Solutions to Econ 422 Problem Set 4

1 a) \$ 23,400 (recall that Price is measured in \$ 1000s)

b) In this case $\Delta BDR = 1$ and $\Delta Hsize = 100$. The resulting expected change in price is $23.4 + 0.156 \times 100 = 39.0$ thousand dollars or \$ 39,000.

c) The loss is \$48,800 (-48.8×1000).

d) Since $\bar{R}^2 = 1 - \frac{n-1}{n-k-1}(1-R^2)$, so $R^2 = 1 - \frac{n-k-1}{n-1}(1-\bar{R}^2)$.

Thus $R^2 = 0.727$ where $n = 220$ and $k = 6$.

2 a) By adding and subtracting $\beta_2 X_{1i}$, we have $\beta_1 X_{1i} + \beta_2 X_{2i}$

$= \beta_1 X_{1i} - \beta_2 X_{1i} + \beta_2 X_{1i} + \beta_2 X_{2i} = (\beta_1 - \beta_2) X_{1i} + \beta_2 (X_{1i} + X_{2i})$

$= \gamma X_{1i} + \beta_2 (X_{1i} + X_{2i})$

Therefore, testing $\beta_1 = \beta_2$ is equivalent to testing $\gamma = 0$

where $Y_i = \beta_0 + \gamma X_{1i} + \beta_2 (X_{1i} + X_{2i}) + u_i$

b) By adding and subtracting $a\beta_2 X_{1i}$, we have $\beta_1 X_{1i} + \beta_2 X_{2i}$

$= \beta_1 X_{1i} + \beta_2 X_{2i} + a\beta_2 X_{1i} - a\beta_2 X_{1i} = (\beta_1 + a\beta_2) X_{1i} + \beta_2 (X_{2i} - aX_{1i})$

$= \gamma X_{1i} + \beta_2 (X_{2i} - aX_{1i})$

Therefore testing $\beta_1 + a\beta_2 = 0$ is equivalent to testing $\gamma = 0$

where $Y_i = \beta_0 + \gamma X_{1i} + \beta_2 (X_{2i} - aX_{1i}) + u_i$

c) ① Subtract X_{1i} on both sides, we have

$Y_i - X_{1i} = \beta_0 + \beta_1 X_{1i} - X_{1i} + \beta_2 X_{2i} + u_i$

② Add and subtract $\beta_2 X_{1i}$ on Right Hand Side, we have

$Y_i - X_{1i} = \beta_0 + \beta_1 X_{1i} - X_{1i} + (\beta_2 X_{1i} - \beta_2 X_{1i}) + \beta_2 X_{2i} + u_i$

③ Rearranging : $Y_i - X_{1i} = \beta_0 + \underbrace{(\beta_1 + \beta_2 - 1)} X_{1i} + \beta_2 (X_{2i} - X_{1i}) + u_i$

c) ③ $Y_i - X_{1i} = \beta_0 + \gamma X_{1i} + \beta_2 (X_{2i} - X_{1i}) + u_i$

Therefore testing $\beta_1 + \beta_2 = 1$ is equivalent to testing $\gamma = 0$

where $Y_i - X_{1i} = \beta_0 + \gamma X_{1i} + \beta_2 (X_{2i} - X_{1i}) + u_i$

3 a) The t-statistic is $\dfrac{0.485}{2.61} = 0.186 < 1.96$. Therefore, the coefficient on BDR is not statistically significantly different from zero.

b) The coefficient on BDR measures the partial effect of the number of bedrooms holding house size (Hsize) constant. Yet, the typical 5-bedroom house is much larger than the typical 2-bedroom house. Thus, the results in (a) says little about conventional wisdom.

c) The 99% C.I. for effect of lot size on price is $2000 \times [0.002 \pm 2.58 \times 0.00048]$ or 1.52 to 6.48 (in thousands of dollars)

d) Choosing the scale of the variables should be done to make the regression results easy to read and to interpret. If the lot size were measured in thousands of square feet, the estimated coefficient would be 2 instead of 0.002.

e) The 10% critical value from the $F_{2,\infty}$ distribution is 2.30. Because $0.08 < 2.30$, the coefficients are not jointly significant at the 10% level.

4 a) i) ln(Earnings) for females are, on average, 0.44 lower for men than for women.

ii) The error term has a standard deviation of 2.65 (measured in log-points)

iii) Yes. But the regression does not control for many factors (size of firm, industry, profitability, experience and so forth).

iv) No. Gender discrimination means that two workers, identical in every way but gender, are paid different wages. Thus, it is also important to control for characteristics of the workers that may affect their productivity (education, years of experience, etc.) If these characteristics are systematically different between men and women, then they may be responsible for the difference in mean wages. (If this were true, it would raise an interesting and important question of why women tend to have less education or less experience than men, but that is a question about something other than gender discrimination.) These are potentially important omitted variables in the regression that will lead to bias in the OLS coefficient estimator for 'Female'. Since these characteristics were not controlled for in the statistical analysis, it is premature to reach a conclusion about gender discrimination.

b i) If MarketValue increases by 1%, earnings increase by 0.37%.

ii) 'Female' is correlated with the two new included variables and at least one of the variables is important for explaining ln(Earnings). Thus the regression in part (a) suffered from omitted variable bias.

c) Forgetting about the effect of 'Return', whose effects seems small and statistically insignificant, the omitted variable bias formula

$$\hat{\beta}_1 \rightarrow \beta_1 + \rho_{Xu}\frac{\sigma_u}{\sigma_X} \quad \text{where } \rho_{Xu} = \operatorname{corr}(X_i, u_i) \text{ suggests}$$

that 'Female' is negatively correlated with ln(MarketValue).

(Coefficient on 'Female' dropped from -0.28 to -0.44 when ln(MarketValue) is omitted from the regression.