Data Appendix to "The Financialization of U.S. Higher Education"

The analysis in this paper uses a unique new database that we have constructed to harmonize financial, enrollment, and other organizational data for all colleges that are eligible to receive Title IV funds under the U.S. Higher Education Act. For this paper, we used the database's variables from the Integrated Post Secondary Education Data System (IPEDS) and the National Association of College and University Business Officers (NACUBO) Endowment Survey. We also use original data gathered by the authors to code the ownership form of all Title IV eligible for-profit colleges form 1997 through 2012. The data on ownership form was created using the Thompson One database of private equity investment, for-profit online course catalogs, SEC 10-K statements, and unpublished documents on for-profit college ownership provided to the authors by the U.S. Senate Committee on Health, Education, Labor, and Pensions. We use IPEDS data only after 2003 because of problems of missing data for key variables in years prior to 2003. IPEDS data was used to estimate total revenue for two-year and above colleges as a key component of all funding for higher education from the federal government, states, municipalities, households, donors, and enterprises operated by colleges. IPEDS data was also used to estimate colleges' total spending, institutional debt, interest spending for institutional debt, full-time-equivalent (FTE) student enrollment, and student loan borrowing by full-time freshmen at four-year colleges. Data from both NACUBO and IPEDS was used for endowment asset analysis and estimates.

This data appendix will explain how we dealt with problems of missing data in IPEDS and NACUBO for the years 2003 and after. We will address these issues for data from each section of the paper in the order that they appear in the paper.

Total Spending on Higher Education

Before proceeding with section-by-section issues, we will first address how we estimated total higher education expenditures, because we use total higher education expenditures as a baseline for comparison throughout the paper. For at least the last 15 years, the U.S. Department of Education and the OECD have used data from IPEDS to estimate total spending on higher education in the U.S. This measure simply takes the sum of all expenditures reported by all Title IV institutions reported in IPEDS. This is a

better measure for total higher education spending than measures of colleges' total revenue which are extremely volatile in years with large swings in endowment asset values. Total college spending thus better represents the amount of aggregate funds from many sources such as tuition, endowment returns, and state appropriations that get expended on higher education. This measure, however, has underrepresented higher education spending for two reasons. First, 5 percent of public colleges and 15 percent of community colleges failed to report total spending in 2003 in part because of changing reporting standards. The percentage of colleges that failed to report declined quickly in the following years. Second, measures of spending by colleges do not account for household spending on interest for student loans.

Undercounting total higher education spending makes spending on higher education financing costs appear larger as a share of total spending because it shrinks the denominator in this equation. So, we developed procedures for improving the estimate of total spending. We addressed the first problem of missing data from public and community colleges with two imputations. First, we used data for total spending beginning in 2002 and made a linear imputation by adding together total spending from year n-1 and year n+1 and dividing that sum by 2 to estimate total spending for a missing observation with year n. Second, we calculated the average change in total spending by sector for each year from 2002 to 2012. We then divided the average rate of change by total spending for year n+1 to estimate total spending for each college with a missing observation for year n beginning with 2011 and proceeding in declining order to 2003. After the imputation procedure, we summed all reported and imputed spending to get total spending by sector and for all colleges. This method therefore assumes that the average rate of change for college spending would be the same for colleges with missing data as those without. We believe this is a better assumption than the assumption that colleges with missing data spent nothing in the years for which they have missing data. This assumption and imputation also have the benefit of making our estimates conservative for measuring financing costs as a share of all spending. The effect of the imputation was to increase the total spending estimate by $9 billion or 2% in 2003. The effect of the imputation declined to nearly 0% for 2005 and all subsequent years. Without this increase in the estimate of total spending for 2003, financing costs as a share of all

spending would have been 6% that year. Their change in the share of spending estimate is negligible in later years and zero in 2012.

For all financial measures by quantile, we use only the first linear imputation described above. Then, in calculating shares of spending by quantile such as in Figure 4, we include only institutions that reported all relevant data for all years so as to hold the sample constant.

We address the second issue of spending on interest for student loans by simply adding our estimate of annual student loan interest to the sum of total spending by colleges from IPEDS. The effect of this adjustment is small and again makes estimates of the relative size of financing costs more conservative. Even in 2012, after student loan interest payments had increased to their largest share of overall spending, this caused a decrease of less than 1 percent in our estimate for financing costs as a share of all spending. We explain the method by which we estimate total student loan interest spending later in this appendix.

The same adjustment to total higher education spending is not necessary for interest on college institutional debts or proprietary college profits. These financing costs are already included in total spending reports in IPEDS. Interest payments for institutional debts are explicitly included in total spending reports. Profits are not explicitly included in spending reports as they may be retained as earnings just as returns from endowments can be retained. Like endowment returns, however, profits may be expended or invested in a college's activities in later years and are at that point represented in total higher education spending.

Endowment Data

In our data harmonization, we give primacy to the endowment data reported to NACUBO. The number of colleges reporting endowment asset values to NACUBO annually ranged from 739 to 842 between 2003 and 2012. By harmonizing NACUBO data with IPEDS data, we were able to obtain data for endowment assets for all years from 2003 for 209 public systems and 871 private systems. This provides us with full endowment asset data for 68 percent of all undergrad-enrolling public systems. We have the same coverage for 69 percent of private systems. This is the most complete data set

for endowment assets that we know of, and it is currently not possible to determine what share of institutions for which we lack data actually operate endowments.

For all statistics and figures on endowment assets by quantile, we used only actual reported data and estimates imputed with the simple linear imputation for years in which a missing observation for a given institution was available in both the previous and following year. We then used only data for institutions for which this method provided full data for all years from 2003 to 2012. In doing so, we hold the sample constant across years so that trends do not reflect changes in sample inclusion. As Table 1 shows, this provided full data on endowment assets by endowment asset wealth quantile for 215 public systems and 871 private systems. Quantiles for all years are based on 2003 endowment asset values.

All endowment *spending* statistics and figures by quantile include actual data and *college level estimates* of endowment spending for all years from 2003 to 2012 for all of the 215 public systems and 871 private systems for which we have full asset data for all years as described in the previous paragraph. The number of public, undergraduate-enrolling systems for which we have data on spending rates from endowments ranges from 117 in 2003 to 132 in 2012. Such data for private systems ranges from 377 systems in 2003 to 434 in 2009. Reported endowment spending rates were between four and six percent for almost all reporting institutions. Spending rates tended to fall between five and six percent at private institutions and between four and five percent at public systems. Given this consistency, we used the average annual spending rate for an institution's sector for the given year if an institution did not report its spending rate for that year but had reported its endowment asset values for all years.

When calculating total endowment asset growth and total spending from endowments as a share of all university funding, we applied to missing endowment data the same broader imputation method used for total spending that is described above in the "Total Spending on Higher Education" section of the appendix. The effect of this imputation on total endowment asset and spending values was negligible. The share of public colleges reporting no endowment assets fell from 16 percent in 2003 to 5 percent in 2012. The share of private colleges reporting no endowment assets fell from 16 percent to 15 percent. After imputation, total endowment asset values increased only 1 percent in

2003 and increased less than 1 percent in 2012. This is because of the relatively small value of endowment assets held by colleges that reported assets some years but not others.

In Figure 11 and the conclusion, we estimate spending from endowments per FTE by the number of students enrolled at all institutions, including students enrolled at proprietary colleges and public, private, and community colleges that did not report endowment spending. We do so to show how rising endowment spending figures in the overall higher education system.

Financing Costs for Colleges Institutional Debts

For publics and non-profits, we calculated total annual interest spending for institutional debt using data from the Integrated Post-Secondary Education Data System. This is measured as total annual expenditure on interest payments. The number of public undergraduate-enrolling systems for which such data is available ranges from 182 or 59 percent of such systems in 2003 to 211 or 69 percent in 2012. It is not known what share of the remaining systems actually issue their own debt as opposed to receiving capital projects funding financed by borrowing by state governments, local governments, tribal authorities, or by federal appropriations in the case of military institutions. The number of private undergraduate-enrolling systems for which such data is available ranges from 806 or 62 percent of such systems in 2003 to 850 or 65 percent in 2012. As in the case of endowments, we use IPEDS data on total spending by college to calculate spending on interest for institutional debts as a share of total spending by colleges.

When we disaggregate interest spending averages by sector and quantiles of endowment wealth, we only use the simple linear imputation. Again, averages by quantile are only for 209 public and 871 private systems for which we have endowment asset values for all years. All but one private system above the 89th percentile reported interest spending data for all years. All public institutions above the 89th percentile reported interest spending for all years. For private institutions below the 90th percentile, 82 percent of institutions reported interest spending for all years. For public systems below the 90th percentile, 87 percent of systems reported interest spending for all years. Given these high rates for reporting of interest spending for these institutions with

endowment assets for all years, our reported results by endowment asset quantile *do not* further exclude institutions for which no interest spending was reported in some years. We found almost no difference in unreported results for interest spending averages by endowment quantile that did exclude institutions for which no interest spending was reported in some years

For missing IPEDS college interest spending and debt data, we use the broader imputation method described in the total spending section of this appendix only when estimating total institutional debt costs as a share of all spending of all types on U.S. higher education. Without this imputation, missing data in earlier years makes the increase in interest spending appear larger as more colleges report their interest spending in later years. Our broader imputation method allows us to estimate interest-spending data from 2003 to 2012 for 92% of public colleges, 91 percent of private colleges, and 79 percent of community colleges. This imputation increased the overall value of interest payments for all colleges by 11% in 2003 and by 1% in 2013. This provides a more conservative estimate of the total increase in interest spending for institutional debt.

In Figure 11 and the conclusion, we estimate interest spending per FTE by the number of students enrolled at all institutions, including students enrolled at proprietary colleges and public, private, and community colleges that did not report interest spending. We do so to show how rising institutional interest spending contributes to the increase in overall financing costs relative to all spending throughout the higher education system. The increase in institutional interest costs per student over its 2003 level is 40 percent when estimated this way. The increase in institutional interest costs is 38 percent when estimated using only enrollments at colleges that reported interest data.

Proprietary College Profits

As we noted in the article, we acquired annual figures on operating profits for publicly traded firms from the income sheets of their fiscal year-end 10-K statements filed with the SEC. In doing so, we found that operating margins reported in IPEDS were closely correlated to operating margins reported by publicly traded firms such as the Apollo Group for whom Title IV institutions make up the overwhelming majority of their business. To be consistent across ownership forms, and because closely held companies

and firms owned by private equity firms do not publicly report income statements, we use IPEDS revenue and expenditure data from IPEDS to estimate operating margins from Title IV activities for all for-profits. We matched individual campus records in the IPEDS data by institutional affiliation, and subtracted total expenses from total current funds revenues.

No imputation was needed for missing IPEDS data for calculations to estimate proprietary college operating profits. Less than 1 percent of proprietary colleges had missing data for either total expenditures or total revenue from 2003 to 2012.

Financing Costs for Student Loans

Annual interest payments on institutional debts are tracked for each college and published in the Integrated Postsecondary Education Data System (IPEDS). Annual interest payments on student loans, however, are not tracked at the college level. And interest payments on Federal Family Education Loans (FFEL), the largest area of student loan origination prior to 2010, have never been tracked at any level. To address this inadequacy, we use data on annual student loan origination by loan type, the annual interest rates for each student loan type, and average time in deference and in repayment for student loans overall to estimate annual interest payments for each student loan cohort by loan type. For each year, total student loan interest payments by loan type are the sum of payments across all cohorts, reported in constant 2012 dollars. Below, we describe the exact procedures for calculating these annual interest payments by loan type which are carried out in the supplemental "Amortization for Student Loan Interest Payments _2015_11_18.xlsx" workbook which is also available online.

Data on loan origination and interest rates came from the following sources. Our loan origination data by loan type for all federal and non-federal loans is by academic year and comes from the College Board.[1] For federal student loans, we use the annual interest rates for each academic year reported by FinAid.org.[2] For non-federal student loans, we used the estimates of average annual private student loan interest rates reported in the Consumer Financial Protection Bureau's 2012 *Private Student Loans* report. The CFPB used 2011 sample lender loan margin and historical LIBOR data to estimate these mean interest rates for private student loans with a standard methodology.[3] We use an

estimate of private student loan interest rates as our estimate of interest rates for all non-federal loans. Some non-federal student loans are issued by states and nonprofits. But the vast majority of non-federal student loans are private student loans issued by banks.

Very little data is available on how quickly or slowly borrowers pay off student loans. We assume that the average time from the origination of a student loan until it enters repayment is two years. Absent better data, we use two years as a conservative estimate given that most borrowing is by four-year degree students for whom the median time to complete a degree is 4.33 years.[4] As national enrollment grew annually from year to year and dropouts thinned second, third, and fourth year cohorts, it was more likely in each year that borrowers would come from 1st and second year cohorts than later cohorts. To be conservative, however, we assume an equal likelihood that borrowers came from a given cohort between one and four.

We further assume that all student loans are paid off at a constant rate over seven years. A seven-year average post-enrollment repayment time is latest estimate available to the authors for the average time to repayment for federal student loans.[5] It is likely that many student loans are paid off more quickly at the end of the repayment period than at the beginning because borrowers are entering the labor market and have more resources to pay down a greater share of the principal. If this is the case, our estimates of interest payments are conservative because they assume that borrowers paid principal down at the rate necessary to generate fixed monthly payments at a given interest rate while the loan was in repayment. To the extent that borrowers actually paid down more principal later in the repayment period, they also paid more interest on that principal earlier in the repayment period.

Using the above data, we used the following process to calculate annual interest payments on all outstanding student loans. For each loan type, we calculated the total interest payments made in each year on the total loan origination for each annual cohort of loan recipients in nominal dollars. For the two years that we assumed student loans remained deferred before entering repayment, we also assumed that interest was not compounded, but was instead paid in the year of its accrual. We do not expect that this assumption holds equally across all loan types, but assuming that this interest is not carried forward yields the most conservative estimate of interest payments overall.

For the seven years that we assumed that the loans were in repayment, we calculated the total amount of principal and interest paid in the current year based on an amortization schedule that would generate fixed monthly payments over the remaining years of the repayment period. Because repayment interest rates varied from year to year, we repeated these calculations for each year based on the principal remaining from the previous year and number of monthly repayments remaining in the seven-year schedule. The result was a series of nine annual payments for each cohort of loans made on the total loan origination for each loan type (except for subsidized Stafford direct loans).

To obtain the total interest payments made in a given year on a given loan type, we summed the total interest payments made on each cohort of loans in a given year. Because we assumed that the period from origination to complete repayment was nine years in total, it was necessary to estimate the interest payments made on all student loans from the 1993 cohort (who paid on their student loans through 2002) through to the 2012 cohort. We then converted the resulting annual total into constant 2012 dollars. Second, we calculated what share of each monthly annuity payment went to principal payments and what share went to interest payments.

The procedures governing the accrual of interest during the initial deferment period differ considerably by loan type, and we modified our calculations for each loan type to reflect these key differences. For subsidized Stafford FFEL loans, we calculated interest payments for the two-year period before the loans entered repayment and include these in our annual totals because they reflect payments by the federal government to the private originators. For direct subsidized Stafford loans, we calculated no interest or principal costs or payments for two years after the origination of a given loan cohort, because the federal government essentially pays itself interest during this period. For unsubsidized loans, which make up all other loan origination, we calculated interest payments for the two-year period before the loans entered repayment. In most cases it is likely that the interest accrued during deferral on unsubsidized loans is not paid and added to the principal at the start of repayment, but assuming it is paid yields a more conservative estimate.

Our calculations also take into account the annual change in interest rates on federal and private variable rate loans. Annual interest rates for federal loans vary based

on the type of loan, whether the loan is in its initial deferment period or in repayment, and the date on which the loan was originated. For subsidized and unsubsidized federal loans, we specified an annual interest rate schedule for each loan cohort to account for these variations. For non-federal loans, detailed information is not available, and therefore we used the same private student loan interest rate estimate from the CFPB report for a given year for each loan cohort in its initial deferment or in repayment during that year.

A further word is in order about the average time to repayment from the end of enrollment. This is an average for all federal student loan types. So it is only appropriate to apply the seven-year average repayment time to national estimates of interest payments across all sectors and across all loan types. For this same reason, we are unable to provide estimates of student loan interest payments by sector. We have requested from the Department of Education data a set of more current estimates of average time in repayment that are broken down by borrowers' risk category. The distribution of borrowers by risk category varies by higher education sectors. So we hope to use this data in the future to estimate variation in annual interest payments on student loans by sector. Given the recent increase in default and deferment rates, we suspect that the average time in repayment has increased. If so, our estimate of annual student loan interest payments is lower than it should be. We will only know for sure after receiving updated data from the Department of Education.

We also use the seven-year average repayment time to estimate annual interest payments on non-federal student loans, a category that includes private student loans. We know of no comparable data that is publicly available for average repayment times for non-federal student loans. It may be possible to estimate average repayment times for private student loans by examining data published for student loan asset backed securities. As noted, the vast majority of non-federal student loans are private student loans. We suspect that average repayment times for non-federal student loans may be higher than the seven-year average because of the higher default rates for private student loans which make up much of the non-federal student loans.[6] If so, our estimate for annual interest payments on non-federal student loans is conservative.

Data Appendix Endnotes

---

[1] See College (The College Board, 2013d). The College Board uses unpublished data from Policy, Budget, and Analysis Staff, U.S. Department of Education, and the National Student Loan Data System (NSLDS) of federal loans. College Board, *Trends in Student Aid:* 2013 reports on page 34 that, "estimates for 2010-11 through 2012-13 provided by the Consumer Bankers Association, MeasureOne, and the Consumer Financial Protection Bureau. Earlier data based on information provided by lenders supplemented by data from annual reports and from NPSAS, 2008. Data on lending also collected from the major credit unions and their associations. Estimates of institutional lending are based on NPSAS, 2008 and 2012, as well as a survey of institutions conducted for the College Board by the National Association of Student Financial Aid Administrators (NASFAA). Data on loans from states are based on information collected from staff of state-sponsored private loan programs or state grant agencies, in addition to NASSGAP."

[2] See http://www.finaid.org/loans/historicalrates.phtml.

[3] See page 14 of *Private Student Loans*.

[4] See http://nces.ed.gov/pubs2011/2011236.pdf

[5] This estimate was provided by David Bergeron, former U.S. Acting Assistant Secretary for Higher Education based on technical briefings provided by the Department of Education prior to 2010.

[6] (Consumer Financial Protection Bureau and U.S. Department of Education, 2012) *Private Student Loans.*