Instructions:

a. You have 3 hours to finish your exam. Write your name and student ID number on the upper left corner of this page, as well as the names of the students seated next to you. And, in case you wish to do so, please sign the Buckley Waiver.

STUDENT CONSENT FOR RELEASE OF STUDENT INFORMATION
(Buckley Waiver)

I hereby authorize the UCSD Economics Department to return my graded final examination/research paper by placing it in a location accessible to all students in the course. I understand that the return of my examination/research paper as described above may result in disclosure of personally identifiable information, that is not public information as defined in UCSD PPM 160-2, and I hereby consent to the disclosure of such information.

Signature

b. Confirm that your test has 13 pages. Make sure you find a sheet with two tables of critical values, in the pages of your exam. The last page can be used as scratch paper.

c. There are two parts to this exam – multiple-choice questions (Part I) and longer questions (Part II). You do not need to justify your answers for the multiple-choice questions. Show ALL your work for the longer questions.

d. Use a pen to write your answers. You give up your right to a regrade if you use a pencil.

e. The table below indicates how points will be allocated on the exam. You can answer the questions in any order you like. Use your time carefully and efficiently.

<table>
<thead>
<tr>
<th>Question</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part I</td>
<td>24</td>
</tr>
<tr>
<td>Part II</td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>8</td>
</tr>
<tr>
<td>2.</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Exam Total</td>
<td>100</td>
</tr>
</tbody>
</table>

f. With you, you should only have a pen, and a basic calculator.

g. You will NOT be allowed to leave the room during the exam. Turn off your cell phone and your IPod, and good luck!
PART I: Multiple-Choice Questions (1.5 points each, 24 points total)

1) You have estimated the following equation: 
   \[ \text{TestScore} = 607.3 + 3.85 \text{Income} - 0.0423 \text{Income}^2 \] 
   where TestScore is the average of the reading and math scores on the Stanford 9 standardized test administered to 5th grade students in 420 California school districts in 1998. Income is the average annual per capita income in the school district, measured in thousands of 1998 dollars. The equation
   a. suggests a steeper effect of income on test scores at low values of income than at higher values of income.
   b. shows a negative relationship until a certain value of income and then a positive relationship after.
   c. does not make much sense since the square of income is entered.
   d. suggests a positive relationship between test scores and income for all the sample.

2) In a model with a binary dependent variable, a predicted value of 0.6 means that
   a. the most likely value the dependent variable will take is 60 percent.
   b. given the values for the explanatory variables, there is a 60 percent probability that the dependent variable will equal one.
   c. the model makes little sense, since the dependent variable can only be 0 or 1.
   d. given the values for the explanatory variables, there is a 40 percent probability that the dependent variable will equal one.

3) For the polynomial regression model,
   a. you need new estimation techniques since the OLS assumptions do not apply any longer.
   b. the techniques for estimation and inference developed for multiple regression can be applied.
   c. you can still use OLS estimation techniques, but the t-statistics do not have an asymptotic normal distribution.
   d. the critical values from the normal distribution have to be changed to 1.96^2, 1.96^3, etc.

4) Having more relevant instruments
   a. is a problem because instead of being just identified, the regression now becomes overidentified.
   b. is like having a larger sample size in that the more information is available for use in the IV regressions.
   c. typically results in larger standard errors for the TSLS estimator.
   d. is not as important for inference as having the same number of endogenous variables as instruments.

5) The logit marginal effect of increasing the value of an explanatory variable
   a. is constant for all values of the explanatory variables.
   b. depends on the standard normal probability density function.
   c. can potentially take a value outside the range 0 to 1.
   d. none of the above.

6) In the regression model \[ Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_i + \beta_3 (X_i \times D_i) + u_i \] where X is a continuous variable and D is a binary variable, to test that the two regressions (one for D_i = 0 and the other for D_i = 1) are identical, you must use the
   a. t-statistic separately for \( \beta_0 = 0 \), \( \beta_1 = 0 \).
   b. F-statistic for the joint hypothesis that \( \beta_0 = 0 \), \( \beta_1 = 0 \).
   c. t-statistic separately for \( \beta_2 = 0 \).
   d. F-statistic for the joint hypothesis that \( \beta_2 = 0 \), \( \beta_3 = 0 \).

7) The following tools from multiple regression analysis carry over in a meaningful manner to the linear probability model, with the exception of the
   a. F-statistic.
   b. significance test using the t-statistic.
   c. 95% confidence interval using \( \pm 1.96 \) times the standard error.
   d. regression R^2.
8) In the fixed effects regression model, using \((n - 1)\) binary variables for the entities (the dummy for the first entity is left out), the coefficient of the binary variable
   a. indicates the level of the fixed effect of the \(i^{th}\) entity.
   b. will be either 0 or 1.
   c. indicates the difference in fixed effects between the \(i^{th}\) and the first entity.
   d. indicates the response in the dependent variable to a percentage change in the binary variable.

9) Negative first autocorrelation in the change of a variable implies that
   a. the variable contains only negative values.
   b. the series is not stable.
   c. an increase in the variable in one period is, on average, associated with a decrease in the next.
   d. the data is negatively trended.

10) With panel data, regression software typically uses an “entity-demeaned” algorithm because
    a. the OLS formula for the slope in the linear regression model contains deviations from means already.
    b. there are typically too many time periods for the regression package to handle.
    c. the number of estimates to calculate can become extremely large when there are a large number of entities.
    d. deviations from means sum up to zero.

11) Weak instruments are a problem because
    a. the TSLS estimator may not be normally distributed, even in large samples.
    b. they result in the instruments not being exogenous.
    c. the TSLS estimator cannot be computed.
    d. you cannot predict the endogenous variables any longer in the first stage.

12) The forecast is
    a. made for some date beyond the data set used to estimate the regression.
    b. another word for the OLS predicted value.
    c. equal to the residual plus the OLS predicted value.
    d. close to 1.96 times the standard deviation of \(Y\) in the sample.

13) Time fixed effects regression are useful in dealing with omitted variables
    a. even if you only have a cross-section of data available.
    b. if these omitted variables are constant over time but vary across entities.
    c. when there are more than 100 observations.
    d. if these omitted variables are constant across entities but not over time.

14) The distributed lag model assumptions include all of the following with the exception of
    a. there is no perfect multicollinearity.
    b. \(X_i\) is strictly exogenous.
    c. \(E(u_i \mid X_{i1}, X_{i2}, \ldots) = 0\).
    d. The random variables \(X_i\) and \(Y_i\) have a stationary distribution.

15) In the case of exact identification
    a. you can use the \(J\)-statistic in a test of overidentifying restrictions.
    b. you cannot use TSLS for estimation purposes.
    c. you must rely on your personal knowledge of the empirical problem at hand to assess whether the instruments are exogenous.
    d. OLS and TSLS yield the same estimate.

16) Stationarity means that the
    a. error terms are not correlated.
    b. probability distribution of the time series variable does not change over time.
    c. random variables become independently distributed when the time separating them gets larger.
    d. the error term has conditional mean zero, given all the regressors and additional lags of the regressors beyond the lags included in the regression.
PART II: Problems (76 points total)
Note: Assume through the end of the exam, that the significance level for any hypothesis testing is 5%.

1) (8 points) A recent article studied the effects of attending a Catholic High School on the probability of attending college. Let \textit{college} be a binary variable that takes a value of one if a student attends college, and zero otherwise. Let \textit{CathHS} be a binary variable equal to one if the student attends a Catholic high school, and zero otherwise. A linear probability model is

\[
\text{college} = \beta_0 + \beta_1 \text{CathHS} + \text{other factors} + u,
\]

where the other factors include gender, race, family income and parents’ education.

a. Why might \textit{CathHS} be correlated with \(u\)? (Hint: think of factors that are most likely included in the error term).

b. Let \textit{CathRel} be a binary variable that takes the value one if the student is a Catholic. Discuss the two requirements needed for this to be a valid instrument for \textit{CathHS} in the preceding equation. Which of the two can be tested, and how?
2) (15 points) Are rents influenced by student population in a college town? Let *rent* be the average monthly rent paid on rental units in a college town in the United States. Let *pop* denote the total city population, *avginc* the average city income, and *pctstu* the student population as a percentage of the total population. We have data for the years 1980 and 1990 for 64 cities in the US.

a. You decide the pool all the observations for the two years, and you estimate the following model which includes a dummy variable D90, indicating the year (=1 if year = 1990 and zero otherwise):

\[
\log(rent) = -0.569 + 0.262 D90 + 0.041 \log(pop) + 0.571 \log(avginc) + 0.0050 \text{pctstu}
\]

\[
(0.851)\quad (0.058)\quad (0.022)\quad (0.098)\quad (0.0011)
\]

\[
n = 128\quad R^2 = 0.861\quad SER = 0.126
\]

What does the estimate on the year dummy variable D90 tell you? Interpret the estimate on *pctstu*.

b. You think more in depth about the problem and conclude that there are unobserved variables that determine the rental rates in city *i*, but that are constant over the 10-year period (city fixed effects). If in fact that is the case, and if some of those unobserved variables are correlated with *pctstu*, can you rely on the estimate you obtained in a.? Explain.
c. To take into account the city fixed effects you decide to estimate the model using changes (difference between the variable in 1990 and the variable in 1980). The results of that estimation are as follows:
\[
\Delta \log(\text{rent}_a) = 0.386 + 0.072 \Delta \log(\text{pop}_a) + 0.310 \Delta \log(\text{avginc}_a) + 0.0112 \Delta \text{pctstu}_a
\]

\[
(0.049) \quad (0.070) \quad (0.089) \quad (0.0029)
\]

\[n = 64 \quad R^2 = 0.322 \quad SER = 0.090\]

Compare your estimate on \textit{pctstu} with that from part a. Does the relative size of the student population appear to affect rental rates?

d. Finally compare the standard error of the estimator of \textit{pctstu} you obtained in c. with the one obtained in a. How do you explain the higher standard error of the regression in part c.?
3) (8 points) We have annual data from 1948 to 2003 on the three-month T-bill rate \((i_{3t})\), the inflation rate based on the consumer price index \((\text{inf}_t)\), and the federal budget deficit as a percentage of GDP \((\text{def}_t)\). The estimated equation is:

\[
i_{3t} = 1.73 + 0.606 \text{inf}_t + 0.513 \text{def}_t \quad n = 56 \quad R^2 = 0.602 \quad SER = 1.843
\]

(0.38) (0.093) (0.160)

Assume that stationarity is respected.

a. In order for us to say that the inflation rate or the deficit are exogenous, what do we need to make sure about the effect of lagged values of those variables?

b. We decide to add a single lag of the inflation rate and of the deficit to the equation. The results are:

\[
i_{3t} = 1.61 + 0.343 \text{inf}_t + 0.382 \text{inf}_{t-1} - 0.190 \text{def}_t + 0.569 \text{def}_{t-1} \quad n = 55 \quad R^2 = 0.685 \quad SER = 1.660
\]

(0.35) (0.130) (0.161) (0.236) (0.185)

What is the impact effect of inflation on the three-month T-bill rate? What is the long-run cumulative dynamic multiplier of the inflation rate?

c. The \(F\)-statistic for significance of \(\text{inf}_{t-1}\) and \(\text{def}_{t-1}\) is about 5.22. Are those variables jointly significant at the 5% significance level?
4) (15 points) You have data on wages for 852 working men, as well as information on other individual and family characteristics. You are interested in finding the returns from education (\( educ \)), but you know that an OLS estimation of \( \log(\text{wage}) \) on education alone suffers from omitted variable bias. You decide to try to do instrumental variables estimation to obtain a consistent estimate of the returns from education.

a. The variable \( brthord \) is birth order (\( brthord \) is one for a first-born child, two for a second-born child, and so on). Explain why \( educ \) and \( brthord \) might be negatively correlated.

b. You use two-stage least squares to estimate the returns from education. The results are as follows:
\[
\text{log}(\text{wage}) = 5.03 + 0.131 \text{ educ},
\]
\[
(0.42) \quad (0.031)
\]
Interpret the estimate of the coefficient on education.

c. Now, suppose you include number of siblings (\( sibs \)) as an explanatory variable in the wage equation; this controls for family background, to some extent:
\[
\text{log}(\text{wage}) = \beta_0 + \beta_1 \text{ educ}, + \beta_2 \text{ sibs}, + u,
\]
Suppose that you want to use the variable \( brthord \) as an IV for education, assuming that \( sibs \) is exogenous. Write down the first stage least squares equation, and tell how you would test the relevance of the instrument.

d. The results of the two-stage least squares regression are:
\[
\text{log}(\text{wage}) = 4.94 + 0.137 \text{ educ}, + 0.0021 \text{ sibs},
\]
\[
(1.08) \quad (0.077) \quad (0.0179)
\]
Comment on the standard errors of the estimator on education. (In particular, I want you to answer the questions: Is the standard error larger? Why?)
5) (12 points) We have a panel data on school districts in Michigan for the years 1992 through 1998. The dependent variable in this question is \( \text{math4} \), the percentage of fourth graders in a district receiving a passing score on a standardized math test. The key explanatory variable is \( \text{rexpp} \), which is real expenditure per pupil in the district. The amounts are in 1997 dollars. The spending variable will appear in logarithm form.

a. We pool all the observations of all the years and regress the following model, using OLS:

|        | Coef. | Robust Std. Err. | t   | P>|t| | [95% Conf. Interval] |
|--------|-------|-------------------|-----|------|---------------------|
| \( \text{math4} \) |       |                   |     |      |                     |
| 1994   | 6.377355 | .7118323         | 8.96 | 0.000 | 4.981676 - 7.773034 |
| 1995   | 18.6502  | .7230283         | 25.77 | 0.000 | 17.231      - 20.0694 |
| 1996   | 18.03336 | .7697507         | 23.43 | 0.000 | 16.52412    - 19.5426 |
| 1997   | 15.34006 | .7934234         | 19.33 | 0.000 | 13.78441    - 16.89571 |
| 1998   | 30.39788 | .7713735         | 39.41 | 0.000 | 28.88546    - 31.91031 |
| \( \text{lrexpp} \) | .5339144 | .0967727         | 5.54 | 0.000 | 4.468277    - 5.53614 |
| \( \text{lrexpp}_1 \) | .061527 | .0967727         | 0.64 | 0.525 | -.1282279   - .2512818 |
| \( \text{lenrol} \) | .2450874 | .0967727         | 2.01 | 0.044 | .0151887    - 1.170135 |
| \( \text{lunch} \) | .5926719 | .2945313         | 2.01 | 0.044 | .0151887    - 1.170135 |
| \( \text{cons} \) | -31.66156| 12.35386         | -2.56 | 0.010 | -55.88358   - 7.439533 |

where \( \text{lrexpp} \) is the log of the first lag of real expenditure per pupil, \( \text{lenrol} \) is the log of total district enrollment and \( \text{lunch} \) is the percentage of students in the district eligible for the school lunch program (\( \text{lunch} \) is a pretty good measure of the district-wide poverty rate). The first available year (the base year) is 1993.

Is the sign of the \( \text{lunch} \) coefficient what you expected? Interpret the magnitude of the coefficient. Would you say that the district poverty has a big effect on test pass rates?

b. We re-estimate the model, assuming that there are district fixed effects. The results are as follows:

|        | Coef. | Robust Std. Err. | t   | P>|t| | [95% Conf. Interval] |
|--------|-------|-------------------|-----|------|---------------------|
| \( \text{math4} \) |       |                   |     |      |                     |
| 1994   | 6.177316 | .6102614         | 10.12 | 0.000 | 4.980697    - 7.373935 |
| 1995   | 18.09267 | .6616857         | 21.00 | 0.000 | 16.40305    - 19.78229 |
| 1996   | 17.9404  | 1.0132777        | 17.71 | 0.000 | 15.95353    - 19.92786 |
| 1997   | 15.19184 | 1.0929922        | 13.90 | 0.000 | 13.04867    - 15.33501 |
| 1998   | 29.88319 | 1.195098         | 25.00 | 0.000 | 27.53981    - 32.22657 |
| \( \text{lrexpp} \) | -.4111804 | 3.873469        | -0.11 | 0.915 | -.8.006393 - 7.184033 |
| \( \text{lrexpp}_1 \) | 7.002988 | 3.480338        | 2.01 | 0.044 | .177461     - 13.82852 |
| \( \text{lenrol} \) | .2450874 | 1.205092        | 0.20 | 0.839 | -2.117903   - 2.608078 |
| \( \text{lunch} \) | -.061527 | .0967727        | 0.64 | 0.526 | -.128279    - .2512818 |
| \( \text{cons} \) | -16.08091| 35.73831         | -0.45 | 0.653 | -.86.15765   - 53.99583 |

\( \text{distid} \) absorbed (550 categories)
What happened to the estimated coefficient of the lagged spending variable. Is it still significant?

c. The F-statistic for the test of joint significance of the enrollment and lunch variables is 0.22 (p-value of 0.8027). Why do you think, in the fixed effects estimation, the enrollment and lunch program variables are jointly insignificant?

6) (12 points) You have time series of monthly data on the index of industrial production (IP) from January of 1947 to June of 1993. You compute annualized rates of monthly percentage changes of the index of industrial production and call it \(pcip\). (Note: \(pcip\) is in percentage points)

a. You estimate a AR(3) model for \(pcip\). You convince yourself that you have included the correct number of lags (in particular, when you add a fourth lag you verify that it is very insignificant). Here are the results of the AR(3) regression:

\[
\begin{align*}
pcip_t &= 1.80 + 0.349 \text{ipcip}_{t-1} + 0.071 \text{pcip}_{t-2} + 0.067 \text{pcip}_{t-3} \\
        &+ (0.062) (0.488) (0.415)
\end{align*}
\]

Forecast the value of \(pcip\) for July of 1993, using the following information for the values of the index of industrial production:

<table>
<thead>
<tr>
<th>Date</th>
<th>Jan93</th>
<th>Feb93</th>
<th>Mar93</th>
<th>Apr93</th>
<th>May93</th>
<th>Jun93</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP</td>
<td>109.3</td>
<td>109.9</td>
<td>110.1</td>
<td>110.4</td>
<td>110.3</td>
<td>110.1</td>
</tr>
</tbody>
</table>

Hint: \(pcip_t = 1200 \times [\ln(IP_t) - \ln(IP_{t-1})]\)

b. If much of the forecast error arises as a result of future error terms (and that source of error dominates the error resulting from estimating the unknown coefficients, since the number of observations is large), then what is your best guess of the RMSFE here?
c. You add three lags of the variable \( pcsp \) to your AR(3) model of part a. \( pcsp \) is the annualized rate of monthly percentage changes in the S&P500 index. How do you test whether \( pcsp \) Granger causes \( pcip \)? (State the hypothesis). If the F-stat of that test is 5.71, what is your conclusion?

7) (6 points) We are trying to explain the standardized score on a final exam (\( stndfnl \)) in terms of the percentage of classes attended (\( atndrte \)), prior college grade point average (\( priGPA \)), and ACT score (\( ACT \)). Using observations on a random sample of 680 students in microeconomics principles, we estimate the following model:

| Coef. | Std. Err. | t | P>|t| | [95% Conf. Interval] |
|-------|-----------|---|------|------------------|
| \( stndfnl \) | \( atndrte \) | \( -0.067129 \) | \( 0.0099434 \) | \( -0.68 \) | \( 0.500 \) | \( 0.262367 \) | \( 0.128109 \) |
| \( priGPA \) | \( -1.62854 \) | \( 0.5046929 \) | \( -3.23 \) | \( 0.001 \) | \( 2.619502 \) | \( 0.637578 \) |
| \( ACT \) | \( -1.280394 \) | \( 0.1038536 \) | \( -1.23 \) | \( 0.218 \) | \( -3.319554 \) | \( 0.758766 \) |
| \( priGPA\_sq \) | \( 0.2959046 \) | \( 0.1032631 \) | \( 2.87 \) | \( 0.004 \) | \( 0.093148 \) | \( 0.4986613 \) |
| \( ACT\_sq \) | \( 0.0045334 \) | \( 0.0022945 \) | \( 1.98 \) | \( 0.049 \) | \( 0.0000281 \) | \( 0.0090386 \) |
| \( GPA\_atndrte \) | \( 0.0055859 \) | \( 0.0040768 \) | \( 1.37 \) | \( 0.171 \) | \( -0.0024189 \) | \( 0.0135907 \) |
| \( stndfnl \) | \( 2.050293 \) | \( 1.396816 \) | \( 1.47 \) | \( 0.143 \) | \( -0.6923489 \) | \( 4.792935 \) |

where \( priGPA\_sq \) is the square of \( priGPA \), \( ACT\_sq \) is the square of \( ACT \), and \( GPA\_atndrte \) is an interaction term = \( priGPA \times atndrte \). A F-test on the coefficients of \( atndrte \) and \( GPA\_atndrte \) is 0.14 (so, the two variables are jointly significant).

a. What was the idea behind introducing the interaction coefficient?

b. The mean value for \( priGPA \) in this sample is 2.59. What is the estimated effect of 10 percentage point increase in attendance on the student’s performance in the final, when \( priGPA \) is at its mean. Explain in words what the estimated effect means.
Appendix: Tables and Formulas

Large Sample Critical Values for the t-statistic from the Standard Normal Distribution

<table>
<thead>
<tr>
<th>Significance Level</th>
<th>10%</th>
<th>5%</th>
<th>1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-Sided Test</td>
<td>1.64</td>
<td>1.96</td>
<td>2.58</td>
</tr>
<tr>
<td>1-Sided Test (&gt;</td>
<td>1.28</td>
<td>1.64</td>
<td>2.33</td>
</tr>
<tr>
<td>1-Sided Test (&lt;</td>
<td>-1.28</td>
<td>-1.64</td>
<td>-2.33</td>
</tr>
</tbody>
</table>

Large Sample Critical Values for the F-statistic from the $F_{q,\infty}$ Distribution

<table>
<thead>
<tr>
<th>Degrees of Freedom (q)</th>
<th>Significance Level</th>
<th>10%</th>
<th>5%</th>
<th>1%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2.71</td>
<td>3.84</td>
<td>6.63</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2.30</td>
<td>3.00</td>
<td>4.61</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2.08</td>
<td>2.60</td>
<td>3.78</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.94</td>
<td>2.37</td>
<td>3.32</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.85</td>
<td>2.21</td>
<td>3.02</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.77</td>
<td>2.10</td>
<td>2.80</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1.72</td>
<td>2.01</td>
<td>2.64</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.67</td>
<td>1.94</td>
<td>2.51</td>
<td></td>
</tr>
</tbody>
</table>

Some formulas:

- \( R^2 = \frac{ESS}{TSS} = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = 1 - \frac{SSR}{TSS} = 1 - \frac{\sum_{i=1}^{n}u_i^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \) and \( SER = s_u \), where \( s_u^2 = \frac{SSR}{n-k-1} \).

- \( \ln(x + \Delta x) - \ln(x) \equiv \frac{\Delta x}{x} \) when \( \frac{\Delta x}{x} \) is small.

- \( Y_i = \beta_0 + \beta_1X_i + u_i \) and \( Z_i \) is a valid instrument, then \( \hat{\beta}_{T\text{SLS}} = \frac{s_{ZY}}{s_{ZX}} \).

- \( j^{th} \) sample autocovariance = \( \text{cov}(Y_{t+j},Y_{t+j}) = \frac{1}{T} \sum_{t=j+1}^{T} (Y_t - \bar{Y}) (Y_{t+j} - \bar{Y}) \), where \( \bar{Y}_{t+j} \) denotes the sample average of \( Y_t \) computed over observations \( t = j+1, \ldots, T \).

- \( j^{th} \) sample autocorrelation = \( \hat{\rho}_j = \frac{\text{cov}(Y_{t+j},Y_{t+j})}{\text{Var}(Y_t)} \).

- \( \text{RMSFE} = \sqrt{\text{E}[(Y_{t+j} - \hat{Y}_{t+j})^2]} \)
SCRATCH PAPER: