Instructions: This exam is closed book/notes and has a total of 120 points. Point values for each question are listed to the right of each question number. This time, 424 and 524 students will answer the same 6 questions. Good luck.

1. (18 points) The following estimated equations represent the investment decisions of General Electric (GE) and Westinghouse (W), where \( \hat{I}_t \) is the predicted value of investment, \( V_t \) is the value of the firm, and \( K_t \) is the stock of capital. Absolute values of t-ratios are listed in parentheses below each parameter estimate. SSE stands for sum of squared errors, and there are 20 annual observations for each firm for a total of 40 observations.

Both: 
\[
\hat{I}_t = 17.8720 + 0.0152 V_t + 0.1436 K_t \\
SSE_{\text{both}} = 777.45 \\
(2.544) (2.452) (7.719) \\
N = 40
\]

It may be that investment strategies differ for GE and for W; the investment model was estimated separately for GE and W as follows. This time, standard errors are listed in parentheses below the coefficient estimates (this is just because of how they were published—don’t get caught up in this).

GE: 
\[
\hat{I}_t = -9.956 + 0.0265 V_t + 0.1517 K_t \\
SSE_{\text{GE}} = 777.45 \\
(31.37) (0.0156) (0.0257) \\
N = 20
\]

W: 
\[
\hat{I}_t = -0.509 + 0.0529 V_t + 0.0924 K_t \\
SSE_{\text{W}} = 104.31 \\
(8.02) (0.0157) (0.0561) \\
N = 20
\]

Given this information, how would you test for significant differences in investment behavior for GE and W? You need not actually calculate the numeric value of the test statistic, just explain what you would do and fill in the values that you can.

2. (32 points) Explain the difference between the following pairs of concepts:

a. perfect and imperfect multicollinearity
b. \( R^2 \) and Adjusted \( R^2 \)
c. weighted least squares estimator and ordinary least squares estimator

d. population and sample
3. (25 points) The following is an estimated regression of the logarithm of the price of wine on the age of the wine. The price data are annual averages per case for 1960-1971, 1973-1976, and 1978. (Note that the output is from a different software package. Some of the statistics are located in different places than in SAS.)

Dependent Variable: lnprice

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>.868465144</td>
<td>1</td>
<td>.868465144</td>
<td>F( 1, 13) = 5.46</td>
</tr>
<tr>
<td>Residual</td>
<td>2.06913692</td>
<td>13</td>
<td>.159164379</td>
<td>Prob &gt; F = 0.036</td>
</tr>
<tr>
<td>Total</td>
<td>2.93760207</td>
<td>14</td>
<td>.209828719</td>
<td>R-squared = 0.2956</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Adj R-squared = 0.2415</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Root MSE = .39895</td>
</tr>
</tbody>
</table>

Parameter Estimates

| Variable | Estimate | Standard Error | t     | P>|t| | [95% Conf. Interval] |
|----------|----------|----------------|-------|-------|----------------------|
| Intercept | 5.418722 | .3652171       | 14.84 | 0.000 | 4.629718 - 6.207725  |
| age      | -.0456397| .0195384       | -2.34 | 0.036 | -.0878499 - .0034295 |

a. Write down the estimated regression equation.
b. Graph the estimated equation. Be sure to label the axes and identify the slope and intercept on the graph.
c. Interpret the R-squared value.
d. Using the prob-value test for the t-ratio, does age significantly affect the log price of wine at the 1% level of significance?
e. Why do you think that the prob-value for the t-test is the same as the prob-value for the F-test in this example?
f. i. Circle the correct answer: Bad vintages are those with errors that are positive or negative
    ii. Explain your answer in part (i).

4. (24 points)
a. What does each letter in the acronym BLUE stand for?
   i. B–
   ii. L–
   iii. U–
   iv. E–
b. Define each term in parts (i)-(iv).
5. (5 points) Indicate the econometric problem apparent in the following graph. If there is no problem, state “No Problem.”

![Residuals Plotted Against Income](image)

6. (16 points) Suppose that you have data for 1000 workers on wages ($W_t$), grades of education completed ($ED_t$), and whether or not they earned a bachelor’s degree. Suppose that you specify the following model:

$$W_t = \beta_1 + \beta_2 ED_t + e_t$$

(1)

a. What is the interpretation of the parameter on $ED_t$, $\beta_2$, in equation (1)?

b. Studies show that wages increase with additional years of education. Some economists argue that there is a “sheepskin effect,” i.e., an additional premium for the receipt of a degree.

i. Define a variable that represents whether or not a person has a bachelor’s degree.

ii. Modify equation (1) to include this bachelor’s degree variable. What signs do you expect on the parameters of the model, assuming that the sheepskin effect is true.

iii. Based on the model in (ii), how would you test for the sheepskin effect, i.e., how would you test to see if there is an additional premium for having a bachelor’s degree