Econometrics – 616
Subal Kumbhakar
Answer Key for Assignment 1

1. For your first regression you should have gotten something similar to the following:

Valid cases:	185	Dependent variable:	FDr
Missing cases:	0	Deletion method:	None
Total SS:	165.132	Degrees of freedom:	180
R-squared:	0.134	Rbar-squared:	0.115
Residual SS:	142.934	Std error of est:	0.891
F (4,180):	6.988	Probability of F:	0.000
Durbin-Watson:	2.062		

Variable	Estimate	Standard Error	t-value	Prob >|t|	Standardized Estimate	Cor with Dep Var
Int.	-0.231943	0.125614	-1.846468	0.066	---	---
lagp	0.016065	0.020034	0.801897	0.424	0.057370	0.019556
fdlagy	18.380743	5.758921	3.191699	0.002	0.231255	0.266486
fdlagr	0.237455	0.074070	3.205812	0.002	0.236410	0.253725
fd2lagr	-0.154017	0.072538	-2.123248	0.035	-0.153180	-0.076320

And for your auxiliary regressions you should have obtained:

Valid cases:	185	Dependent variable:	FDR
Missing cases:	0	Deletion method:	None
Total SS:	165.132	Degrees of freedom:	181
R-squared:	0.131	Rbar-squared:	0.117
Residual SS:	143.445	Std error of est:	0.890
F(3,181):	9.121	Probability of F:	0.000
Durbin-Watson:	2.047		

Variable	Estimate	Standard Error	t-value	Prob >|t|	Standardized Estimate	Cor with Dep Var
Int	-0.158011	0.085232	-1.853896	0.065	---	---
FDLAGY	17.475550	5.641636	3.097603	0.002	0.219866	0.266486
FDLAGR	0.243729	0.073583	3.312290	0.001	0.242656	0.253725
FD2LAGR	-0.147164	0.071962	-2.045018	0.042	-0.146364	-0.076320

Valid cases:	185	Dependent variable:	LAGP
Missing cases:	0	Deletion method:	None
Total SS:	2105.973	Degrees of freedom:	181
R-squared:	0.060	Rbar-squared:	0.045
Residual SS:	1978.563	Std error of est:	3.306
F(3,181):	3.885	Probability of F:	0.010
Durbin-Watson:	0.395		

Variable	Estimate	Standard Error	t-value	Prob >|t|	Standardized Estimate	Cor with Dep Var
Int	4.602092	0.316544	14.538566	0.000	---	---
FDLAGY	-56.346248	20.952572	-2.689228	0.008	-0.198509	-0.169888
FDLAGR	0.390527	0.273282	1.429025	0.155	0.108874	0.095533
FD2LAGR	0.426592	0.267261	1.596165	0.112	0.118805	0.137823

Valid cases:	185	Dependent variable:	EHAT
Missing cases:	0	Deletion method:	None
Total SS:	143.445	Degrees of freedom:	183
R-squared:	0.004	Rbar-squared:	-0.002
Residual SS:	142.934	Std error of est:	0.884
F(1,183):	0.654	Probability of F:	0.420
Durbin-Watson:	2.062		

Variable	Estimate	Standard Error	t-value	Prob >|t|	Standardized Estimate	Cor with Dep Var
VHAT	0.016065	0.019869	0.808552	0.420	0.059663	0.059663

Notice that the estimate of VHAT is the same as the estimate of lagp from the first regression that you ran. Checking the residuals is up to you. But this should give you a taste of the FWL Theorem. See me with any questions about this if you are still not convinced (Chris not Subal).

2. The answer to this question is very algebra intensive. Here goes: $\sum \hat{v}_{2i} = 0, \sum \hat{v}_{3i} = 0, \sum \hat{w}_i = 0 \sum \hat{v}_{2i}\hat{w}_i = 0, \sum \hat{v}_{3i}\hat{w}_i = 0$ which all follow from the Gauss-Markov assumptions from the 4 regressions. We will use these properties to derive a relationship b/w the coefficients from the first regression and the last. $\sum \hat{v}_{2i}Y_i = \hat{\delta}_2 \sum \hat{v}_{2i}^2 + \hat{\delta}_3 \sum \hat{v}_{2i}\hat{v}_{3i}$ follows if we multiply the last regression by the residuals from the 2nd regression and sum over all observations. Next we will substitute in the definition of the residuals from the 3rd regression. $\hat{v}_{3i} = X_{3i} - \hat{c} - \hat{d}X_{2i}$ which leads to $\sum \hat{v}_{2i}Y_i = \hat{\delta}_2 \sum \hat{v}_{2i}^2 + \hat{\delta}_3 \hat{d} \sum \hat{v}_{2i} X_{2i}$ using the

above properties as well as $\sum \hat{v}_{2i} X_{3i} = \sum \hat{v}_{3i} X_{2i} = 0$. Now we will substitute the definition of X_{2i} into our equality and simplify. This gives us the following: $\sum \hat{v}_{2i} Y_i = \left(\hat{\delta}_2 - \hat{\delta}_3 \hat{d}\right) \sum \hat{v}_{2i}^2 \Rightarrow \sum \hat{v}_{2i} Y_i \Big/ \sum \hat{v}_{2i}^2 = \left(\hat{\delta}_2 - \hat{\delta}_3 \hat{d}\right)$ and from our first regression we know that $\hat{\beta}_2 = \sum \hat{v}_{2i} Y_i \Big/ \sum \hat{v}_{2i}^2 \Rightarrow \hat{\beta}_2 = \hat{\delta}_2 - \hat{\delta}_3 \hat{d}$ and using a similar display of algebra it may be shown that $\hat{\beta}_3 = \hat{\delta}_3 - \hat{\delta}_2 \hat{b}$. See me (Chris) if you cannot figure out the algebra for this one.

3. To show that the slope coefficient can be interpreted as a residual interpretation we will first need to regress X on a constant. If we do this we get the following: $\hat{X}_i = \overline{X} \quad \hat{v}_i = X_i - \hat{X}_i = X_i - \overline{X}$ and now if we regress Y on the residuals we get the following: $\hat{\beta} = \sum \hat{v}_i Y_i \Big/ \sum \hat{v}_i^2 = \sum \left(X_i - \overline{X}\right) Y_i \Big/ \sum \left(X_i - \overline{X}\right)^2 = \hat{\beta}_{OLS}$ The intercept can also be interpreted from a residual point of view as well. We regress a vector of ones on our X variable to get $\hat{\beta} = \sum X_i \Big/ \sum X_i^2 \quad \hat{w}_i = 1 - \left(n\overline{X} \Big/ \sum X_i^2\right) X_i$ and if we regress Y on these residuals we get the following: $\hat{\alpha} = \dfrac{n\overline{Y} - n\overline{X} \sum X_i Y_i \Big/ \sum X_i^2}{n\left[1 - \left(\sum X_i\right)^2 \Big/ n\left(\sum X_i^2\right)\right]} = \dfrac{n\overline{Y} - n\overline{X} \sum X_i Y_i \Big/ \sum X_i^2}{n\left[\sum \left(X_i - \overline{X}\right)^2 \Big/ \sum X_i^2\right]}$ and we can simplify this to get $\hat{\alpha} = \dfrac{\overline{Y}\left(\sum \left(X_i - \overline{X}\right)^2 + n\overline{X}^2 \overline{Y}\right) - \overline{X} \sum X_i Y_i}{\sum \left(X_i - \overline{X}\right)^2}$ and still simplifying more we get $\hat{\alpha} = \overline{Y} - \overline{X} \dfrac{\sum X_i Y_i - n\overline{X}\overline{Y}}{\sum \left(X_i - \overline{X}\right)^2} = \overline{Y} - \overline{X} \hat{\beta}_{OLS} = \hat{\alpha}_{OLS}$ so the intercept also has a residual interpretation although it was a little trickier to get. To answer the third part of the question we take the variance of our original definition of the slope coefficients as follows: $V\left(\hat{\beta}\right) = V\left(\sum \hat{v}_i Y_i \Big/ \sum \hat{v}_i^2\right) = \sigma^2 \Big/ \sum \hat{v}_i^2 = \sigma^2 \Big/ \sum \left(X_i - \overline{X}\right)^2$ which is the definition of the variance for the slope coefficient from a simple regression. As for the intercept we use the same procedure: $V(\hat{\alpha}) = V\left(\sum \hat{w}_i Y_i \Big/ \sum \hat{w}_i^2\right) = \sigma^2 \Big/ \sum \hat{w}_i^2 = \dfrac{\sigma^2 \sum X_i^2}{n \sum \left(X_i - \overline{X}\right)^2}$ which is also the definition of the variance for the intercept from the simple regression.

4. The Intuitive way: When you add more variable to a regression there is no way to negatively explain something, either you have explanatory power or you do not. If you have explanatory power with an additional regressor then your residual sum of squares must become smaller as you are explaining more.
   The Mathematical Way: OLS tries to minimize the sum of squares by choosing the appropriate coefficients for the model.
   $RSS_{OLS} = \min_{\beta} \sum \left(Y_i - \alpha - \beta_1 X_{1i} - \cdots \beta_K X_{Ki}\right)^2$ and so if I add more regressors to the model then under this framework I could in essence make their parameter estimates zero and the original parameter estimates those obtained by OLS which would mean that the residual sum of squares would be minimized if the true parameter estimates of the additional variable were really zero. However if they were non-zero then the residual sum of squares would have to be smaller with OLS since it is minimizing the sum of squared errors for the old variable and the new ones added to the model. There is an even more mathematical way to show this with matrix algebra.
   $(1) \quad Y = X\beta + u \quad RSS_1 = Y' M_x Y$
   $(2) \quad Y = X\beta + Z\gamma + v \quad RSS_2 = Y' M_{X,Z} Y$ Where the M matrix is an idempotent matrix that projects orthogonally off of the space spanned by M's sub-script. We can use the results of the FWL theorem to complete the proof. By the FWL Theorem we know that $M_X Y = M_X Z\gamma + v \quad RSS_2 = Y' M_X M_{M_X Z} M_X Y$
   $RSS_2 = Y' M_X \left(I - M_X Z\left(Z' M_X Z\right)^{-1} Z' M_X\right) M_X Y = Y' M_X Y - Y' M_X P_{M_X Z} M_X Y$
   $RSS_2 = RSS_1 - \left\| P_{M_X Z} Y \right\|^2 \le RSS_1$ where the final result holds due to the fact that the P matrix is idempotent and the squared norm of P is always positive.

5. If we regress the residuals from a two variable regression on a constant and the two variables the coefficient of determination will be zero since the residuals are uncorrelated with any of the exogenous variables from the regression. Having a positive coefficient of determination would go against one of the Gauss-Markov assumptions. If we regress Y on its fitted value and a constant then we will see that the intercept estimate will be very near zero and the slope coefficient will be approximately one. The coefficient of determination will be the same from the model that yielded the fitted Y values. If we reverse the procedure and regress Y on the residuals from the regression we will see that the estimated intercept is the mean value of Y and that the slope estimate is one once again. The coefficient of determination will not be the same though. It will be equal to one minus the old coefficient of determination because now we are substituting the residual sum of squares for the explained sum of squares for this regression. It needs to be backwards to preserve the relationship between Y and its fitted value. We will need math to answer the next part of this question.
   $Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \quad \hat{\beta}_3 = \sum \hat{r}_{3_i} Y_i \Big/ \sum \hat{r}_{3i}^2$
   $e = \lambda + \delta_1 X_1 + \delta_2 X_2 + \delta_3 X_3 \quad \hat{\delta}_3 = \sum \hat{r}_{3i} e_i \Big/ \sum \hat{r}_{3i}^2$

$e = Y - \alpha - \beta_1 X_1 - \beta_2 X_2 \quad \hat{\delta}_3 = \sum \hat{r}_{3i} \left(Y_i - \alpha - \beta_1 X_{1i} - \beta_2 X_{2i} \right) \Big/ \sum \hat{r}_{3i}^2$

$\hat{\delta}_3 = \sum \hat{r}_{3i} Y_i \Big/ \sum \hat{r}_{3i}^2 = \hat{\beta}_3 \quad \sum \hat{r}_{3i} = \sum \hat{r}_{3i} X_{1i} = \sum \hat{r}_{3i} X_{2i} = 0$

The r-hat in the equations is the residual from the regression of the new X variable on a constant and the pre-existing X variables. If we look at the definition for the variance of an estimator from a multiple regression (see Wooldridge, 2002, pg. 96) we can see that taking the variance of either one of the above estimators, since they are both the same, will yield the same formula from any undergraduate econometrics text book. We can also answer this question using partitioned matrix results. If you remember from class, I derived the formula for the slope coefficient for the simple linear regression model using a partitioned matrix. Refer to the notes if you don't remember. I can use the same argument here as I did there to show my result. In this case I will partition the matrix so that not only the constant but the other X's besides the new X are what I call X_1. And the coefficient vector is partitioned likewise so that the intercept and the pre-existing slope coefficients are now what I call b_2. Using this transformation and making the appropriate change in the A matrix the result will follow.