Answer for Homework #4

6.5.
(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.
(d) From the text $\bar{R}^2 = 1 - \frac{n-1}{n-k-1}(1-R^2)$ so $R^2 = 1 - \frac{n-k-1}{n-1}(1-\bar{R}^2)$ thus, $R^2 = 0.727$.

6.6.
(a) There are other important determinants of a country's crime rate, including demographic characteristics of the population.
(b) Suppose that the crime rate is positively affected by the fraction of young males in the population, and that counties with high crime rates tend to hire more police. In this case, the size of the police force is likely to be positively correlated with the fraction of young males in the population leading to a positive value for the omitted variable bias so that $\hat{\beta}_1 > \beta_1$.

6.10.

(a)
$$\sigma^2_{\hat{\beta}_1} = \frac{1}{n}\left[\frac{1}{1-\rho^2_{X_1,X_2}}\right]\frac{\sigma^2_u}{\sigma^2_{X_1}}$$

Assume X_1 and X_2 are uncorrelated: $\rho^2_{X_1X_2} = 0$

$$\sigma^2_{\hat{\beta}_1} = \frac{1}{400}\left[\frac{1}{1-0}\right]\frac{4}{6}$$
$$= \frac{1}{400} \times \frac{4}{6} = \frac{1}{600} = 0.00167$$

(b) With $\rho_{X_1,X_2} = 0.5$

$$\sigma^2_{\hat{\beta}_1} = \frac{1}{400}\left[\frac{1}{1-0.5^2}\right]\frac{4}{6}$$
$$= \frac{1}{400}\left[\frac{1}{0.75}\right]\frac{4}{6} = .0022$$

(c) The statement correctly says that the larger is the correlation between X_1 and X_2 the larger is the variance of $\hat{\beta}_1$, however the recommendation "it is best to leave X_2 out of the regression" is incorrect. If X_2 is a determinant of Y, then leaving X_2 out of the regression will lead to omitted variable bias in $\hat{\beta}_1$.

6.11.

(a) $\sum (Y_i - b_1X_{1i} - b_2X_{2i})^2$

$$\frac{\partial \sum (Y_i - b_1X_{1i} - b_2X_{2i})^2}{\partial b_1} = -2\sum X_{1i}(Y_i - b_1X_{1i} - b_2X_{2i})$$

(b) $$\frac{\partial \sum (Y_i - b_1X_{1i} - b_2X_{2i})^2}{\partial b_2} = -2\sum X_{2i}(Y_i - b_1X_{1i} - b_2X_{2i})$$

(c) From (b), $\hat{\beta}_1$ satisfies $\sum X_{1i}(Y_i - \hat{\beta}_1X_{1i} - \hat{\beta}_1X_{2i}) = 0$ or $\hat{\beta}_1 = \dfrac{\sum X_{1i}Y_i - \hat{\beta}_2\sum X_{1i}X_{2i}}{\sum X_{1i}^2}$ and the result follows immediately.

(d) Following analysis as in (c)

$$\hat{\beta}_2 = \frac{\sum X_{2i}Y_i - \hat{\beta}_1\sum X_{1i}X_{2i}}{\sum X_{2i}^2}$$

and substituting this into the expression for $\hat{\beta}_1$ in (c) yields

$$\hat{\beta}_1 = \frac{\sum X_{1i}Y \frac{\sum X_{2i}Y_i - \hat{\beta}_1\sum X_{1i}X_{2i}}{\sum X_{2i}^2}\sum X_{1i}X_{2i}}{\sum X_{1i}^2}.$$

Solving for $\hat{\beta}_1$ yields:

$$\hat{\beta}_1 = \frac{\sum X_{2i}^2\sum X_{1i}Y_i - \sum X_{1i}X_{2i}\sum X_{2i}Y_i}{\sum X_{1i}^2\sum X_{2i}^2 - (\sum X_{1i}X_{2i})^2}$$

(e) The least squares objective function is $\sum (Y_i - b_0 - b_1X_{1i} - b_2X_{2i})^2$ and the partial derivative with respect to b_0 is

$$\frac{\partial \sum (Y_i - b_0 - b_1X_{1i} - b_2X_{2i})^2}{\partial b_0} = -2\sum (Y_i - b_0 - b_1X_{1i} - b_2X_{2i}).$$

Setting this to zero and solving for $\hat{\beta}_0$ yields: $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}_1 - \hat{\beta}_2\bar{X}_2$

(f) Substituting $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}_1 - \hat{\beta}_2\bar{X}_2$ into the least squares objective function yields $\sum (Y_i - \hat{\beta}_0 - b_1X_{1i} - b_2X_{2i})^2 = \sum \left((Y_i - \bar{Y}) - b_1(X_{1i} - \bar{X}_1) - b_2(X_{2i} - \bar{X}_2) \right)^2$, which is identical to the least squares objective function in part (a), except that all variables have been replaced with deviations from sample means. The result then follows as in (c).

Notice that the estimator for β_1 is identical to the OLS estimator from the regression of Y onto X_1, omitting X_2. Said differently, when $\sum (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2) = 0$, the estimated coefficient on X_1 in the OLS regression of Y onto both X_1 and X_2 is the same as estimated coefficient in the OLS regression of Y onto X_1.