Merging Data to Facilitate Analyses

Robert Kelchen

Associate Professor, Department of Education Leadership. Management and Policy

Seton Hall University

robert.kelchen@shu.edu

This is the originally submitted version (December 2018) of an article published in *New Directions for Institutional Research* in November 2019. The final version can be accessed at https://authorservices.wiley.com/api/pdf/fullArticle/16581951.

Keywords: IPEDS, merging data, institutional research, higher education policy

Abstract: In order to demonstrate the value of higher education and evaluate the effectiveness of various policies and practices, researchers are increasingly being expected to merge data from the Integrated Postsecondary Education Data System (IPEDS) with a range of other federal and private-sector data sources. In this article, I detail a number of the most important higher education data sources and explain how to properly merge them in order to conduct analyses. I conclude by providing some examples of how researchers have merged together multiple datasets in order to conduct policy analyses.

Although the typical person sees significant economic and noneconomic benefits from a college education (Doyle & Skinner, 2017; Webber, 2016), the American public has become increasingly skeptical of the value of higher education since the Great Recession began. A number of surveys have found concerns about whether higher education was worth the price tag or even important for economic success (Schleifer & Silliman, 2016; Strada Education Network & Gallup, 2018; Taylor, 2011), while first-year college students report that getting a good job was the most important reason why they attended college (Eagan, Stolzenberg, Zimmerman, Aragon, Sayson, & Rios-Aguilar, 2017).

With heightened accountability pressures for nearly all colleges and universities (Kelchen, 2018), it is the job of institutional researchers to harness data to show the power of higher education. Institutions must now demonstrate their own value or otherwise face the risk of having an accountability system placed upon them. As a result, researchers are increasingly being asked to conduct more sophisticated analyses to help answer important institutional and public policy questions. Thanks to new datasets such as those from the federal government and private sources such as Opportunity Insights, colleges can now benchmark their performance to other institutions on metrics such as the share of first-generation students, earnings of former students, social mobility rates, and the number of military students served. Additionally, there are a number of state-level and local-level data sources that can be used to provide important context in institution-level analyses or to improve benchmarking.

In this paper, I discuss how to use a range of new and existing data sources that can be used in conjunction with data from the Integrated Postsecondary Education Data System (IPEDS) to research a number of important higher education topics. I detail how to combine these datasets with IPEDS and provide information about the strengths and limitations of these

merges. I conclude by showing some analyses that are possible after merging together multiple data sources and how they can be used to improve the state of higher education research.

Federal Data Sources

The federal government provides a number of data sources about student outcomes, institutional finances, and how colleges perform on accountability metrics in addition to IPEDS. Most of these data are compiled by the U.S. Department of Education's Office of Federal Student Aid (FSA), which collects data on colleges receiving federal financial aid under Title IV of the Higher Education Act. FSA data classify institutions based on their Office of Postsecondary Education identification number (OPEID), which is based on the unit of analysis under which a program participation agreement is entered upon with the Department of Education (Office of Federal Student Aid, 2017).

Merging FSA data onto IPEDS is tricky because there is not a one-to-one relationship between IPEDS UnitIDs and FSA OPEIDs. College and university systems often operate under the same program participation agreement, and thus all report data together to FSA under what is often called a 'parent-child' agreement (Jaquette & Parra, 2014). However, seemingly similar university systems differ in whether institutions report separately or together to FSA. For example, Indiana University and University of Wisconsin campuses report separately, while Ohio State University and Rutgers University report as systems. Further complicating this merge is that colleges that share the same program participation agreement can report certain IPEDS data elements (such as finance and completions) at the OPEID level while reporting other elements at the UnitID level (National Center for Education Statistics, 2018).

The way to identify potential parent-child reporting relationships is to look at the eight-digit OPEID number. The first six digits represent the unique program participation agreement number, while the final two digits indicate whether the campus is the parent (00) or child (01-99 or an alphanumeric combination) institution. Of the 7,128 Title IV-participating institutions in the 2017 IPEDS Directory Information survey, 5,691 (79.8%) were parent institutions and the remainder were child institutions. The parent-child issue is particularly acute in the for-profit sector, with 36.3% of institutions being child institutions; this compares to 5.9% of public institutions and 8.4% of nonprofit institutions (author's calculations using IPEDS data).

Once the parent and child institutions are identified, analysts must make decisions about how to treat these data elements. One option is to simply drop all institutions with parent-child issues, which results in a reasonably representative sample in the public and private nonprofit sector but eliminates most of the large for-profit college chains. Another option is to aggregate IPEDS data to the OPEID level and treat multiple institutions as one in some cases. Finally, analysts could assume that parent and child institutions are similar enough that using the same data elements is a reasonable step to take (while also clustering standard errors at the OPEID level). None of these solutions are perfect, so analysts should carefully consider the pros and cons of each strategy.

College Scorecard

The College Scorecard was first proposed by President Obama during a 2012 speech at the University of Michigan (The White House, 2012), and soon a website was created that showed basic information on college prices and graduation rates. But the Scorecard did not contain any data that were not already on other federal websites until a September 2015 relaunch

that came out of the Department of Education's aborted effort to rate colleges (Kelchen, 2018). The updated version of the College Scorecard tied together financial aid data from FSA with earnings data from the Department of the Treasury to provide the first national look at how colleges performed on a range of metrics (Council of Economic Advisers, 2017). Scorecard data can be accessed in three forms: by downloading the full dataset from the 1996-97 academic year to the present, by accessing a slimmed-down dataset of key metrics for the most recent year available, or by using the consumer-facing dashboard that has a small number of metrics for the most recent year.

Some of the most prominent new elements in the full Scorecard dataset are the following:

- Earnings are measured 6, 8, and 10 years after college entry for more recent cohorts and 7 and 9 years after college entry for older cohorts. The mean, median, and 10^{th}, 25^{th}, 75^{th}, and 90^{th} percentile distributions are available for all colleges along with a measure of the percentage of students earning above $25,000 or $28,000 (the average earnings of a high school graduate), depending on the cohort. Earnings are also available by gender, family income tercile, and dependency status during college.
- Undergraduate federal student loan burdens (excluding private and PLUS loans) are measured upon leaving college. The mean, median, and percentile distributions are all available, as are debt burdens by gender, family income tercile, dependency status, completion status, and first-generation status.
- There is a student loan repayment rate defined as the percentage of borrowers repaying at least $1 in principal 1, 3, 5, and 7 years after entering repayment. Loan repayment rates are available for the same subgroups as student loan debt, as well as by Pell/non-Pell status. Analysts using data from the fall 2015 or fall 2016 data releases should download

the most recent data due to a coding error corrected in January 2017 that significantly overstated repayment rates in earlier releases (Fain, 2017).

- The percentage of first-generation students (defined as the share of students who do not have at least one parent with some college experience) is the final important new measure.

One important limitation of the Scorecard is that only students who ever received federal financial aid are included because it is based on FSA data from the National Student Loan Data System. This covers roughly 70% of all students, but coverage rates are far lower in community colleges and highly-selective four-year institutions than for-profit colleges and broad-access four-year colleges (Council of Economic Advisers, 2017).

Analysts should also watch out for two other issues when working with College Scorecard data. The first issue is the sheer size of the full Scorecard datasets, with the 21 annual data files (as of the fall 2018 data release) checking in at a 240 megabyte .zip file. Not all computers and statistical software packages can handle working with the full data files due to the large number of variables included, so check your software's operating capacity and close all other processes before beginning to use Scorecard data.

Second, the rapidly changing landscape of closures, mergers, and acquisitions makes it difficult to link UnitIDs and OPEIDs across time. Without taking this into account, there is the possibility that colleges will inappropriately be dropped from an analyst's sample or that colleges that have consolidated will not have that consolidation reflected when using old OPEIDs. The .zip file with the full College Scorecard datasets contains crosswalk files for each year of data with information about how UnitIDs and OPEIDs have changed across time. Analysts interested

in panel analyses, particularly in the for-profit sector or in states like Georgia and Indiana with frequent higher education mergers or restructurings, should familiarize themselves with how OPEIDs and UnitIDs have changed to facilitate accurate matching.

Federal Student Aid Data

Federal Student Aid produces a number of institution-level datasets that detail how much money colleges (at the OPEID level) receive from federal Title IV financial aid programs and how institutions fare on a range of federal accountability provisions. FSA's Title IV Program Volume Reports webpage (Office of Federal Student Aid, 2018a) contains annual datasets on federal grant and loan programs going back to the 1999-2000 award year along with data on the campus-based aid programs (Federal Work-Study, the now-expired Perkins Loan program, and the Supplemental Educational Opportunity Grant) beginning in the 2001-02 award year.

Analysts should be aware of several concerns in working with the Title IV volume datasets. First, spreadsheets are available for the non-campus-based programs on a quarterly basis, with tabs for "Quarterly Activity" and "Award Year Summary." To get a complete look at funds received in a given year, analysts must use the "Award Year Summary" in the Q4 spreadsheet for each year. Unfortunately, the "Quarterly Activity" spreadsheet is the default tab, meaning that additional care must be exercised (and the "Award Year Summary" must be specified when importing the data into a statistical software program).

The loan data have two additional concerns that analysts must address. First, the loan data include both the amount of loans originated and the amount disbursed. My recommendation is to use the amount disbursed because not all loans that are initiated are actually given to students, while the disbursement number includes all actual awards. Second, prior to the end of the 2009-

10 award year, federal loans were awarded through both the now-defunct Federal Family Education Loan Program and the Direct Loan Program. To get a complete total of loan volume, analysts must add together the disbursement amounts from the two programs. Finally, the campus-based aid data include columns for both the federal award (funds provided to the college to run the program) and the amount disbursed because campuses provide matching funds for these programs. Analysts must choose which figure is more relevant for their needs.

FSA also maintains a number of databases for accountability purposes, all of which only include a subset of Title IV-eligible colleges. Colleges participating in the federal student loan program must pass a cohort default rate (CDR) metric in order to maintain eligibility to offer students federal loans and/or any Title IV financial aid, depending on default rates. FSA also maintains a list of colleges facing heightened cash monitoring (HCM) sanctions due to financial, accreditation, or institutional capacity concerns. The CDR dataset is updated annually, while the HCM dataset is updated quarterly. Both are included in the College Scorecard and updated there on an annual basis, but analysts looking for the most up-to-date data must go to the FSA webpages.

Certain types of colleges are required to meet three additional accountability requirements in order to be eligible to receive financial aid, and FSA maintains separate datasets for each policy that can be merged onto IPEDS and other data sources. One policy that affects private nonprofit and for-profit colleges is a financial responsibility requirement, which requires colleges to meet a benchmark score of financial health in order to receive federal funds or face additional requirements. For-profit colleges must pass the 90/10 rule, which requires institutions to get at least 10% of their total revenue from non-federal sources (excluding veterans' benefits). Finally, gainful employment regulations require nearly all programs at for-profit colleges and

nondegree programs at other institutions to pass a debt-to-earnings test in order to receive federal funds. The gainful employment regulations, which operate at the program level instead of the institutional level, are currently in limbo but one year of data was released in early 2017 (Kreighbaum, 2018). For more details on each of these accountability policies and how they affect colleges, see Kelchen (2018).

Military and Veterans' Benefits Data

Analysts interested in examining military and veterans' benefits by college can turn to three different federal data sources. The IPEDS student financial aid survey contains information on the number of students receiving Post-9/11 GI Bill benefits or Department of Defense Tuition Assistance Program benefits as well as the total amount of funds, broken down by undergraduate and graduate students. These data are typically available at the UnitID level and were first collected in the 2013-14 academic year. FSA's Title IV volume reports contain details at the OPEID level over time on the Iraq/Afghanistan Service Grant Program, which is a smaller grant program (57 students received it in 2017-18, according to FSA data).

A new longitudinal database on Post-9/11 GI Bill benefits was first released by the Department of Veterans Affairs (2018) in March 2018. This dataset contains information on the number of students receiving benefits and the amount of benefits that went to pay tuition and fees beginning in Fiscal Year 2009 (the 2008-09 academic year). The data appear to be presented at the institutional level, but the facility code included in the dataset does not match either UnitIDs or OPEIDs. As a result, analysts must create their own crosswalk between the VA data and IPEDS data.

Research Productivity Data

Given the federal government's large investment in sponsored research and supporting doctoral students, it is not surprising that the federal government also collects a substantial amount of institution-level data to check in on how its resources are used. There are two main federal sources of research productivity data that are infrequently used by analysts in spite of being around for many years.

The first is the Higher Education Research and Development Survey (HERD), which is sponsored by the National Science Foundation (NSF), contains annual data on sources and uses of research dollars going back to Fiscal Year 1972 (National Science Foundation, 2018). Since Fiscal Year 2010, IPEDS UnitIDs are provided in the dataset for ease of merging, but prior years include FICE codes (a precursor to UnitIDs) as the identifier. The FICE variable is then renamed inst_id beginning in 2010, allowing for a crosswalk to be created for ease of analysis. One additional analytic challenge is that the data are presented in long format (with each variable having a separate row), meaning that analysts will need to reshape the data into wide format for many analyses.

The second data source is the Survey of Earned Doctorates (SED), which is a census of research doctorate recipients that has been conducted annually and sponsored by NSF since 1957. The survey contains information on the number of research doctorates by field of study (excluding most professional doctorates) awarded by each university, as well as the baccalaureate origins of new doctoral degree recipients (National Center for Education Statistics, 2017). The data can be downloaded going back to 1958 using the NSF's data tool at https://ncsesdata.nsf.gov/ids/sed. Analysts should note that UnitIDs are not included in the dataset, which makes the merging process more tedious and requires analysts to also download

the state along with the name (for cases such as Wheaton College where multiple institutions share the same name).

Opportunity Insights Data

The U.S. Department of Education has made great strides in recent years to make institution-level data on social mobility public by including earnings, debt burdens, and loan repayment rates by family income in the College Scorecard and publishing Pell Grant recipients' graduation rates in IPEDS. But these data pale in granularity to the dataset Raj Chetty of Harvard University and his colleagues released under the Equality of Opportunity Project banner in 2017 and renamed Opportunity Insights in 2018. Chetty's team was able to construct a dataset following traditional-age students from high school into young adulthood using IRS data on parental earnings, tuition payments to colleges, and student earnings (Chetty, Friedman, Saez, Turner, & Yagan, 2017). They created social mobility rates, which examine the percentage of students who come from various income quintiles and end up in higher quintiles. In addition, they also created median earnings for each cohort as well as tracking the percentage of students who were observed as being married in tax data.

The Opportunity Insights dataset covers students who were born between 1980 and 1991 and were tracked through 2014. This allows the earliest (1980-82) cohorts to be tracked through ages 32-34, with the focal estimates being on these cohorts since the earnings distribution stabilizes by approximately age 30. But there are separate datasets for each of the 12 birth cohorts for analysts who are interested in changes in access and social mobility among students who started college during the late 1990s and 2000s.

There are two main drawbacks to the Opportunity Insights dataset. The first is that unlike most of the federal datasets (with gainful employment being the possible exception), it is far from clear whether this dataset will continue to be updated in the future. The initial dataset was compiled through a data agreement with the United States Treasury, which would be difficult to renegotiate without significant support from the federal government. The data still provide an important look at social mobility in the United States, but the value to colleges will decline over time as the youngest birth cohort started college about ten years prior to the publication of this article.

The second drawback is due to how colleges report tuition data to the IRS for tax filing purposes. This unit of analysis varies across institutions, with some reporting at the UnitID level, some reporting at the OPEID level, and some reporting as broader (and occasionally unclear) systems under what the researchers classify as a "super OPEID." The Opportunity Insights dataset contains information on 2,461 reporting units, with many smaller colleges being excluded due to sample size restrictions. Of these units, 2,143 have unique UnitIDs, 222 are at the OPEID level, and there are 96 super OPEID clusters (for a list, see Kelchen (2017)). Some of the super OPEID clusters are logical (all three public universities in Arizona report together, as does the University of Wisconsin System), while others appear haphazard at best ("University Of Maryland System (Except University College) And Baltimore City Community College", "Certain Colorado Community Colleges").

Table 11 of the Opportunity Insights dataset provides information on which super OPEIDs match onto each OPEID. Super OPEIDs of less than 1000 (or the multi variable being equal to 1) indicate that multiple OPEIDS match to a super OPEID, while super OPEIDs of greater than 1000 match to the first four digits of the main OPEID. This means that to merge

onto IPEDS data, multiply the OPEID in the Opportunity Insights data by 100. For most analyses, colleges in which the super OPEID encompasses multiple OPEIDs will likely be dropped, but it is possible to pool IPEDS or other data up to the super OPEID level if desired.

State and Local Data Sources

In addition to the institution-level data sources provided above, a number of resources exist that help provide state-level and local-level context about how colleges operate. These data elements can be used to help better identify relevant comparison institutions, and they can also help to explain factors underlying changes in colleges' actions. They can be merged onto IPEDS using the existing state/FIPS or county codes. I discuss some of the key state and local data sources in this section.

The two best sources for longitudinal state higher education finance data come from the State Higher Education Executive Officers Association's annual State Higher Education Finance (SHEF) report and the National Association of State Student Grant and Aid Programs (NASSGAP)'s annual state financial aid report. The SHEF report includes state-level data on sources and uses of state funding for higher education as data on full-time equivalent enrollment and net tuition revenue for public colleges and universities. In recent years, some of these data are also available separately for two-year and four-year colleges. NASSGAP's annual report includes data on the amount of state need-based aid, merit-based aid, student loans, and other aid, with grant aid data being broken down between undergraduate and graduate students.

These data sources allow for the creation of two different types of higher education funding effort metrics when combined with Census Bureau data on state populations. One type

of metrics examines total state appropriations per college-age student, while the other looks at need-based grant aid per college-age student. Because Census Bureau data can be broken down for each age (such as 18-year-olds or 45-year-olds), the measures can be customized to meet the analyst's preferences. For example, college-age student could be defined as ages 18-24 or ages 18-65, if desired.

There are a number of federal and private-sector data sources that provide useful information on state economic, demographic, and political characteristics, with the Correlates of State Policy Project at Michigan State University serving as a repository for many of these variables over time (Jordan & Grossman, 2017). Economic characteristics of interest include unemployment rates (Bureau of Labor Statistics), median household income (Bureau of Economic Analysis), and the percentage of residents living in poverty (Census Bureau). Demographic characteristics of interest include educational attainment rates and the racial/ethnic makeup of the state (both from the Census Bureau), while political characteristics include partisan control of the state House, Senate, and governor's office (National Conference of State Legislatures). Many of the economic and demographic characteristics are also readily available at the county level through sources such as the Bureau of Labor Statistics and the Census Bureau's American Community Survey (ACS). However, county-level ACS data only go back to 2005 and in some cases estimates are averaged across multiple years of data.

Examples of Data Merges with IPEDS

In published academic research, it is common to see IPEDS data merged with other data sources in order to answer questions relevant to researchers and practitioners. In this section, I briefly discuss how researchers have used data merges to answer an important policy question: the various factors that are associated with whether students are able to avoid defaulting on their loans and begin repaying principal.

Because data on student loan default and repayment rates are not contained in IPEDS, analysts must merge other federal datasets into IPEDS in order to conduct analyses. Before the creation of the College Scorecard, analysts merged FSA data on default rates into IPEDS in order to examine factors which affected default rates. For example, Hillman (2015) merged these two datasets together with an accreditation database maintained by the U.S. Department of Education to explore the extent to which student demographics, institutional sector, and accreditation agency were associated with having default rates that resulted in federal sanctions. Webber and Rogers (2014) used IPEDS and FSA data to study how institutional finances, selectivity, location, and institutional type were related to default rates among four-year colleges.

To this point, the only research that has used College Scorecard data to examine factors affecting student loan repayment rates is Kelchen and Li (2017). In that research, we merged College Scorecard data on default rates, repayment rates, and the percentage of first-generation students with IPEDS data on institutional characteristics, student demographics, graduation rates, and student financial aid characteristics. Because student loan repayment may be affected by economic conditions that students face, we also used Bureau of Labor Statistics data on state unemployment rates and household income, educational attainment, and poverty data from the Census Bureau as additional controls. Many of these state-level characteristics explained

differences in shorter-term and longer-term repayment rates, highlighting the importance of merging together multiple datasets.

Concluding Remarks

Given the growing importance of rigorous empirical analyses to demonstrate the value of higher education and to figure out whether commonly-adopted policies are effective in improving student outcomes, analysts are increasingly being asked to construct datasets using more than IPEDS data. In this article, I have detailed some of the most prominent institution-level data sources and the challenges that come with merging them into IPEDS. Many of the non-IPEDS datasets have been used relatively infrequently by researchers, in part due to their complexity and in part due to their relative newness. For example, few published journal articles to this point have used College Scorecard or Opportunity Insights data, even though both sources have the potential to improve the body of knowledge. Additionally, I highlighted additional sources of state-level and county-level data that provide important context for analyses and can also be used for benchmarking purposes.

Ultimately, the future of higher education policy research (from the perspective of both faculty members and institutional researchers) will require merging IPEDS datasets with both federal sources such as the College Scorecard and policy datasets compiled by the analyst in order to answer important policy questions. Many important federal and state policies are not covered by the standard datasets and thus require additional data collection or the use of third-party datasets. For example, Klasik and Hutt (2018) combined data on accreditation actions compiled by the Center for American Progress with IPEDS and College Scorecard data, while

research on state performance-based funding policies relies on researcher-developed measures of these policies merged with IPEDS data and state policy characteristics (see Li (2018) for an example). In order to do these types of research, analysts must become comfortable merging data across a number of sources while collecting their own institution-level data as needed.

References

Chetty, R., Friedman, J., Saez, E., Turner, N., & Yagan, D. (2017). *Mobility report cards: The role of colleges in intergenerational mobility.* Cambridge, MA: National Bureau of Economic Research Working Paper 23618.

Council of Economic Advisers (2017). *Using federal data to measure and improve the performance of U.S. institutions of higher education*. Washington, DC: Executive Office of the President.

Department of Veterans Affairs (2018, March 12). *VA releases data of educational organizations receiving GI Bill payments.* Accessed 13 December 2018 from https://www.blogs.va.gov/VAntage/46298/va-releases-data-educational-organizations-receiving-gi-bill-payments/.

Doyle, W. R., & Skinner, B. T. (2017). Does postsecondary education result in civic benefits? *The Journal of Higher Education, 88*(6), 863-893.

Eagan, K., Stolzenberg, E. B., Zimmerman, H. B., Aragon, M. C., Sayson, H. W., & Rios-Aguilar, C. (2017). *The American freshman: National norms fall 2016*. Los Angeles, CA: Cooperative Institutional Research Program.

Fain, P. (2017, January 16). College Scorecard screwup. *Inside Higher Ed*. Accessed 11 December 2018 from https://www.insidehighered.com/news/2017/01/16/feds-data-error-inflated-loan-repayment-rates-college-scorecard.

Hillman, N. W. (2015). Cohort default rates: Predicting the probability of federal sanctions. *Educational Policy, 29*(4), 559-582.

Jaquette, O., & Parra, E. E. (2014). Using IPEDS for panel analyses: Core concepts, data challenges, and empirical applications. Pp. 467-533 in M. B. Paulsen (Ed.), *Higher education: Handbook of theory and research (Vol. 29).* Dordrecht, The Netherlands, Springer.

Jordan, M. P., & Grossman, M. (2017). *The correlates of state policy project v.2.1.* East Lansing, MI: Institute for Public Policy and Social Research.

Kelchen, R. (2017). *Beware OPEIDs and super OPEIDs.* Accessed 14 December 2018 from https://robertkelchen.com/2017/08/21/beware-opeids-and-super-opeids/.

Kelchen, R. (2018). *Higher education accountability.* Baltimore, MD: Johns Hopkins University Press.

Kelchen, R., & Li, A. Y. (2017). Institutional accountability: A comparison of the predictors of student loan repayment and default rates. *The ANNALS of the American Academy of Political and Social Science, 671,* 202-223.

Klasik, D., & Hutt, E. L. (2018). Bobbing for bad applies: Accreditation, quantitative performance measures, and the identification of low-performing colleges. *The Journal of Higher Education.* doi: 10.1080/00221546.2018.1512804.

Kreighbaum, A. (2018, October 4). Missed deadline stalls DeVos agenda. *Inside Higher Ed*. Accessed 13 December 2018 from https://www.insidehighered.com/news/2018/10/04/education-department-misses-deadline-its-overhaul-student-loan-rules.

Li, A. Y. (2018). Performance funding policy impacts on STEM degree attainment. *Educational Policy.* doi: 10.1177/0895904818755455.

National Center for Education Statistics (2017). *Survey of earned doctorates.* Washington, DC: Author.

National Center for Education Statistics (2018). *2018-19 survey materials.* Accessed 10 December 2018 from https://surveys.nces.ed.gov/ipeds/visresults.aspx.

Office of Federal Student Aid (2017). *2017-2018 Federal Student Aid handbook.* Washington, DC: U.S. Department of Education.

Office of Federal Student Aid (2018a). *Title IV program volume reports.* Accessed 13 December 2018 from https://studentaid.ed.gov/sa/about/data-center/student/title-iv.

Schleifer, D., & Silliman, R. (2016). *What's the payoff? Americans consider problems and promises of higher education.* New York: Public Agenda.

Strada Education Network & Gallup (2018). *2018 Strada-Gallup alumni survey.* Indianapolis, IN and Washington, DC: Authors.

Taylor, P., Parker, K., Fry, R., Cohn, D., Wang, W., Velasco, G., & Dockterman. D. (2011). *Is college worth it? College presidents, public assess value, quality, and mission of higher education.* Washington, DC: Pew Research Center.

The White House (2012, January 27). FACT SHEET: President Obama's blueprint for keeping college affordable and within reach for all Americans. Accessed 11 December 2018 from https://obamawhitehouse.archives.gov/the-press-office/2012/01/27/fact-sheet-president-obama-s-blueprint-keeping-college-affordable-and-wi.

Webber, D. A. (2016). Are college costs worth it? How ability, major, and debt affect the returns to schooling. *Economics of Education Review, 53*, 296-310.

Webber, K. L., & Rogers, S. L., (2014). Student loan default: Do characteristics of four-year institutions contribute to the puzzle? *Journal of Student Financial Aid, 44*(2), 99-124.