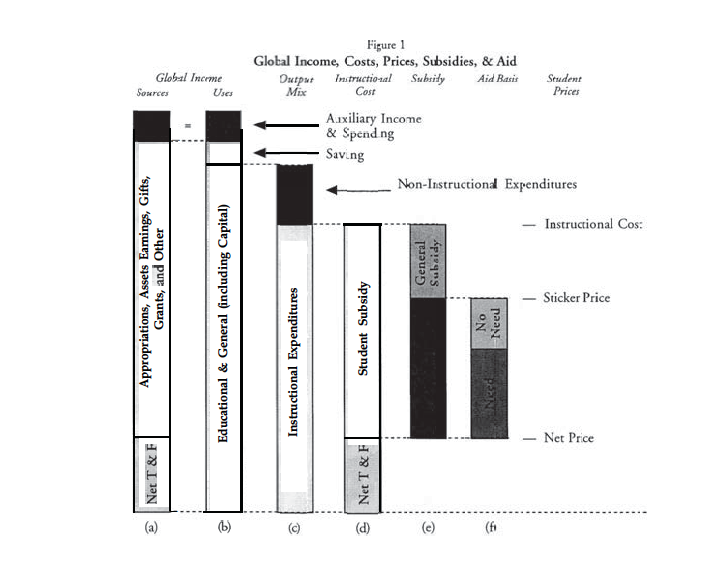
Data Description:

The dataset for this research project is sourced from the National Center for Education Statistics (NCES). NCES is a federal entity which is the principally responsible organization for collecting and analyzing data related to U.S. education. NCES offers a number of data products which are relevant to research in higher education. For the purposes of this project, financial data was pulled from NCES for 2,386 institutions across the years 2004-2015. The data set was truncated at 2015 because of changes in accounting standards and changes in the NCES collection process. To ensure that all relevant metrics were reported in the same way, this paper focuses on data that was collected under identical methodologies. The data set included data on all private not-for-profit institutions as well as public institutions that report under the accounting standards established by the Financial Accounting Standards Board (FASB). The data includes revenue sources, expenditures by function, endowment assets, and assets and liabilities. The data is helpful from our perspective because it allows us to look at our primary variables: background risk and general subsidy levels.

This model explores a number of variables in seeking to understand the data generating process for general subsidies. All data was rescaled to be interpreted as per full time enrolled student. This is a way to account for size at the institutional level without controlling it through a proxy such as total assets. To create the dataset, every survey year of relevance was downloaded in a CSV format, renamed, and merged such that we had a consistent data frame across years. First let’s explore the distribution of institutions across years. As you can see in the plot, there is a fairly even distribution across years. This was an important check in our merging process. If variables had not mapped correctly, I would have expect to see more variation between years.

 Our variable of interest for this research project is general subsidy levels. To compute this variable we borrow the methodology from Winston and Yen (1995) in their paper *Costs, Prices, Subsidies, and Aid in U.S. Higher Education*. I will note that the aforementioned paper was conducted using data from the 1990’s. The NCES has evolved since that point, along with accounting practices and standards. As a result some of the variables that Winston and Yen describe do not exist in the same format. Nonetheless, the variables of importance were recreated under the same methodology, instead having to roll together more variables as the summary variables used in Winston and Yen (1995) no longer are reported. The following visual helps act as a guide for understanding how the variable of interest, general subsidy, came to be. It will also help highlight differences between this paper and Winston and Yen (1995). We begin with global income, which can be then broken down into uses and sources. Winston and Yen (1995) decide to ignore auxiliary revenue as they point out it often cancels out auxiliary expenses. However, I find in this dataset that this feature does not remain the same, and that auxiliary revenue can be seen as another important source of revenue. The remainder of the sources of global income can be broken down to: Net Tuition and Fees, Appropriations, Assets Earnings, Gifts, Grants, and Other. Winston and Yen (1995) then break down the global income uses to instructional vs. non-instructional expense. General subsidy is then created as a derivative. General subsidy is equal to Instructional Expense less Net Tuition and Fees less individual aid (Pell Grants, etc.).

 To be more granular, let’s dive into exactly how the variables from Winston and Yen (1995) were created. Instructional E&G&K expense is defined as the sum of instruction expense, academic support expense, student services expense, institutional support expense, and net grant aid. Net tuition revenue is defined as the total tuition and fees line item reported in the financial statements. Total tuition and fees is the tuition total minus financial aid provided by the institution. General subsidy then becomes Instructional E&G&K expense minus net tuition revenue minus student grants (Pell Grants, etc.). Let’s take a look at the distribution below. We can see that general subsidy seems to take a student t distribution. It is centered to the left of zero and has very long tails. This seems to fall in line with the hypothesis of this paper, and within the framework laid out in Winston and Yen (1995). For one it appears that on average, schools are granting a negative subsidy level to their students. However, as outlined by the hierarchy of education discussed in the literature review, there are schools whom due to their wealth and size, are able to continue to grant positive general subsidy levels. It will be interesting to see how this holds over time as well.

 As is shown in the chart, there seemed to be a slightly positive general subsidy, on average, prior to the 2007-2008 financial crisis. However, after that point, the average general subsidy has continued to drive lower and has turned negative. The huge explosions on the tail seem unreasonably high, however when investigating the specific schools on the tails of these distributions, we can see that intuitively it makes sense. Yale grants the largest positive subsidies, this is driven by tremendous alumni donations and phenomenal performance in their massive endowment portfolio. Though we do not want to eliminate this outlier, it is important to understand that there are heuristic differences between some schools that cannot be captured in financial data.

On the whole we will analyze how the various sources of financial income and revenue effect the general subsidy levels. This paper theorizes that not all revenue sources will be viewed with the same certainty by institutions, and subsidy levels will tend to be more negatively impacted by volatile revenue sources, and positively impacted by stable revenue sources.