The University of Illinois at Chicago

Economics 346: Econometrics

The Saving-Income Relationship

We will download data from the most recent *Economic Report of the President* on Personal Income and Saving to see if the relationship between these 2 variables has changed in response to the Reagan revolution.

To Download Data and Import Data into Excel

1. Click on Netscape, then open http://www.access.gpo.gov/eop (or, alternatively, try http://www.whitehouse.gov/WH/EOP/CEA/html/CEA.htm or http://www.access.gpo.gov/su_docs/budget/index.html)
2. Follow instructions to *Economic Report of the President* Statistical Tables in Appendix B. We want Table B-30 "Disposition of Personal Income". Do not take the PDF version.
3. Click on file/Save As and save the Table B-30 on your disk in the A drive. **Save it as a text (.txt) file.** The computer will try to name it waisgate.cgi, so edit the name to waisgate.txt (or bill.txt or whatever you want)
4. Exit Netscape.
5. Open Excel, and open your file in Excel. When the conversion box appears, use the arrow keys to move down past the headings to the data. Choose the Fixed Width option, click on Next, click between series to create separate columns in Excel.
6. When the Excel file is ready, save it as an Excel file (change the file type at the bottom of the Save As box). If you are likely to use your data in another program later, save it as Excel 5, not Excel 97.
7. The ERP tables have blank rows for easier reading. Edit out the blank rows in the annual data (Edit/Delete/Entire Row/OK).
8. The 1998 line will likely be out of alignment (some digits or numbers in the wrong columns). Fix these by replacing the cell or editing a cell's contents in the formula bar (by the = sign at the top of the screen).

First Regressions: Has there been a significant change in the savings-income relationship since 1981?

1. Plot your data. Plot personal saving on the Y-axis and disposable personal income on the X -axis from 1959 to 1998, annual data. Probably, disposable income will be in column E and personal saving in column J. It is convenient to put the income and saving data in adjacent columns, such as income in O and savings in P.
2. Regression A: Whole period from 1959-1998. Regress saving on income. Call the residual sum of squares from this regression RSS1, the number of observations n1, and the degrees of freedom left in the residuals df1.

Calculate the residuals, residual plots, standardized residuals, and line fit plots.

3. Regression B: First half. Regress saving on income from 1959-1980. Call the residual sum of squares RSS2, the number of observations n2, and the degrees of freedom left in the residuals df2.
4. Regression C: Second half. Regress saving on income from 1981-19980. Call the residual sum of squares RSS3, the number of observations n3, and the degrees of freedom left in the residuals df3.

T-Tests and F-Tests

1. Using Regression B standard error of the slope coefficient, construct a t-test of the null hypothesis that the 2 regressions have the same slope. Another way to say this is, test the hypothesis that the slope coefficient in Regression B does not equal the estimated value for the slope coefficient in Regression C. (Hint: Your usual t-statistic tests the hypothesis that the slope coefficient does not equal 0, so just substitute the estimated value for Regression C for 0 in the formula.)
    a. What is your null hypothesis? What is your alternative hypothesis?
    b. Is this a 1-tailed or 2-tailed test?
    c. Find the t critical value for 95% confidence from Function Wizard/TINV where x=.05, the number of tails is your answer in (b) and df=df2.
    d. Do you reject, or fail to reject the null hypothesis?
2. Now look at Regression C. The R-Square is small. Test the hypothesis that this R-Square is equal to zero. Using a 95% confidence level, do we reject or fail to reject?
3. F tests are versatile. We can use a type of F-test to see if the slopes of 2 regressions are statistically different for the 2 periods. This F test is called a Chow test.
    a. Add RSS2 +RSS3 to get RSS4, with degrees of freedom df2+df3.
    b. Subtract RSS4 from RSS1 to get RSS5.
    c. Your F statistic is (RSS5/(k+1))/[RSS4/(n2+n3-2k-2)]
    d. Find the significance level of your F with k+1 degrees of freedom in the numerator and (n2+n3-2k-2) degrees of freedom in the denominator. (In Excel, use the Function Wizard/FDIST. Your x is the number you have calculated).
    e. Test the null hypothesis that there is no change, or that the marginal propensity to save out of income was the same from 1959-1980 and 1981-1998 with 99% confidence. (To find the F critical value, use Function Wizard/FINV with x=.01 and the same numerator and denominator degrees of freedom as in (d).)
    f. Do you reject, or fail to reject?
    g. In words, what is the conclusion? Did the relationship between saving and disposable income change in 1981?

An Alternative Story: Different Functional Forms

Perhaps the relationship between saving and income is not linear. Then a linear regression is not appropriate. We will transform the data and check other common functional relationships.

1. **Double-log Form: ln(Personal Saving) = a + b[ln(Personal disposable income)]**
2. **Semi-Log Form 1: Personal Saving = a + b[ln(Personal disposable income)]**
3. **Semi-Log Form 2: ln(Personal Saving) = a + b(Personal disposable income)**
4. **Polynomial Form: Personal Saving = a + b1(Personal disposable income) + b2(Squared Personal disposable income)**
5. **Inverse Form: Personal Saving = a + b[1/(Personal disposable income)]**

You are assigned 1 of the above functional relationships. Everyone will run at least one of the above relationships, then transform the results so they are comparable.

1. Transform your data and run the regression. To take the natural log of a number in Excel, use the Function Wizard/ln or type in the command. If you savings data are in column P, starting row 22, and you want the log of savings in column Q, select Q22 and type =LN(p22), since (LN) is the usual sign for natural log.

2. Run your assigned regression over the whole period, from 1959-1998. Save the residuals, standardized residuals, line fit plots, and residual plots.

3. You can't directly compare the R-squared values if the dependent variables are different (for example, savings and log of savings), but you can compare adjusted R-squared values if the X-variables are different. So we need comparable R-squares, called quasi-R-squares.

   a. If your dependent variable is in level ($) form, not logged, your quasi-R-square will be from predicted saving transformed to log form, which we do by taking the natural log. The Excel command is =LN(a number or cell address).

   b. If your dependent variable is in log form, you quasi-R-square will be from predicted saving transformed back to level ($) form, which we do by taking the anti-log. The Excel command is =EXP(a number or cell address).

4. **Transform** your predicted saving and (if necessary) the personal saving data, and subtract the transformed saving from the (possibly transformed) actual data to get transformed residuals. For example, if your assigned regression was double-log, your predicted personal saving and residuals are also in log form. To compare your regression's R-square with the

original from Regression A, we must transform your predictions and residuals back to dollars from logs. Take the anti-log (EXP) of each of your predictions of saving in log form to get predicted saving in dollars for each year. Subtract each prediction from the personal saving data to get new residuals. If your regression was in level form (dollars) then take the log (LN) of each prediction and of each actual observation, and subtracted the logs of the predictions from the logs of the actuals to get residuals translated into log form. NOTE FOR INVERSE FUNCTION: Some predicted values are negative numbers. You cannot log negative numbers. Shift actuals and predicteds up (add a constant to all values) until all are nonzero.

5. Calculate quasi-R-square for your functional form

a. Create new Total Sum of Squares (TSS)
    1. Find the mean of your transformed saving series.
    2. Subtract the mean from each transformed point.
    3. Square these deviations from the mean and sum them up.

b. Create new Residual Sum of Squares (RSS)
    1. Square the transformed residuals.
    2. Sum them up.

c. The quasi-R-square is 1 - (RSS/TSS)

The University of Illinois at Chicago

Economics 346: Econometrics

The Saving-Income Relationship

	Functional Form	Name
1.	Log-Log	_____
2.	Semi-log 1	_____
3.	Semi-log 2	_____
4.	Polynomial	_____
5.	Inverse	_____

Compare R-squares and quasi-R-squares of different functional forms.

Number	Functional Form	R-square or quasi-R-square: **dollars**	R-square or quasi-R-square: **logs**
1.	Log-log		
2.	Semi-log 1		
3.	Semi-log 2		
4.	Polynomial		
5.	Inverse		

The above functional forms are nonlinear alternatives for the saving-disposable income relationship. Which has the highest R-square measured in dollars?

_____

Which has the highest R-square or quasi-R-square measured in logs?

_____

Compare the line fit plots and residual plots.

Number	Functional Form	Line fit plots: Do the functions match? If not, describe how.	Residual plots: Do the assumptions hold? If not, which are broken?
1.	Log-log		
2.	Semi-log 1		
3.	Semi-log 2		
4.	Polynomial		
5.	Inverse		

Of these 5 forms, which do you prefer and why?