Risk Aversion Behavior in Higher Education:

A Bayesian Approach to Linking Background Risk to General Subsidy Levels at Higher Education Institutions in the United States

Kyle Persaud Davis
Masters Candidate, Columbia University, 2019

Abstract:

This paper seeks to explore the link between background risk and general subsidy levels at higher education institutions. Higher education Institutions, and the non-profit sector as a whole, operate in an informationally asymmetric market with a non-distribution constraint. There are limited strategic decisions for higher education entities to make, and fewer so that can be used to manage revenue risk. This paper theorizes that general subsidy levels are influenced by the background risk perceived by the institution. This paper employs a Bayesian approach to understanding the data generating process of the general subsidy distribution in the United States. The research is produced using the Stan programming language and the R package BRMS. Bayesian analysis allows us to gain a more complete understanding of the posterior predictive distribution for general subsidy and allows us to account for uncertainty in our estimates of the parameters across the sample space. This paper will shed light on the vast dispersion within general subsidy levels in higher education.

A special thanks to my parents, Mark and Donna, my brothers, Kevin and Kieran, and my partner, Shayleen Reynolds, for always supporting me. A special thanks to Benjamin Goodrich for giving me the tools and confidence to pursue a project in Bayesian Statistics.

Introduction:

In 1980 Henry Hannsman published a seminal paper entitled *The Role of Nonprofit Enterprise.* The paper was powerful in that it drew to the attention of lawyers and economists a very eminent reality. Nonprofit organizations had become an increasingly prevalent part of the modern economy, yet the existing literature at the time had largely overlooked the role of the nonprofit corporation in the modern economy. Out of Hannsman's work one entity structure drew a heightened amount of focus, higher education institutions.

Higher education institutions operating in a particular market that economists sometimes refer to as "trust markets." They are described in this way for the asymmetric information available to each party entering a transaction. In informational asymmetric markets, buyers are vulnerable to a supplier's opportunism. The non-profit structure of higher education institutions encourages honest, if not profit-sacrificing, behavior that can create trust with the buyers and become the preferred suppliers in the "trust" market. It is important to note that non-profit does not mean no profit. No matter the market, or market structure, entities operating at a loss will disappear. A non-profit enterprise is differentiated from other entity structures in that it cannot distribute profits to stakeholders. This key legal and economic characteristic is a non-distribution constraint (Hansmann, 1980). However, the non-distribution constraint can be bent to benefit personnel within the institution (Winston, 1999). In higher education this may take the form of tuition revenue supporting administrators' perks. While it seems to be an oversimplification, profit maximization is an umbrella utility function that can be used to describe the behavior of higher education institutions. However, the non-distribution constraint blurs the lines as to the beneficiaries of this behavior. This foundation lays the groundwork for the research presented in this paper.

Student loan debt in the United States has continued to rise, rounding out at $1.56 trillion as of February 2019. Camilo Maldonado, a Forbes writer, reported in July of last year that the price of college is increasing at a rate nearly 8 times that of wages. According to the National Center for Education Statistics, between 1989 and 2016 the cost of a four-year degree rose by 2.6% a year. Compared to the Federal Reserve Bank of St. Louis report outlining that wages have only grow 0.3% per year over the same time period. This has implied, that on some level, each successive cohort of college graduates is worse off. Higher costs, coupled with stagnant wages and large debt burdens calls into question the overall structure of higher education and how it operates within the modern economy.

This research paper seeks to understand the data generating process for an indicator of payout to students from higher education institutions. Particularly this paper will investigate general subsidy levels in higher education to understand how the distribution of these subsidy levels is produced. This paper theorizes that increases in background risk will decrease subsidy levels. More specifically, revenue sources that are more volatile will have d to approach this research a Bayesian framework will be used. Bayesian analysis has the advantage over frequentist statistics in being able to more fully understand the data generating process of our outcome variable. In taking this approach we will be able to visualize the entirety of the posterior distribution and more adequately handle the uncertainty that should surround any research question. This paper will focus more particularly on private higher education institutions as their entity structure may have more volatile sources of revenue compared to state institutions.

Literature Review:

To truly understand the research that will be laid out, one must first transition to thinking about higher education institutions as constrained businesses. There is a great degree of differentiation between colleges. Though there are comparable substitutes, in truth each institution is highly unique to its own entity, incentive, and offering structure. Largely, this differentiation is driven by non-tuition differences (quality of life, location, etc.). Gordon and Yen (1995) propose that the schools then use their non-tuition resources to subsidize their students. Doing so allows the buyer, to have a more complete understanding of the "bargain" they are receiving. Interestingly enough, Gordon and Yen provide a framework where the rest of the model of higher education is largely a run off of this effect.

Any investment decision is filled with uncertainty around the potential for future realization of hoped-for gains. Perhaps the investment decisions with the most amount of uncertainty revolving around future gains is investment into human capital. Individuals investing in human capital through a purchase of higher education don't know exactly what they are buying. As was eluded to in the introduction, higher education institutions operate in an informationally asymmetric market. Their non-profit structure can cause buyers to trust that they will get what they pay for. Education is a large, "one-time", expenditure that will not realize gains for years and years. Often, buyers of higher education won't and can't know what they paid for until it is far too late to do anything about it (Litten, 1980; Winston, 1988). If the non-distribution constraint is blurred in an inappropriate manner (larger payouts for collegiate sports coaches, added perks for deans or board members, etc.) the buyers, in this case students and their benefactors, will be worse off.

As colleges and universities compete for market share, a hierarchy emerges. The cause of such a caste system comes from the donative wealth and an institutions ability to cultivate past and future donative resources. The differences in wealth between institutions strongly enhances their positioning within the market place. Institutions that collect more non-tuition revenue can in turn offer their product (one unit of education) at either a lower cost or at a higher quality. The uneven playing field of this market place make it difficult for smaller, less profitable institutions to grow in size, scope, and quality.

Winston and Yen (1995) contextualize the strategic decision making of these institutions. Their paper posited that an anomaly of higher education, and the non-profit sector as a whole, is they offer their product at a price that is less than the average cost of production. They describe this phenomenon as a general subsidy, that nearly all institutions provide to their students. Subsidies involve a fascinating set of strategic decisions that do not have a parallel in the for-profit space. Despite an institution's decision to grant a subsidy, they still make a non-distributive profit through other sources of revenue. Schools with substantial non-tuition revenue resources grant larger subsidies. The "better bargain" flows through to other parts of the institution's ability to be profitable. The large subsidy increases student demand, and given the restricted supply allows institutions to increase student quality through selective admissions. Differences in institutions are amplified by the winner-take-all market of higher education, in which subsidy levels, limited supply, and student quality act as positive feedback mechanisms.

This paper will explore how the structure laid out by previous work holds today. Between 2007 and 2013, state level funding for higher education dropped by 14.1 billion dollars. This paper would anticipate that the general subsidy level for many institutions has turned negative since the previous literature explored the subject. As non-tuition resources have become more volatile, institutions have been forced to change their strategic decision framework. While this paper will use Winston's methodology to assess the general subsidy levels, it in whole does not agree that general subsidies are positive for all institutions. This paper hypotheses that general subsidy levels are no longer consistently positive and instead have a set of drivers. More specifically they are driven by the background risk at an institution.

The underlying generation of general subsidy levels explored in this paper is background risk. In Machina's (1982) seminal analysis, it was established that anyone with an everywhere-concave utility function would satisfy the preference ordering of second-order stochastic dominance, and would be strictly worse off. This result was driven by the addition of any independent background risk. Constant risk aversion (Safra and Segal, 1998; Quiggin and Chambers, 1998), occurs when preference rankings are unaffected by incremental additions of risk. Higher education institutions experience constant risk aversion in that they are extremely option limited. It is difficult for institutions to add positive convexity to their product portfolios, and as a result they are risk vulnerable. Gollier and Pratt (1996) describe increasing sensitivity as a result of independent background risk as a sufficient condition for risk vulnerability. The primary result of previous literature has been, a seemingly obvious, yet important result. The addition of independent (or quasi-independent) risk reduces overall welfare. The question at hand becomes focused on whether risk aversion should increase in the presence of an independent background risk. In the context of higher education, this paper seeks to investigate whether the volatility of non-tuition revenue sources can act as a mechanism for understanding background risk at the institutional level. This research will investigate if risk aversion to negative net tuition revenue is increased or decreased in correlation with volatile sources of non-tuition revenue.

There are several advances in previous literature which led the concept of background risk to be included in this paper. Guiso and Paiella (2008) provided empirical evidence that absolute risk tolerance is an increasing function of a consumers' resources. Their evidence helped reject constant absolute risk aversion preferences. The historic definition of risk aversion and tolerance developed by Arrow (1970) and Pratt (1964) assumes that initial wealth in nonrandom. Updates to these definitions showed how constraints from a liquidity perspective also change attitudes towards risk taking. Gollier (2000) shows that consumers who are credit constrained, or perceive themselves to be at some point in the future, are less willing to bear risk in the present. Risk aversion is positively affected by background risk and by the threat of liquidity constraints. These concepts can be applied to the higher education strategic decisions model. Though background risks at higher education entities are often correlated (not independent as discussed in the literature), the effect on decisions remains the same. In the presence of background risks, and with the additional threat of liquidity (or rather lack thereof), institutions are more risk averse. However, risk tolerance increases for the largest most prominent institutions who are wealthy enough to weather a storm.

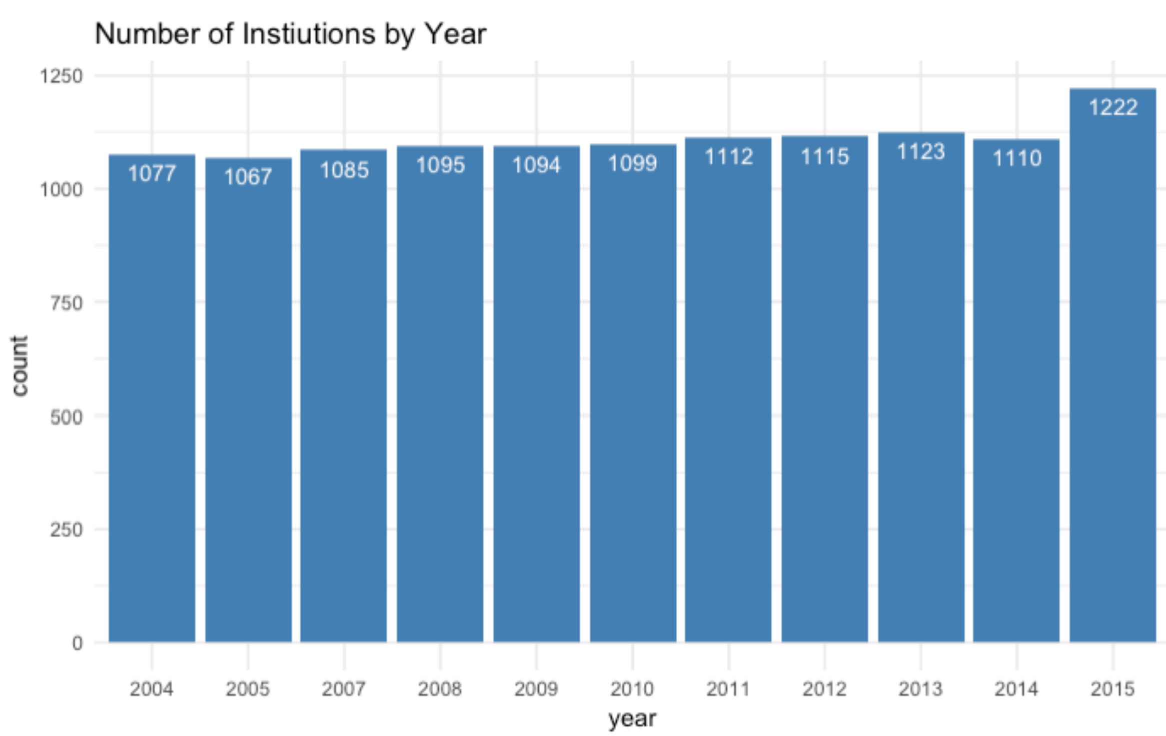
For certain utility functions, the existing literature predicts that in the presence of exogenous, non-insurable, risks, agents should react by reducing the exposure to other risk sources. This idea has been explored in the context of higher education and their asset allocation in the endowment. Though this research looks at endowment returns as a source of background risk, the ideas applied to endowment management in terms of background risk are still applicable. Hansmann (1990) stated that endowments are meant to serve as a cushion against financial distress. If this is so, then the risk of financial distress driven by background risk, should affected endowment allocation decisions. Hansmann (1990) and Winston (1999) note that there is not a well-defined and generally accepted theory of university objective functions. It remains unclear if Universities, or any non-profits for that matter, will consider background risk when making decisions. Merton (1992) provides a closed form model of endowment investment decisions and finds that income risk should result in more conservative asset allocations. Dimmock (2012) operationalizes background risk in higher education by defining it as the standard deviations of a universities' nonfinancial income. While in theory this concept makes great sense, in practice data is reported too infrequently to correctly measure a universities' true volatility in its non-financial income. Higher education institutions produce financial statements annually, which do not capture the intra-year volatility of these risk sources. It may be better, and more transparent, to leave the revenue sources in dollar terms and instead rely on the notion of perceived volatility. There are revenue sources that higher education *feels* are more unpredictable than others. For example, government support programs often pledge multi-year commitments and require extensive bureaucratic motion to be altered. In this way any changes in these funding programs would be known in advance of their occurrence, making them appear to be less volatile and in theory a lower source of background risk. This is quite different from the view on endowment returns as a form of background risk as market returns can be very unpredictable year after year. Universities are particularly vulnerable to financial shocks for three reasons: their assets are poor collateral as they are highly specific to each institution, universities cannot issue equity like public for profit companies, and higher education costs structure are hard to change for structural reasons such as tenure.

The culmination of the prior literature has laid a groundwork from which two separate concepts in higher education research can be linked. The general subsidy levels of higher education are largely driven by each institutions' perception of the volatility of their background risks. Background risk can be defined as the risk associated with non-financial income sources. Due to the limited number of strategic decisions that an institution can make, institutions take advantage of the one where they have the most ability to do so. Higher education institutions change their general subsidies to students based on their perceived revenue risk.

Data Description:

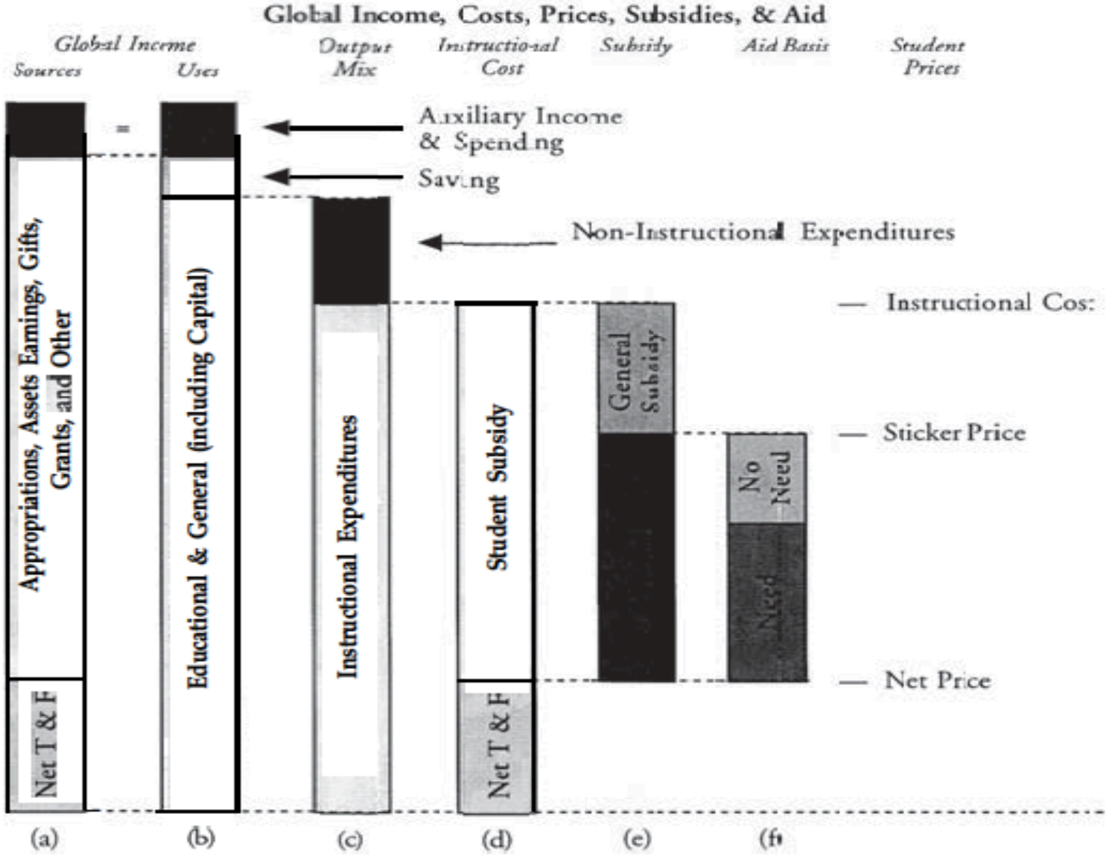
The dataset for this research project is sourced from the National Center for Education Statistics (NCES). NCES is a federal entity which is the principally responsible organization for collecting and analyzing data related to U.S. education. NCES offers a number of data products which are relevant to 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 expected to see more variation between years.



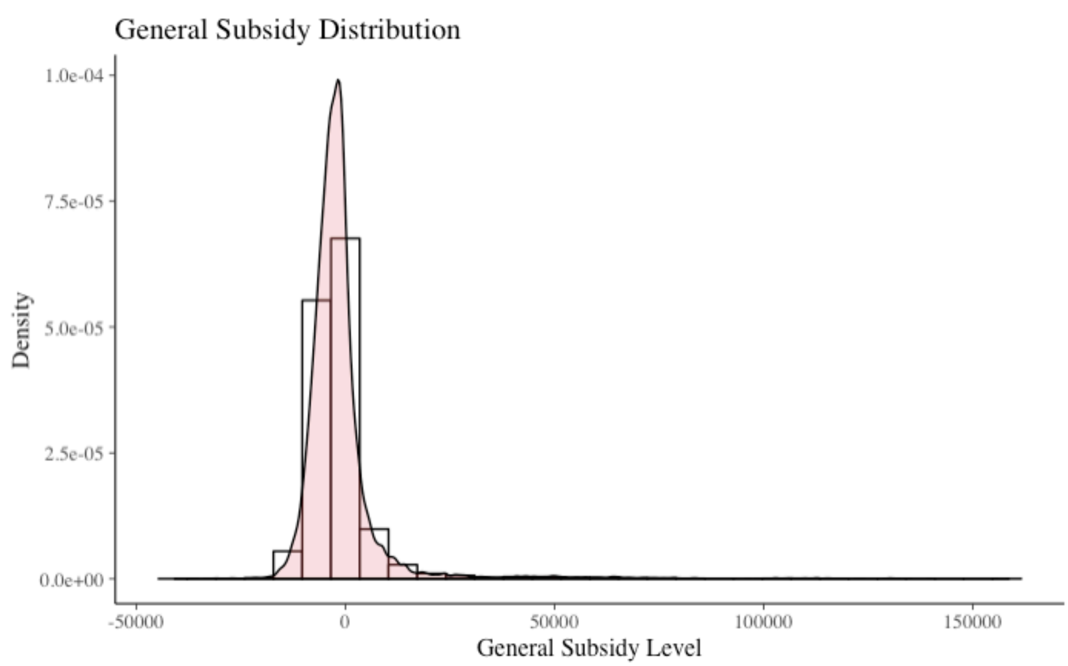
Our variable of interest for this research project is general subsidy levels. To compute this variable, we borrow the methodology from Winston and Yen (1995) in their paper *Costs, Prices, Subsidies, and Aid in U.S. Higher Education*. I will note that the aforementioned paper was conducted using data 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

Figure 1
Global Income, Costs, Prices, Subsidies, & Aid

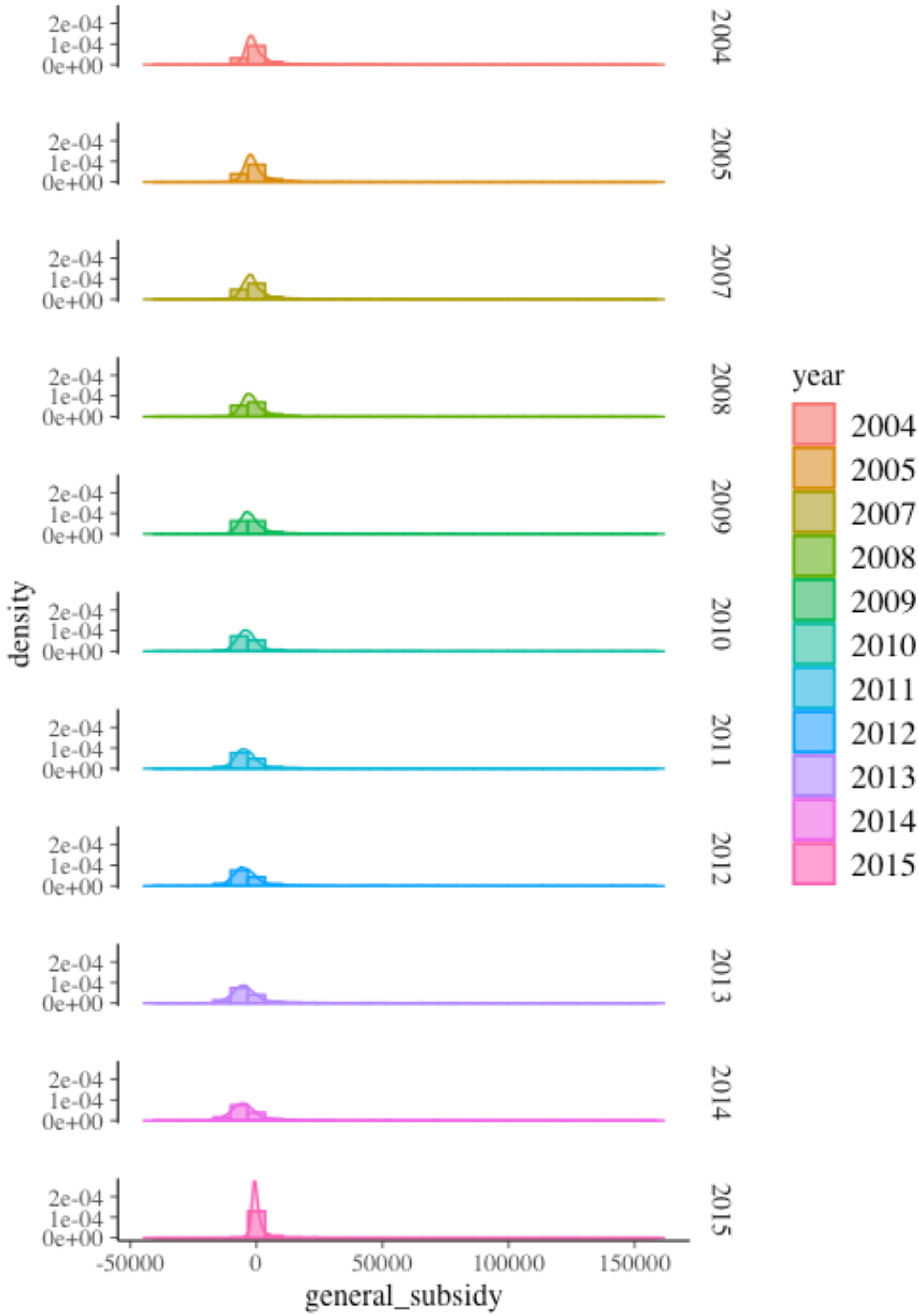


describe do not exist in the same format. Nonetheless, the variables of importance were recreated under the same methodology, instead having to roll together more variables as the summary variables used in Winston and Yen (1995) no longer are reported. The following visual helps act as a guide for understanding 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 are 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.

Methodology:

Introducing Stan and the Bayesian Framework:

To approach the research question at hand, this paper employs a Bayesian framework. As Bayesian analysis is still in the minority to frequentist statistics, this paper will spend a portion of the methodology section explaining the framework, and why it is an improvement on other approaches.

To implement the Bayesian approach, we will take advantage of the Stan programing language. Stan is a probabilistic high-level programming language that is used to specify statistical models. A Stan program defines a log probability function over parameters conditioned on specified data and constraints. Stan provides the user with the ability to conduct full Bayesian inference through Markov chain Monte Carlo methods such as No-U-Turn sampling. Stan is set up such that densities, gradients, and Hessians, along with intermediate quantities are easily accessible. Prior to Stan it was difficult to accurately compute Bayes rule and apply it to research settings.

Stan utilizes Markov chain Monte Carlo simulation to get draws from the posterior predictive distribution. A Markov process is a sequence of random variables with a particular dependence structure where the future is conditionally independent of the past given the present, but noting is marginally independent of anything else. We can construct a Markov process such that the marginal distribution of a random variable is a given target distribution as the number of simulations moves to infinity. As a result, we can get a random draw, or a set of dependent draws, from the target distribution by letting the Markov process to run for many iterations.

Stan specifically uses a form of Markov chain Monte Carlo simulation called No-U-Turn Sampling. First, a quick review of ancient Markov chain Monte Carlo samplers. Metropolis-Hastings only requires user to specify the numerator of Bayes Rule. However, only 22 percent of proposals ideally get accepted to get relatively big jumps in the sampling process. The effective sample size, as a result, can be essentially zero. Gibbs sampling, in general, forces a user to work out all full-conditional distributions. Jumps are always accepted in the sampling process, but they might not be very big. Effective sample size is low if the parameters are highly correlated. Stan, like Metropolis-Hastings, only requires the user to specify the numerator of Bayes Rule. Like Metropolis-Hastings, but unlike Gibbs sampling, proposals are joint. Unlike

Metropolis-Hastings, but like Gibbs sampling, proposals are always accepted and tuning of proposals is (often) not required. Unlike both Metropolis-Hastings and Gibbs sampling, the effective sample size is typically 25 percent to 125% of the nominal number of draws from the posterior distribution. As mention, Stan utilizes No U-Turn Sampling (NUTS) as its sampler for the Markov Chain Monte Carlo Simulation. The location of Θ moving according to Hamiltonian physics at any instant would be a valid draw from the posterior distribution. The challenge became then determining when to stop as Θ moves indefinitely (in the absence of friction). Hoffman and Gelman (2014) proposed stopping when there is a U-Turn. In the sense that the footprints found in the Monte Carlo simulation turn around and start to head in the direction they just came from. After the U-Turn, one footprint is selected with probability proportional to the posterior kernel to be the realization of Θ on iteration *s* and the process repeats itself S times. NUTS discretizes a continuous-time Hamiltonian process in order to solve a system of Ordinary Differential Equations (ODEs). These ODEs require a step size that is also tuned during the warmup phase of the Monte Carlo simulation.

Model Methodology:

To explore the data generating process for the variable of interest, general subsidy levels, I will use the Bayesian regression framework. Bayesian regression differs from frequentist analysis in that it does not rely on a single point estimate for a parameter. Instead it looks at the parameter across the entire sample space and is able to generate a posterior distribution, or a range of values, within which the parameter might probabilistically fall. This is far more helpful in understanding exactly the values that the parameter could take when faced with new data. In this way we can gain a more complete understanding of the data generating process. We will build 5 models in our process, compare them using leave one out cross validation, and use projection prediction to see if our more complicated models can be expressed with fewer variables.

Our five models will grow in complexity by adding more potential variables that potentially aid to the data generating process for general subsidy levels. In model one we will only include the intercept and investment returns, which are defined as the change in market value of the endowment portfolio between time equal to t-1 and time equal to t. In model 2 we will also include non-instructional expenses, which is defined as total expenses less instructional E&G&K expenses. In model 3 we will include the categorical variable year to see if any differences in general subsidy are generated based on the year the data was collected. In model 4 we will include a more granular break out of non-tuition revenue sources including: total gift revenue, total grants and appropriations revenue, net auxiliary revenue, net hospital revenue, and net other revenue.

To compare the models presented in the paper we will use leave one out cross validation. Penalty functions that are used in supervised learning make for poor priors because they are intended to shift the mode rather than reflect the authors original prior beliefs. Supervised learning does not attempt to quantify the uncertainty in parameters or model choice. This differs sharply from the Bayesian approach which seeks to track, as closely as possible, the types of uncertainty that are being introduced during the model process. While supervised learning attempts to optimize for a single point estimate, there is a more complete and improved way to approach the model validation problem. For Bayesians, we estimate over an entire function, in this case the log of the predictive density function is most appropriate.

We will seek to choose a model that maximizes the expectation of the log predictive density function over future data.

Leave one out cross validation can be described with the following procedure. Obtain draws from the posterior PDF and condition of those posterior draws to draw from the posterior predictive distribution for the outcome variable of interest. Omit the i^{th} observation and draw from the posterior PDF evaluating the log-likelihood for the i^{th} observation. Repeat this process across the posterior draws of Θ to compute the expected log predictive density. Though in absolute terms this does not mean much, it does allow us to compare between models. We are looking to see how concentrated the prediction density function is over the distribution of the real data. Leave one out cross validation utilizes Pareto smoothed importance sampling. Pareto smoothed importance sampling provides a move accurate and reliable estimate by fitting a Pareto distribution to the upper tail of the importance weights distribution. Pareto smoothed importance sampling allows us to compute a leave one out cross validation using importance weights that would otherwise be unreliable and noisy.

Results:

Model 1:

Prior Distributions for Model 1:

$$
\begin{gathered}
General\ Subsidy \sim Student\_T(\mu_i, \sigma, nu) \\
\mu_i = \alpha + \beta x_i \\
\alpha \sim Normal(0{,}10000) \\
\beta \sim Normal(0{,}5) \\
\sigma \sim HalfCauchy(0{,}1) \\
nu \sim gamma(4{,}1)
\end{gathered}
$$

Model 1 Summary Output:

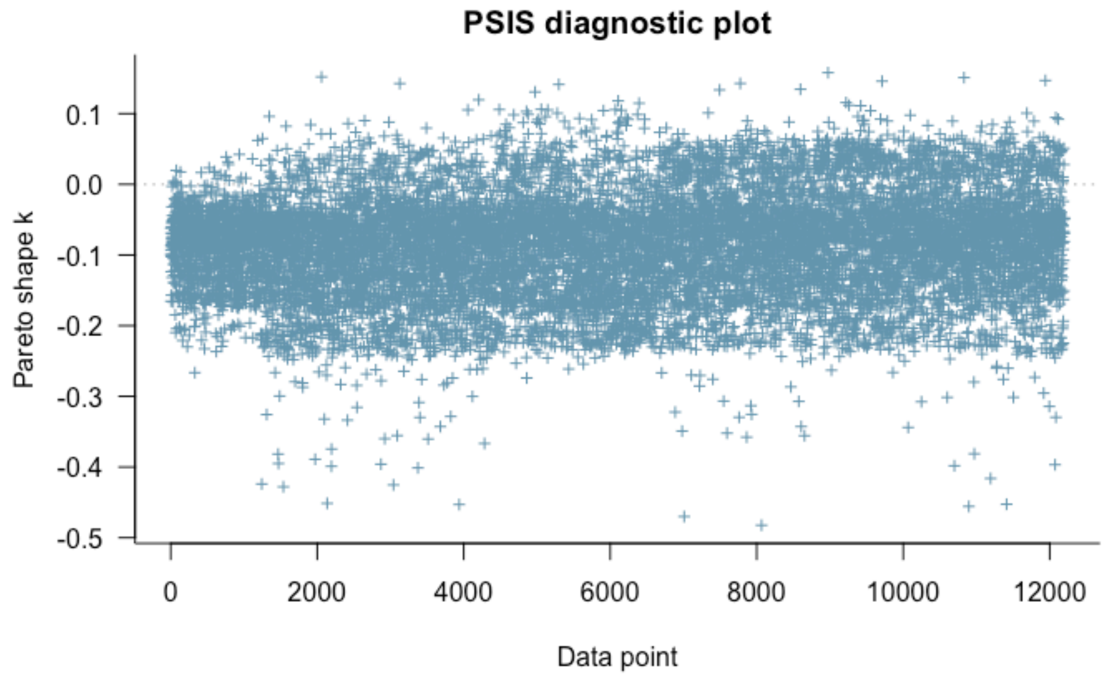
```
 Family: student
  Links: mu = identity; sigma = identity; nu = identity
Formula: general_subsidy ~ 1 + investment_return
   Data: final_regdata_FTE (Number of observations: 12199)
Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup samples = 4000

Population-Level Effects:
                  Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
Intercept         -3018.95     38.82 -3097.07 -2942.94       3364 1.00
investment_return     0.15      0.00     0.14     0.16       3625 1.00

Family Specific Parameters:
      Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
sigma  3329.58     38.52  3255.11  3404.23       2747 1.00
nu        2.16      0.05     2.06     2.25       2258 1.00

Samples were drawn using sampling(NUTS). For each parameter, Eff.Sample
is a crude measure of effective sample size, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

In the table above we see the results for the first Bayesian model. This model estimates the posterior distribution of the parameters that generate our outcome variable *General Subsidy*. The prior on the intercept is normally distributed with mean 0 and standard deviation $10,000. This was selected to allow for the rather long tails that tend to be associated with this variable. While the majority of institutions hover around 0, there are some institutions that sit far out in the right tail (Yale University is one such example). The intercept had an estimated mean of -$3,018.95 with estimated error of $38.82. 95% of our distribution on the intercept parameter fall between -$3,097.07 and -$-2942.94. The coefficient on investment return is estimated to be 0.15 with an estimated standard error of 0. 95% of the distribution for the investment return coefficient parameter falls between 0.14 and 0.16. Both intercept and investment return coefficient parameters had an Rhat of 1 indicating that the sampling chains converged.



The Pareto Smoothed Importance Sampling plot for model 1 looks to pass our checks. All of the shape K parameters are below our threshold, giving us confidence in the stability of our estimates, particularly when looking at the expected log predictive density of our posterior predictive distribution. We will use this when applying leave one out cross validation.

Model 2:

Prior Distributions for Model 2:

$$General\ Subsidy \sim Student\_T(\mu_i, \sigma, nu)$$
$$\mu_i = \alpha + \beta_1 Investment\_Return_i + \beta_2 Non\_Instructional\_Expense_i$$
$$\alpha \sim Normal(0{,}10000)$$
$$\beta_1 \sim Normal(0{,}5)$$
$$\beta_2 \sim Normal(0{,}1)$$
$$\sigma \sim HalfCauchy(0{,}1)$$
$$nu \sim gamma(4{,}1)$$

Model 2 Summary Output:

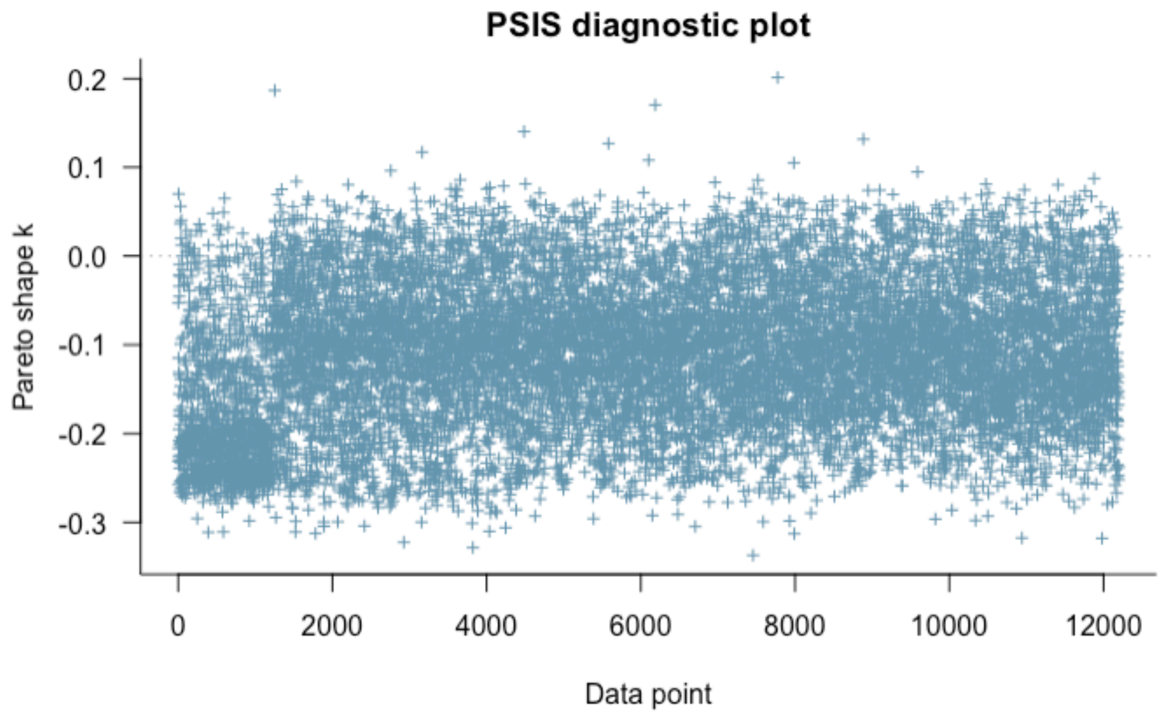
```
 Family: student
  Links: mu = identity; sigma = identity; nu = identity
Formula: general_subsidy ~ 1 + investment_return + non_instructional_exp
   Data: final_regdata_FTE (Number of observations: 12199)
Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup samples = 4000

Population-Level Effects:
                      Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
Intercept             -3803.45     44.82 -3891.59 -3716.09       4173 1.00
investment_return         0.12      0.01     0.11     0.13       2006 1.00
non_instructional_exp     0.12      0.00     0.12     0.12       4516 1.00

Family Specific Parameters:
      Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
sigma  3501.36     39.54  3425.92  3579.74       1977 1.00
nu        2.45      0.06     2.34     2.57       1631 1.00

Samples were drawn using sampling(NUTS). For each parameter, Eff.Sample
is a crude measure of effective sample size, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

The table above specifies the results of the second Bayesian model. In this instance the variable non-instructional expense was also included. With the inclusion of this added variable we see that the intercept parameter estimate moves to -$3,803.45.57 and the standard error on this estimate moves as well to $44.82. 95% of the distribution for the intercept parameter falls between -$3,891.59 and -$3,716.09. The coefficient parameter on investment return dropped slightly to 0.12, while the standard error is now at 0.01. 95% of the distribution for the coefficient parameter on investment return falls between 0.11 and 0.13. The newly introduced variable non-instructional expense has a coefficient parameter estimate of 0.12 with a standard error of 0. 95% of the distribution for the coefficient parameter on non-instructional expense falls between 0.12 and 0.12. All Rhat values are equal to 1, showing that there is convergence in the simulated chains.



Again, the visual analysis of our pareto smoothed importance sampling shows good results. All of our k shape parameters fall below the threshold of 0.7, indicating reliable and stable estimates of our importance weighting. This gives us more confidence with using our leave one out cross validation method which will be applied later in this section.

Model 3:

Prior Distributions for Model 3:

$$General\ Subsidy \sim Student\_T(\mu_i, \sigma, nu)$$
$$\mu_i = \alpha + \beta_1 Investment\_Return_i + \beta_2 Non\_Instructional\_Expense_i + \beta_3 Year_i$$
$$\alpha \sim Normal(0{,}10000)$$
$$\beta_1 \sim Normal(0{,}5)$$
$$\beta_2 \sim Normal(0{,}1)$$
$$\beta_3 \sim Normal(0{,}1)$$
$$\sigma \sim HalfCauchy(0{,}1)$$
$$nu \sim gamma(4{,}1)$$

Model 3 Summary Output:

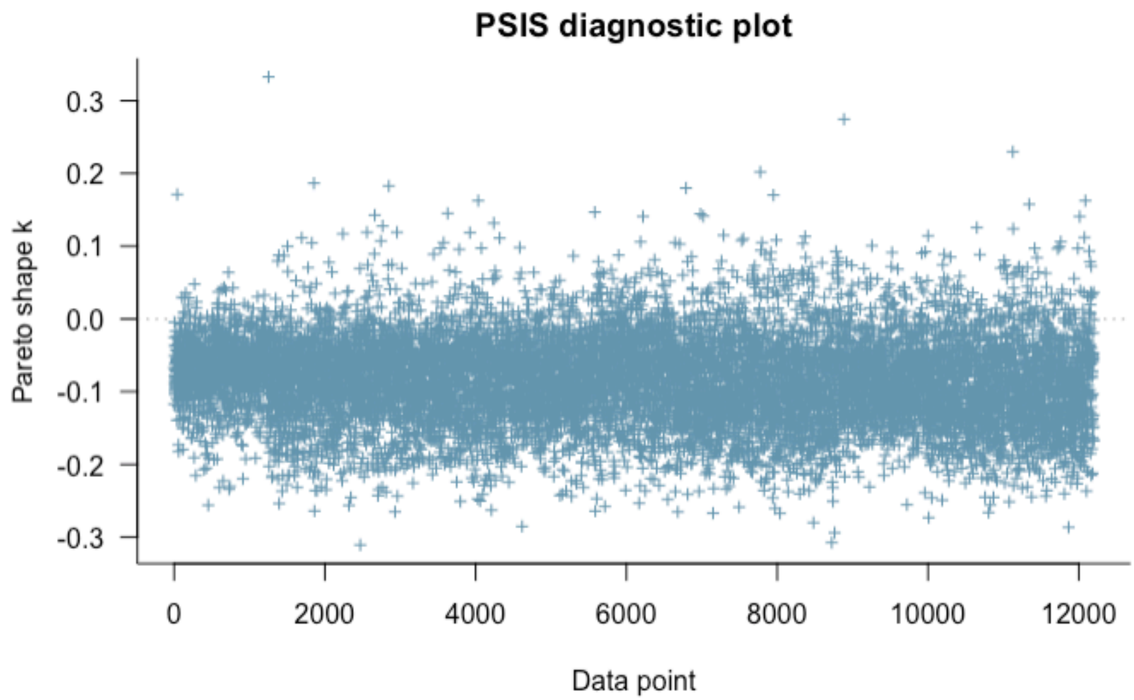
```
 Family: student 
  Links: mu = identity; sigma = identity; nu = identity 
Formula: general_subsidy ~ 1 + investment_return + non_instructional_exp + year 
   Data: final_regdata_FTE (Number of observations: 12199) 
Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup samples = 4000

Population-Level Effects: 
                      Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
Intercept             -3801.42     43.83 -3888.28 -3716.03       5461 1.00
investment_return         0.12      0.00     0.11     0.13       6265 1.00
non_instructional_exp     0.12      0.00     0.12     0.12       5031 1.00
year2005                  0.09      0.99    -1.85     1.99       4952 1.00
year2007                  0.02      1.00    -1.94     2.00       5409 1.00
year2008                  0.03      0.99    -1.88     1.98       4838 1.00
year2009                  0.02      0.96    -1.86     1.95       4755 1.00
year2010                 -0.08      1.01    -2.06     1.85       4977 1.00
year2011                 -0.12      0.98    -2.01     1.77       5003 1.00
year2012                 -0.11      0.99    -2.04     1.81       5152 1.00
year2013                 -0.13      0.99    -2.09     1.87       4912 1.00
year2014                 -0.13      1.02    -2.09     1.86       5180 1.00
year2015                  0.28      0.98    -1.57     2.22       6088 1.00

Family Specific Parameters: 
      Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
sigma  3502.07     41.37  3418.58  3580.90       4106 1.00
nu        2.46      0.06     2.34     2.58       3819 1.00

Samples were drawn using sampling(NUTS). For each parameter, Eff.Sample 
is a crude measure of effective sample size, and Rhat is the potential 
scale reduction factor on split chains (at convergence, Rhat = 1).
```

The table above shows the summary output for the third Bayesian model. This model introduces the factor variable year. With the inclusion of this added variable we see that the intercept parameter estimate moves to -$3,801.42 and the standard error on this estimate drops as well to $43.83. 95% of the distribution for the intercept parameter falls between -$3,888.28 and -$3,716.03. The newly introduced variable has coefficient parameter estimates ranging from -0.13 to 0.28. Though the distributions for these variables are not exactly the same, all of the distributions have their upper and lower 95% confidence interval bands on opposite sides of 0. The importance of this will be discussed in the results section. All Rhat values are equal to 1, showing that there is convergence in the simulated chains.



The Pareto Smoothed Importance Sampling plot looks good, as all shape k parameters are below the threshold of 0.7.

Model 4:

Prior Distributions for Model 4:

$$General\ Subsidy \sim Student\_T(\mu_i, \sigma, nu)$$
$$\mu_i = \alpha + \beta_1 Investment\_Return_i + \beta_2 Non\_Instructional\_Expense_i + \beta_3 Year_i + \beta_4 Gifts_i + \beta_5 Grants\_Appropriations_i + \beta_6 Net\_Auxilary\_Revenue_i + \beta_7 Net\_Hospital\_Rev_i + \beta_8 Net\_Other\_Rev_i$$
$$\alpha \sim Normal(0{,}10000)$$
$$\beta_1 \sim Normal(0{,}5)$$
$$\beta_2 \sim Normal(0{,}1)$$
$$\beta_3 \sim Normal(0{,}1)$$
$$\beta_4 \sim Normal(0{,}5)$$
$$\beta_5 \sim Normal(0{,}5)$$
$$\beta_6 \sim Normal(0{,}1)$$
$$\beta_7 \sim Normal(0{,}10)$$
$$\beta_8 \sim Normal(0{,}0.5)$$
$$\sigma \sim HalfCauchy(0{,}1)$$
$$nu \sim gamma(4{,}1)$$

Model 4 Summary Output:

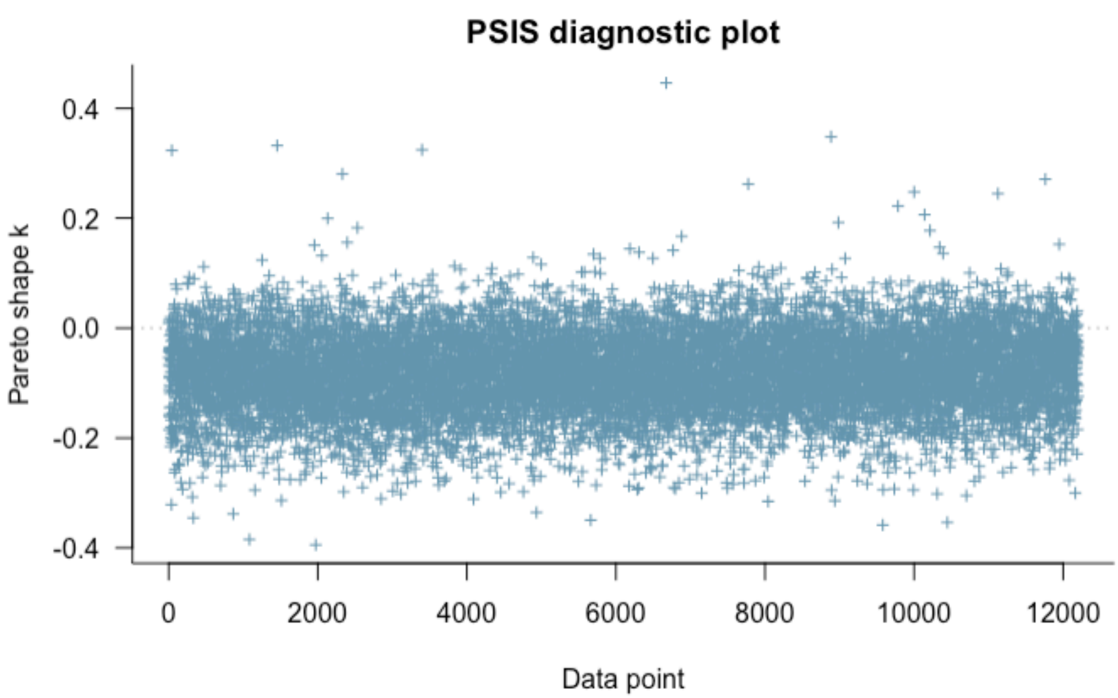
```
 Family: student 
  Links: mu = identity; sigma = identity; nu = identity 
Formula: general_subsidy ~ 1 + investment_return + non_instructional_exp + year + gifts_total + grants_appr + net_aux_rev + 
net_hospital_rev + net_other_rev 
   Data: final_regdata_FTE (Number of observations: 12199) 
Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup samples = 4000

Population-Level Effects: 
                      Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
Intercept             -3488.90     59.89 -3605.06 -3368.79       5082 1.00
investment_return         0.07      0.01     0.06     0.08       5638 1.00
non_instructional_exp    -0.02      0.00    -0.03    -0.01       4984 1.00
year2005                  0.06      0.99    -1.93     1.99       6965 1.00
year2007                  0.02      1.00    -1.87     2.02       5287 1.00
year2008                  0.02      1.00    -1.99     1.97       6371 1.00
year2009                  0.05      1.00    -1.91     2.04       7425 1.00
year2010                 -0.07      0.98    -2.00     1.83       5933 1.00
year2011                 -0.13      1.05    -2.24     1.95       6466 1.00
year2012                 -0.09      1.00    -1.98     1.81       5151 1.00
year2013                 -0.12      1.01    -2.11     1.87       6426 1.00
year2014                 -0.12      1.00    -2.05     1.82       6698 1.00
year2015                  0.31      1.02    -1.67     2.32       7132 1.00
gifts_total               0.50      0.01     0.47     0.53       6171 1.00
grants_appr               0.91      0.02     0.88     0.94       3841 1.00
net_aux_rev              -0.52      0.04    -0.60    -0.45       6909 1.00
net_hospital_rev          0.41      0.04     0.33     0.50       3663 1.00
net_other_rev             0.32      0.02     0.29     0.35       5219 1.00

Family Specific Parameters: 
      Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
sigma  3308.56     36.53  3238.81  3382.68       5386 1.00
nu        3.06      0.09     2.89     3.23       5113 1.00

Samples were drawn using sampling(NUTS). For each parameter, Eff.Sample 
is a crude measure of effective sample size, and Rhat is the potential 
scale reduction factor on split chains (at convergence, Rhat = 1).
```

In the fourth Bayesian model we add several net revenue sources including: gifts, grants/appropriations, auxiliary revenue, hospital revenue, and other revenue. All other variables were maintained. The estimate on the intercept parameter dropped to -$3,488.90 with a standard error of $59.89. 95% of the distribution on this estimate falls between -$3,605.06 and -$3,368.79. The estimate of the coefficient parameter on investment return has moved to 0.07, with 95% of the distribution falling between 0.06 and 0.08. The estimate on the coefficient parameter of non-instructional expenses turned negative to -0.02, with 95% of the distribution falling between -0.03 and -0.01. The various levels of year all have similar properties associated with their coefficient parameter estimates. Most notably all of the distributions have a lower and upper band of the 95% confidence interval on opposite sides of zero. We will discuss this significance further in the discussion section. The coefficient parameter on gifts is estimated at 0.50 with standard error of 0.01. 95% of the distribution falls between 0.47 and 0.53. The coefficient parameter on grants and appropriations is estimated at 0.91 with standard error 0.02. 95% of the distribution on this parameter falls between 0.88 and 0.94. The coefficient parameter estimate for net auxiliary revenue is -0.52 with standard error of 0.04. 95% of the distribution on this parameter falls between -0.60 and -0.45. The coefficient parameter estimate for net hospital revenue is 0.41 with standard error 0.04. 95% of the distribution falls between 0.33 and 0.50. The coefficient parameter estimate on net other revenue is 0.32 with standard error 0.01. 95% of the distribution on this parameter fall between 0.29 and 0.35. All Rhat values are equal to 1, showing that there is convergence in the simulated chains.



The Pareto Smoothed Importance Sampling plot looks good, as all shape k parameters are below the threshold of 0.7.

Model 5:

Prior Distributions for Model 5:

$$
\begin{gathered}
General\ Subsidy \sim Student\_T(\mu_i, \sigma, nu) \\
\mu_i = \alpha + \beta_1 Investment\_Return_i + \beta_2 Non\_Instructional\_Expense_i + \beta_3 Gifts_i \\
+ \beta_4 Grants\_Appropriations_i + \beta_5 Net\_Auxilary\_Revenue_i \\
+ \beta_6 Net\_Hospital\_Rev_i + \beta_7 Net\_Other\_Rev_i \\
\alpha \sim Normal(0,10000) \\
\beta_1 \sim Normal(0,5) \\
\beta_2 \sim Normal(0,1) \\
\beta_3 \sim Normal(0,5) \\
\beta_4 \sim Normal(0,5) \\
\beta_5 \sim Normal(0,1) \\
\beta_6 \sim Normal(0,10) \\
\beta_7 \sim Normal(0,0.5) \\
\sigma \sim HalfCauchy(0,1) \\
nu \sim gamma(4,1)
\end{gathered}
$$

Model 5 Summary Output:

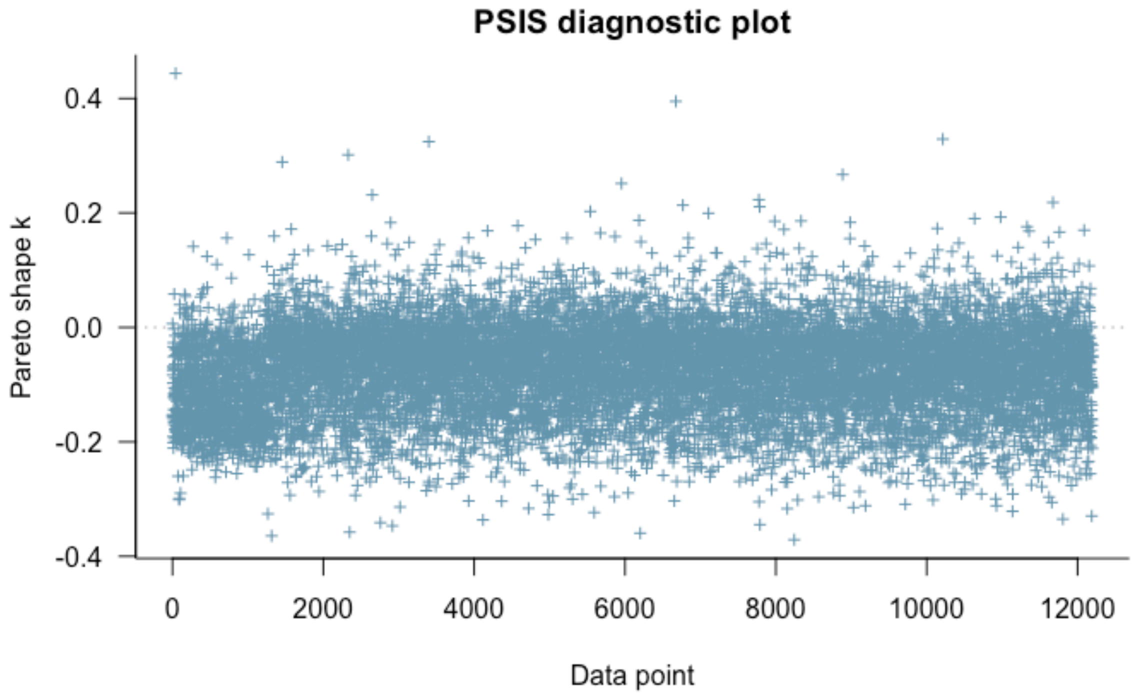
```
 Family: student 
  Links: mu = identity; sigma = identity; nu = identity 
Formula: general_subsidy ~ 1 + investment_return + non_instructional_exp + gifts_total + grants_appr + net_aux_rev + 
net_hospital_rev + net_other_rev 
   Data: final_regdata_FTE (Number of observations: 12199) 
Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup samples = 4000

Population-Level Effects: 
                      Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
Intercept             -3486.54     60.49 -3607.33 -3365.53       3555 1.00
investment_return         0.07      0.00     0.06     0.08       3997 1.00
non_instructional_exp    -0.02      0.00    -0.03    -0.01       3878 1.00
gifts_total               0.50      0.01     0.47     0.53       3403 1.00
grants_appr               0.91      0.02     0.88     0.94       3217 1.00
net_aux_rev              -0.52      0.04    -0.60    -0.44       4072 1.00
net_hospital_rev          0.41      0.04     0.33     0.50       2715 1.00
net_other_rev             0.32      0.02     0.29     0.35       2799 1.00

Family Specific Parameters: 
      Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
sigma  3307.15     37.03  3234.37  3380.60       3340 1.00
nu        3.05      0.09     2.89     3.23       3172 1.00

Samples were drawn using sampling(NUTS). For each parameter, Eff.Sample 
is a crude measure of effective sample size, and Rhat is the potential 
scale reduction factor on split chains (at convergence, Rhat = 1).
```

In the fifth Bayesian model we drop the categorical variable year, for reasons to be discussed later. The intercept parameter estimate shifted to -$3,486.54 with standard error $60.49. 95% of the distribution on this parameter falls between -$3,365.33 and -$3,365.53. The distributions for the coefficient parameters remain unchanged from model 4. All Rhat values are equal to 1, showing that there is convergence in the simulated chains.

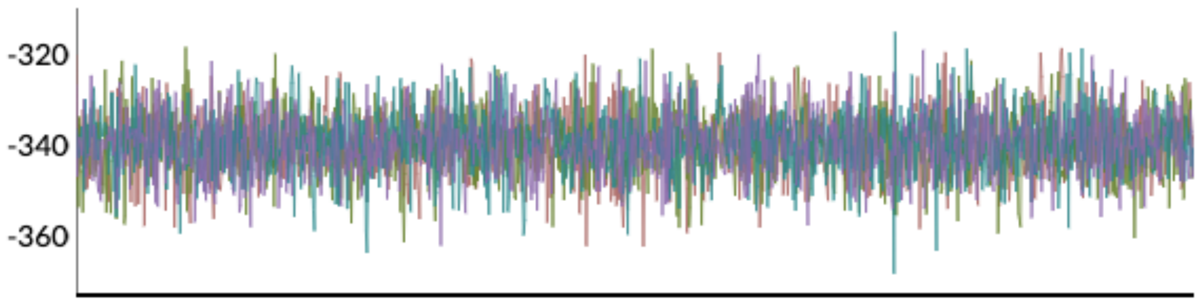


The Pareto Smoothed Importance Sampling plot looks good, as all shape k parameters are below the threshold of 0.7.

Model Diagnostics:

In order to discuss the results of the model, we must also explore the diagnostics affiliated with them. A true advantage of the Bayesian framework is the ability to predict well out of sample relative to other techniques (supervised learning, traditional regression analysis). To diagnose the model, we will use several measures including: chain consistency/convergence, energy information, MCMC chain autocorrelation, and review the posterior predictive distribution.

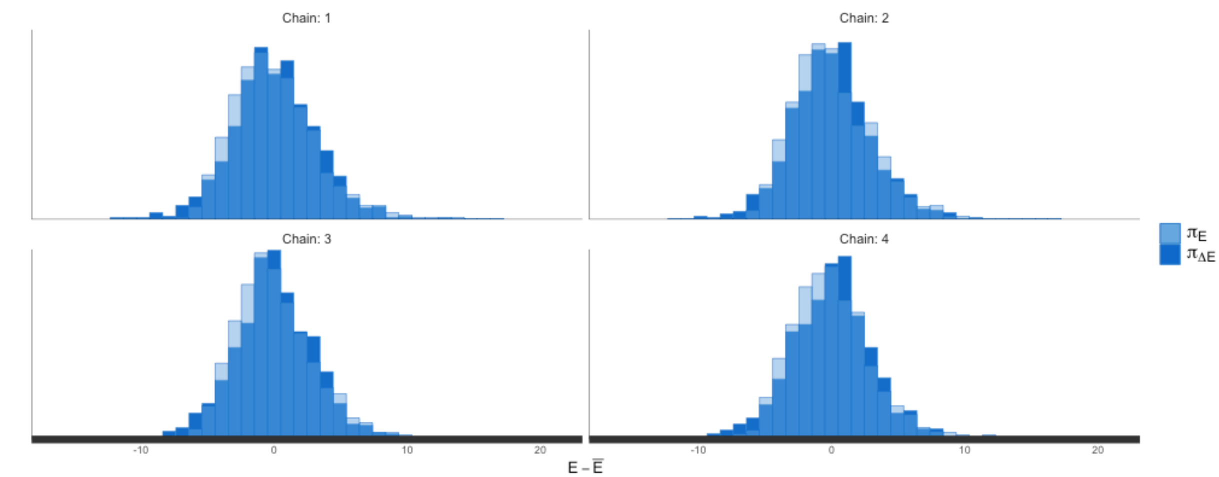
Checking the MCMC Chains:



Parameter: b_Intercept

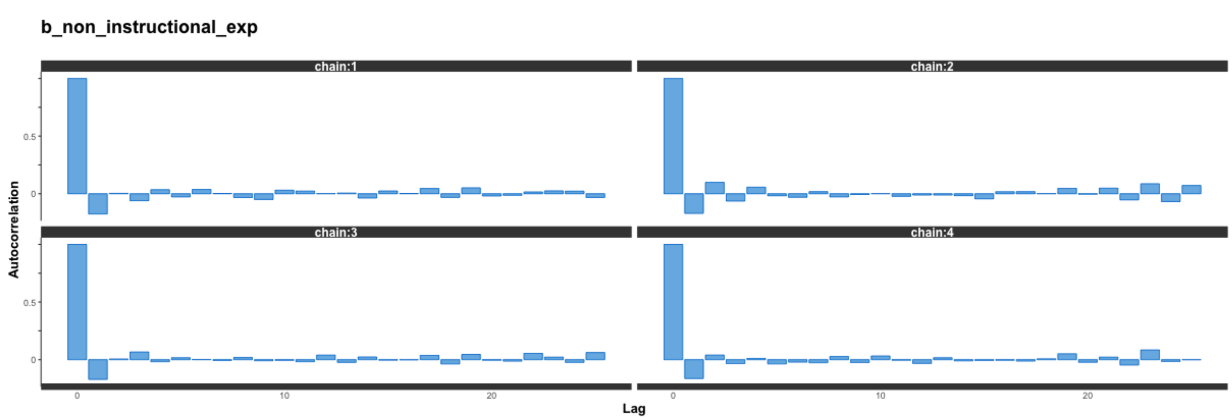
A Markov process is a sequence of random variables with a specific dependence structure. The future is conditionally independent of the past given the present. However, nothing is marginally independent of anything else. Stan utilizes a specific method of Markov chain Monte Carlo simulation called No U-Turn Sampling (NUTS). NUTS discretizes a continuous-time Hamiltonian process in order to solve a system of Ordinary Differential Equations. Above is a visual representation of the four MCMC chains and is a way of checking how NUTS performed during our simulation. From a visual perspective, NUTS seems to have performed well, though we will need to examine other parts of the posterior distribution to confirm this. The consistency of the random samples throughout the simulation process shows that the chains converged and worked well.

Energy Information:



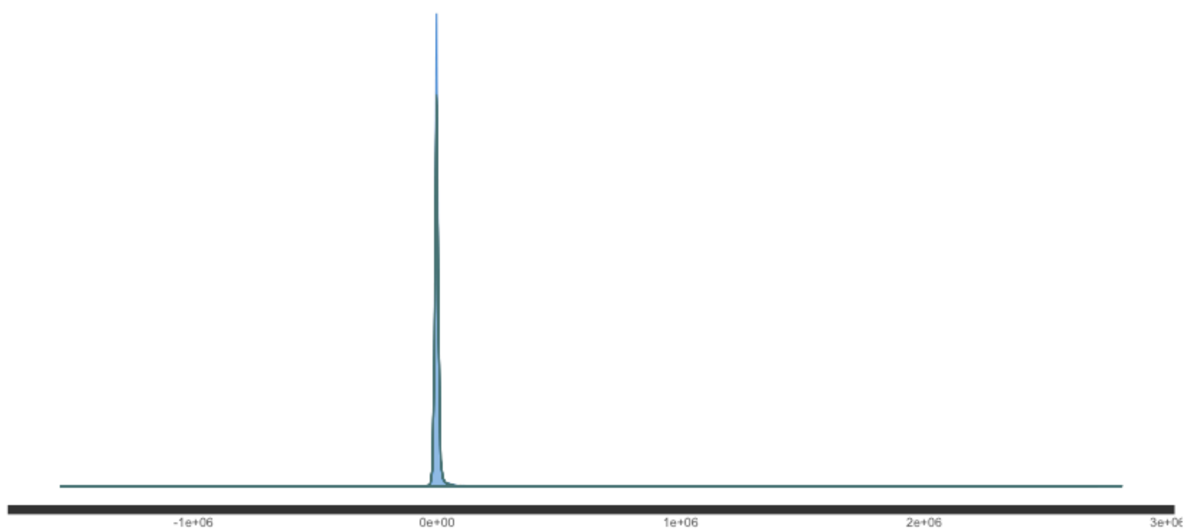
The energy information visual above represents a way to assess the low Bayesian fraction of missing information (LBFMI). When the tails of the posterior probability density function are very light, NUTS can have difficulty moving through Θ efficiently. This will result in an unreliable estimate of our effective sample size. This is not issue in our model as is indicated by the energy information plots. We want to see that the distribution of energy (π_{ε}) is the same as the distribution for the change in energy($\pi_{\varepsilon i}$). Visually we can see that this is the case for all four of our chains.

Autocorrelation:



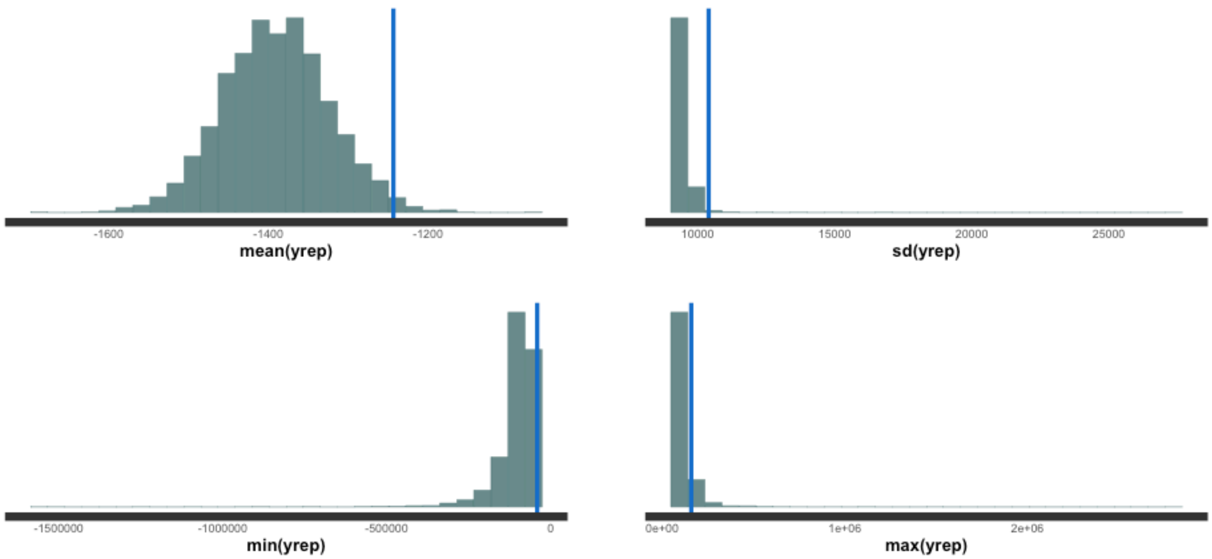
In the plot above we examine the autocorrelation between draws in each of four chains. Specifically looking at the various draws for the coefficient parameter on non-instructional expense, though all the coefficient parameters experienced similar trends. We initially started out with some autocorrelation between draws, but this quickly went to zero. This is an indication that our model was well specified and that NUTS did not struggle with sampling our model.

Posterior Distributions vs. Actual Distribution of Data:



The chart above is a way of comparing our posterior predictive distribution to the actual data. This allows us to get a sense of how well we are approximating the generating process, and which portions of the actual outcome distribution is our model missing. The light blue distribution is the actual distribution of general subsidy as seen in the dataset. The darker lines are the predicted posterior distributions. The posterior predictive distribution does a good job of capturing the actual data. It is centered well around the mean but allows for draws from the extreme values at the tails. A student t prior was a good choice in allowing Stan to take draws that were far away from the center.

Distribution of Test Statistics:



Exploring the posterior predictive distribution further, the above plot shows how well the posterior distribution estimates the mean, standard deviation, minimum, and maximum of the actual general subsidy distribution. The posterior estimates the minimum and the maximum of general subsidy distribution well. The posterior struggles more with the tails of the distribution. The model consistently overestimates the mean of general subsidy. This can be attributed to our student t priors that we have placed on our posterior predictive distribution. The need to allow for large tails can also influence our estimates of the mean. That being said, on the whole, given how long the tails of the general subsidy distribution are, this model seems to do a good job of approximating the entirety of the posterior distribution.

Leave One Out Cross Validation:

To evaluate which model does the best job of predicting future data, we use the leave one out cross validation approach discussed in the methodology section. This approach allows us to evaluates the expected log predictive density of each model.

	elpd_diff	se_diff
model4	0.0	0.0
model5	-0.3	0.1
model3	-1766.0	87.4
model2	-1766.2	87.4
model1	-1919.9	94.4

The fourth model seems to do best in generating the outcome variable general subsidy. As expressed, we measured this in expected log predictive density. We can see that model 4 and 5 have very similar expected log predictive density, with model 5 having a difference of -0.3 with standard error 0.1. The first three models have a significantly lower expected log predictive density. This may indicate that it is the combination of all the sources of background risk that contribute to a school's strategic decision to set a subsidy level. As opposed to any one source of background risk dominating the decision process. Despite the slight edge in expected log predictive density between models 4 and 5, the parameter estimates in model 5 all contained distributions to one side of zero. This indicates that there is less uncertainty that the estimates are exactly zero. Due to this advantage, and the relative predictive density of the two models, model 5 seems as a clear choice for both understanding and predicting general subsidy levels in higher education.

Discussion and Conclusion:

Our fifth Bayesian model appears to have a solid fit of the data. We will now discussion the implications of the results as well as discuss the model diagnostics to check validity. The intercept parameter in model 5 is estimated to be centered around -$3,486.54. This starkly differs from previous research which found the average general subsidy to be positive. Winston and Yen (1995) estimated that in 1991 the average higher education institution produced a unit of education for $10,653, sold it for $3,101, creating an average general subsidy of $7,551. This paper went on to say that one of the fundamental parts of the economics of higher education is the permanent feature of general subsidies. Perhaps this difference is caused by the differences in time periods explored. Where Winston and Yen (1995) looked at data from the 1991 NCES IPEDS database. Though I source my data from the same organization, my data spans 2004-2015. Perhaps there were shifts in the way that higher education institutions were run during that time period versus the one relevant to this paper. The revelation is important however, as the higher education market continues to increase in competition the effects of general subsidies on the institution should grow as well. We also saw that by adding new variables to the model we were able to tighten the distribution for the intercept parameter. By adding additional variables to our Bayesian model, we were able to reduce some of the uncertainty around this estimate.

This paper has theorized that increases in revenue sources not affiliated with Educational and General activities would have a relationship with the outcome variable general subsidy. However, this relationship would vary based on the perceived volatility of these sources. In creating this theory, we marry the concepts of background risk with that of the general subsidy. The decision to use net revenue variables in several cases was made in order to capture both the revenue and expenditure effects of these variables. Starting with net auxiliary revenue, we see that there is a negative estimated coefficient of -0.52. There is limited uncertainty around this coefficient estimate (standard error = 0.04) and lower and upper confidence intervals of -0.60 and -0.44 respectively. This result is expected as auxiliary expenses and revenues generally net out to zero. Auxiliary revenue is also tied to student consumption activity. A larger general subsidy would perhaps lead to less out of pocket spend on auxiliary services at the institution by students. Net hospital revenue shows a positive relationship with the outcome variable with an estimated coefficient parameter of 0.41 (standard error = 0.04). There is limited uncertainty around this parameter as well, giving us increased confidence in the fact that a relationship may exist. This is an expected result as hospital revenues tend to have stable revenue sources. I would imagine that the estimated coefficient parameter is not higher because so many institutions do not generate hospital revenues. The estimated coefficient parameter on net other revenue is 0.32 (standard error = 0.02). There is again limited uncertainty around this estimated parameter, giving us more confidence in the positive relationship. The magnitude of this revenue source parameter estimate is smaller because of the volatility in this measure. Net other revenue encompasses many different revenue sources on an annual basis, and therefore can bounce around quite a bit. One would expect that there is some positive correlation between net other revenue and general subsidy, however one would not expect this relationship to be stronger than is grants/appropriations. The coefficient parameter estimate on grants/appropriations has the largest magnitude. This is to be expected as appropriation and grant resources tend to have low volatility. Most grants and appropriations can span multiple years, and the organizations that issue the tend to only change their policies with sufficient warning. In another sense, this is one of the revenue sources that an institution should be able to forecast with some amount of certainty.

The results seem to match well with the proposed hypothesis. Generally speaking, increased revenue led to an increase in the general subsidy that a school offers. The source of that increased revenue does seem to have an impact on the magnitude of the relationship. As discussed previously, revenue sources that have less background risk associated with them (i.e. they are more stable and predictable) have a larger effect on general subsidy. This is highlighted the most by the two ends of the range of coefficient parameters. Auxiliary revenue and other revenue sources tend to be more volatile than other revenue sources. As a result, the magnitude of other revenue sources coefficient parameter is less than that of grants/appropriations, hospital revenue, and gifts. Auxiliary revenue, which is dependent on more unknown parameters than the other predictors in our model, has a negative relationship with general subsidy. While we explored the specifics of this relationship already, the point remains the same: volatile revenue sources have less (to a negative) impact on general subsidy than more stable revenue sources.

Works Cited:

D. Hoffman, Matthew & Gelman, Andrew. (2011). The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. Journal of Machine Learning Research. 15.

Dekel, E.: Asset demands without the independence axiom. Econometrica 57 (1), 163-169 (1989)

Dimmock, Stephen G. "BACKGROUND RISK AND UNIVERSITY ENDOWMENT FUNDS." *The Review of Economics and Statistics*, vol. 94, no. 3, 2012, pp. 789–799., www.jstor.org/stable/23261478.

Donaldson, D., Weymark, J. A.: A single-parameter generalization of the Gini indices of inequality. Journal of Economic Theory 22 (1), 67-86 (1980)

Hansmann, Henry B. "The Role of Nonprofit Enterprise." *The Yale Law Journal*, vol. 89, no. 5, 1980, pp. 835–901. *JSTOR*, www.jstor.org/stable/796089.

Gilboa, I., Schmeidler, D.: Maxmin expected utility with non-unique prior. Journal of Mathematical Economics 18 (2), 141-153 (1989)

Gollier, C.: The economics of risk and time. Cambridge, MA: MIT Press 2000

Gollier, C., Pratt, J. W.: Risk vulnerability and the tempering effect of background risk. Econometrica 64 (5), 1109-1123 (1996)

Grant, S., Kajii, A.: A cardinal characterization of the Rubinstein-Safra-Thomson axiomatic bargaining theory. Econometrica 63, 1241-1249 (1995)

Grant, S., Karni, E.: A theory of quantifiable beliefs. Working Papers in Economics and Econometrics No. 388, Australian National University (2000)

Gui, F.: A theory of disappointment aversion. Econometrica 59 (3), 667-686 (1991)

Kimball, M. S.: Standard risk aversion. Econometrica 61 (3), 589-611 (1993)

Litten, Larry H. "Marketing Higher Education: Benefits and Risks for the American Academic System." *The Journal of Higher Education*, vol. 51, no. 1, 1980, pp. 40–59. *JSTOR*, www.jstor.org/stable/1981124.

Mark Machina, (1982), "Expected Utility Analysis without the Independence Axiom", *Econometrica*, 50, (2), 277-323

Pratt, J.: Aversion to one risk in the presence of others. Journal of Risk and Uncertainty 1 (4), 395-414 (1988)

Pratt, J. Zeckhauser, R.: Proper risk aversion. Econometrica 55 (1), 143-154 (1987)
Quiggin, J., Chambers, R. G.: Risk premiums and benefit measures for generalized expected utility theories. Journal of Risk and Uncertainty 17 (2), 121-37 (1998)

Rubenstein, A., Safra, Z., Thomson, W.: On the interpretation of the Nash bargaining solution and its extension to non-expected utility preferences. Econometrica 60 (5), 1171-1186 (1992)

Safra, Z., Segal, U.: Constant risk aversion. Journal of Economic Theory 83 (1), 19-42 (1998)

Weymark, J.: Generalized Gini inequality indices. Mathematical Social Sciences 1, 409-430 (1981)

Winston, Gordon C., and Ivan C. Yen. "Costs, Prices, Subsidies, and Aid in U.S. Higher Education." *Williams Project on the Economics of Higher Education*, July 1995, sites.williams.edu/wpehe/files/2011/06/DP-32.pdf.

Winston, Gordon C. "Subsidies, Hierarchy and Peers: The Awkward Economics of Higher Education." *The Journal of Economic Perspectives*, vol. 13, no. 1, 1999, pp. 13–36. *JSTOR*, www.jstor.org/stable/2647135.

Yaari, M.: The dual theory of choice under risk. Econometrica 55 (1), 95-115 (1987)