8.2 With Var(u|*inc,price,educ,female*) = $\sigma^2 inc^2$, $h(\mathbf{x}) = inc^2$, where $h(\mathbf{x})$ is the heteroskedasticity function defined in equation (8.21). Therefore, $\sqrt{h(\mathbf{x})} = inc$, and so the transformed equation is obtained by dividing the original equation by *inc*:

$$\frac{beer}{inc} = \beta_0(1/inc) + \beta_1 + \beta_2(price/inc) + \beta_3(educ/inc) + \beta_4(female/inc) + (u/inc).$$

Notice that β_1, which is the slope on *inc* in the original model, is now a constant in the transformed equation. This is simply a consequence of the form of the heteroskedasticity and the functional forms of the explanatory variables in the original equation.

8.4 (i) These coefficients have the anticipated signs. If a student takes courses where grades are, on average, higher – as reflected by higher *crsgpa* – then his/her grades will be higher. The better the student has been in the past – as measured by *cumgpa* – the better the student does (on average) in the current semester. Finally, *tothrs* is a measure of experience, and its coefficient indicates an increasing return to experience.

The t statistic for *crsgpa* is very large, over five using the usual standard error (which is the largest of the two). Using the robust standard error for *cumgpa*, its t statistic is about 2.61, which is also significant at the 5% level. The t statistic for *tothrs* is only about 1.17 using either standard error, so it is not significant at the 5% level.

(ii) This is easiest to see without other explanatory variables in the model. If *crsgpa* were the only explanatory variable, H_0: $\beta_{crsgpa} = 1$ means that, without any information about the student, the best predictor of term GPA is the average GPA in the students' courses; this holds essentially by definition. (The intercept would be zero in this case.) With additional explanatory variables it is not necessarily true that $\beta_{crsgpa} = 1$ because *crsgpa* could be correlated with characteristics of the student. (For example, perhaps the courses students take are influenced by ability – as measured by test scores – and past college performance.) But it is still interesting to test this hypothesis.

The t statistic using the usual standard error is $t = (.900 - 1)/.175 \approx -.57$; using the heteroskedasticity-robust standard error gives $t \approx -.60$. In either case we fail to reject H_0: $\beta_{crsgpa} = 1$ at any reasonable significance level, certainly including 5%.

(iii) The in-season effect is given by the coefficient on *season*, which implies that, other things equal, an athlete's GPA is about .16 points lower when his/her sport is competing. The t statistic using the usual standard error is about –1.60, while that using the robust standard error is about –1.96. Against a two-sided alternative, the t statistic using the robust standard error is just significant at the 5% level (the standard normal critical value is 1.96), while using the usual standard error, the t statistic is not quite significant at the 10% level ($cv \approx 1.65$). So the standard error used makes a difference in this case. This example is somewhat unusual, as the robust standard error is more often the larger of the two.

8.5 (i) No. For each coefficient, the usual standard errors and the heteroskedasticity-robust ones are practically very similar.

(ii) The effect is $-.029(4) = -.116$, so the probability of smoking falls by about .116.

(iii) As usual, we compute the turning point in the quadratic: $.020/[2(.00026)] \approx 38.46$, so about 38 and one-half years.

(iv) Holding other factors in the equation fixed, a person in a state with restaurant smoking restrictions has a .101 lower chance of smoking. This is similar to the effect of having four more years of education.

(v) We just plug the values of the independent variables into the OLS regression line:

$$\widehat{smokes} = .656 - .069 \cdot \log(67.44) + .012 \cdot \log(6,500) - .029(16) + .020(77) - .00026(77^2) \approx .0052.$$

Thus, the estimated probability of smoking for this person is close to zero. (In fact, this person is not a smoker, so the equation predicts well for this particular observation.)

8.6 (i) The proposed test is a hybrid of the BP and White tests. There are $k + 1$ regressors, each original explanatory variable and the squared fitted values. So, the number of restrictions tested is $k + 1$, and this is the numerator *df*. The denominator *df* is $n - (k + 2) = n - k - 2$.

(ii) For the BP test, this is easy: the hybrid test has an extra regressor, $\hat{y}^2$, and so the R-squared will be no less for the hybrid test than for the BP test. For the special case of the White test, the argument is a bit more subtle. In regression (8.20), the fitted values are a linear function of the regressors (where, of course, the coefficients in the linear function are the OLS estimates). So, we are putting a restriction on how the original explanatory variables appear in the regression. This means that the R-squared from (8.20) will be no greater than the R-squared from the hybrid regression.

(iii) No. The F statistic for joint significance of the regressors depends on $R^2_{\hat{u}^2}/(1 - R^2_{\hat{u}^2})$, and it is true that this ratio increases as $R^2_{\hat{u}^2}$ increases. But, the F statistic also depends on the *df*, and the *df* are different among all three tests: the BP test, the special case of the White test, and the hybrid test. So we do not know which test will deliver the smallest p-value.

(iv) As discussed in part (ii), the OLS fitted values are a linear combination of the original regressors. Because those regressors appear in the hybrid test, adding the OLS fitted values is redundant; perfect collinearity would result.

C8.4 (i) The estimated equation is

$$\widehat{voteA} = \underset{(4.74)}{37.66} + \underset{(.071)}{.252}\, prtystrA + \underset{(1.407)}{3.793}\, democA + \underset{(0.392)}{5.779} \log(expendA)$$

$$- \underset{(0.397)}{6.238} \log(expendB) + \hat{u}$$

$$n = 173, \quad R^2 = .801, \quad \bar{R}^2 = .796.$$

You can convince yourself that regressing the $\hat{u}_i$ on all of the explanatory variables yields an R-squared of zero, although it might not be exactly zero in your computer output due to rounding error. Remember, OLS works by choosing the estimates, $\hat{\beta}_j$, such that the residuals are uncorrelated in the sample with each independent variable (and the residuals have a zero sample average, too).

(ii) The B-P test entails regressing the $\hat{u}_i^2$ on the independent variables in part (i). The F statistic for joint significant (with 4 and 168 *df*) is about 2.33 with p-value $\approx .058$. Therefore, there is some evidence of heteroskedasticity, but not quite at the 5% level.

(iii) Now we regress $\hat{u}_i^2$ on $\widehat{voteA_i}$ and $(\widehat{voteA_i})^2$, where the $\widehat{voteA_i}$ are the OLS fitted values from part (i). The F test, with 2 and 170 *df*, is about 2.79 with p-value $\approx .065$. This is slightly less evidence of heteroskedasticity than provided by the B-P test, but the conclusion is very similar.