Exercise Class - Econometrics Class 3

Instructor: Irene Iodice
Email: irene.iodice@malix.univ-paris1.fr

Ex.1: Recall on what we have already done on heteroskedasticity
Consider the consumption function:

$$cons_i = \beta_0 + \beta_1 inc_i + u_i \qquad u_i = e_i\sqrt{inc_i} \tag{1}$$

where e_i is a random variable with $E(e_i) = 0$ and $Var(e) = \sigma_e^2$. Assume that e_i is independent of inc_i.

1. *Explain the intuition behind the structure of the error.*
The error is an increasing (not linear but square root) function of the income, thus the larger the income of a person, the bigger the deviation from the regression line or the error we make in estimating his consumption.

2. *Show that HP3 is satisfied, which is that the zero conditional mean assumption holds.*
Recall that HP3 states that $E(u_i|x_i) = 0$. In our case we have $E(e_i\sqrt{inc_i}|inc_i) = \sqrt{inc_i}E(e_i|inc_i) = \sqrt{inc_i} \times 0 = 0$, thus it is because (i) when we condition on the variable inc_i in computing an expectation, its value becomes a constant, (ii) e_i is independent of inc_i thus its conditional expectation is equal to the unconditional one, which in turn is equal zero by Ass.

3. *Show that HP5 is violated, which is that the error terms and the regressor are not independent.*
Again, when we condition on inc_i in computing a variance, whatever function of inc becomes a constant. So $Var(u_i|inc_i) = Var(e_i\sqrt{inc_i}|inc_i) = (\sqrt{inc_i})^2 Var(e_i|inc_i) = inc_i\sigma_e^2$ because $Var(e_i|inc_i) = \sigma_e^2$.

Ex.2: Consequences of heteroskedasticity
Which of the following are consequences of heteroskedasticity?

1. The OLS estimators are inconsistent.

2. The usual F statistic no longer has an F distribution.

3. The OLS estimators are no longer BLUE.

Parts (2) and (3). The homoskedasticity assumption plays no role in showing that OLS is consistent (first 4 assumptions of OLS are necessary). 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 (all 5 assumptions of OLS), OLS is no longer BLUE.

Ex.3: Example on inference with different standard errors
The following equation was estimated for the fall and second semester students:

$$\begin{aligned}\widehat{trmgpa} = {} & \underset{\substack{(.55)\\ [.55]}}{-2.12} + \underset{\substack{(.175)\\ [.166]}}{.900}\, crsgpa + \underset{\substack{(.064)\\ [.074]}}{.193}\, cumgpa + \underset{\substack{(.0012)\\ [.0012]}}{.0014}\, tothrs \\ & + \underset{\substack{(.0002)\\ [.0002]}}{.0018}\, sat - \underset{\substack{(.0018)\\ [.0019]}}{.0039}\, hsperc + \underset{\substack{(.085)\\ [.079]}}{.351}\, female - \underset{\substack{(.098)\\ [.080]}}{.157}\, season\end{aligned}$$

$$n = 269,\ R^2 = .465.$$

Here, trmgpa is term GPA, crsgpa is a weighted average of overall GPA in courses taken, tothrs is total credit hours prior to the semester, sat is SAT score, hsperc is grad- uating percentile in high school class, female is a gender dummy, and season is a dummy variable equal to unity if the student's sport is in season during the fall. The usual and heteroskedasticity-robust standard errors are reported in parentheses and brackets, respectively.

1. Test whether there is an in-season effect on term GPA, using both standard errors. Does the significance level at which the null can be rejected depend on the standard error used?
   In both cases we are testing the same hypothesis, under same assumptions, what changes is just the t^{act} we compute with the data at hand, in particular:
   $$t_1^{act} = \frac{\hat{\beta}_7 - 0}{SE(\hat{\beta}_7)} = \frac{0.157}{0.098} = 1.60 \tag{2}$$
   $$t_2^{act} = \frac{\hat{\beta}_7 - 0}{SE(\hat{\beta}_7)} = \frac{0.157}{0.080} = 1.96 \tag{3}$$
   The pvalue in the first case is $Pr(t > t_1^{act}) = 1 - 0.9452 = 0.0548$ while in the second is $Pr(t > t_2^{act}) = 1 - 0.975 = 0.025$. Then, since it is a test at two tail, we would reject the second, while not the first, both at 10% and at 5% significance level.
2. Comment the differences between the standard errors and the robust ones.
   Some robust standard errors are larger and other are smaller wrt to the conventional ones.

   Note that from a simple regression model the nominator of the standard error can be rewritten as:
   $$Var(u_i(x_i - \overline{x})) = E[u_i^2(x_i - \overline{x})^2] = E[u_i^2]E[(x_i - \overline{x})^2] + Cov[u_i^2, (x_i - \overline{x})^2]$$
   where the covariance is zero if we assume the errors are iid. Then the sign of the covariance between the error term and X_i tells us whether the robust are greater or smaller than the conventional standard errors.

Ex.4: Introduction to GLS
Consider a linear model to explain monthly beer consumption

$$beer = \beta_0 + \beta_1 inc + \beta_2 age + \beta_3 educ + \beta_4 female + u$$

with $E(u|inc, age, educ, female) = 0$ and $Var(u|inc, price, educ, female) = \sigma^2 inc^2$. Write the transformed equation that has a homoskedastic error term and comment.
With $Var(u|inc, price, educ, female) = \sigma^2 inc^2$, $f(x) = inc^2$, where $f(x)$ is the heteroskedasticity function. Therefore, $\sqrt{f(x)} = inc$, and so the transformed equation is obtained by dividing the original equation by inc

$$\frac{beer}{inc} = \frac{\beta_0}{inc} + \beta_1 + \beta_2\frac{age}{inc} + \beta_3\frac{educ}{inc} + \beta_4\frac{female}{inc} + \frac{u}{inc} \tag{4}$$

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.

Ex.5: Implication of GLS
Consider the model

$$y_i = \beta_0 + \beta_1 x_i + u_i \tag{5}$$

with heteroskedastic variance $Var(u_i) = \sigma_i^2$ and its transformed homoskedastic version $y_i^* = b_0\sigma_i^{-1} + b_1 x_i^* + u_i^*$ where $y_i^* = \sigma_i^{-1} y_i$, $x_i^* = \sigma_i^{-1} x_i$ and $u_i^* = \sigma_i^{-1} u_i$.

1. Write the formula for the estimator of b_1 and b_0 You can recall the various steps from the derivation we did in the class of imetrics 1 to work out the following formula for $\hat{b}_1$ and $\hat{b}_0$
   $$\hat{b}_1 = \frac{\sum x_i^* y_i^* - n\overline{x^* y^*}}{\sum_i (x_i^*)^2 - n\overline{x}^{*2}} \tag{6}$$
   $$\hat{b}_0 = \frac{\sum_i y_i^* - b_1 \sum_i x_i^*}{\sum \sigma_i^{-1}} \tag{7}$$
2. Show that the estimators of b_1 and b_0 are equal to the least squares estimators of β_1 and β_0 when $Var(u_i|x_i) = \sigma^2$. That is, the error variances are constant. We transform y^*, x^* into y_i

and x_i

$$\hat{b}_1 = \frac{\sum x_i y_i(\sigma_i^{-2}) - \frac{1}{n}\sum_i x_i\sigma_i^{-1} \times \sum y_i\sigma_i^{-1}}{\sum_i x_i^2\sigma_i^{-2} - \left(\sum x_i\sigma_i^{-1}\right)^2} \tag{8}$$

Now we assume that $\sigma_i = \sigma$, then we can bring σ out of the summation and collect it so that:

$$\hat{b}_1 = \frac{\sigma^{-2}\left(\sum x_i y_i - n\overline{xy}\right)}{\sigma^{-2}\left(\sum_i (x_i)^2 - n\overline{x}^2\right)} = \hat{\beta}_1 \tag{9}$$

$$\hat{b}_0 = \frac{\sigma^{-1}\left(\sum_i y_i - \hat{b}_1\sum_i x_i\right)}{n\sigma^{-1}} = \frac{\left(\sum_i y_i - \hat{\beta}_1\sum_i x_i\right)}{n} = \overline{y} - \hat{\beta}_1\overline{x} = \hat{\beta}_0 \tag{10}$$

3. Does a comparison of the formulas for b_0 and b_1 with those for β_0 and β_1 suggest an interpretation for the new estimators?
The least squares estimators β_0 and β_1 are functions of the following aggregators: $\overline{x} = \frac{1}{N}\sum x_i$, $\overline{y} = \frac{1}{N}\sum y_i$, $\frac{1}{N}\sum y_i x_i$, $\frac{1}{N}\sum x_i^2$. For the generalized least squares estimator for b_0 and b_1, these unweighted are replaced by the weighted aggregators. In particular, imagine to divide each term in the formulas by $\sum \sigma_i^2$, so that $\sigma_i^{-2}/\sum_i \sigma_i^{-2}$ can be interpreted as a weight. Then we have that the elements are: $\overline{x} = \frac{\sum \sigma_i^{-2}x_i}{\sum \sigma_i^{-2}}$, $\overline{y} = \frac{\sum \sigma_i^{-2}y_i}{\sum \sigma_i^{-2}}$, $\frac{\sum \sigma_i^{-2}x_i y_i}{\sum \sigma_i^{-2}}$, $\frac{\sum \sigma_i^{-2}x_i^2}{\sum \sigma_i^{-2}}$. From this it is evident that each observation is weighted by the inverse of the error variance. Reliable observations with small error variances are weighted more heavily than those with higher error variances that make them more unreliable. NOte that for this reason the GLS is said to be a particular case of WLS.

Ex.6: Example of GLS

Consider the model:

$$y_i = \beta_0 + \beta_1 x_i + u_i \tag{11}$$

where the u_i are independent errors with $E(u_i) = 0$ and $var(u_i) = \sigma^2 x_i^2$. Suppose that you have the following five observations $y = (4, 3, 1, 0, 2)$ and $x = (1, 2, 1, 3, 4)$. Find the generalized least squares estimates of β_0 and β_1.
For the model $y_i = \beta_0 + \beta_1 x_i + u_i$ where $var(u_i) = \sigma^2 x_i^2$, the transformed model that gives a constant error variance is $y_i* = \beta_0 x_i^* + \beta_1 + u_i^*$ where $y_i^* = \frac{y_i}{x_i}$, $x_i^* = \frac{1}{x_i}$, and $u_i^* = \frac{u_i}{x_i}$. This model can be estimated by least squares with the usual simple regression formulas, but with β_0 and β_1 reversed. Thus, the generalized least squares estimators for β_0 and β_1 are:

$$\hat{\beta}_0 = \frac{\sum_i x_i^* y_i^* - n\overline{x^* y^*}}{\sum_i (x^*)^2 - n\overline{x^*}^2} \quad \hat{\beta}_1 = \overline{y}^* - \hat{\beta}_1\overline{x}^* \tag{12}$$

Using the observation, we get:

$$\overline{x^*} = \frac{1}{5}(1 + \frac{1}{2} + 1 + \frac{1}{3} + \frac{1}{4}) = \frac{37}{60} = 0.61 \tag{13}$$

$$\overline{y^*} = \frac{1}{5}(4 + \frac{3}{2} + 1 + 0 + \frac{1}{2}) = \frac{7}{5} = 1.4 \tag{14}$$

$$\sum x_i^* y_i^* = \frac{4}{1^2} + \frac{3}{2^2} + \frac{1}{1^2} + \frac{0}{3^2} + \frac{2}{4^2} = \frac{47}{8} = 5.875 \tag{15}$$

$$(\overline{x^*})^2 = (\frac{37}{60})^2 = \frac{1369}{3600} = 0.38 \tag{16}$$

$$\sum_i (x^*)^2 = 1 + \frac{1}{4} + 1 + \frac{1}{9} + \frac{1}{16} = \frac{154}{64} = 2.41 \tag{17}$$

Then we get that

$$\hat{\beta}_0 = \frac{5.875 - 0.61(1.4 \times 5)}{2.41 - 5 \times 0.38} = 3.15 \quad \hat{\beta}_1 = 1.4 - 3.15 \times 0.6 = -0.49 \tag{18}$$

Ex.7: Clusters, robust and standard errors
You want to estimate the following model:

$$sales_{it} = \beta_0 + \beta_1 solvency\_rate_{it} + u_{it} \tag{19}$$

where i is a firm identifier and t represent the year. Now you estimate the model with artificial data provided by Petersen (2009) through the 'Sandwich' package of R, making different assumption on the standard errors. Below the graph of data and the estimated regression line and the table summarizing the results.

1. Compare the estimated coefficients of running OLS with classical and robust standard errors, have they changed?
The estimated coefficient have not changed since the formulas of the OLS estimators does not contain the variance of the estimator.

2. Compare the robust standard error with the classical ones (do they increase or decrease?). By looking at the graph try to comment whether you would expect that direction of change.
Conventional standard errors are biased up, this happens whenever observations on x_i far from the mean (observations with high leverage) are associated with lower variance residuals, that is that the covariance between x_i and u_i is negative. An equivalent result holds in multivariate regressions. This follows easily by first partialling out any additional regressors from both x_i and y_i . All the results above also hold for these residuals.

3. Compare the clustered standard errors.

   Note that the standard errors clustered by firm are different from the classical OLS standard errors. On the other hand, the standard errors clustered by year does not change much with respect to the conventional ones. This suggest that the residual and the independent variable both contain a firm effect, but no year effect. The last type of cluster is year and firm, this is a two-way clustering. Note that the residuals in this case are not far from the ones computed using only firms clusters, then this reinforces the idea that there is mainly a firm effect that relates the residuals to the the solvency rate of an observation.

Figure 1: Scatterplot with regression line

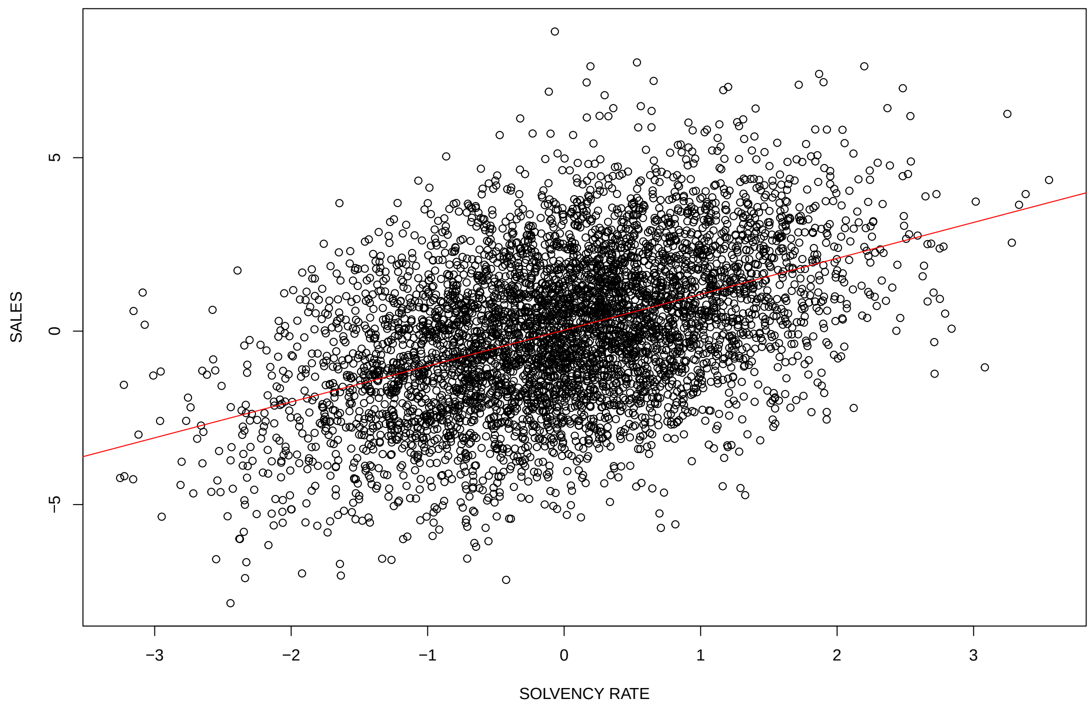

Figure 2: OLS results with classical standard errors

```
Call:
lm(formula = sandw$y ~ sandw$x)

Residuals:
    Min      1Q  Median      3Q     Max 
-6.7611 -1.3680 -0.0166  1.3387  8.6779 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.02968    0.02836   1.047    0.295    
sandw$x      1.03483    0.02858  36.204   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.005 on 4998 degrees of freedom
Multiple R-squared:  0.2078,    Adjusted R-squared:  0.2076 
F-statistic:  1311 on 1 and 4998 DF,  p-value: < 2.2e-16
```

Figure 3: OLS results with robust standard errors

```
> coeftest(est1, vcov = vcovHC(est1, "HC1"))

t test of coefficients:

            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 0.029680   0.028361  1.0465   0.2954    
sandw$x     1.034833   0.028395 36.4440   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 4: Standard errors clustered by groups: (i) firm, (ii) years and (ii) firm*year

	classical	Firm-cluster	Year-cluster	Firm*Year-cluster
(Intercept)	0.02835932	0.06701270	0.02338672	0.06506392
x	0.02858329	0.05059573	0.03338891	0.05355802

Standard Normal Probabilities

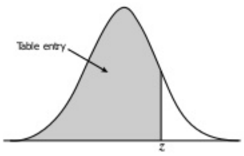

Table entry for z is the area under the standard normal curve to the left of z.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998