Production and Consumption of Tertiary Education Services

Takashi Yamashita (Bureau of Economic Analysis, United States)

Paper prepared for the 34th IARIW General Conference

Dresden, Germany, August 21-27, 2016

Session 6E: Research Statistics Required to Support Higher Education Decision Making

Time: Thursday, August 25, 2016 [Afternoon]

Production and Consumption of Tertiary Education Services

Takashi Yamashita
Economist, National Income Accounts Research
Bureau of Economic Analysis
Department of Commerce

July 2016

Abstract

This paper estimates production of education services in tertiary education by credit hours earned by students. I then compare the growth of credit hours to the growth of household consumption expenditures in the National Income and Product Accounts and analyze sources of discrepancies. Between 2000 and 2013, the volume output of credit hours grew faster than the real household consumption expenditures on post-secondary education services in the NIPAs. The increases in other volume measures, such as the number of full-time-equivalent students, are consistent with the increase in credit hours index, while the quantity index of higher education services that households consume stays relatively flat throughout this period. I then construct a price index from net tuition payment by students after major grants are taken into consideration. My estimate shows that the two series, the PCE deflator for higher education and the price index per credit hour for private non-profit universities diverged throughout the period, particularly after 2010 when the Federal government expanded Pell grants. This research points to further areas of possible improvements in how we measure education services.

JEL Classification: E01, E24, H52, I22, I23
Keywords: education services, post-secondary education, volume output,

What do educational institutions produce and how do we measure the output? These questions are fundamental to national income accounts and pose serious challenges to capturing education output correctly. Traditionally, education services are mostly produced by government and non-profit institutions. SNA recommends education services provided by government entities and nonprofit institutions serving households (NPISH) at "economically not significant prices" to be valued by the cost of production, i.e., the sum of intermediate consumption, compensation of employees, consumption of fixed capital and taxes on production. Most statistical agencies in the world follow this tradition.

The American system of higher education has a number of characteristics that distinguish it from those of other countries. Most of all, three main types of institutional control and ownership coexist: government-owned public entities, private, not-for-profit institutions, and private, for-profit institutions with the substantially larger presence of the for-profit sector than other countries. There is also a great variation in size; from small, specialized colleges with a few hundred students to University of Phoenix with enrollment close to a half million students. In terms of what they offer, some provide liberal arts education and world-class science education while others offer specialized vocational training such as culinary arts and fashion design, which in some countries may not confer a university degree.

Co-existence of for-profit and not-for-profit institutions creates inconsistency in measuring the sector's output in the standard national accounting methodology. While education services provided by government entities and non-profit institutions are measured by the cost of direct inputs, value added of the for-profit sector is measured by subtracting cost of intermediate input from total revenue. Information on the education sector, while important in a nation's economy, is scattered in various parts of the national income statistics and is hard to come by.

Without a consistent accounting method that cuts across all three types of ownership, we do not have a good measure of output of the sector.

Several approaches have been proposed to measure output of the education sector. Two most prominent methods are the cost method and the income method (see, for example, Abraham, 2010 for a concise summary and critique). Similar to the method in SNA, Kendrick (1976) recommends to measure the output of the education sector by the cost of input. In addition to cost of market input as suggested by SNA, Kendrick suggest to value the cost of nonmarket input, including the opportunity cost of attending school and time spent on studies outside of class.

One of the major drawbacks of measuring output by input is that this method renders the measure of productivity meaningless. Since productivity measures how efficiently input is used to produce output, productivity is not defined when input is the same as output. Furthermore, there is at best a tenuous relationship between input and outcome of education; increase in input does not necessarily translate into improvement in outcome (e.g., learning, skills, earnings, and eventually, health outcomes). By measuring the output of the sector by the cost of input and by confusing *output* and *outcomes*, one may erroneously believe that we are producing better educated citizenry just because we spend more money on schools. Hanushek (1996, 2003) and Hanushek and Rivkin (1997) convincingly illustrate how increased school spending has failed to improve educational outcomes.

Considering that education is investment in human capital, Jorgenson and Fraumeni in pioneering work (1989, 1992) propose to measure stock of human capital by the discounted present value of future income streams and then take the change in the level of stock to measure investment. This approach, by measuring education output by labor market outcomes, avoids one of the main issues associated with the cost method. Following this pioneering work, many applications and updates have been presented for a number of countries (see, for example, Fraumeni, at al 2015; Gu and Wong 2012, 2015; Wei 2007).

The income method is not without criticism, however. First, as human capital accumulation involves substantial investment at home and outside of classroom, this method expands the production boundary defined in SNA, which explicitly excludes home production. Related to the first point, secondly, since earnings from human capital reflect cumulative investment made one's lifetime and beyond, any change in earnings arising from other factors (e.g., pre-natal environment, informal early childhood education) cannot be distinguished from improvement in school quality or output of the education sector. Third, it is extremely difficult to distinguish change in the quantity of human capital from change in the price of human capital not related to education. Return on education could change substantially because of temporary demand shocks (e.g., Goldin and Margo 1992) or change in cohort size (supply) (e.g., Card and Lemieux 2001). In sum, the income method is not a measure of school output but a measure of broader human capital investment and casts too wide a net to measure output of the more narrowly defined education sector.

Another way to measure output of education services is by the volume output educational institutions produce. A measure of education services can be physical quantities such as hours of instructions pupils and students receive, the number of students enrolled in schools, or the number of diplomas, certificates, and degrees awarded. They are pure physical measures of output, just like how many barrels of crude oil or how many vehicles are produced in a given year. They are thus comparable across countries and across time, save the problems associated with measuring quality differences. This approach is less ambitious than the income and cost methods and imposes less onerous data requirements. There is no difficult choice of discount rate, or deciding the opportunity cost vs replacement cost of foregone earnings. The volume measure would thus provide a straightforward metric to evaluate the growth of the education sector.

This paper presents estimates of volume measures of education services produced by colleges and universities (tertiary educational institutions). I use credit hours taken by students as a primary measure of education services output, supplemented by the number of full-time equivalent (FTE) students enrolled at schools. I then compare the growth of the volume measures of output to the real consumption expenditures from the National Income and Product Accounts (NIPAs). I find substantial differences in the growth of higher education services between the two measures and consider possible reasons why such differences arise.

There have been a handful of studies that attempt to measure volume output of U.S. educational institutions. Previous studies include Fraumeni et al. (2009), O'Mahony and Stevens (2009), and Powers (2016). Fraumeni et al (2009) construct indices of volume output of primary and secondary education services provided by government, using enrollment (the number of pupils) data from the Current Population Survey (CPS). Powers (2016), using two surveys from the National Center for Education Statistics (NCES) covering public and private schools, estimates attendance-adjusted number of students at the primary and secondary education institutions to measure productivity. O'Mahony and Stevens (2009) also use CPS and expand the coverage to measure enrollment at the post-secondary level. Contrast to the above studies, I obtain enrollment and credit hour data from the regulatory data compiled by the overwhelming majority of colleges and universities and measure the volume output by credit hours as well as enrollment. To the best of my knowledge, this is the first attempt to measure volume output at the post-secondary level in the U.S. using the data set analogous to the "establishment survey" of college and universities.

I. Volume Output of Education Services

There is an emerging consensus among national income statisticians to measure production of education services by output volumes. Metrics of volume suggested include the number of pupils/students and the number of classroom hours. As a starting point, OECD (Schreyer 2010) and Eurostat (Smith and Street 2007) suggest a Laspeyres index of the form:

$$L_{qt} = \frac{\sum_i q_{it} p_{i0}}{\sum_i q_{i0} p_{i0}} = \sum_i \left(\frac{q_{i0} p_{i0}}{\sum_i q_{i0} p_{i0}}\right) \frac{q_{it}}{q_{i0}} \tag{1}$$

in which q_{i0} is the quantity of service produced by institution *i* in base period (t = 0) and p_{i0} is price per unit at institution *i* in base period. The weight (p_{i0}) could either be cost of production or price that consumers pay. In case of higher education, cost of production is extremely difficult to obtain because many universities cross-subsidize instructional activities by other activities (e.g., sales of medical services and income from endowments) and there is no clear-cut way to divide faculty and staff salaries into instructional vs non-instructional activities. On the other hand, the data on expenditures by purchasers of educational services (i.e., revenues from tuition) is readily available. In implementation, I use both the number of credit hours and the FTE student count to represent q_{it} and net student tuition after institutional grants and Federal, state and local government grants to represent expenditures ($q_{i0}p_{i0}$).

A number of countries have implemented or experimented with the volume-based measures of education services. Australia, Belgium, France, the Netherland, and New Zealand have long-established series of volume output measures counting number of pupils/students at various levels of education. See table 2.4 in Schreyer (2010) for an overview of country practices.

A simpler measure of volume output could be just the aggregate number of credit hours produced or the total number of students. In the United States, National Research Council (NRC) (Sullivan et al. 2012) recommends credit hours adjusted for completion as the primary measure of output in calculating productivity of the higher education sector. Aggregating credit hours across institutions, however, assumes quality of education services across institutions is reasonably similar. Recognizing this point, NRC suggests to group institutions into groups with similar characteristics: public research, public master's, public community colleges, private nonprofit research, private nonprofit master's, private nonprofit bachelor's and for-profit institutions.

One crucial challenge of measuring volume output of services is how to account for quality change. Several approaches have been proposed in this regard. For primary and secondary education, Fraumeni et al. (2009) adjust quality by the student-pupil ratio and Powers (2016) uses test scores as a proxy for quality. Some countries (e.g., Italy, Poland, Spain) use class size (pupil-teacher ratio) while others (e.g., France, Sweden, U.K.) take into account performance measures such as scores of standardized exams or student promotion record to account for quality change. A good teacher could make a great deal of difference on students' performance (Chetty et al 2011): Denmark uses the proportion of teacher credentials to adjust for quality.

The above attempts to adjust for quality change in education are laudable but the question of how these measures affect student outcomes is far from clear. The estimates on the effects of class size on student performance vary substantially and there is no consensus whether or not smaller class size improves student outcomes (see discussion between Krueger 2002 and Hanushek 2002). Test scores can also be easily tailored to show "improvement" (Kortez 2002) and could produce misleading results when input quality is not accounted for. Castro and Coen-Pirani (2016) show the change in average cohort ability best explains recent declines in U.S. male educational attainment. When the ability of average students declines as more students are enrolled in colleges, which in turn results in lower test scores and higher drop-out ratio, would we conclude that the quality of education has deteriorated? Many teachers would vehemently object to such ideas.

In this paper, I do not attempt adjustments for quality. The purpose of tertiary education is to provide specialized training and there exist no standardized exams that measure student performance across disciplines at the post-secondary level. Information on class size is not available either. In the data, it is possible to adjust for a fraction of full-time faculty among total instructional staff. However, it is far from clear if a full-time, tenured faculty member is a better instructor than a part-time untenured instructional staff. Given the difficulty in quality adjustment and in obtaining appropriate data, a number of countries (e.g., Australia, Belgium, New Zealand, to name a few) do not attempt to adjust for quality changes.

By focusing on credit hours and student enrollment, the volume output considered here does not include other products and services produced by universities. A tertiary educational institution is generally a multi-product firm producing education, research, and community and alumni services. Universities with medical schools and hospitals participate in substantial market activities selling medical services (and earn a considerable fraction of total revenues from medical services). To the extent that instruction and research are compliments, research affects education through quality of instruction. On the other hand, both research and teaching compete for faculty and staff time and not taking into account time spent on research may distort measures of productivity as research-intensive universities may hire more faculty members to teach the same number of students as less research-intensive ones.[1] Since I do not attempt quality adjustment, not considering research at this stage would be justified. Furthermore, my focus in this paper is on output measures and not productivity; assessing productivity of post-secondary education is beyond the scope of this paper.

II. International Standard Classification of Education System and Definition of Tertiary Education

In compiling national statistics, defining what constitutes post-secondary education presents its own challenges. In the popular parlance, at least in the United States, higher education, post-secondary education, and tertiary education are often used interchangeably. An unaccredited non-degree-granting institution is free to call itself a “college” while institutions with similar names may provide totally different types of education[2]. To complicate the matter further, the variation in naming and differences in duration of programs make international comparison of education output at the post-secondary level extremely difficult.

An attempt to standardize different cross-country educational systems is made by UNESCO’s Institute for Statistics through its International Standard Classification of Education System (ISCED). ISCED 2012, the most recent version, divides different stages of education from early-childhood to doctoral education into nine levels based on program orientation, complexity, duration, and entry requirements. ISCED specifically distinguishes tertiary

1 Webber and Ehrenberg (2010) show that increases in per-student expenditures on sponsored research are associated with lower graduation rates after controlling for instructional expenditures per student.

2 For example, Massachusetts Institute of Technology provides world-class education from bachelor’s to doctoral level while SAE Institute of Technology offers short-term training for operating audiovisual devices.

education from post-secondary education and defines the former as a subset of the latter. ISCED Level 4 refers to post-secondary non-tertiary education, which provides vocational training aimed at direct labor market entry and offers programs that are shorter than two years. In the United States, examples of post-secondary non-tertiary institutions include beauty and cosmetology schools and schools for massage therapists and practical nurses.

ISCED classifies post-secondary tertiary education into four levels: from Level 5 (short-cycle tertiary education) to Level 8 (doctoral level). In this paper, I focus on credit hours earned in colleges and universities corresponding to ISCED's levels 5 to 8. The levels correspond to community and junior colleges (Level 5), four-year colleges (Level 6), and graduate programs (Levels 7 and 8) including courses with strong academic orientation as well as professional programs such as law, medical, and business schools. Given data limitations, however, I combine ISCED Levels 7 (master's level) and 8 (doctoral level) credit hours. Source data do not distinguish credit hours for master's level versus doctoral level studies. In some fields (e.g., economics) in the United States, students are matriculated into longer-term doctoral programs directly after completing a bachelor's degree and without a master's level preparation. In such programs, it is impossible to distinguish master's level courses from doctoral-level instructions.

Classification of the U.S. post-secondary educational institutions into the ISCED levels may not be perfect. ISCED classifications depend on the duration of the programs and their complexity. Many vocational schools that focus on business, health, or technician certificates have started as short-term programs (ISCED level 4) but progressed into two-year, or sometimes four-year degree-granting institutions. When an institution transition to longer duration, would the complexity of instruction necessarily increase? It is hard to judge by looking at the duration and degree-granting status. To distinguish ISCED Level 4 institutions from more traditional colleges and universities, I rely on Carnegie Classification of Institutions of Higher Education. I exclude private, for-profit, two-year institutions that do not have Carnegie Classifications.[3] I limit my exclusions only to for-profit institutions. Some public and non-profit institutions would meet the criteria for higher ISCED levels but may not have Carnegie Classifications. Carnegie Foundation does not classify religious institutions such as seminaries and schools for rabbis as well as medical internship and residency programs at hospitals. As most such programs are either in the public sector or private non-profit sector, I include them in my sample and count their enrollment and credit hours.

Finally, I do not count credits earned in Advanced Placement (AP) courses by high-school students if such courses are taken at high school campuses, although credits of AP courses are often used to fulfill college requirements. Similar to AP courses, some high-school students earn college credits by taking courses at local colleges and universities. As long as such credits are reported in the data, they are counted towards total credits earned by students. The unit of observation in my data set is a "campus" (or physical location) of a post-secondary institution. For example, University of Phoenix has many locations throughout the United States and each one of them is counted as a separate institution as they are reported in the data.

III. Integrated Postsecondary Education Data System

I use the longitudinal data from the Integrated Postsecondary Education Data System (IPEDS) collected by the National Center for Education Statistics. IPEDS is a longitudinal

[3] I limit the exclusion criteria to only for-profit two year institutions because they are most likely the ones that do not meet the ISCED Levels above 5. I include public and private, nonprofit institutions that do not have Carnegie Classifications because they are often new (e.g., University of California, Merced) or part of a larger institution (e.g., private part of Cornell University is a Research Intensive Doctoral institution but the public part, Cornell University-New York State Statutory Colleges do not report its Carnegie Classification; after 2004, only the private, non-profit part of Cornell University compile combined numbers of these two institutions).

survey collecting data on institution, student enrollment and their demographics, and institution finances. Completing IPEDS surveys is required for all institutions that participated in any federal financial assistance program under Title IV of the Higher Education Act of 1965. Since Economic Census, conducted every five years, does not cover colleges and universities, IPEDS is the most comprehensive data set available on most institutions in post-secondary education. I use the 2000-2013 longitudinal data compiled from IPEDS by the American Research Institute for the Delta Cost Project, supplemented with the calendar system information obtained directly from IPEDS.

Table 1 show the number of institutions in IPEDS by type. Private, non-profit institutions account for the majority of four-year colleges and universities while public institutions make up the majority among community colleges. The number of public and private, not-for-profit institutions stay relatively constant throughout the period, except private, non-profit two-year institutions which substantially shrank in this time period.[4] On the other hand, private, for-profit institutions have grown substantially, growing from 13.9% of all tertiary institutions in 2000 to 24.3% in 2013. The growth seems particularly pronounced after 2010 (for-profit two-year institutions grew by 43% from 2009 to 2010; the growth of four-year institutions was 15.4% from 2010 to 2011). This growth comes from many institutions lengthen their program durations (e.g., from less-than-two-years to two-year, and from two-year to four-year). Curiously, the lengthening of program durations at for-profit institutions seems to have taken place as the Department of Education has intensified scrutiny of for-profit institutions and required disclosure of gainful employment statistics.

4 Note that some of two-year institutions convert to offer four-year degrees in selected programs while keeping the two-year degree options; thus the decrease of two-year institutions in this table does not necessarily mean that they exit the industry.

Expansion of for-profit post-secondary education is studied extensively in recent years. Deming et al (2012) analyze the growth of for-profit institutions over a longer period looking into how they are organized and what types of students they attract. They document that for-profit colleges focus on career-oriented fields such as business, accounting, and nursing and enroll far greater number of minorities and first-time college students. Cellini (2009, 2010) attributes part of this growth to public-sector funding constraints and wider availability and more generous provision of federal and state financial aid. Gilpin et al (2015) show for-profit two-year institutions are more responsive to changes in local labor market conditions while public two-year institutions remain unresponsive to employment growth and wage changes in related occupations.

One main challenge of using IPEDS for calculating credit hours is that the data set seems fraught with inconsistency and reporting errors. While the overwhelming majority of institutions report their data correctly, some institutions report no or unusually low credit hours in a year when there are hundreds of full-time students enrolled. On the other hand, a small number of institutions report only a handful of students while reporting hundreds of thousands of credit hours.

To ameliorate the problems of misreporting and extreme values by entry error, I apply the algorithm as exhibited in Figure 1. I first convert the quarter and other calendar systems to semester hours, using the guidelines in IPEDS to calculate FTE status. I then calculate the (semester) credit hours per FTE student and estimate the median of per-student credit hours by institution. If an institution reports credit hours per student exceeding twice the median or smaller than half of the median, then I replace the reported number with the median plus an error term $\sim N(0, 1)$. I use credit hours per FTE student because they are very stable for most institutions and variation across institutions is also small; most institutions report about 30 semester credit hours per FTE student with the standard deviation of the median of all institutions less than one. However, this procedure does not preclude the estimated credit hours per FTE student from becoming negative as the median credit hours could be very small when reporting errors are persistent. When the median credit hours is smaller than three, I use the institution's mean to calculate the replacement value. Once credit hours per FTE student is estimated this way, I multiple the adjusted credit hours by the number of FTE students to arrive at adjusted total credit hours.[5]

Two technical details warrant special mention: Conversion of credit hours to semester equivalent and calculation of FTE students. Various institutions use different calendar systems to calculate credit hours. Sometimes, the same university employs different systems depending on the program. To make different ways of calculating credit hours across institutions as comparable as possible, I convert reported credit hours to semester-equivalent units. Specifically, one quarter credit hour is equivalent to two-thirds of a semester credit hour. When an institution uses more than one calendar system, I convert reported credit hours by multiplying 4/5 to arrive at the semester-equivalent credit hours. To convert contact hours to semester-equivalent, I divide reported contact hours by 37.5.

To convert the number of part-time students to FTE, I use the factors used in the IPEDS guidelines. IPEDS specifies the factors to multiply the number of part-time students by types of institutions; for example, for undergraduate part-time students at public 4-year universities, a factor of 0.403543 is applied and added to the total number of full-time students to arrive at the FTE student count. The factors are specific to institution-type and level: private 4-year, both not-for-profit and for-profit (0.392857), public 2-year (0.335737), and all other undergraduate

5 Far more institutions report FTE student counts than those reporting total credit hours.

institutions (0.397058). For graduate enrollment, they are 0.361702 for public universities and 0.382059 for private universities.

These operations produce quite robust credit hours per FTE student. For undergraduate programs, the median of all institutions is 31.7 semester credit hours per FTE student while the mean is 33.6 hours (interquartile range of 29.7 hours and 35.5 hours). For graduate programs, comparable statistics are: 26.6 (median), 26.9 (mean), 20.2 (25 percentile), and 32.0 (75 percentile).

IV. Credit Hours Produced by Different Types of Institutions

I compute total credit hours produced by tertiary education institutions by aggregating all institutions by type (four-year or above and two-year institutions) and by institutional ownership/control (public, private non-profit, private for-profit). Grouping here is different from the NRC recommendation for simplicity. Following the NRC recommendation of seven sectors and dividing credit hours for graduate and undergraduate studies would require at least 18 columns. Instead, I present graduate and undergraduate credit hours by institutional control. NRC also recommends to adjust total credit hours for completion and to add an extra year of credit hours (30 semester-equivalent credit hours) based on an institution's completion rate. NRC justifies this adjustment on the basis of the well-known "sheepskin" effect of college degrees, i.e., return on college education is non-linear and those who complete a degree earn substantially more than those who have just 15-years of schooling as well as those who have 16-years of schooling but have not completed a degree (see Wood 2009 for a non-technical review).

I do not make this adjustment for the sheepskin effect. I argue that completion does not necessarily increase quantity of human capital but changes its price of by signaling something else. My argument is similar to Heckman and Rubinstein (2001) who demonstrate that completion of high-school diplomas signals noncognitive skills of diploma holders compared to those with GED certificate. Heckman and Rubinstein show that GED recipients, who presumably have the same level of cognitive skills as high-school graduates, earn substantially less than high-school graduates. Furthermore, they earn less than other high school dropouts, who have lower cognitive skills as evidenced by lower test scores, when years of schooling and AFQT scores are controlled for. They attribute this difference, i.e., GED recipients have higher cognitive skills but earn less than high-school dropouts, to differences in noncognitive skills such as dependability, tenacity, and stability which employers value.

A similar argument can be made to consider earnings premium college graduates enjoy over dropouts. Consider, for example, two otherwise identical individuals with the same number of credit hours. However, one has earned a college degree while the other has not completed because his credit hours do not satisfy diversification requirements or requirements for physical education courses. It is plausible that the one who has completed the degree requirements exhibits better organizational skills and planning, which are not directly related to what a college or university produces but represent her noncognitive abilities. In such a case, we could think the price of human capital is higher due to her noncognitive skills although two individuals have accumulated equal amount of human capital through learning. Since my focus in this paper is volume output of education services, not considering the price effect in estimates is justified.

Table 2 exhibits semester-equivalent credit hours by level and type of institutions and panel (a) of figure 2 presents the same information graphically. In 2013, American higher-educational institutions produced 47 million credit hours of graduate studies and 425 million hours of undergraduate studies. Although public institutions account for less than 20 percent of all four-year colleges and universities, they produce overwhelming majority of undergraduate credit hours and nearly half of graduate credit hours. Credit hours produced at all levels of institutions grew substantially over years; the growth of graduate hours accelerated from the academic ending in June 2009 to 2010 possibly reflecting substitution effects arising from lowered opportunity cost during the Great Recession. As the economy recovered, total credit hours taken by graduate students declined after 2011. Similar patterns are found in undergraduate credit hours although the peak of undergraduate hours was 2012 at four-year institutions and 2011 at two-year institutions. In sum, growth of credit hours has been substantial in the past decade.

We observe similar trends in other metrics of higher education output. Table 3 and panel (b) of figure 2 show the number of FTE students by sector of institutions. In 2000, about 11 million students were enrolled in American higher-education. By 2011, the number has grown to 16 million students while the recent years saw a slight decline in student enrollment. As in the credit hours, the majority of students are educated at public institutions, while private, for-profit institutions have shown the fastest growth in student enrollment. Table 4 shows the total number of various degrees conferred by tertiary educational institutions. The number of degrees awarded grew on average 3 to 4 percent between 2000 and 2013. In 2013, nearly 1 million associate degrees and 1.8 bachelor degrees were awarded while over 900,000 graduate degrees were conferred. Since it normally takes a few years to earn a degree, the number of degrees conferred does not show the declines in credit hours and FTE students of the past two years.

V. Comparison with the NIPA Household Consumption Expenditures on Education

(a) Credit Hours and Quantity Index

I've shown that simple physical measures (credit hours and FTE student count) of higher education services have consistently increased between 2000 and 2011 and then declined slightly during the last few years as the economy recovered. How do my estimate compare to BEA's estimates of consumption of higher education services? Since education services produced by schools must be consumed by students, volume output would ideally match the quantity of education services consumed by households. Is this really the case?

I construct the Laspeyres quantity indices (equation (1)) of credit hours and student count and compare them to the NIPAs' estimate of household consumption of higher education services. Figure 3 compares my estimates of credit hours and FTE students to the real household consumption expenditures (quantity index) on higher education services (NIPA Table 2.4.3 line 101). Since my estimates based on IPEDS are for the academic years ending in June, I adjust NIPA series by taking quarterly average from the third quarter of the previous year to the second quarter of the current year (quarterly details are available from underlying detail tables). Surprisingly, both growth of total credit hours and FTE students far exceeded the growth of the NIPA index of household consumption of higher education services. To reconcile a 60% increase of the index for FTE student count and a mere 10% increase of the NIPA quantity index, per-student consumption of higher education services must have decreased by nearly one-third, which seems extremely implausible. At least, the NIPA quantity index captures the peak of credit hour production and the decline thereafter correctly.

Figure 4 splits the total credit hour growth into two sectors, proprietary and public institutions (line 286) and non-profit institutions serving households (NPISH) (line 287) of the

Table 2.4.3U. Again, total credit hour index estimated from IPEDS grows far faster than the NIPA estimates of real household consumption expenditures in both sectors. In particular, NIPA estimate shows that household consumption of higher education from NPISH is lower in real terms in 2012 than the 2000 level while the volume index from IPEDS doubled during the same period. The decline of consumption of higher education services in the NIPAs is starkly different from the picture presented by IPEDS data.

One major difference between the NIPA series and the estimates from IPEDS is that NIPA real quantity index could be adjusted for quality changes while my indices do not take into account possible changes in instructional quality.[6] Could quality change be a reason to explain the divergence? I find this story very unlikely.

If the constant-quality index stays flat while quality-unadjusted series grows rapidly, then there must be a massive quality deterioration in higher education. I find such a scenario is extremely hard to fathom. The decade after 2000 is the time when colleges and universities have invested heavily in emerging instructional technologies. Many universities have adopted WebCT and Blackboard to better manage and monitor student progress. Growth of the Internet has enabled university instructors to learn what is taught at far away colleges and universities and share problem sets and assignments. Innovative companies such as Aplia have produced appealing course content, which freed instructors from grading and creating new problem sets and spend more time on students. The story of quality deterioration at public and NPISH universities would thus be a hard sell to many practitioners of higher education.

---

6 The consumer price index (CPI) for college tuition and fees, which is the basis for the PCE deflator for higher education, does not seem to be adjusted for quality changes, however. When reported price changes, BLS field staff asks respondents if there is a cause for a change. For college tuition, the majority of price changes are not given specific reasons by responding institutions. When no specific reasons ("causes") are given to explain tuition changes and the characteristics for the identified tuition items remain unchanged, the adjustment for quality is not made in the college tuition index (http://www.bls.gov/cpi/cpifacct.htm).

Could the change in composition explain the decline in quality? The time period under study is when for-profit colleges and universities expanded rapidly. If quality of instruction at for-profit universities is inferior, then growth of the sector would affect quality of the entire sector negatively by composition change. Again, this story seems unlikely. If for-profit colleges institutions negatively affect overall quality, then the effect would be bigger in the proprietary and public institutions than the NPISH series. However, the divergence between the NIPA and IPEDS series is larger between the NPISH series than the proprietary and public institution series.

(b) Net Student Tuition and PCE Deflator

Another possibility for the divergence between NIPA and IPEDS series is that the PCE deflator to deflate the nominal household expenditures for higher education may not reflect the actual price changes of higher education. I explore this possibility in this section.

The rise of college cost is probably one of the most pressing issues today facing young Americans. The popular press report story after story depicting how young adults are accumulating ever-higher amount of student debt because of rising cost of college and hardship caused by mounting debts. We hear friends and relatives spending a fortune sending their children to schools. We seldom doubt that cost of attending college is skyrocketing.

The price index for college tuition and fees published by statistical agencies confirms such views. Bureau of Labor Statistics' Consumer Price Index (CPI) for college tuition and fees rose from 332.8 in June 2000 (the first year of my data) to 732.0 in June 2013, implying an annualized growth rate of 6.3% p.a. Only a few items in the CPI show greater increases during the same period (delivery services, tobacco and smoking products, and fuel oil).[7] BEA's PCE

[7] Looking at the period after July 2013, the CPI for fuel oil decreased substantially while delivery service's index plateaued. On the other hand, the index for college tuition and fees kept increasing.

deflator for higher education is based on data from CPI and two series track each other very closely.

There are two issues with respect to measuring price of higher education. First, college and university courses are priced non-linearly. For example, at many universities, a student attains a full-time status once she takes 12 or more credits per semester. Once she pays the tuition for full-time students, she can take courses up to 18 credits without paying additional tuition and fees. Some students take advantage of this system and enroll in as many courses as possible within budget and obtain a bachelor's degree in less than four years.

Second, colleges and universities, operating in monopolistic competitive markets, practice price discrimination, probably more so than any other sectors in the economy. Tuitions and fees paid by students vary greatly even within the same institution as colleges and universities offer various financial aid packages to attract students they want. Furthermore, students can obtain grants and scholarships from other sources such as private foundations. Estimating the price of college education thus requires tremendous amount of detailed data on student payments.

IPEDS provides useful data in understanding how the price of attending colleges and universities. Institutions report detailed data on revenues as well as institutional grants given to students as scholarships and grants (but not including loans to students). Unlike the CPI, which surveys a sample of a few hundreds of colleges and universities, IPEDS provides consistent reporting from a bigger universe of higher educational institutions.[8]

I estimate the price of a unit of college education, namely how much students pay for one semester-equivalent credit hour. This measure might give a different picture of price increase of

---

8 BLS used to use the "sticker price" of tuitions and fees collected from college and university catalogues. Starting in 2003, BLS includes financial aid and scholarships to calculate the CPI for college tuition and fees.

higher education than the CPI. I use the net student tuition reported by institutions, which is the total amount universities receive from students after institutional grants, Pell and other Federal grants, and grants from state and local government. It does not include loans taken by students or grants students may obtain from private foundations. Using this variable, I calculate the price of an credit hour weighted by the institution's credit hours produced in year *t* as:

$$L_{pt} = \frac{\sum_i q_{i0} p_{it}}{\sum_i q_{i0} p_{i0}} = \sum_i \left(\frac{q_{i0} p_{i0}}{\sum_i q_{i0} p_{i0}} \right) \frac{p_{it}}{p_{i0}} \quad (2)$$

in which p_{it}, the price per credit hour at institution *i* in year *t*, is calculated as $p_{i,t} = \left(\frac{Net\ Student\ Tuition_{i,t}}{Total\ Credit\ Hours_{i,t}}\right)$. Net student tuition may be negative when students receive more aid than the full amount of tuition, such as stipends for graduate students and aid to cover rooms, board and textbooks. Since grants for such purposes do not reflect the price of credits, I replace the value with zero when the reported net student tuition is negative. I also calculate price per FTE student in an analogous manner.

Figure 5 presents the comparison of higher education PCE deflator and two measure of higher education price indices estimated from IPEDS. As a reference, the PCE deflator for services is also plotted. Compared to the PCE deflator, net tuition per credit hour increased at a slower pace. Furthermore, as the government considerably expanded the federal grant programs after the financial crisis of 2008/2009, net student tuition declined considerably. In recent years, the growth of net student tuition per credit hour accelerated but it is still lower than the level projected based on the pre-2009 trend line. The figure also plots net student tuition per FTE student, which is smoother than the price index of credit hours. The decline of this measure after 2009 was less than that of per-credit price indicating that students may have been taking heavier course load during the recession.

Figure 6 splits the series to different sectors: Public and Proprietary schools (line 286) and NPISH sector (line 287) in NIPA table 2.4.4U and price per credit hour for public, NPISH, and for-profit institutions from IPEDS. The two series from the NIPAs are almost identical and thus do not show well in this graph. The increase of net tuition at public universities tracks the growth of PCE deflator except a few years after 2009. On the other hand, the price index of credit hour at NPISH institutions grew only half as fast as the level indicated by the PCE deflator. With respect to for-profit institutions, per-credit hour tuition increased rapidly in the early 2000s then slowed down substantially in later years. One possible reason why net student tuition grew slower at NPISH colleges and universities may be financial aid pledges some have made earlier in the decade. In 2001, Princeton University pledged no-loan policy for its students and converged previous loans of the current students to grants (Rubinstein and Rouse, 2011). Following this decision, many of Princeton's peer institutions, most of which are private non-profit institutions, have made no-loan pledges to all or some of their financial aid recipients.[9] On the other hand, budget crisis and cut in state appropriations propelled public universities to increase tuitions and increase the share of students who pay higher out-of-state tuitions. The strikingly different trends in cost of attending college between public and private institutions point to the need for closer examination of how PCE deflators are calculated in the two sectors of higher education institutions.

VI. Conclusions

This paper presents volume output of educational services provided by institutions of higher education in the United States by level and institutional ownership. First, I present simple

9 The Institute for College Access & Success publishes a list of colleges and universities that have made financial aid pledges. As of 2009/10 academic year, there are 64 such institutions, including all Ivy League schools, Stanford, Caltech, top liberal arts colleges and flagship state universities.

measures of output by credit hours and FTE student count, and then compare the Laspeyres quantity indices based on credit hours and FTE student count with the NIPA quantity index for higher education. The new estimates show that volume output of tertiary educational institutions has grown substantially between 2000 and 2013, compared to the quantity index in the NIPAs. Price index from the IPEDS data show a wide divergence of tuition growth between public and NPISH universities, which the PCE deflators in the NIPAS do not capture.

This paper points to further areas of research so that we can measure output of education services in the United States. Research topics include refinement of the index measures presented here and measurement of productivity of tertiary educational institutions suggested by the NRC report. More research is needed to produce similar indices for primary and secondary education.

References

Abraham, Katharine G. “Accounting for Investments in Formal Education,” *Survey of Current Business*, June 2010.

Card, David and Lemieux, Thomas, “Can Falling Supply Explain the rising Return to College for Younger Men?” *Quarterly Journal of Economics*, Vol. 116, Issue 2, May 2001.

Castro, Rui, and Coen-Pirani, Daniele, “Explaining the Evolution of Educational Attainment in the United States,” *American Economic Journal: Macroeconomics*, Vol 8 No. 3, July 2016.

Cellini, Stephanie Riegg, “Crowded Colleges and College Crowd-Out: the Impact of Public Subsidies on the Tow-Year College Market,” *American Economic Journal: Economic Policy*, Vol. 1 Issue 2, August 2009.

Cellini, Stephanie Riegg, “Financial Aid and For-Profit Colleges: Does Aid Encourage Entry?” *Journal of Policy Analysis and Management*, Vol. 29 Issue 3, Summer 2010.

Chetty, Raj, Friedman, John N., Hilger, Nathaniel, Saez, Emmanuel, Whitmore Schanzenbach, Diane, and Yagan, Danny, “How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR,” *Quarterly Journal of Economics*, Vol. 126 Issue 4, November2 011.

Deming, David J., Goldin, Claudia, and Katz, Lawrence F., “The For-Profit Postsecondary School Sector: Nimble Critters or Agile Predators?” *Journal of Economic Perspectives*, Vol 26 No. 1, Winter 2012.

Fraumeni, Barbara M., Christian, Michael S., and Samuels, Jon D. “Price and Real Output Measures for the Education Function of Government: Exploratory Estimates for Primary and Secondary Education.” In Diewert, W. Erwin; Greenless, John S., and Hulten, Charles R. (eds.) *Price Index Concepts and Measurement*, University of Chicago for NBER, 2009.

Fraumeni, Barbara M., Reinsdorf, Marshall, Robinson, Brooks B., and Williams, Matthew P. “Price and Real Output Measures for the Education Function of Government: Exploratory Estimates for Primary and Secondary Education.” In Diewert, W. Erwin; Greenless, John S., and Hulten, Charles R. (eds.) *Price Index Concepts and Measurement*, University of Chicago for NBER, 2009.

Gilpin, Gregory A., Saunders, Joseph, and Stoddard, Christina, “Why Has For-Profit Colleges’ Share of Higher Education Expanded So Rapidly? Estimating the Responsiveness to labor Market Changes,” *Economics of Education Review*, Vol. 45, April 2015.

Goldin, Claudia, and Margo, Robert A. "The Great Compression: The Wage Structure in the United States at Mid-Century," *Quarterly Journal of Economics*, Vol. 107, Issue 1, February 1992.

Gu, Wulong and Wong, Ambrose, *Measuring the Economic Output of the Education Sector in the National Accounts*, Statistics Canada Working Paper, 11F0027M No. 080 October 2012.

Gu, Wulong and Wong, Ambrose, "Productivity and Economic Output of the Education Sector," *Journal of Productivity Analysis*, Vol. 43 Issue 2, April 2015.

Hanushek, Eric A. "The Productivity Collapse in Schools," *Development in School Finance*, 1996.

Hanushek, Eric A. "Evidence, Politics, and the Class Size Debate," in Lawrence Mishel and Richard Rothstein (eds.) *The Class Size Debate*, Economic Policy Institute, 2002.

Hanushek, Eric A. "The Failure of Input-Based Schooling Policies," *Economic Journal*, Vol. 113, No. 485: F64-F98, February 2003.

Hanushek, Eric A. and Rivkin, Steven G. "Understanding the Twentieth-Century Growth in U.S. School Spending," *Journal of Human Resources*, Vol. 32 No. 1, Winter 1997.

Heckman, James J. and Rubinstein, Yona, "The Importance of Noncognitive Skills: Lessons from the GED Testing Program," *American Economic Review*, Vol. 91 No. 2, May 2001.

Jorgenson, Dale and Fraumeni, Barbara M., "The Accumulation of Human and Nonhuman Capital," in Robert E. Lipsey and Helen Stone Tice (eds.) *The Measurement of Saving, Investment, and Wealth*, University of Chicago for NBER, 1989.

Jorgenson, Dale and Fraumeni, Barbara M., "The Output of Education Sector," in Zvi Griliches (ed.) *The Output Measurement in the Service Sectors*, University of Chicago for NBER, 1992.

Kendrick, John W., *The Formation and Stocks of Total Capital*, Columbia University Press, 1976.

Kortez, Daniel M. "Limitations in the Use of Achievement Tests and measures of Educator's Productivity," *Journal of Human Resources*, Vol. 37 Issue 4, Autumn 2002.

Kruger, Alan B. "Understanding the Magnitude and Effect of Class size on Student Achievement," in Lawrence Mishel and Richard Rothstein (eds.) *The Class Size Debate*, Economic Policy Institute, 2002.

Moretensen, Jorgen, O'Mahony, Mary, Pastoer, Jose Manuel, Serrano, Lorenzo, and Stokes, Lucy, "Measuring Education Input, Output, and Outcomes: State of the Art and Data Availability," INDICSER Review Paper, January 2011.

O'Mahony, Mary and Stevens, Philip, "Output and Productivity Growth in the Education Sector: Comparisons for the US and UK," *Journal of Productivity Analysis*, Vol. 31 Issue 3, June 2009.

Powers, Susan. "Labor Productivity Growth in Elementary and Secondary School Services: 1989-2012*" Monthly Labor Review*, forthcoming, 2016.

Rubinstein, Jesse, and Rouse, Cecilia, "Constrained after College: Student Loans and Early-Career Occupational Choices," *Journal of Public Economics*, Vol 95, Issue 1-2, February 2011.

Schreyer, Paul, *Towards Measuring the Volume Output of Education and health Services: A Handbook* Statistics Directorate Working Paper No. 31, Organization for Economic Co-Operation and Development, Paris, 2010.

Smith, Peter C. and Street, Andrew, "The Measurement of Non-Market Output in Education and Health," *Economic & Labor Market Review*, Vol. 1 No. 6, June 2007.

Sullivan, Teresa A., Mackie, Christopher, Massy, William F., and Sinha Esha (eds.) *Improving Measurement of Productivity in Higher Education*, Panel on Measuring Higher Education Productivity: Conceptual Framework and Data Needs, The National Academies Press, 2012.

UNESCO Institute for Statistics, "ISCED: International Standard Classification of Education 2011," UNESCO Institute for Statistics, Montreal, Quebec 2012.

Webber, Douglas A., and Ehrenberg, Ronald D. "Do Expenditures Other than Instructional Expenditures Affect Graduation and Persistence Rates in American Higher Education?" Economics of Education Review, Vol. 29 Issue 6, December 2010.

Wei, Hui, "Measuring Australia's Human Capital Development: The Role of Post-School Education and the Impact of Population Ageing," *Statistical Journal of the IAOS*, Vol. 24 Issue 3/4, April 2007.

Wood, Tom, "The Sheepskin Effect," National Association of Scholars at https://www.nas.org/articles/The_Sheepskin_Effect, July 2009.

Table 1 Number of Institutions in IPEDS by Type, 2000-2013

Year	Four-Year or Above			Two-Year			Total
	Public	Private, Non-Profit	Private, For-Profit	Public	Private, Non-Profit	Private, For-Profit	
2000	524	1,827	332	981	390	271	4,328
2001	516	1,525	298	961	226	256	3,782
2002	514	1,504	299	959	223	248	3,747
2003	515	1,516	311	964	219	245	3,770
2004	515	1,510	326	960	214	244	3,769
2005	515	1,501	342	956	209	371	3,894
2006	516	1,507	384	958	189	350	3,904
2007	514	1,505	392	961	191	351	3,913
2008	516	1,509	431	959	176	339	3,930
2009	515	1,501	463	962	181	330	3,952
2010	521	1,510	500	956	172	473	4,132
2011	521	1,506	577	962	167	438	4,171
2012	524	1,523	623	959	178	406	4,213
2013	520	1,524	637	951	172	380	4,184

Year denotes the academic year ending in June of each year. All tables are based on author's calculation from IPEDS data sets obtained from Delta Cost Project. Note that random checks of schools that appear in the 2000 data but not in 2001 reveal many of such institutions are not accredited and thus are not eligible for financial aid under Title IV.

Table 2 Total Credit Hours by Level and Type of Institution, 2000-2013
(in thousands of semester credit hours)

Year	Graduate Hours			Undergraduate Hours at 4-Year Institution			Undergraduate Hours at 2-Year Institution		
	Public	Private, Non-Profit	Private, For-Profit	Public	Private, Non-Profit	Private, For-Profit	Public	Private, Non-Profit	Private, For-Profit
2000	18,630	14,143	751	130,355	56,689	7,480	99,292	1,410	3,743
2001	18,909	14,683	946	131,260	59,872	9,053	103,765	1,590	4,986
2002	20,065	15,960	1,058	137,338	61,665	9,310	111,974	1,530	5,538
2003	20,801	16,397	1,319	141,803	63,862	10,690	119,002	1,618	6,146
2004	20,578	16,468	1,972	143,747	65,100	12,675	120,818	1,621	6,503
2005	21,073	16,815	2,381	144,413	66,401	14,740	120,848	1,640	9,094
2006	21,193	16,887	2,721	146,445	67,057	16,581	119,343	1,188	9,032
2007	21,399	17,238	2,947	149,638	68,572	17,651	121,264	1,403	8,758
2008	22,334	18,300	3,428	152,656	69,770	21,028	127,638	1,206	8,359
2009	23,706	19,518	4,298	155,799	70,935	26,421	136,490	1,316	8,502
2010	26,687	22,084	5,187	162,004	72,889	33,731	150,717	1,351	13,737
2011	27,189	22,962	5,768	165,723	74,653	36,066	155,608	2,328	13,771
2012	22,981	19,891	5,472	169,470	75,805	33,593	147,332	1,564	11,434
2013	22,250	19,928	4,870	166,738	76,325	29,397	141,854	1,463	9,365

Table 3 Total Number FTE Students by Level and Type of Institution, 2000-2013

(in thousands)

Year	Total	Graduate Students			Undergraduates at 4-Year Institutions			Undergraduates at 2-Year Institutions		
		Public	Private, Non-Profit	Private, For-Profit	Public	Private, Non-Profit	Private, For-Profit	Public	Private, Non-Profit	Private, For-Profit
2000	10,761	792	592	33	4,240	1,838	209	2,913	47	93
2001	11,216	805	625	39	4,313	1,925	230	3,117	47	114
2002	11,709	831	634	49	4,453	1,972	271	3,325	45	130
2003	12,293	882	677	69	4,596	2,029	309	3,547	47	136
2004	12,617	903	709	89	4,670	2,070	357	3,591	48	150
2005	12,983	901	732	121	4,762	2,117	443	3,640	47	220
2006	13,159	902	741	137	4,823	2,150	544	3,610	37	216
2007	13,346	909	761	146	4,890	2,183	557	3,645	44	210
2008	13,714	933	792	152	4,984	2,214	641	3,762	36	201
2009	14,286	955	811	186	5,089	2,243	792	3,954	37	219
2010	15,450	993	855	212	5,293	2,292	1,018	4,397	38	351
2011	15,849	1,014	883	223	5,409	2,349	1,059	4,512	37	361
2012	15,704	1,009	893	222	5,487	2,389	1,040	4,315	43	306
2013	15,383	1,002	896	204	5,482	2,410	917	4,179	41	253

Table 4 Total Number of Degrees Conferred by Level and Type of Institution, 2000-2013
(in thousands)

Year	Graduate Degrees			Bachelor Degrees			Associate Degrees		
	Public	Private, Non-Profit	Private, For-Profit	Public	Private, Non-Profit	Private, For-Profit	Public	Private, Non-Profit	Private, For-Profit
2000	304	262	11	811	406	21	445	47	59
2001	307	269	14	812	408	24	448	47	70
2002	311	276	16	841	427	27	464	48	73
2003	328	292	18	875	445	32	490	48	81
2004	350	310	29	906	455	42	516	48	83
2005	359	316	34	932	461	45	539	48	97
2006	364	324	47	955	471	62	552	46	109
2007	365	333	53	976	481	70	574	49	110
2008	377	344	56	996	491	74	569	45	121
2009	387	361	67	1,019	497	84	603	47	134
2010	403	374	72	1,049	504	94	628	47	159
2011	423	388	85	1,088	511	111	683	47	194
2012	436	403	87	1,131	527	132	733	53	203
2013	436	407	80	1,161	535	133	755	53	172

Figure 1 Algorithm to Calculate Adjusted Credit Hours per FTE Student

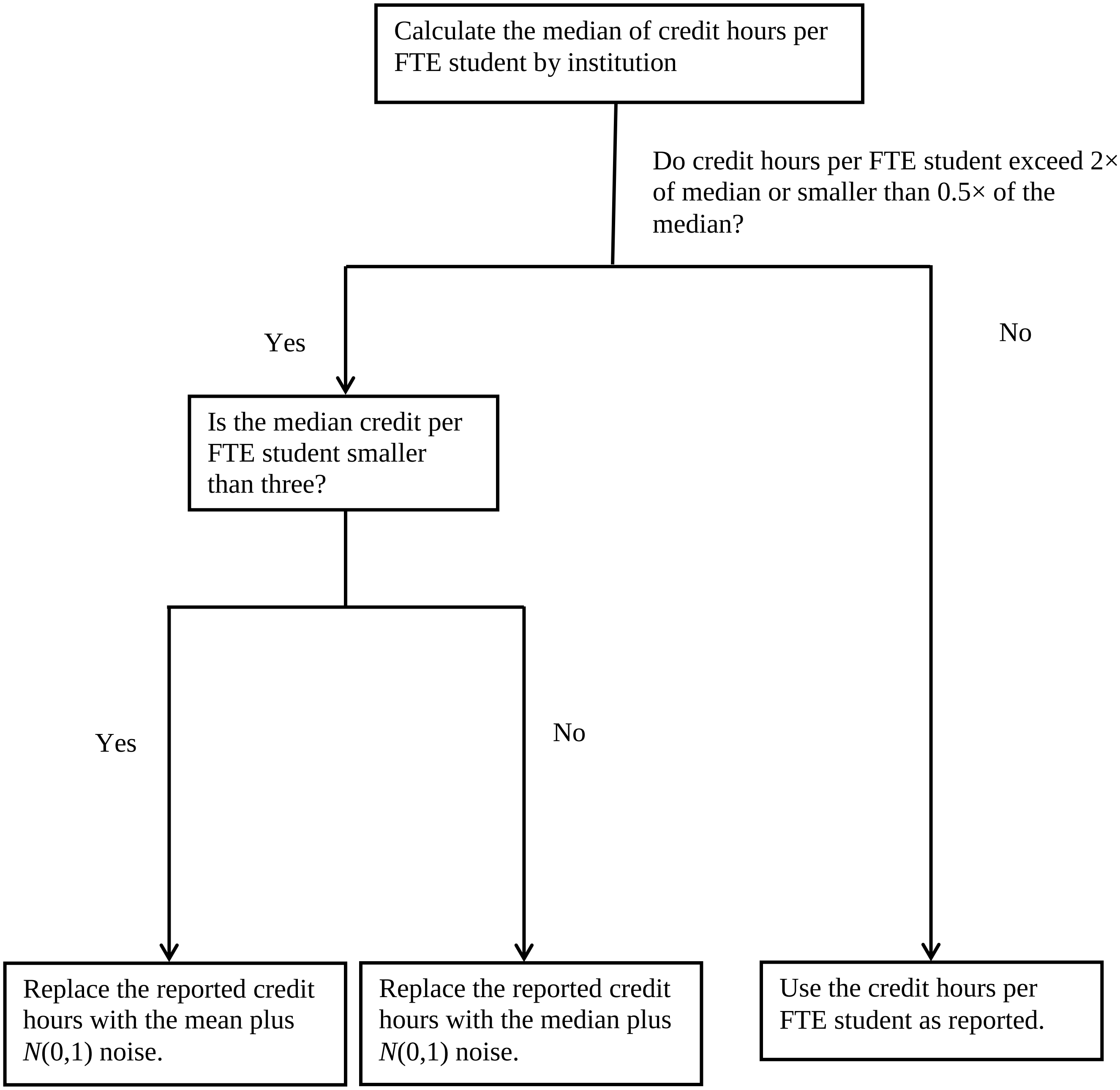
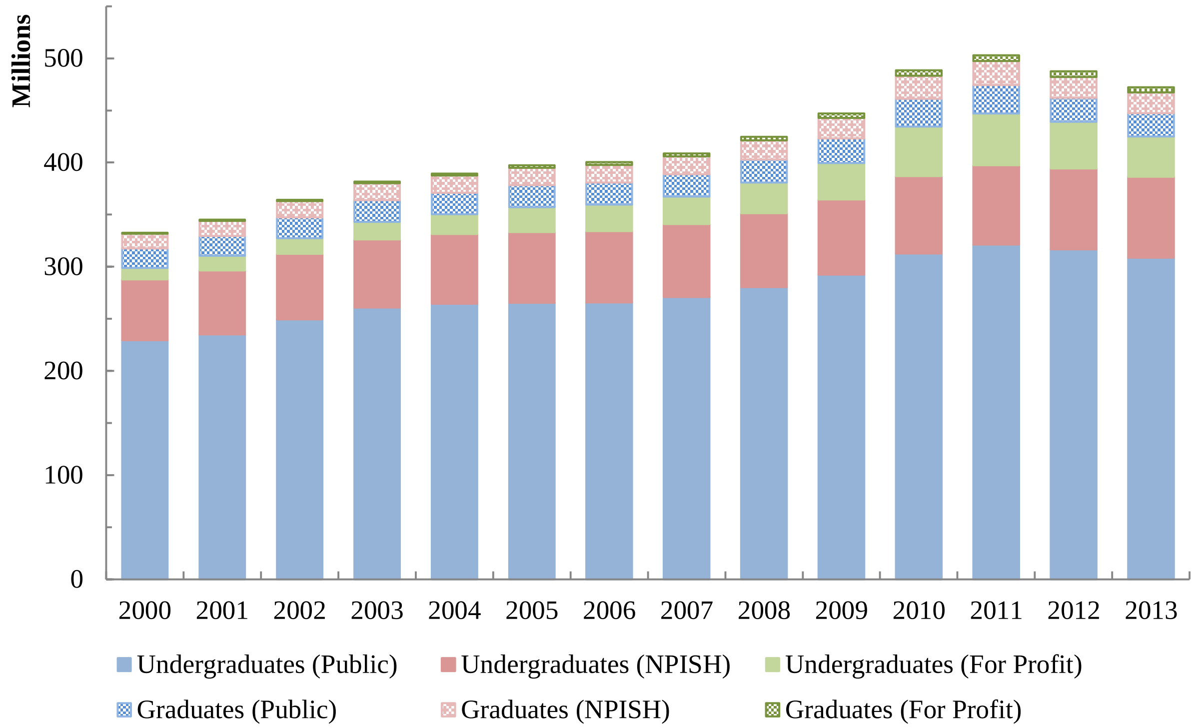


Figure 2 Total Credit Hours and FTE Students by Sector, 2000-2013

(a) Total Credit Hours by Type of Institution



(b) Total FTE Students by Type of Institution

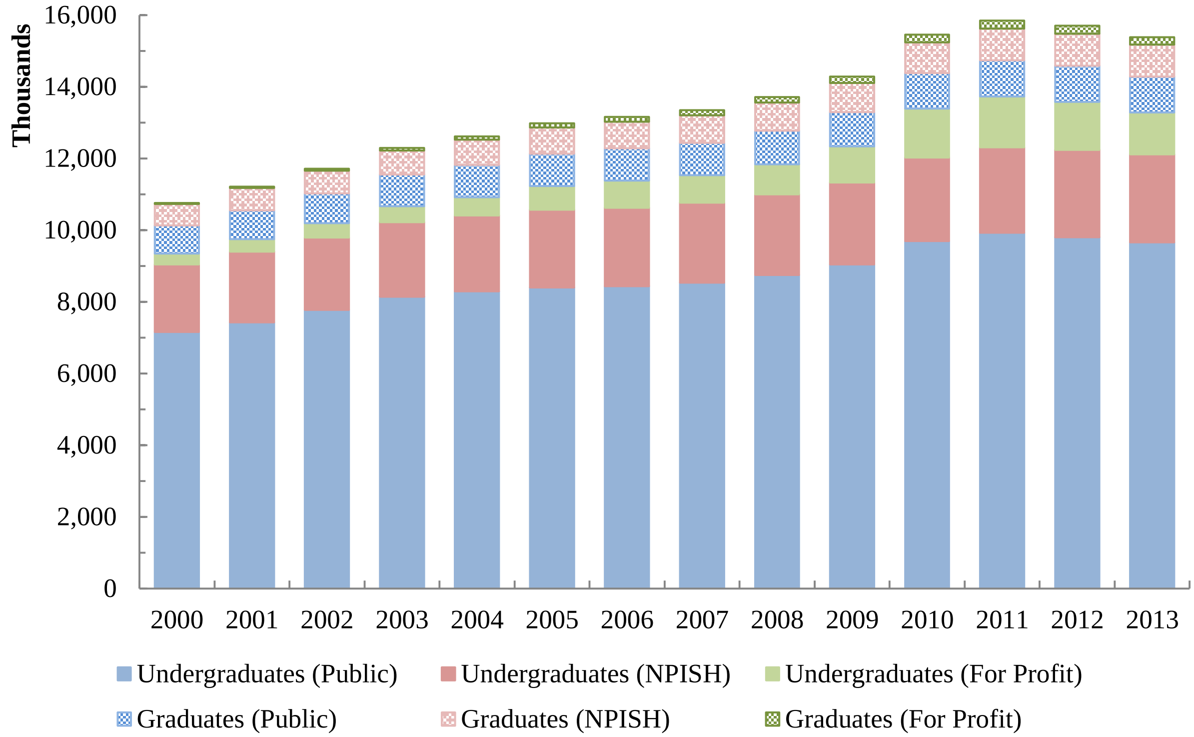


Figure 3 Comparison of IPEDS Quantity Indices and NIPA Quantity Index for Higher Education, 2000-2013 (2000=100)

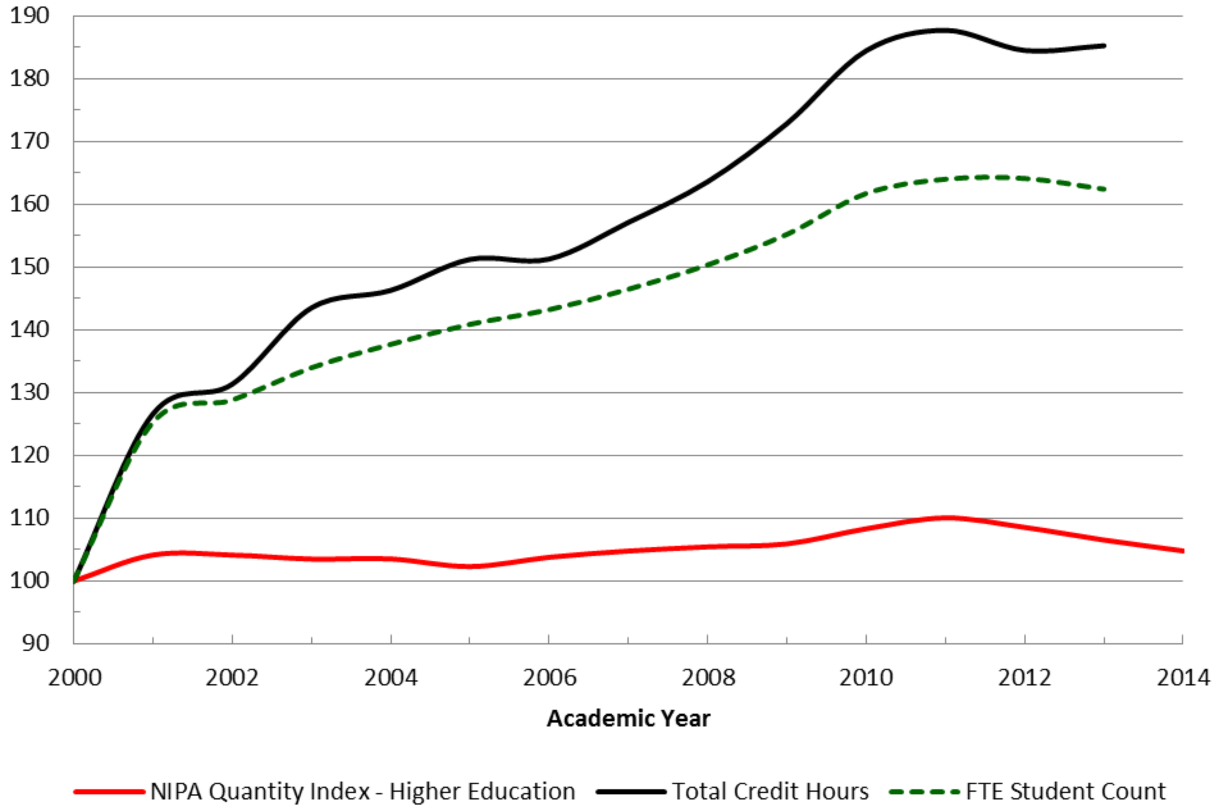


Figure 4 IPEDS Credit Hours vs NIPA Quantity Index by Sector of Higher Education 2000-2013 (2000=100)

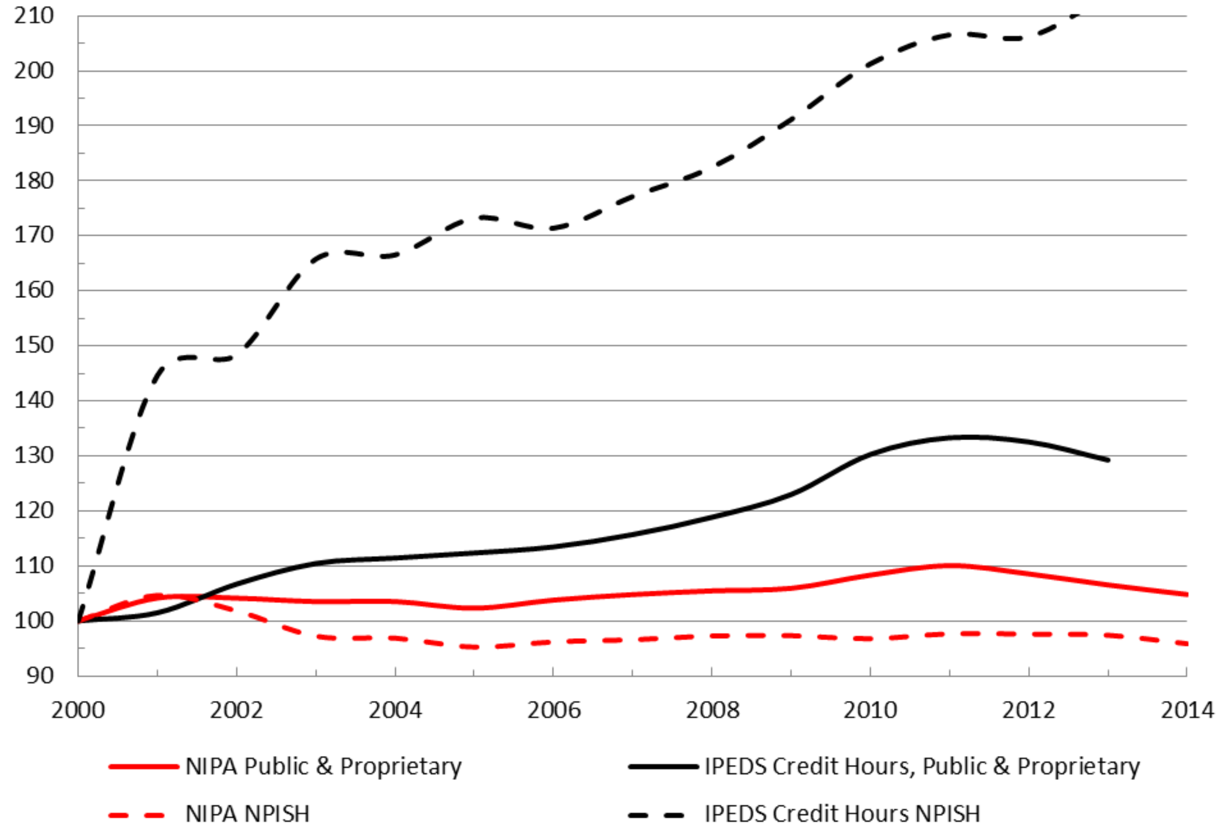


Figure 5 Comparison of Net Student Tuition Price Index and NIPA PCE Deflator for Education Services, 2000-2013 (2000=100)

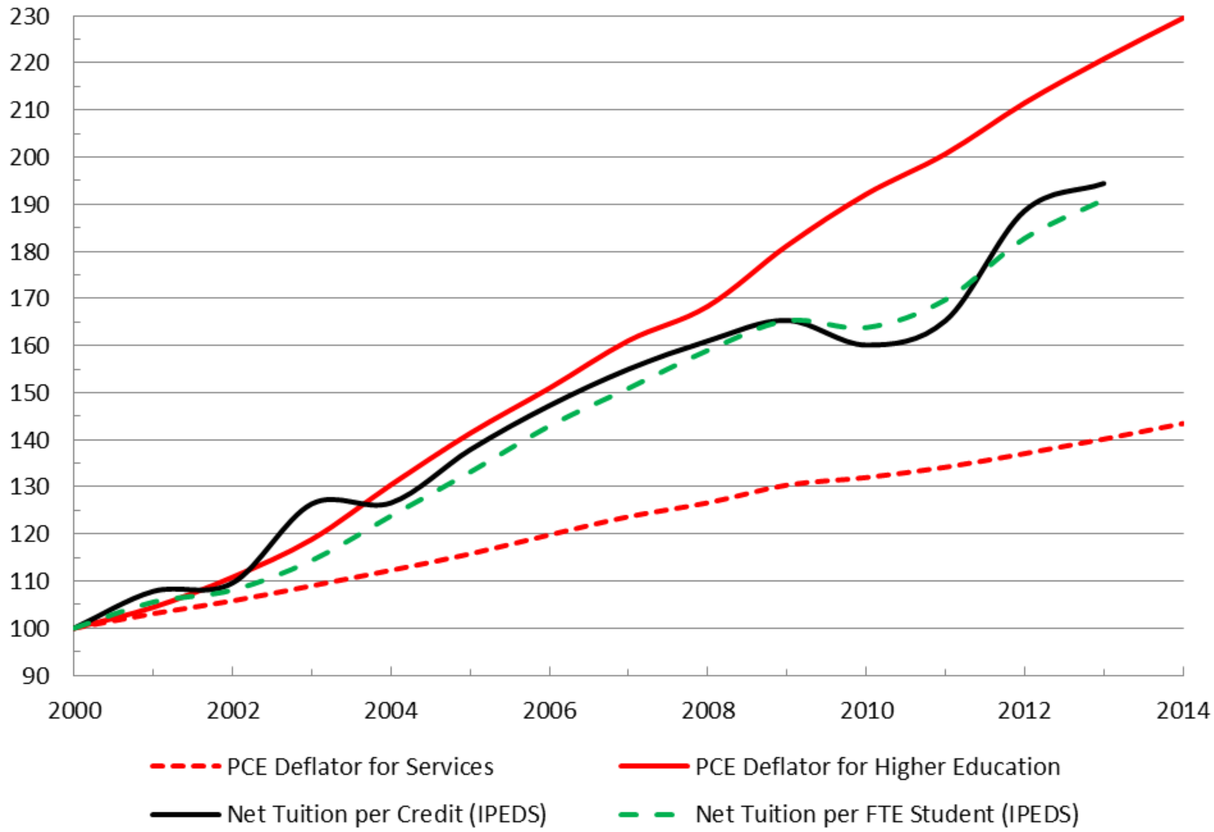


Figure 6 Net Student Tuition Index vs NIPA PCE Deflator by Sector of Higher Education, 2000-2013 (2000=100)