In a recent study, David Card examined the role of regulatory reform in the airline industry on the wages of airline workers. In his data, he has wage data and personal characteristics of workers inside and outside of the airline industry.

In one regression, Card regresses the log of hourly wages on an intercept and a dummy indicating whether the worker is in the airline industry. The coefficient on the airline dummy was .458 and the standard error on the coefficient was .004.

a. What is your interpretation of the coefficient .458? Be precise in explaining what it tells you about the effect of being in the airline on the wage (not log-wage) and remember that the dependent variable is measured as the log of the hourly wage.

b. Based on the information provided, is there a statistically significant difference between the log(hourly wage) between airline and other workers? What do you base your answer upon?

In a second regression, Card regresses the hourly wage on the airline dummy and adds controls for worker education and experience. The coefficient on the airline dummy drops from .458 to .310.

c. Given that both education and experience have positive effects on earnings, what could explain the drop in the coefficient on the airline dummy when these variables are added to the regression. Be sure to explain why the coefficient would drop and just why it would change.

Suppose that in Card’s data, experience is measured with error. Moreover, suppose that experience is positively correlated with being male (i.e. male workers have more years of experience). If Card regression includes a dummy variable indicating whether a person is male along with the experience variable,

d. how will the measurement error affect the coefficient on experience? Justify your answer.

e. how will the measurement error affect the coefficient on the male dummy? Justify your answer.

Card was interested in how deregulation of the airline industry affected the wages of airline workers relative to the rest of the population. To test for the effect of deregulation, he had data for years before and after deregulation occurred and defined a dummy variable (call it DEREG) that was unity after deregulation and zero before.

f. Explain how Card could test whether the DEREG period caused airline wages to fall relative to other workers’ wages. Be sure to define any additional variables that he would have to create, the regression that he would have to run, and the appropriate test statistic for the null hypothesis.

2. (30 points) Using data from the Health and Retirement Survey, I estimated a regression of wealth on race dummies. Wealth includes checking and savings accounts, home equity, and pension wealth. The results are listed below:

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>3</td>
<td>9.1412769E12</td>
<td>3.0470923E12</td>
<td>24.118</td>
<td>0.0001</td>
</tr>
<tr>
<td>Error</td>
<td>5192</td>
<td>6.5596211E14</td>
<td>126340930971</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C Total</td>
<td>5195</td>
<td>6.6510339E14</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Parameter Estimates

| Variable | DF | Parameter Estimates | Standard Error | T for H0: Parameter=0 | Prob > |T| |
|----------|----|---------------------|----------------|------------------------|--------|---|
| INTERCEP | 1  | 196367              | 5483.3253132   | 35.812                 | 0.0001 |
| BLACK    | 1  | -109439             | 13183.301131   | -8.301                 | 0.0001 |
| INDIAN   | 1  | -97507              | 54481.446767   | -1.790                 | 0.0736 |
| ASIAN    | 1  | 29574               | 42246.917152   | 0.700                  | 0.4839 |

The four races are White, Black, Indian, and Asian. White is the omitted dummy variable.

Based on the above regression results,

- **a.** what is the difference in the average wealth between Whites and Blacks? **No justification required.**
- **b.** what is the difference in the average wealth between Blacks and Asians? **No justification required.**

If the reference group was changed from Whites to Blacks,

- **c.** what would the coefficient on the White dummy be? **No justification required.**
- **d.** what would the coefficient on the Indian dummy be? **No justification required.**

**e.** Suppose that I expressed my wealth measure in thousands of $. How would this affect (i) the coefficient on the race dummies; (2) the t-statistics on the race dummies? Give as precise an answer as possible. **No justification required.**

**f.** Suppose that you wanted to test the null hypothesis that wealth was identical for Blacks, Indians, and Asians (notice I did not include Whites here). Describe how you could construct an F-test of the null. **Define the restricted and unrestricted regression models, how the f-stat would be constructed, and the number of degrees of freedom for the f-stat. Also, on the basis of the f-stat, how would you determine whether to reject the null?**
3. (30 points) In an article in the January 1999 issue of the *Industrial and Labor Relations Review*, Linda Babcock and John Engberg examine the effect of bargaining unit composition on the returns to education and tenure in school districts. Their main interest is whether the composition of the school district employees (i.e. the bargaining unit) affects the returns to education or tenure. In particular, they are interested in whether school districts that have a employees with higher levels of education also negotiate a contract that pays a high return to education; and whether school districts with employees that have a high level of tenure (years employed with the district) negotiate a high return to tenure.

For this problem, let's focus on their estimates of the return to education. They measure the returns to education (EDRET) in a given school district as the the percentage increase in the salary paid for the receipt of a Masters degree. They regress EDRET on the percentage of school teachers in that district with a Masters degree (MAPCT). The coefficient estimate on MAPCT is .2697 and it has a standard error of .0918. Thus, increasing MAPCT from 40 to 60 would increase the premium that a teacher with a Masters degree receives by 20*.2697=5.4%.

In summarizing their results, they conclude that school districts that have more educated teachers negotiate a higher salary increase for a Masters degree. They argue that this is consistent with a median voter model of union behavior.

The authors point out one potential problem with this analysis. They note that "teachers in districts with ... higher returns to education are more likely to obtain a Master's degree". They go on to suggest that, if this is the case, their regression estimate may be biased due to "reverse causation".

a. Given the concern they raise, would the estimated return to a Masters degree be upward or downward biased given the nature of the "reverse causation" they describe. Justify your answer.

b. Explain how the model could be properly estimated with 2 stage least squares (2SLS). In your answer,
   i. describe what additional data (variables) would be necessary for "identification" of the EDRET equation and be sure to define the criteria you would use for selecting such "identifiers"
   ii. describe the 2 stages of 2SLS. Briefly describe how the 2nd stage avoids the problem of simultaneity bias that occurs in OLS.

c. Explain why the second stage of 2SLS would "break down" if there were no identifying variables. Be as precise as possible.

d. Explain why the precision of the 2SLS estimates would improve with more "identifying" variables.