1. For each of the following estimated models provide the simplest possible explanation of the effect of IQ (intelligent quotient has mean of 100) on EARNS (annual earnings in $). Hint: For models with logs use elasticities or semi-elasticities.

(a) $EARNS = 1000 + 500 \times IQ$.

(b) $EARNS = 20000 + 200 \times IQ + 2 \times IQ^2$

(c) $EARNS = -20000 + 50000 \times ln(IQ)$

(d) $EARNS = 45000 + 10000 \times d, where\ d = 1\ if\ IQ > 100\ and\ d = 0\ if\ IQ \leq 100$.

(e) $ln(EARNS) = 10 + 0.010 \times IQ$.

(f) $ln(EARNS) = 5 + 0.90 \times ln(IQ)$.

(g) For the model in (c) give the marginal effect at the mean if $\overline{IQ} = 110$.

2. Use data in file ADVERTISING.DTA at the course website.

In answering the following it can be helpful (though not necessary) to use **estimates store** and **estimates table** to present model results side by side.

(a) Fit linear, linear-log, log-linear and log-log models for regression of sales on tv.

(b) Which model, if any, do you prefer? Explain.

(c) Compute the average marginal effect (AME) and the marginal effect at the mean (MEM) for the linear-log model.

(d) Use command **predict** after the linear regression to predict sales from the linear model. Does the prediction on average equal sales on average?

(e) Use command predict after the log-linear regression to predict log-sales from the log-linear model. Then create a crude prediction of sales. Does the prediction on average equal sales on average?

(f) The notes in Chapter 12.5 suggest multiplying the preceding prediction by $\exp(s_{e2}/2)$ where s_e is

the standard error of the regression of lny on x. Does this lead to a better prediction? Explain.

(g) Plot the three different predictions from parts (d), (e) and (f) against tv on the same graph. Comment on these plots. Suppose the predictions are called p1, p2 and p3.

3. Use data in file ADVERTISING.DTA at the course website.

(a) Obtain summary statistics for sales, tv, radio, and newspaper. Are the data in the expected ranges? HINT: Use command describe to get variable descriptions.

(b) Test the claim that the population mean sales are less than 13000 at level 0.05.

(c) Which variable(s) do you think will be important predictors of sales?

(d) Which variables are most highly correlated with sales? Are there any surprises?

(e) Which variables are statistically significant (two-sided test at 5%)? Use the individual p-values.

(f) State which, if any, of the regressors do not have the sign you expect a priori and give an explanation for why you expected a different sign.