Collin DeVore
Time Series Analysis 2 Project

Abstract

This analysis attempts to explain the differing effects of predictor variables on the number of female and male bachelor's level enrollments each year. Four independent variables are chosen: personal income per capita, the elasticity of college tuition and other fees, the unemployment rate, and the amount of money given in scholarships and other such aid. Personal income per capita, college tuition, and the average number of scholarships per student are all adjusted to real 2018 dollar values. After testing for cointegration, the variables are separated into a male enrollment and a female enrollment regression, which are then put into a seemingly unrelated regression structure. The personal income per capita and the log of the college tuition are then omitted from the overall regression, since they are not found to be significant. Theil's F-test is performed on the remaining two regressors for male and female enrollments. The study then finds that neither the amount of scholarship aid per student nor the unemployment rate affects males and females differently.

Introduction

The number of students that enroll full time at a four-year institution appears to be on the rise in present day society, though the cost of tuition and other such fees have continued to increase drastically. Nevertheless, males and females appear to be enrolling in college in different numbers, almost as if they are responding differently to these sorts of stimuli. Many studies have attempted to observe the variation and the way in which future students react to these different stimuli within the enrollment process. Studies have been performed, for instance, that have analyzed the impact of the unemployment rate on community college enrollments, finding that the amount of community college students enrolled depends on and reacts heavily to the unemployment rate (Betts & McFarland 1995). Though the authors of this article appear to suggest that this may not be generalizable to the entire college system, they do find that the amount of revenue for the community colleges decreases in response to the unemployment rate as the government begins cutting its funding (Betts & McFarland 1995). For this reason, though not stated by the authors, it could be possible that the unemployment rate works as a predictor of the amount of full-time college enrollment at the national level, especially when the type of college is not differentiated. Another article, continuing the research of Betts and McFarland, examined the same phenomenon while accounting for tuition, employment, and income, among other such variables (Hillman & Orians 2013). This article confirmed the previous article's conclusion using panel data while controlling for other factors (Hillman and Orians 2013). It also included the results from the Great Recession (Hillman & Orians 2013). Other studies, such as the study performed by James Wetzel, Dennis O'Toole, and Steven Peterson, have analyzed the effect of cost on collegiate enrollment rates. Here, Wetzel, O'Toole, and Peterson analyze how costs affect the enrollment rates of minority students, particularly of white and black students (Wetzel, O'Toole, and Peterson 1998). It appears that minority students have a larger response to the costs of college than majority students do (Wetzel, O'Toole, and Peterson 1998). The study does conclude, however, that aggregate enrollment rates do not respond much when the cost of college changes (Wetzel, O'Toole, and Peterson 1998). They also suggest that the response of minorities to the changes in cost imply that the amount of financial aid given will affect the minorities more than others (Wetzel, O'Toole, and Peterson 1998). Braunstein, Mcgrath, and Pescatrice studied a more generalized version of this phenomenon, in which the amount of scholarships given per student was analyzed so that the impact on the enrollment rates could be measured (Braunstein, Mcgrath, and Pescatrice 1999). These researchers found that the amount of aid given did influence enrollment rates positively, though the amount given would depend on the characteristics of the market and of the college itself (Braunstein, Mcgrath, and Pecatrice 1999). Much of the enrollments and the aid appear to depend heavily on the outside characteristics and labor markets. There do not appear to be many studies, however, that examine the differing effects of the enrollment process on men and women, though there are researchers such as Ayako Kando who have studied the impact that entering the labor force has on men and women, and across race (Kando 2015). Kando has shown in his analysis that entering the labor force during a recession affects men more than women, though the races do not appear to differ (Kando 2015). Though these interrelated analyses are beneficial to an overall understanding of how college employments work, they are incomplete without an analysis of how male and female enrollment rates respond to changes within the predictor variables.

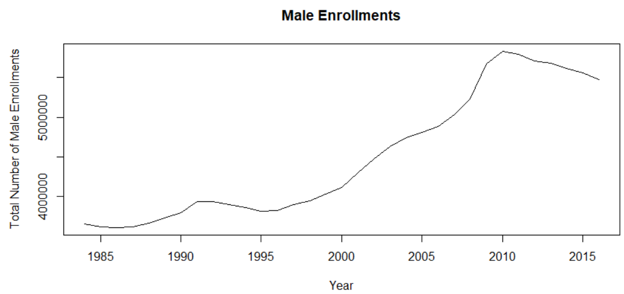
Method

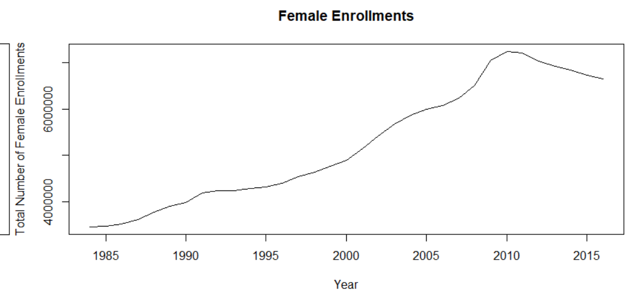
This paper seeks to address and compare the way in which males and females react to different stimuli when enrolling full time in a four-year institution. Specifically, this paper attempts to measure and compare the differing dynamics that personal income per capita, college tuition and fees, the unemployment rate, and the amount of scholarships and other aid per student have on male and female enrollment rates over time. These four variables serve as a sort of intertemporal measure of the demand of college, which, ideally, will help to explain some of the differences between the rates of enrollment between the sexes. After the dynamics of the systems are analyzed and first differences are taken of any nonstationary variables, cointegration is measured between the enrollment rate of each of the three variables by the Engle – Granger two step cointegration test, in order to examine if any of these variables cointegrate individually with either regressor. After the possibility of cointegration has been analyzed, seemingly unrelated regression models are set up for both males and females utilizing the three regressors. Any of the three regressors that are not significant in the regression equations are omitted in order to ensure that the models have the least amount of bias possible. Once these regressions are set up, Theil's F – Test calculation is used in order to determine if any of the coefficients of the remaining regressors are statistically the same for both males and females.

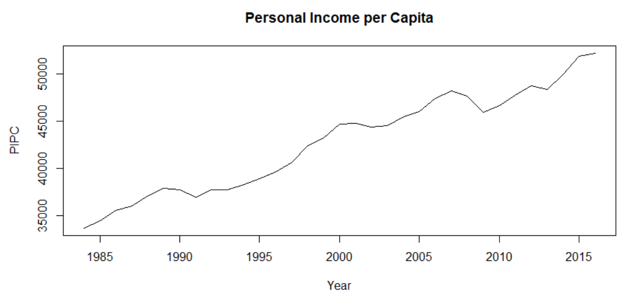
Data

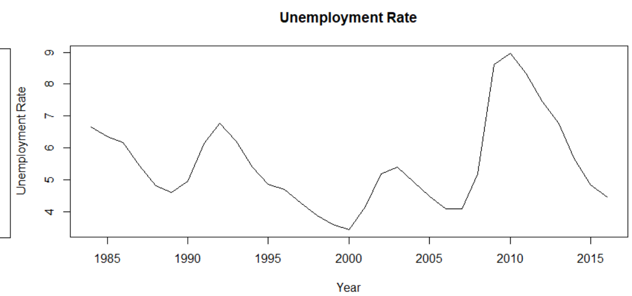
The college enrollment data is gathered from the Digest of Education Statistics reported by the National Center for Education Statistics. This is a digest of data published yearly regarding the state of education within the United States. For this analysis, the total number of males and females enrolled full time at the undergraduate level in the fall semester are the only dependent variables considered. This dataset ranges from 1984 to 2016, creating 33 separate datapoints. The data for 2002 was never reported, so the values from 2001 and 2003 are averaged together to estimate this point in time. The unemployment data and the measure of the personal income per capita are given by the Saint Louis Federal Reserve (FRED) system. Personal income per capita is chosen for its ability to measure income outside of regular wages so that the estimate of income for individuals is more accurate. Since unemployment is not reported yearly, the monthly unemployment rate was taken and averaged together for each year. The other two variables, tuition and fees and scholarships and other aid, is obtained from the organization Collegeboard. The graph of each of these variables is provided below. Due to the nature of the data, the measure of personal income per capita is adjusted to reflect real dollar values based in the year 2018. The data for the amount of tuition and fees has already been adjusted to real values based in the year 2018 by Collegeboard, and the variable for scholarships and other aid has been adjusted to reflect real values based in the year 2017 by Collegeboard. Using the Consumer Price Index with a base year around 1984 provided by usinflationcalculator.com, the variable for scholarships and other aid is adjusted to reflect real 2018 dollar values. Once this has been performed, each of the monetary values can be said to be adjusted to show real 2018 dollar values. Because of the logarithmic nature of the data, only the log of the cost of tuition and fees for each year will be analyzed. Thus, this variable will show how well the elasticity of the amount of tuition of and fees predicts the aggregate number of males and females enrolled in college in the fall semesters.

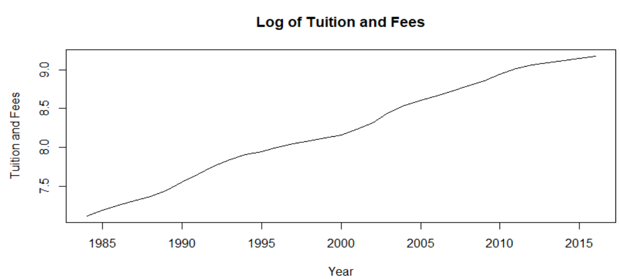
Graphs of the Variables

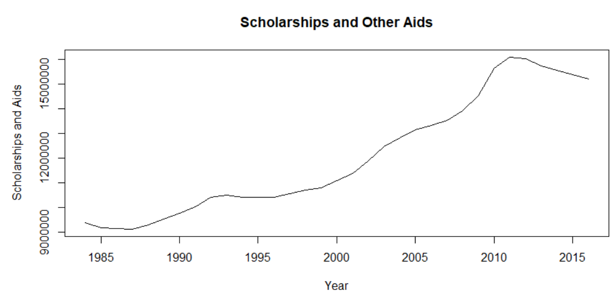












Analysis

In order to analyze the data, each of the datasets must first be made stationary. The number of male enrollments, female enrollments, personal income per capita, the amount of tuition and other fees, and the amount of scholarships and other aid per student can be observed to follow a clear upward trend. Unemployment does not necessarily seem to follow this trend but can instead be observed as a series of peaks and troughs that may follow no trend whatsoever. It does appear, based on sight alone, that the personal income per capita, unemployment rate, and the cost of tuition and fees are affecting each other or the dependent variable in any way, though there could be the influence of some variable not measured that is affecting each of these variables individually. In 2008, for instance, there is a clear rise in the number of both males and females enrolling in college while the unemployment rate begins to skyrocket and the personal income per capita starts to decrease. It is almost as if these people were trying to beat an oncoming financial crisis by enrolling in college. As the unemployment rate begins severely decreasing in 2010 and the personal income per capita begins increasing at nearly the same rate that it was before 2008, the amount of people enrolling for college decreases for the first time since about 1996, during which the unemployment rate was also quickly decreasing rapidly. There is a small change in the amount of tuition in fees from 2008 to 2010, but what occurred is hard to understand on sight alone. On the other hand, the amount of scholarships and other aid per student look as though they could be superimposed onto male and female enrollments and still provide an approximate fit. It follows the number of enrollments extremely closely. For many of these events, males appear to respond more sharply, since they have a curve that is less smooth than the females. In order to analyze that, however, the Dickey – Fuller Tests must be performed, and the nonstationary variables must be differenced. A table summarizing the outcomes of the Dickey – Fuller tests is given in Appendix A of this analysis.

The Dickey – Fuller tests appear to show that the number of male enrollments, the number of female enrollments, the personal income per capita, and the unemployment rate are nonstationary if they have a drift term and therefore must be first differenced. Though variables such as personal income per capita appear to be stationary when looking at the tests with no trend or drift and the test with both trend and drift, for the purposes of this paper the variables will be first differenced in order to ensure stationarity. It is thus assumed here that each of these variables has a drift term within the calculation. This would suggest that unemployment between 1984 and 2016 did, in fact, have a trend, though it is not as prominent as the trend of the other variables. The tuition and fees, on the other hand, show that the null of a unit root must be rejected. This may be because the cost of tuition follows an extremely linear trend. Nonetheless, the trend must be taken out for each variable, so each of them must be first differenced before their dynamics can be analyzed. The graphs of the autocorrelation and partial autocorrelation functions are given in Appendix A of this analysis so that the possibility of trend can be further analyzed.

The autocorrelation and partial autocorrelation functions seem to show strong persistence for every variable except for the unemployment rate. This discrepancy with the Dickey - Fuller tests stem from the fact that there are so few data points, which can make the standard errors large. Since all variables excluding tuition are shown to be nonstationary by the Dickey – Fuller Tests, each of these are first differenced. Likewise, since the tuition variable has a clear trend in the graph of the data and strong persistence in the autocorrelation functions, this variable is first – differenced. Therefore, all the variables provided here are first differenced. The graphs of the differenced variables are provided in Appendix A.

Upon observing the differenced variables, it appears as though many of the graphs are following the same sorts of patterns. The differenced male enrollments, the differenced female enrollments, and the differenced unemployment variable seem to follow the same patterns. The differenced tuition and fees may follow the same patterns, since it has peaks and troughs at about the same areas, but it appears to have much larger troughs than the other graphs. The differenced personal income per capita graph looks as though it is following the inverse pattern of what the other graphs are following, with peaks occurring around the points where there are troughs in other graphs and troughs occurring where there are peaks in other graphs. Nonetheless, it appears as though more testing is needed.

Now that the variables have been made stationary, the differenced unemployment variable, the differenced personal income per capita variable, and the differenced tuition and fees can be tested singularly for cointegration with both the male and female enrollment variables. To do this, the Two – Step Engle – Granger method is used. First, each of the variables are regressed on the male and female enrollment variable. Next, the Dickey – Fuller test is performed to examine if the residuals are nonstationary. The regression equations and the final Dickey – Fuller test are given in Appendix B.

For most of the tests, the conclusion is reached that the null of nonstationarity must be rejected utilizing the test statistic. The exception to this rule is the effect of scholarships and other aid on male enrollment. This would suggest cointegration between these two variables, however, there is not enough evidence to conclude with 95% certainty that the enrollment rate of females is cointegrated with the amount of scholarships and other aid given per student each year. Though the p – value is close to 0.05, since it is not at or below this value, the idea that the number of females enrolling is cointegrated with this variable is not considered. Furthermore, to make the setup of the two equations match more closely together, the long – run effects of males and scholarships and aid are not considered for the final model. In this way, only the short run effects of the variables will be considered in the final model so that the equations for the number of enrollments for men and women will maintain the same structure. The rest of the results indicate that none of the residuals of the regressions are stationary, indicating that there is no singular cointegration between any of these variables and male and female enrollment rates. This means that there are no other error correction or cointegration effects that must be considered in calculating the final model. Furthermore, it appears as though the best model to estimate this data will be a basic dynamic linear regression.

Now that the lack of a necessity to factor in cointegration has been established, a seemingly unrelated regression can be run. Two models are set up of the following forms.

$$\Delta male_t = \beta_1 \Delta pipc_t + \beta_2 \Delta unemp_t + \beta_3 \Delta tuitfee_t + \beta_4 \Delta \mathrm{scholarships_t} + \varepsilon_t$$
$$\Delta \mathrm{fe}male_t = \beta_1 \Delta pipc_t + \beta_2 \Delta unemp_t + \beta_3 \Delta tuitfee_t + \beta_4 \Delta \mathrm{scholarships_t} + \varepsilon_t$$

These two models are then run under the seemingly unrelated regression model, producing the following results.

<u>Output</u>

```
systemfit results
method: SUR

        N DF          SSR              detRCov  OLS-R2 McElroy-R2
system 64 54 237788881575 5994356455023388672 0.76003   0.669073
```

```
    N DF          SSR        MSE    RMSE       R2   Adj R2
m1 32 27  74634475297 2764239826 52576.0 0.800880 0.771380
f1 32 27 163154406278 6042755788 77735.2 0.735177 0.695944

The covariance matrix of the residuals used for estimation
           m1         f1
m1 2764239826 3272502063
f1 3272502063 6042755788

The covariance matrix of the residuals
           m1         f1
m1 2764239826 3272502063
f1 3272502063 6042755788

The correlations of the residuals
         m1       f1
m1 1.000000 0.800709
f1 0.800709 1.000000


SUR estimates for 'm1' (equation 1)
Model Formula: maled ~ pipcd + unempd + ltuitfeedd + scholarshipsd

                     Estimate      Std. Error t value   Pr(>|t|)
(Intercept)     39821.7196779   14769.9714822 2.69613 0.01192975 *
pipcd               2.5574094      19.0280317 0.13440 0.89408142
unempd          85680.2391329   20449.1334801 4.18992 0.00026732 ***
ltuitfeedd     165418.3622748 1234992.5878451 0.13394 0.89444122
scholarshipsd       0.1194021       0.0415413 2.87430 0.00780106 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 52576.038514 on 27 degrees of freedom
Number of observations: 32 Degrees of Freedom: 27
SSR: 74634475297.1815 MSE: 2764239825.82154 Root MSE: 52576.038514
Multiple R-Squared: 0.80088 Adjusted R-Squared: 0.77138


SUR estimates for 'f1' (equation 2)
Model Formula: femaled ~ pipcd + unempd + ltuitfeedd + scholarshipsd

                    Estimate     Std. Error  t value  Pr(>|t|)
(Intercept)     78743.239247   21837.821832  3.60582 0.0012431 **
pipcd              -1.252226      28.133485 -0.04451 0.9648253
unempd          97317.151811   30234.623953  3.21873 0.0033390 **
ltuitfeedd     594151.313398 1825971.575501  0.32539 0.7473948
scholarshipsd       0.159764       0.061420  2.60117 0.0148930 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 77735.164424 on 27 degrees of freedom
Number of observations: 32 Degrees of Freedom: 27
SSR: 163154406277.992 MSE: 6042755788.07379 Root MSE: 77735.164424
Multiple R-Squared: 0.735177 Adjusted R-Squared: 0.695944
```

The results of each step of the seemingly unrelated regression analysis are given in Appendix C. Based on the output, it becomes clear that the regression equations can be improved upon so that the male and female enrollment rate is better predicted by the regression equations, which should further the significance and understanding of the model. This is inferred because the differenced personal income per capita (pipcd) and the twice differenced elasticity of the amount of tuition and fees (ltuitfeedd) appear to be highly insignificant when predicting both the number of female and male enrollments. This can be observed through the large p – values for both the male enrollments equation (with p – values of 0.89408142 for personal income per capita and 0.89444122 for the elasticity of tuition) and the female enrollments equation (with p – values of 0.9648253 for personal income per capita and 0.7473948 for the elasticity of tuition). For this reason, and because extracting one variable will not create a large enough impact to make the other significant, both variables must be extracted from the model. This gives the final output shown below.

Final Output

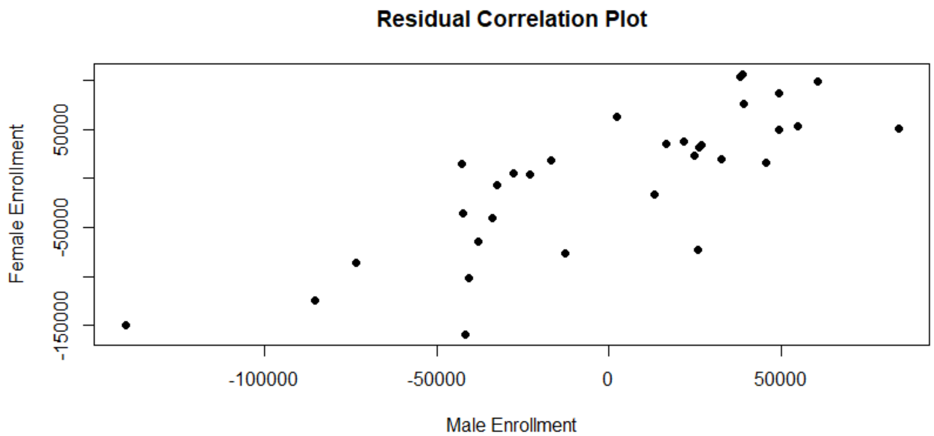
```
systemfit results
method: SUR

        N DF          SSR             detRCov   OLS-R2 McElroy-R2
system 64 58 238547418880 5237448021568674816 0.759264   0.668267

    N DF          SSR        MSE    RMSE       R2   Adj R2
m3 32 29  74752880129 2577685522 50770.9 0.800564 0.786810
f3 32 29 163794538751 5648087543 75153.8 0.734138 0.715803

The covariance matrix of the residuals used for estimation
           m3         f3
m3 2577685522 3053120611
f3 3053120611 5648087543

The covariance matrix of the residuals
           m3         f3
m3 2577685522 3053120611
f3 3053120611 5648087543

The correlations of the residuals
         m3       f3
m3 1.000000 0.800163
f3 0.800163 1.000000


SUR estimates for 'm3' (equation 1)
Model Formula: maled ~ unempd + scholarshipsd

                   Estimate    Std. Error t value      Pr(>|t|)
(Intercept)   40638.5355159 11707.9258260 3.47103     0.0016449 **
unempd        83997.7942077 12849.8506036 6.53687 0.00000036937 ***
scholarshipsd     0.1218895     0.0382899 3.18333     0.0034626 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 50770.912161 on 29 degrees of freedom
Number of observations: 32 Degrees of Freedom: 29
SSR: 74752880129.4501 MSE: 2577685521.70518 Root MSE: 50770.912161
Multiple R-Squared: 0.800564 Adjusted R-Squared: 0.78681


SUR estimates for 'f3' (equation 2)
Model Formula: femaled ~ unempd + scholarshipsd
```

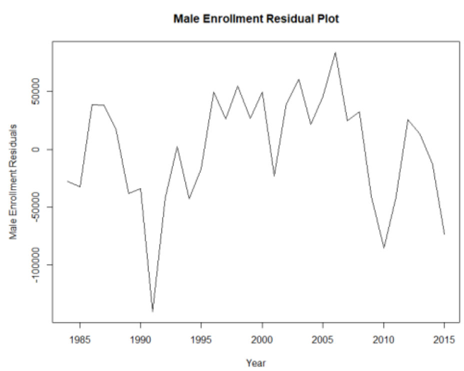
```
                 Estimate    Std. Error t value    Pr(>|t|)
(Intercept)   77324.2724275 17330.6840955 4.46170  0.00011287 ***
unempd        99475.1343452 19021.0208705 5.22975 0.000013394 ***
scholarshipsd     0.1625278     0.0566787 2.86753  0.00763154 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 75153.759341 on 29 degrees of freedom
Number of observations: 32 Degrees of Freedom: 29
SSR: 163794538751.039 MSE: 5648087543.13928 Root MSE: 75153.759341
Multiple R-Squared: 0.734138 Adjusted R-Squared: 0.715803
```

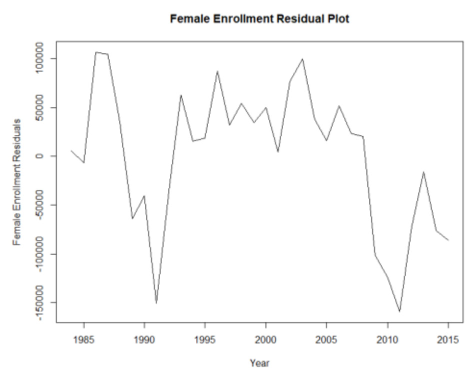
For this final model, both the differenced unemployment variable and the differenced amount of scholarships per student awarded each year are shown to be highly significant predictors of the amount of male and female enrollments each year. Put more simply, this output shows that the following equations hold under a seemingly unrelated regression structure.

$$\Delta males_t = 40{,}638.5355159 + (83{,}997.7942077 * \Delta unemployment_t) + (0.1218895 * \Delta scholarships_t) + \varepsilon_t$$
$$\Delta females_t = 77{,}324.2724275 + (99{,}475.1343452 * \Delta unemployment_t) + (0.1625278 * \Delta scholarships_t) + \varepsilon_t$$

Here, the correlation between the residuals has been incorporated into the estimations of the models. The correlation here is shown above by the correlation matrix within the output of the model. Below, the residual plots are given, along with a plot that shows the correlation between the residuals. For each of these residuals, the Dickey – Fuller test has been calculated. The Dickey – Fuller test for the residuals of the male variable is given by a p – value of 0.02002, suggesting that the residuals are stationary with over 95% confidence. The residuals of the equation for females gives a Dickey – Fuller test p – value of 0.0737, suggesting that the residuals cannot be considered stationary with 95% confidence. Upon observing the residual plots, both sets look stationary, though this may not be the case for the females.







One issue that this dynamic linear model raises within the residuals is the fact that the residuals do not fully create a pattern of white noise. This can be shown by the fact that the autocorrelation and partial autocorrelation plots of the residuals of the seemingly unrelated regression have significant lags for both functions. The graphs of these autocorrelation functions and partial autocorrelation functions are shown in Appendix C. Instead of following white noise, they appear to follow an AR(1) process, in which the two sets of residuals can be represented by the following two equations.

$$MaleError_t = -2{,}694.2080188 + 0.4739946 * MaleError_{t-1} + \varepsilon_t$$
$$FemaleError_t = -3{,}768.2519971 + 0.6228232 * FemaleError_{t-1} + \varepsilon_t$$

After this AR(1) process is fit to the residuals, the Ljung – Box p – values will increase significantly and the autocorrelation and partial autocorrelation functions will return only white noise. Though this autoregressive process occurs, because the model for male enrollments is stationary, and the model for female enrollments is as close to stationary as can be achieved with the data, this model is nonetheless assumed to be a fair estimation of the correct model.

The last question that must be considered is whether or not the coefficients of the unemployment rate and the amount of scholarships per student can be said to be statistically different for males and females. Now that the data has been set up into a seemingly unrelated regression model, it is possible to test this. First, tests on the coefficients of the unemployment rate are run. Using Theil's F test, the null hypothesis is tested in which male unemployment is equal to female unemployment. If this proves false, then it can be said that changes in the unemployment rate influence the enrollment rate of females more than males. The F – statistic given by Theil's F test is 1.7644, while the p – value is 0.1893. This means that there is not enough evidence to say that changes in the unemployment rate have differing effects on male and female enrollment rates. The next test has the null hypothesis that the coefficient of the amount of scholarships per student is the same for both the male and female enrollment rates equations. Theil's F test gives an F – statistic of 1.3699 and a p – value of 0.2466. This shows that the coefficient for the amount of scholarships per student in the male enrollment equation cannot be said to be statistically different than the coefficient for the amount of scholarships per student given in the female equation.

Conclusion and Discussion

Though the coefficients for both unemployment and scholarships are higher for the equation of the total number of enrollments for females, the results of this analysis suggest that male and female enrollments respond the same to changes in the unemployment rate and changes in the amount of scholarships per student. Furthermore, out of the four different predictor variables suggested, only the unemployment rate and the amount of scholarships and other aid per student are shown to be significant predictor variables for the number of enrollments of males and females, while the adjusted personal income per capita and the adjusted elasticity of tuition and fees are not. It is interesting to note here that the coefficients for the unemployment rate in the final models are positive, which seems to suggest that larger unemployment rates lead to more college enrollment. A quick Granger Causality test reveals that the unemployment rate is not granger causal of the amount of males or females enrolled, nor is the reverse true. The Granger Causality test does show, however, that the amount of scholarships per student is granger caused by the unemployment rate with a value of 44.242 and a p – value of 0.0000003231. This corresponds well to the phenomena observed within the previous literature on this subject. Lastly, it has been established that the enrollment of males is cointegrated with the amount of scholarships per student using the Engle – Granger two step cointegration test, while the possibility of the cointegration of the enrollment of females and the amount of scholarships per student must be rejected at the 95% confidence level.

For further analysis, this paper has a few limitations that could be addressed. To begin, in later years, it may become easier to gather more data regarding the number of full-time enrollments at public bachelor's institutions. While the thirty-three-year period is sufficient for this analysis, it may be better to use more data to observe the possible differences between the coefficients of the equations of male and female enrollments. The lack of data for this analysis leads to higher standard errors that could lead to a false conclusion of the equality of male and female enrollment data, or it could lead to incorrect coefficients, especially if the coefficients for each model are supposed to be the same. For example, earlier in this analysis, it is noted that the coefficient of the differenced unemployment rate is about 15,477 units greater for the prediction of the amount of female enrollments than it is for the amount of male enrollments, yet Theil's F test shows that these two values are the same. This is due to a lack of data before the year 1984.

Another suggestion for further analysis would be to analyze the differences between the models in the long run. Earlier, it was established that male enrollment is cointegrated with the amount of scholarships per student. It could not be said, however, with 95% confidence that female enrollment is cointegrated with the amount of scholarships per student, since the p – value of the Dickey – Fuller test is 0.06984. For this reason, long – run effects were omitted from this analysis. If the test is switched to an augmented Dickey – Fuller test in which the number of lags is selected using the Akaike Information Criterion, however, the p – value decreases to 0.05634, while the t – statistic becomes -2.0433 and the t – critical value becomes -2.93. This is still not quite enough to suggest that they are cointegrated with 95% confidence, but it is very close to the cutoff for the assumption of cointegration. It is highly probable that both dependent variables are cointegrated with the amount of scholarships and other aid given per student, but it is also possible that females are becoming more susceptible to accepting aid as time progresses. The data analyzed here shows that the amount of females enrolling full time in college began to outnumber the amount of males enrolling in college around the year 1987. Nonetheless, it is historically possible that, for some of these years, women were not responding as heavily to scholarships and other aid given per student as much as they would in more recent years. For this reason, the probability of cointegration is likely increasing on a continual basis. It may be interesting to examine in later years whether or not the strength of the cointegration and long run effects of female enrollments with scholarships is the same as the strength of the cointegration of male enrollments and scholarships. In this way, it is possible to determine if males or females will respond more to scholarships in the long run.

The last limitation with this analysis is the nonstationarity of the model for females and the AR(1) process that the residuals of both models follow. It could theoretically be possible to fit these models with the following form in order to address these issues, but it comes at a cost.

$$\Delta males_t = \beta_0 + (\beta_1 * \Delta males_{t-1}) + (\beta_2 * \Delta unemployment_t) + (\beta_3 * \Delta scholarships_t) + \varepsilon_t$$
$$\Delta females_t = \beta_0 + (\beta_1 * \Delta females_{t-1}) + (\beta_2 * \Delta unemployment_t) + (\beta_3 * \Delta scholarships_t) + \varepsilon_t$$

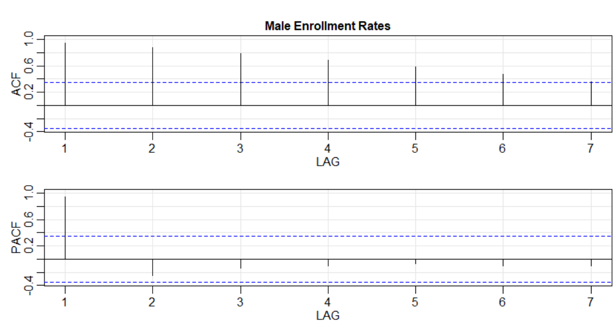
The issue with these two models is that most of the strength of either model is coming from the lagged variable, rather than the unemployment and scholarship variables. In this way, running the model creates a risk of falsely inflating the diagnostics of the fit of the variables, rather than showing how well they predict the number of enrollments. Including the lagged variables will also lead to lower coefficients, that should not necessarily be lower in value. It is possible that lagging these variables could even lead to the inclusion of weak predictor variables that should not be run in the regression. Since the point of this paper is to address if males or females react more to changes in the predictor variables, this structure has been omitted from this analysis. If a better way could be found to ensure the stationarity of both variables with 95% confidence and that filters out the autoregressive process without compromising the diagnostics of the predictor variables, it may be found to be better suited to check the results of this analysis. It could also be possible that the inclusion of long – run dynamics could fix this problem as well, which could be better performed at a later date. Since the probability of stationarity is still high and the predictor variables appear to address the number of male and female enrollments adequately, there appears to be no reason to alter the structure of the model.

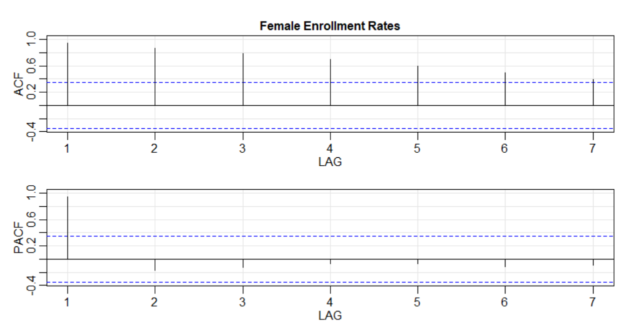
Appendices

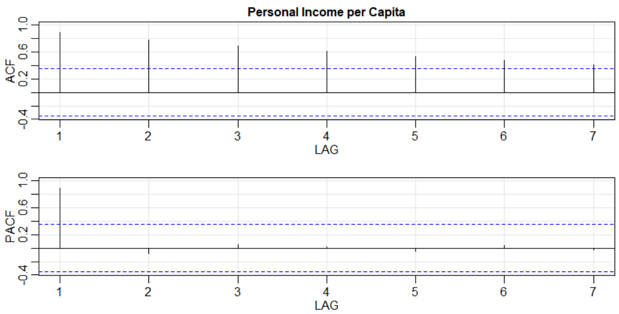
Appendix A (Initial Diagnostics)

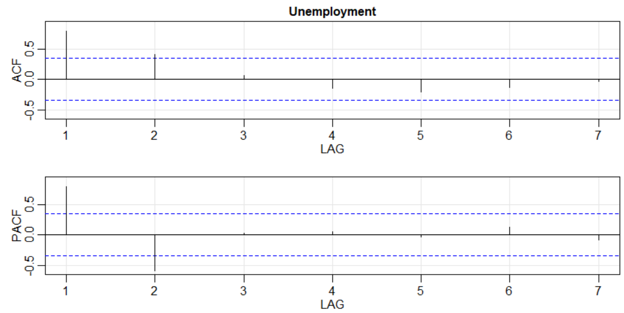
Dickey – Fuller Test								
	Overall Value of the Test Statistic			5 pct Critical Value			Significance Level	Reject/Fail to Reject
	Tau	Phi1	Phi2	Tau	Phi1	Phi2	P – Value	Null
Male Enrollments (No trend or drift)	2.852	N/A	N/A	-1.95	N/A	N/A	0.007666	Reject
Male Enrollments (with drift)	-0.1226	4.1779	N/A	-2.93	4.86	N/A	0.9032	Fail to Reject
Male Enrollments (with trend)	-1.4709	3.6723	1.1574	-3.50	5.13	6.73	0.3284	Fail to Reject
Female Enrollments (No trend or drift)	3.545	N/A	N/A	-1.95	N/A	N/A	0.00127	Reject
Female Enrollments (with drift)	-0.9144	8.4377	N/A	-2.93	4.86	N/A	0.3678	Fail to Reject
Female Enrollments (with trend)	-0.4781	5.4805	0.4464	-3.50	5.13	6.73	0.6442	Fail to Reject
Personal Income per Cap (No trend or drift)	4.0722	N/A	N/A	-1.95	N/A	N/A	0.0002988	Reject
Personal Income per Cap (with drift)	-0.3189	8.5183	N/A	-2.93	4.86	N/A	0.752	Fail to Reject
Personal Income per Cap (with trend)	-2.1863	7.9333	2.3949	-3.50	5.13	6.73	0.109	Fail to Reject
Unemployment (No trend or drift)	-0.8669	N/A	N/A	-1.95	N/A	N/A	0.3927	Fail to Reject
Unemployment (with drift)	-1.7945	1.7167	N/A	-2.93	4.86	N/A	0.08281	Fail to Reject
Unemployment (with trend)	-1.8099	1.1622	1.6398	-3.50	5.13	6.73	0.2115	Fail to Reject
Log of Tuition and Fees (No trend or drift)	12.8377	N/A	N/A	-1.95	N/A	N/A	0.00000000	Reject
Log of Tuition and Fees (with drift)	-2.1056	112.8256	N/A	-2.93	4.86	N/A	0.04372	Reject
Log of Tuition and Fees (with trend)	-0.1637	72.7102	2.1429	-3.50	5.13	6.73	0.1355	Fail to Reject
Scholarships and Aid (No trend or drift)	3.4407	N/A	N/A	-1.95	N/A	N/A	0.001679	Reject
Scholarships and Aid (with drift)	0.0469	5.9903	N/A	-2.93	4.86	N/A	0.9629	Fail to Reject
Scholarships and Aid (with trend)	-1.7818	5.4844	1.742	-3.50	5.13	6.73	0.193	Fail to Reject
Differenced Log Tuition and Fees (No trend or drift)	-1.023	N/A	N/A	-1.95	N/A	N/A	0.3142	Fail to Reject
Differenced Log Tuition and Fees (with drift)	-1.9313	1.9469	N/A	-2.93	4.86	N/A	0.06294	Fail to Reject
Differenced Tuition and Fees (with trend)	-2.2275	1.8444	2.683	-3.50	5.13	6.73	0.08529	Fail to Reject

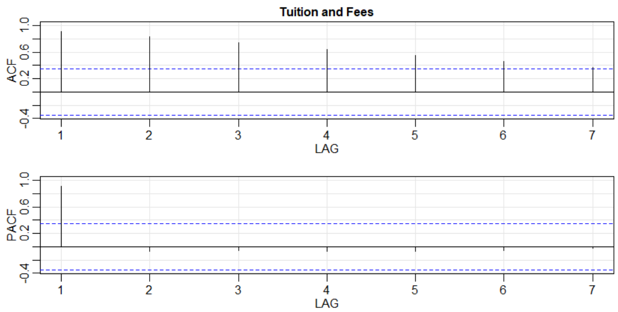
Graphs of the Autocorrelation and Partial Autocorrelation Functions

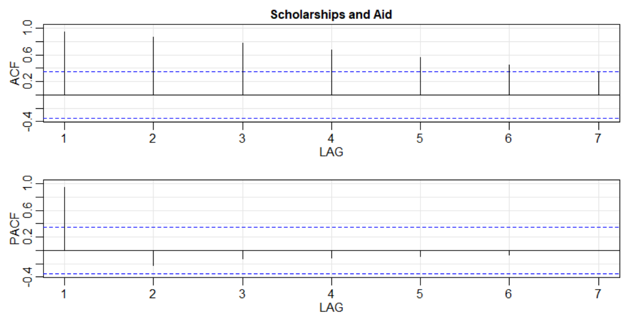




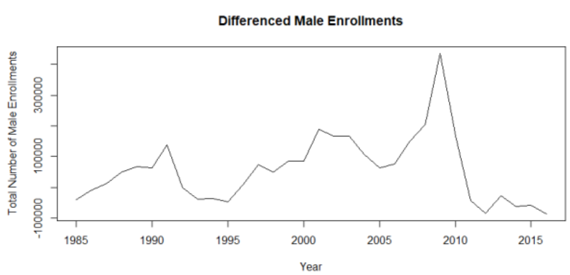


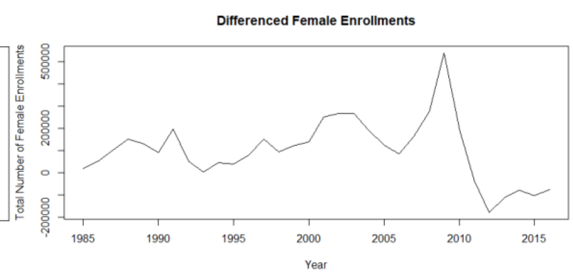


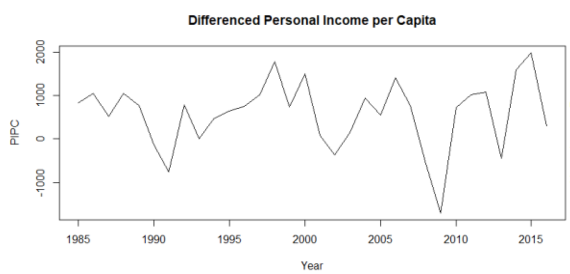


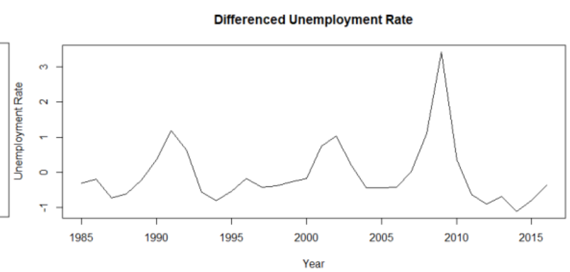


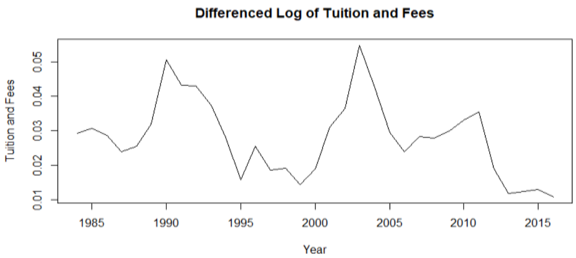
<u>Graphs of the Differenced Variables</u>

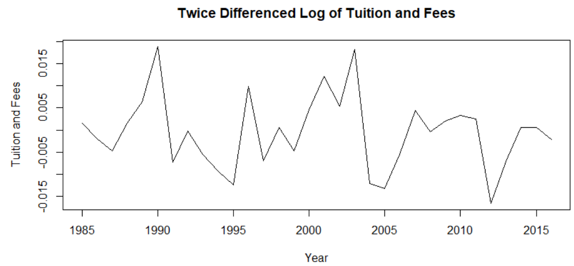


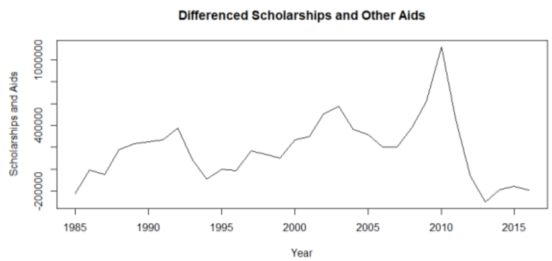







Differenced Log of Tuition and Fees
Tuition and Fees
0.01
0.02
0.03
0.04
0.05
1985
1990
1995
2000
2005
2010
2015
Year


Twice Differenced Log of Tuition and Fees
Tuition and Fees
-0.015
-0.005
0.005
0.015
1985
1990
1995
2000
2005
2010
2015
Year


Differenced Scholarships and Other Aids
Scholarships and Aids
-200000
400000
1000000
1985
1990
1995
2000
2005
2010
2015
Year

Appendix B (Engle – Granger Test Results)

Engle – Granger Two – Step Cointegration Test

Step One: Regression Equations

$$Male = -1{,}288{,}470.0 + 135.1 * pipc + \varepsilon_t$$
$$Female = -4{,}418{,}902.9 + 225.6 * pipc + \varepsilon_t$$
$$Male = 3{,}219{,}389 + 232{,}396 * unemp + \varepsilon_t$$
$$Female = 3{,}701{,}397 + 280{,}243 * unemp + \varepsilon_t$$
$$Male = -5{,}073{,}892 + 1{,}162{,}836 * \log(tuitfee) + \varepsilon_t$$
$$Female = -10{,}598{,}058 + 1{,}924{,}747 * \log(tuitfee) + \varepsilon_t$$
$$Male = 494{,}006.2621 + 0.3318 * scholarships + \varepsilon_t$$
$$Female = -1{,}122{,}687.9384 + 0.5277 * scholarships + \varepsilon_t$$

Step Two: Dickey – Fuller Test Results

	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value	P – Value	Reject/Fail to Reject
Personal Income per Capita (Males)	-1.645	-3.58	-2.93	-2.60	0.1104	Fail to Reject
Personal Income per Capita (Females)	-1.5054	-3.58	-2.93	-2.60	0.1427	Fail to Reject
Unemployment (Males)	-0.243	-3.58	-2.93	-2.60	0.8096	Fail to Reject
Unemployment (Females)	-1.165	-3.58	-2.93	-2.60	0.2532	Fail to Reject
Tuition and Fees (Males)	-1.7581	-3.58	-2.93	-2.60	0.08893	Fail to Reject
Tuition and Fees (Females)	-1.2043	-3.58	-2.93	-2.60	0.2379	Fail to Reject
Scholarships and Aid (Males)	-2.292	-3.58	-2.93	-2.60	0.0291	Reject
Scholarships and Aid (Females)	-1.8801	-3.58	-2.93	-2.60	0.06984	Fail to Reject

Appendix C (SUR Results and Residuals)

SUR Output

Initial Model (All Variables)

```
systemfit results
method: SUR

        N DF          SSR            detRCov  OLS-R2 McElroy-R2
system 64 54 237788881575 5994356455023388672 0.76003   0.669073

    N DF          SSR        MSE    RMSE       R2   Adj R2
m1 32 27  74634475297 2764239826 52576.0 0.800880 0.771380
f1 32 27 163154406278 6042755788 77735.2 0.735177 0.695944

The covariance matrix of the residuals used for estimation
           m1         f1
m1 2764239826 3272502063
f1 3272502063 6042755788

The covariance matrix of the residuals
           m1         f1
m1 2764239826 3272502063
f1 3272502063 6042755788

The correlations of the residuals
         m1       f1
m1 1.000000 0.800709
f1 0.800709 1.000000


SUR estimates for 'm1' (equation 1)
Model Formula: maled ~ pipcd + unempd + ltuitfeedd + scholarshipsd

                    Estimate      Std. Error t value   Pr(>|t|)
(Intercept)     39821.7196779   14769.9714822 2.69613 0.01192975 *
pipcd               2.5574094      19.0280317 0.13440 0.89408142
unempd          85680.2391329   20449.1334801 4.18992 0.00026732 ***
ltuitfeedd     165418.3622748 1234992.5878451 0.13394 0.89444122
scholarshipsd       0.1194021       0.0415413 2.87430 0.00780106 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 52576.038514 on 27 degrees of freedom
Number of observations: 32 Degrees of Freedom: 27
SSR: 74634475297.1815 MSE: 2764239825.82154 Root MSE: 52576.038514
Multiple R-Squared: 0.80088 Adjusted R-Squared: 0.77138


SUR estimates for 'f1' (equation 2)
Model Formula: femaled ~ pipcd + unempd + ltuitfeedd + scholarshipsd

                   Estimate     Std. Error  t value  Pr(>|t|)
(Intercept)     78743.239247   21837.821832  3.60582 0.0012431 **
pipcd              -1.252226      28.133485 -0.04451 0.9648253
unempd          97317.151811   30234.623953  3.21873 0.0033390 **
ltuitfeedd     594151.313398 1825971.575501  0.32539 0.7473948
scholarshipsd       0.159764       0.061420  2.60117 0.0148930 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 77735.164424 on 27 degrees of freedom
Number of observations: 32 Degrees of Freedom: 27
SSR: 163154406277.992 MSE: 6042755788.07379 Root MSE: 77735.164424
Multiple R-Squared: 0.735177 Adjusted R-Squared: 0.695944
```

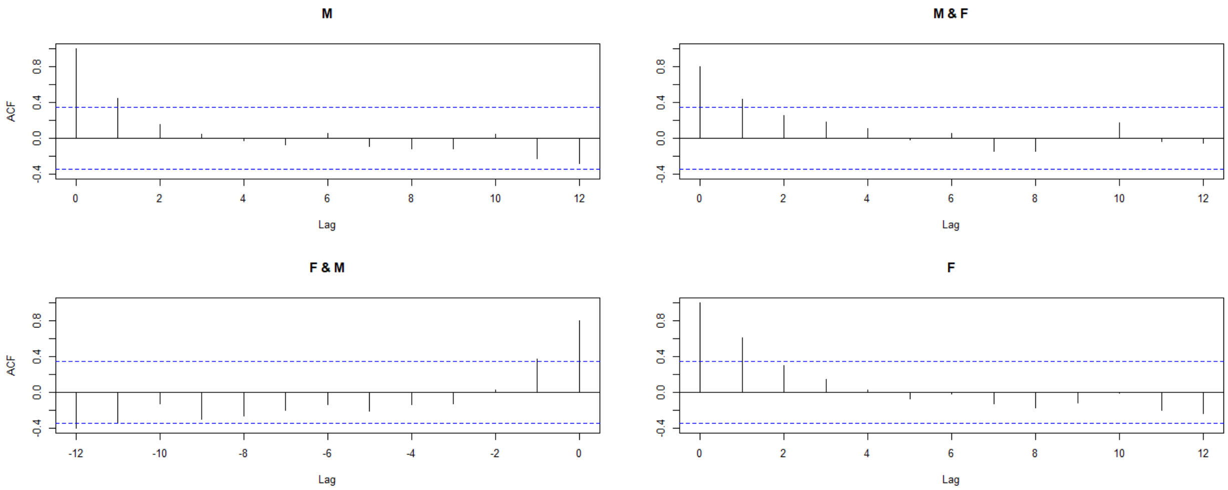
Intermediate Model (Log of Tuition and Fees Omitted for Males, Personal Income Per Capita Omitted for Females)

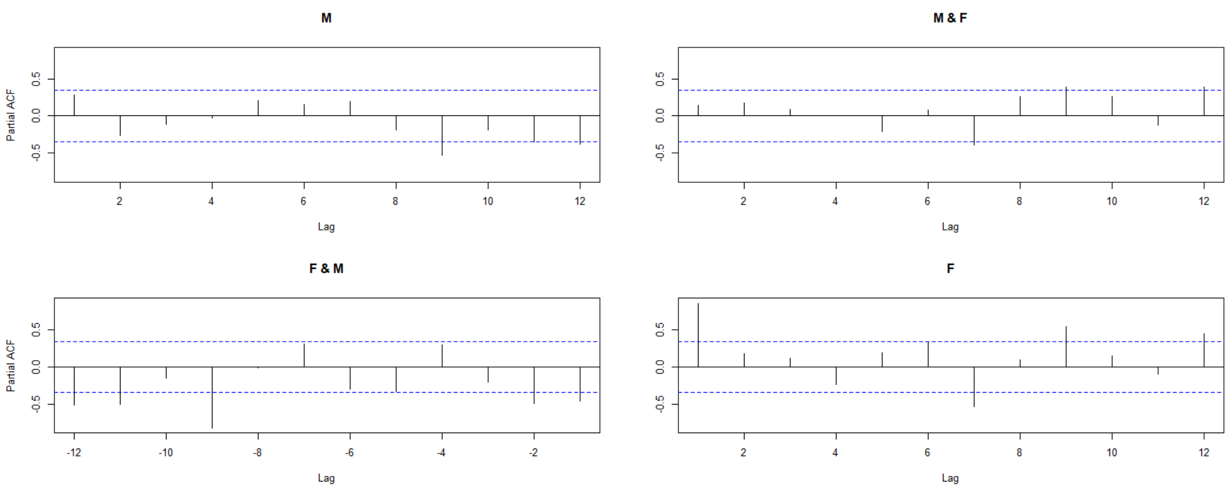
```
systemfit results
method: SUR

        N DF          SSR             detRCov   OLS-R2 McElroy-R2
system 64 56 237915895217 5577695341667731456 0.759902   0.668884

    N DF          SSR        MSE    RMSE       R2   Adj R2
m2 32 28  74685409509 2667336054 51646.3 0.800744 0.779395
f2 32 28 163230485708 5829660204 76352.2 0.735054 0.706667

The covariance matrix of the residuals used for estimation
           m2         f2
m2 2667288129 3155765744
f2 3155765744 5827370640

The covariance matrix of the residuals
           m2         f2
m2 2667336054 3157842222
f2 3157842222 5829660204

The correlations of the residuals
         m2       f2
m2 1.000000 0.800811
f2 0.800811 1.000000


SUR estimates for 'm2' (equation 1)
Model Formula: maled ~ pipcd + unempd + scholarshipsd

                   Estimate    Std. Error t value     Pr(>|t|)
(Intercept)   39229.9870619 12785.3958403 3.06834    0.0047403 **
pipcd             3.3776689    11.1514907 0.30289    0.7642130
unempd        86651.5976209 15736.0852691 5.50655 0.0000069589 ***
scholarshipsd     0.1198927     0.0395037 3.03498    0.0051508 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 51646.258857 on 28 degrees of freedom
Number of observations: 32 Degrees of Freedom: 28
SSR: 74685409509.0065 MSE: 2667336053.89309 Root MSE: 51646.258857
Multiple R-Squared: 0.800744 Adjusted R-Squared: 0.779395


SUR estimates for 'f2' (equation 2)
Model Formula: femaled ~ unempd + ltuitfeedd + scholarshipsd

                     Estimate      Std. Error t value    Pr(>|t|)
(Intercept)     77921.2443577   17677.4907689 4.40794  0.00013958 ***
unempd          98693.5295733   19435.8655743 5.07791 0.000022413 ***
ltuitfeedd     395523.0677531 1069805.3481673 0.36971  0.71437652
scholarshipsd       0.1601953       0.0579159 2.76600  0.00993463 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 76352.211519 on 28 degrees of freedom
Number of observations: 32 Degrees of Freedom: 28
SSR: 163230485708.048 MSE: 5829660203.85885 Root MSE: 76352.211519
Multiple R-Squared: 0.735054 Adjusted R-Squared: 0.706667
```

Final Model (Personal Income per Capita and Log of Tuition and Fees Omitted)

```
systemfit results
method: SUR

        N DF          SSR             detRCov   OLS-R2 McElroy-R2
system 64 58 238547418880 5237448021568674816 0.759264   0.668267

    N DF          SSR        MSE    RMSE       R2   Adj R2
m3 32 29  74752880129 2577685522 50770.9 0.800564 0.786810
f3 32 29 163794538751 5648087543 75153.8 0.734138 0.715803

The covariance matrix of the residuals used for estimation
           m3         f3
m3 2577685522 3053120611
f3 3053120611 5648087543

The covariance matrix of the residuals
           m3         f3
m3 2577685522 3053120611
f3 3053120611 5648087543

The correlations of the residuals
         m3       f3
m3 1.000000 0.800163
f3 0.800163 1.000000


SUR estimates for 'm3' (equation 1)
Model Formula: maled ~ unempd + scholarshipsd

                   Estimate    Std. Error t value      Pr(>|t|)
(Intercept)   40638.5355159 11707.9258260 3.47103     0.0016449 **
unempd        83997.7942077 12849.8506036 6.53687 0.00000036937 ***
scholarshipsd     0.1218895     0.0382899 3.18333     0.0034626 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 50770.912161 on 29 degrees of freedom
Number of observations: 32 Degrees of Freedom: 29
SSR: 74752880129.4501 MSE: 2577685521.70518 Root MSE: 50770.912161
Multiple R-Squared: 0.800564 Adjusted R-Squared: 0.78681


SUR estimates for 'f3' (equation 2)
Model Formula: femaled ~ unempd + scholarshipsd

                   Estimate    Std. Error t value    Pr(>|t|)
(Intercept)   77324.2724275 17330.6840955 4.46170  0.00011287 ***
unempd        99475.1343452 19021.0208705 5.22975 0.000013394 ***
scholarshipsd     0.1625278     0.0566787 2.86753  0.00763154 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 75153.759341 on 29 degrees of freedom
Number of observations: 32 Degrees of Freedom: 29
SSR: 163794538751.039 MSE: 5648087543.13928 Root MSE: 75153.759341
Multiple R-Squared: 0.734138 Adjusted R-Squared: 0.715803
```

Autocorrelation and Partial Autocorrelation Functions of the SUR Model Residuals





Bibliography

Betts, J., & McFarland, L. (1995). Safe Port in a Storm: The Impact of Labor Market Conditions on Community College Enrollments. *The Journal of Human Resources, 30*(4), 741-765. doi:10.2307/146230

Braunstein, A., Mcgrath, M., & Pescatrice D. (1999). Measuring the Impact of Income and Financial Aid Offers on College Enrollment Decisions. *Research in Higher Education*, *40*(3), pp 247–259.

Collegeboard. (2018). *Average Aid per Student over Time: All Postsecondary Students, Undergraduate Students, and Graduate Students* [Data File]. Retrieved from https://trends.collegeboard.org/student-aid/figures-tables/average-aid-student-over-time-postsecondary-undergraduate-graduate

Collegeboard. (2019). *Tuition and Fees and Room and Board over Time* [Data File]. Retrieved from https://trends.collegeboard.org/college-pricing/figures-tables/tuition-fees-room-and-board-over-time

Digest of Education Statistics. (1990-2017). *Total Fall Enrollment in Degree-Granting Postsecondary Institutions, by Attendance Status, Sex, and State or Jurisdiction* [Table]. Retrieved from https://nces.ed.gov/pubsearch/getpubcats.asp?sid=091#061

Digest of Education Statistics. (1987-1989). *Total Fall Enrollment in Degree-Granting Postsecondary Institutions, by Attendance Status, Sex, and State or Jurisdiction* [Table]. Retrieved from https://catalog.hathitrust.org/api/volumes/oclc/3133477.html

Hillman, N.W., & Orians, E.L. (2013). Community Colleges and Labor Market Conditions: How Does Enrollment Demand Change Relative to Local Unemployment Rates. *Research in Higher Education*, *54*(7), pp. 765-780.

Kondo, A. (2015). Differential Effects of Graduating During a Recession Across Gender and Race. *IZA Journal of Labor Economics*, *4*(23).

U.S. Bureau of Economic Analysis, Personal income per capita [A792RC0A052NBEA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/A792RC0A052NBEA, April 26, 2019.

U.S. Bureau of Labor Statistics, Unemployment Rate: 20 years and over [LNS14000024], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/LNS14000024, April 26, 2019.

US Inflation Calculator. (2019). *Consumer Price Index Data from 1913 to 2019* [Table]. Retrieved from https://www.usinflationcalculator.com/inflation/consumer-price-index-and-annual-percent-changes-from-1913-to-2008/

Wetzel, J., Dennis, O., & Peterson, S. (1998). An Analysis of Student Enrollment Demand. *Economics of Education Review*, *17*(1), pp. 47-54.