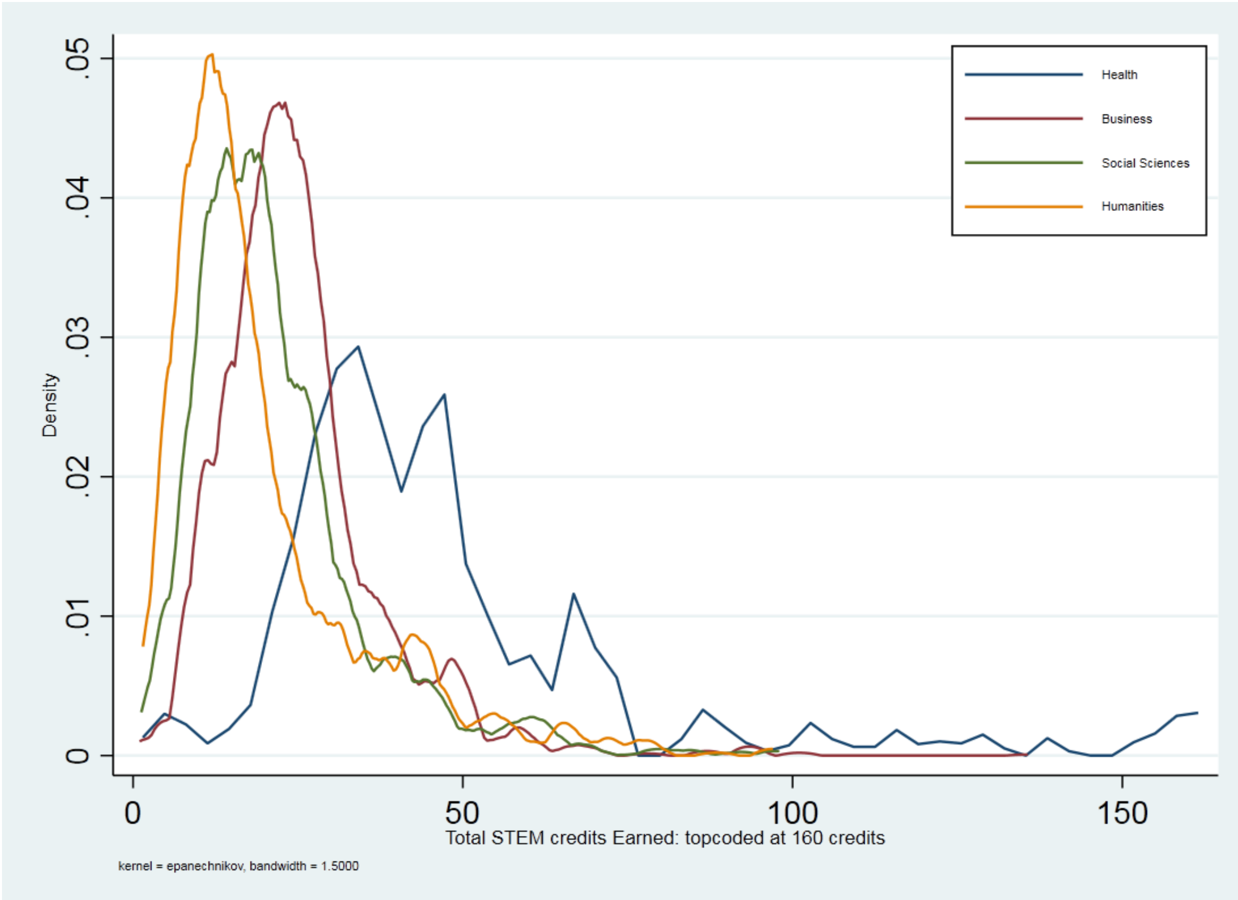

Figure 3. STEM credits earned by detailed Non-STEM major, CNCC definition

Source: BPS 04/09 PETS data, Authors' calculations.

Beyond attempts to increase the number of STEM majors, there is ongoing interest in expanding the extent of STEM education among non-STEM majors (e.g., National Academy of Sciences, 2018). Our analysis of the process and classification of course coding shows the errors that lead to a systematic and significant undercounting of STEM credits taken by non-STEM students. The actual STEM course taking distribution by major is shown in Figure 3, using our corrected classification and top coded at 100 credits to better display the central tendencies of the distribution. As suggested in Table 2, Health majors earned a higher mean number of STEM credits, and their distribution has peaks at approximately 35 and 45 credits with a substantial right tail that extends well beyond 100 credits. Perhaps more importantly, the plot identifies many non-STEM majors, in addition health majors have earned substantial numbers of STEM credits, notably business and social science majors.

Discussion and Conclusion

This paper utilizes the College Course Map (Bryan & Simone 2012) and the BPS 04-09 Transcript file to develop a more rigorous picture of STEM course taking among bachelor's graduates. When compared to NCES derived transcript variables, and despite removing double-counted courses, our Comprehensive Net Course Count (CNCC) measures identified greater numbers of earned STEM credits among bachelor's degree completers in the BPS 04-09. CNCC identified additional STEM credits earned by students of all majors, but the difference was most pronounced among non-STEM majors. In particular, CNCC identified additional mathematics courses taken by non-STEM majors. Density plots using CNCC measures identified heterogeneity among STEM degree holders in terms of the number of earned STEM credits, and identified a substantial population of non-STEM degree holders with substantial STEM credits earned.

The CNCC measure decreased STEM course counts by eliminating duplicates and repeated courses, and offset the increases introduced by our more comprehensive definitions and errors corrected in the data set. Using the CNCC measure, our more detailed analyses of mathematics coursetaking (Douglas and Salzman, 2020) uncovered substantial changes in particular subgroups and provided analyses of differentials by gender and advanced mathematics coursetaking. We found that gender does not account for differences in mathematics coursetaking, and that female students in the most math-intensive majors took and passed more advanced mathematics courses. Other papers based on the CNCC measure are in progress.

A STEM major should not be the only way to identify STEM talent. As noted by the National Academies of Science, Engineering, and Medicine, measuring the quality of STEM education, and changes therein over time, requires a diverse set of indicators (NASEM, 2018). The approach authored here relies on STEM credit earning. Our analysis here suggests that STEM talent exists well beyond those who complete STEM degrees.

References

References
Adelman, C. (1990). *A College Course Map: Taxonomy and Transcript Data.* Washington, DC: U.S. Department of Education. https://eric.ed.gov/ED326153.pdf

Adelman, C. (1995, 1999). *The New College Course Map and Transcript Files: Changes in Course-Taking and Achievement, 1972-1993*. Washington, DC: U.S. Department of Education. https://files.eric.ed.gov/fulltext/ED460653.pdf

Adelman, C. (2004). *The Empirical Curriculum: Changes in Postsecondary Course-Taking, 1972- 2000*. Washington, DC: U.S. Department of Education. https://www2.ed.gov/rschstat/research/pubs/empircurr/empircurric.pdf

Anderson, S. (2021, January 21) "DHS makes more international students eligible for STEM OPT." *Forbes*. https://www.forbes.com/sites/stuartanderson/2022/01/21/dhs-makes-more-international-students-eligible-for-stem-opt/

Baker, M. (2016). Is there a reproducibility crisis? A Nature survey lifts the lid on how researchers view the 'crisis rocking science and what they think will help. Nature, 533(7604), 452+. https://link.gale.com/apps/doc/A454730385/AONE?u=vuw&sid=googleScholar&xid=97bd121a

Billings, L. (2018). Cosmic Conflict: Diverging Data on Universe's Expansion Polarizes Scientists. *Scientific American.* https://www.scientificamerican.com/article/cosmic-conflict-diverging-data-on-universes-expansion-polarizes-scientists1/

British Science Council (2009). Our definition of science. https://sciencecouncil.org/about-science/our-definition-of-science/

Bryan, M. & Simone, S. (2012). *2010 College Course Map.* (NCES 2012-162rev). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. https://nces.ed.gov/pubs2012/2012162rev.pdf

Bush, V. (1945). *Science: The Endless Frontier*. Washington, DC: US Government Printing Office. Cited from National Science Board (2015)

Douglas, D. & Attewell, P. (2017). School Mathematics as Gatekeeper. *The Sociological Quarterly* 58(4): 648-669. DOI: 10.1080/00380253.2017.1354733

Douglas, D. & Salzman, H. (2020). Math Counts: Major and Gender Differences in College Mathematics Coursework, *The Journal of Higher Education*, 91 (1): 84-112. DOI: 10.1080/00221546.2019.1602393

Salzman, H & Douglas, D. (2022). "STEM Workforce Development." *International Encyclopedia of Education 4th Edition.* Amsterdam: Elsevier.

Engineers Council for Professional Development. (1947). Canons of ethics for engineers. New York: ECPD. https://www.worldcat.org/title/canons-of-ethics-for-engineers/oclc/26393909?referer=brief_results

Fanelli, D. (2010). "Positive" results increase down the hierarchy of the sciences. *PLoS One* 5(4): e10068. doi:10.1371/journal.pone.0010068

Feibleman (1961). "Pure Science, Applied Science, Technology, Engineering: An Attempt at Definitions." *Technology and Culture*, 2(4): 305-317. https://www.jstor.org/stable/3100886

Frey, R. E. (1991). Another Look at Technology and Science. *Journal of Technology Education*, *3*(1), 1–12. DOI: http://doi.org/10.21061/jte.v3i1.a.2

Hallinen, J. (2015) STEM education curriculum. https://www.britannica.com/topic/STEM-education/STEM-education

Handel, M.J. (2016). What do people do at work?. *Journal of Labour Market Research* 49: 177–197. https://doi.org/10.1007/s12651-016-0213-1

Humphreys, P. (1990). A conjecture concerning the ranking of the sciences. *Topoi* 9: 157-160. https://doi.org/10.1007/BF00135895

Ioannidis, J.P.A. (2005). "Why Most Published Research Findings Are False." *PLoS Med* 2(8): e124. https://doi.org/10.1371/journal.pmed.0020124

Kuhn, T. (1962, 1996). *The Structure of Scientific Revolutions*. University of Chicago Press.

Landivar, L.C. (2013). "The Relationship Between Science and Engineering Education and Employment in STEM Occupations," American Community Survey Reports, ACS-23, U.S. Census Bureau, Washington, DC

Layton (1974). "Technology as Knowledge." *Technology and Culture*, 15(1): 31-41. https://www.jstor.org/stable/pdf/3102759.pdf

Malitz, Gerald. (1987). A Classification of Instructional Programs (CIP). Washington, DC: Center for Education Statistics, Office of Educational Research and Improvement.

Mathematics. *Encyclopedia of Mathematics.* http://encyclopediaofmath.org/index.php?title=Mathematics&oldid=23895

National Academy of Sciences, National Academy of Engineering, and Institute of Medicine. (2007). *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter*

Economic Future. Washington, DC: The National Academies Press. https://doi.org/10.17226/11463.

National Academies of Sciences, Engineering, and Medicine. (2018). Indicators for Monitoring Undergraduate STEM Education. Washington, DC: The National Academies Press. https://doi.org/10.17226/24943.

National Center for Education Statistics. (2021). "Bachelor's degrees conferred by postsecondary institutions, by race/ethnicity and field of study: 2018-19 and 2019-20." Integrated Postsecondary Education Data System (IPEDS), Fall 2019 and Fall 2020, Completions component. https://nces.ed.gov/programs/digest/d21/tables/dt21_322.30.asp

National Research Council. (2013). *Next Generation Science Standards: For States, By States.* Washington, DC: The National Academies Press. https://doi.org/10.17226/18290

National Science Board. (2015). *Revisiting the STEM Workforce, A Companion to Science and Engineering Indicators 2014*. (NSB-2015-10). Arlington, VA: National Science Foundation. National Center for Science and Engineering Statistics (NCSES). 2022. Federal Funds for Research and Development: Fiscal Years 2020–21. NSF 22-323. Alexandria, VA: National Science Foundation. https://ncses.nsf.gov/pubs/nsf22323/

National Science Foundation Act 1950, § 247, 81st Congress. (1950). https://www.nsf.gov/about/history/legislation.pdf

Open Science Collaboration (2015). "Estimating the Reproducibility of Psychological Science." *Science, 349(6251), [aac4716]. https://doi.org/10.1126/science.aac4716*

Path, B. (2016, November 21). "Is the 'T' in STEM Supposed to be Silent." *Forbes. https://www.forbes.com/sites/gradsoflife/2016/11/21/is-the-t-in-stem-supposed-to-be-silent/*

President's Council of Advisors on Science and Technology. (2012). Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering and Mathematics. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/pcast-executive-report-final_2-13-12.pdf

Prinz, F., Schlange, T., & Asadullah, K. (2011). "Believe it or not: how much can we rely on published data on potential drug targets?" *Nature Reviews Drug Discoveries.* https://www.nature.com/articles/nrd3439-c1

Rothwell, J. (2013). *The Hidden STEM Economy*. Washington, DC: Brookings Institution. https://www.brookings.edu/wp-content/uploads/2016/06/thehiddenstemeconomy610.pdf

Salzman, & Benderly, B. L. (2019). STEM performance and supply: assessing the evidence for education policy. In Journal of Science Education and Technology (Vol. 28, Issue 1, pp. 9–25). https://doi.org/10.7282/t3-7551-p083

Teitlebaum, M.S. (2014). *Falling Behind? Boom, Bust & the Global Race for Scientific Talent.* Princeton University Press.

Wine, J., Janson, N., and Wheeless, S. (2011). 2004/09 Beginning Postsecondary Students Longitudinal Study (BPS:04/09) Full-scale Methodology Report (NCES 2012-246). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Appendix A. Approach to Classifying Student Majors.

In order to assess students' math coursetaking by major field of study, we first needed to accurately classify students' majors. In the BPS data, major field of study can come from two sources: institutional administrative data (transcripts) and student-report (surveys). Conflict between the administrative data and the interview data were non-trivial as were missing data fields. Thus, we created a single major field of study variable by following three general steps:

1) For students whose transcript- and interview-reported majors aligned, we kept the transcript value (i.e., there was no conflict).
2) For students who were missing one or the other data point, we substituted the non-missing value.
3) For those whose interview- and transcript-reported majors did not align, we selected the major that accorded with their number of earned STEM credits. This last decision rule connects with our larger effort to assess the STEM content of earned bachelor's degrees. Depending on their particular aims, other researchers could certainly decide on different decision rules. We developed the following sub-rules to determine how to resolve interview and transcript reports.
    a. We calculated the 25^{th} percentile of earned STEM credits by major, separately for interview- and transcript-report variables.
    b. If a student had both a STEM and non-STEM reported major, but fell above the 25^{th} percentile for earned STEM credits for the STEM major, they were coded with that major.
    c. If a student had two different reported STEM majors, but fell below the 25^{th} percentile for one of them, they were coded with the other major.
    d. For students with two different reported STEM majors whose STEM credits fell below either majors' 25^{th} percentile for earned STEM credits, we manually browsed their course history to make decisions based the timing of courses taken and the totality of their coursetaking history. Only about 0.5% of the sample (representing 1.5% of the weighted cases) was classified using this procedure.

Table C. Survey and Interview Major Disagreement in the BPS 04/09 sample.

Match type	N	%	weighted N	weighted %
Transcript and Survey Major Agree	3,920	70.4%	729,100	69.9%
Transcript Major Only	580	10.4%	114,900	11.0%
Survey Major Only	390	7.0%	82,400	7.9%
Transcript and Survey Major do not agree	650	11.7%	110,100	10.6%
Missing in both	30	0.5%	6,700	0.6%
Total	5,570	100%	1,043,300	100.0%

Source: U.S. Department of Education, National Center for Education Statistics, Beginning Postsecondary Students Longitudinal Study (BPS), 2004/2009

Appendix B. Approach to Resolving Double-Counted Courses.

Two important issues emerged in the process of coding and cleaning the NCES transcript files: removing duplicate course records and removing courses in which students did not earn credits. Since our ultimate aim is to look at the STEM content of degrees, operationalized as the number of STEM credits earned in that degree, we focus on unique earned credits in STEM and non-STEM fields.

Unidentified Transfer Courses. The first problem of duplicated courses occurs when students transfer from one institution to another and the course is counted twice, on both the sending and receiving institutions. Although there is a Transcript file flag to indicate a duplicate course, a large number of duplicate courses not flagged.

Unidentified Repeat Courses. A more significant issue concerned repeated courses for which students received credit. When a student repeats a course after earning a failing grade, course credits accrue for only courses with passing grades. However, many programs require students to earn a minimum grade, such as a C or a B, to continue in course sequences. In these cases, students who receive passing grades that accrue credit on the transcript would still repeat courses to receive grades that enabled them to satisfy major requirements or continue in a course sequence. Such repeated courses that accrue credit in both instances are often included in BPS aggregated credit counts.

This may be a more substantial problem for math courses than other subjects because math courses serve as pre-requisites, or gatekeeper courses, in many programs. Thus, a weakness in math would necessitate re-taking a math course for progression in non-math majors whereas low performance in other courses would be less likely to result in large numbers of repeated courses. To assess the overall STEM content of degrees, we count each unique course type only once; students who, for example, took calculus I twice—earning a D on the first attempt and a B-minus on the second—would only show the credit equivalent of one calculus I course in our estimates. Without this correction, the BPS derived variables double-count repeat courses, perhaps showing poor-performing students who repeated courses as having more mathematics education than stronger students who took each course only once.

Appendix C. CNCC Classification Tables

The extended table below provides a comprehensive listing of the Comprehensive Net Course Count approach. Specifically it lists the College Course Map course codes that counted as (S)cience, (T)echnology, (E)ngineering, (M)athematics, STEM-(X), and (N)on-STEM. It also provides the CCM definition of the course code, and an indication of whether and how the CNCC course classification varies from the original NCES classification.