Problemløsninger PS2 – BA-BMECV1031U

1 Parts (b) and (c). The homoskedasticity assumption playes no role in showing that OLS is consistent. But we know that heteroskedasticity causes statistical inference based on the usual *t* and *F* statistics to be invalid, even in large samples. As heteroskedasticity is a violation of the Gauss-Markov assumptions, OLS is no longer BLUE.

2 Var(u|$inc,price,educ,female$) = $\sigma^2 inc^2$, $h(\mathbf{x}) = inc^2$, where $h(\mathbf{x})$ is the heteroskedasticity function. 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.

3 False. The unbiasedness of WLS and OLS hinges crucially on Assumption MLR.4, and, this assumption is often violated when an important variable is omitted. When MLR.4 does not hold, both WLS and OLS are biased. Without specific information on how the omitted variable is correlated with the included explanatory variables, it is not possible to determine which estimator has a small bias. It is possible that WLS would have more bias than OLS or less bias. Because we cannot know, we should not claim to use WLS in order to solve "biases" associated with OLS.

4 (a) Robust standard errors are slightly larger than the usual standard errors but no change in the level of significance (except for the constant).

(b) After estimating the equation, we obtain the squared OLS residuals $\hat{u}^2$. The full-blown White test is based on the *R*-squared from the auxiliary regression (with an intercept),

$$\hat{u}^2 \text{ on } llotsize,\ lsqrft,\ bdrms,\ llotsize^2,\ lsqrft^2,\ bdrms^2,$$
$$llotsize \cdot lsqrft,\ llotsize \cdot bdrms, \text{ and } lsqrft \cdot bdrms,$$

where "*l*" in front of *lotsize* and *sqrft* denotes the natural log. With 88 observations the *F* version of the White statistic is 1.05, and this is the outcome of an (approximately) F_(9,78) random variable. The *p*-value is about .41, which provides little evidence against the homoskedasticity assumption.

5 (a) We now get R^2 = .0527, but the other estimates seem okay.

(b) One way to ensure that the unweighted residuals are being provided is to compare them with the OLS residuals. They will not be the same, of course, but they should not be wildly different.

(c) The *R*-squared from the regression $\breve{u}_i^2$ on $\breve{y}_i$, $\breve{y}_i^2$, $i = 1,...,807$ is about .027. This gives F = 11.15, and so the p-value is essentially zero.

(d) The substantial heteroskedasticity found in part (c) shows that the feasible GLS procedure described does not, in fact, eliminate the heteroskedasticity. Therefore, the usual standard errors, t statistics, and F statistics reported with weighted least squares are not valid, even asymptotically.