Econ 141: Solutions to Homework 6

Jon Schellenberg

Problem 1

Table is below. Note that the return variable, as given, is in percentage terms.

	(1)	(2)	(3)	(4)	(5)
	$return_t$	$return_t$	$return_t$	$\widehat{u}_t^2$	$\widehat{u}_t^2$
$return_{t-1}$	-0.0355	-0.0315	-0.0355	-0.257***	-0.213***
	(0.0257)	(0.0259)	(0.0359)	(0.0652)	(0.0649)
$return_{t-1}^2$		0.0138			0.150***
		(0.0102)			(0.0255)
Constant	0.0256	0.0102	0.0256	1.117***	0.949***
	(0.0271)	(0.0294)	(0.0273)	(0.0687)	(0.0737)
R^2	0.00126	0.00247	0.00126	0.0102	0.0325
RSS	1678.2	1676.2	1678.2	10768.0	10525.4
n	1511	1511	1511	1511	1511
F-stat	1.906	1.868	0.980	15.58	25.34
Standard Errors	Regular	Regular	Robust	Regular	Regular

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

(a) This coefficient is -0.0355, meaning that a 1% increase in the previous period's return will lead to a 0.036% drop in the current period's return. However, the coefficient is not statistically significant ($p = 0.168$ for the two-tailed test, so we fail to reject that it is 0 at the 5% level). Thus, we fail to conclude that current stock returns can be predicted from past returns.

(b) With this model, I would need to test that the coefficients on $return_{t-1}$ and $return_{t-1}^2$ are *jointly* 0 (with the alternative being that at least one is not 0). Doing this in Stata, we get an F-stat of 1.87, and the associated p-value is 0.155, so again, we'd fail to reject at the 5% level and would not be able to reject the efficient market hypothesis.[1]

(c) I'm using the regression (2) for my RESET test, namely:

$$return_t = \beta_0 + \beta_1 return_{t-1} + \beta_2 return_{t-1}^2 + u_t$$

I obtained the predicted values from the previous regression, and then I estimated the same regression while also including the previous squared predicted values on the right hand side. Formally, I ran the following regression for my RESET test:

$$return_t = \alpha_0 + \alpha_1 return_{t-1} + \alpha_2 return_{t-1}^2 + \alpha_3 \widehat{return}_t^2 + \epsilon_t$$

Formally, the RESET test has been written out below. You can either use a t-test or an F-test, since you're doing a two-tailed test with one restriction. I have written out both here, either would be acceptable.

(i) $H_0 : \alpha_3 = 0$

[1] Alternatively, instead of using the p-value, we could find the critical value in the table. Here, $F_{0.05,2,1507} \approx F_{0.05,2,1000} = 3.00$. Recall we reject iff $F > F_{0.05,2,1507}$, and since $1.87 < 3.00$, we would fail to reject.

(ii) $H_1 : \alpha_3 \neq 0$

(iii) Assuming the null is true, I know that $t = \frac{\widehat{\alpha}_3 - 0}{se(\widehat{\alpha}_3)} \sim T_{n-(k+1)} \sim T_{1507}$. We could also use the F-stat here - I also know that if my null is true,
$F \equiv \left(\boldsymbol{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)' \left(\boldsymbol{R}\widehat{\boldsymbol{\Sigma}}_{\widehat{\boldsymbol{\beta}}}\boldsymbol{R}'\right)^{-1} \left(\boldsymbol{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)/q \sim F_{q,n-(k+1)} \sim F_{1,1507}$

(iv) In our case, our t-stat is 3.64 and our F-stat is 13.25. The p-values for both tests are identical with $p = 0.0003$

(v) I can use three different rejection rules:

I. I will reject iff $p < 0.05$.

II. I will reject iff $|t| > |t_{0.05,1507}|$, the two-tailed critical value for the t-test. From our t-table, we find that $t_{0.05,1507} = 1.96$

III. I will reject iff $F > F_{0.05,1,1507}$, the critical value for the F test. from our F-table, we find $F_{0.05,1,1507} = 3.85$.

(vi) Testing any of my three rejection rules (we will come to the same conclusion in all cases):

I. $p = 0.0003 < 0.05$, so we reject H_0.

II. $|t| = |3.64| > |1.96|$, so we reject H_0.

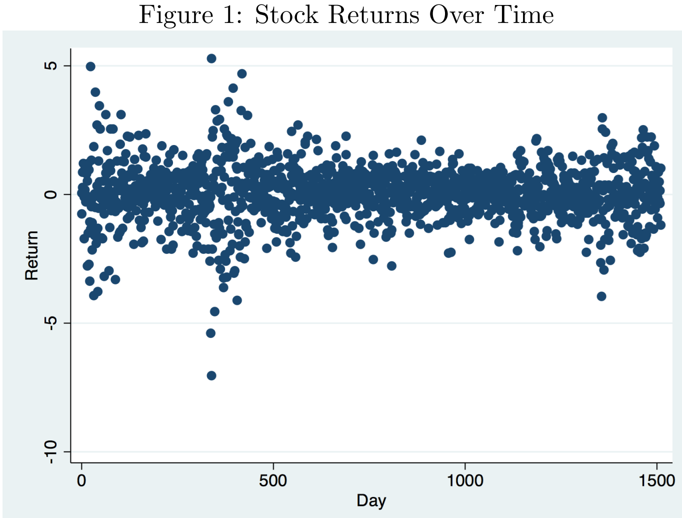
III. $F = 13.25 > 3.85$, so we reject H_0

(vii) Therefore, we conclude that $\alpha_3 \neq 0$, and hence, our nonlinear model is misspecified.

(d) Note that the coefficient estimates are the same in regressions (1) and (3), but the standard errors are different - they are larger in the latter. Thus, since our coefficient was not statistically significant before, we know our coefficient won't be significant in regression (3) either.

(e) Graph below. There really doesn't seem to be much of a trend, which matches our lack of statistically significant findings from our previous regressions.

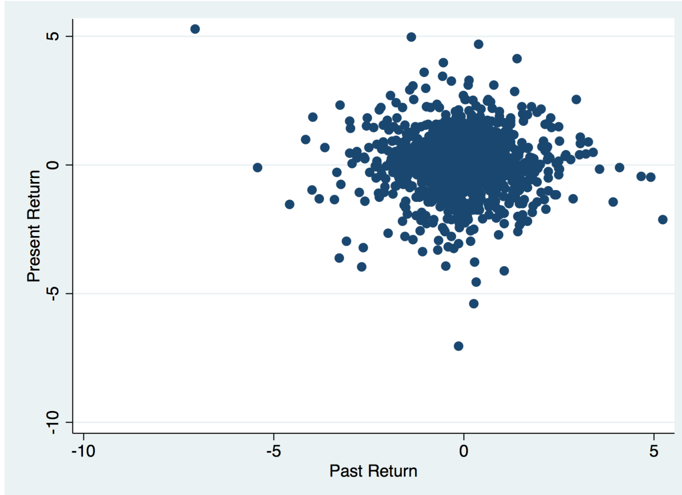
Figure 1: Stock Returns Over Time



(f) Recall that you can plot $return_t$ vs $return_{t-1}$ to look for heteroskedasticity. If the spread of points around the best fit line is relatively uniform, the errors are likely homoskedastic, and if they're not, the data may be heteroskedastic. You don't need to include this graph for the problem, I just thought it might be helpful to see.
Clearly, the spread of points is not uniform around the best fit line - the "eye" test suggests that our errors are heteroskedastic Now, let's formally test for it - to see how this was done in Stata, please see the uploaded do file. I basically took the residuals from regression (1) $\widehat{u}_t$, squared them, and ran the following regression:

$$\widehat{u}_t^2 = \gamma_0 + \gamma_1 return_{t-1} + \epsilon_t$$

Figure 2: Current Stock Returns Vs Past Returns



Formally, the Breusch-Pagan test has been written out below. You can either use a t-test or an F-test, since you're doing a two-tailed test with one restriction. I have written out both here, either would be acceptable.

(i) $H_0 : \gamma_1 = 0$

(ii) $H_1 : \gamma_1 \neq 0$

(iii) Assuming the null is true, I know that $t = \frac{\widehat{\gamma}_1 - 0}{se(\widehat{\gamma})} \sim T_{n-(k+1)} \sim T_{1509}$. We could also use the F-stat here - I also know that if my null is true,
$F \equiv \left(\boldsymbol{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)' \left(\boldsymbol{R}\widehat{\boldsymbol{\Sigma}}_{\widehat{\boldsymbol{\beta}}}\boldsymbol{R}'\right)^{-1} \left(\boldsymbol{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)/q \sim F_{q,n-(k+1)} \sim F_{1,1509}$

(iv) In our case, our t-stat is -3.95 and our F-stat is 15.58. The p-values for both tests are identical with $p = 0.0001$

(v) I can use three different rejection rules:

I. I will reject iff $p < 0.05$.

II. I will reject iff $|t| > |t_{0.05,1509}|$, the two-tailed critical value for the t-test. From our t-table, we find that $t_{0.05,1509} = 1.96$

III. I will reject iff $F > F_{0.05,1,1509}$, the critical value for the F test. from our F-table, we find $F_{0.05,1,1509} = 3.85$.

(vi) Testing any of my three rejection rules (we will come to the same conclusion in all cases):

I. $p = 0.0001 < 0.05$, so we reject H_0.

II. $|t| = |-3.95| > |1.96|$, so we reject H_0.

III. $F = 15.58 > 3.85$, so we reject H_0

(vii) Therefore, we conclude that $\gamma_1 \neq 0$, and hence, our errors are heteroskedastic.

(g) The slope $\widehat{\gamma}_1 = -0.257$, which means that a 1% increase in our past period's return will lead to a 0.257 unit decrease in our squared residuals. This is evidence that our mean squared error, and thus, variance of our true errors, is not constant (as it depends on lagged returns), which is why we conclude that our errors are heteroskedastic.

(h) We would use the results from regression (3) instead of regression (1). Our conclusions will still be the same, although our statistical significance will be (even) lower than before.

(i) Done in Stata. We see that we get some negative predicted values for (4) but not for (5). We would hope that *none* of these would be negative (they're predicting a squared residual, which *must* be nonnegative), so it means your regression may be a poor predictor of our squared residuals. Also, conceptually, negative weights make no sense, so we *cannot* use these predicted values as weights for regression (4). We can for (5) though, since all the weights are postiive.

(j) Done in Stata. See table below to compare our OLS and WLS estimates.

	(1) OLS	(2) WLS
$return_{t-1}$	-0.0355	-0.0182
	(0.0257)	(0.0300)
Constant	0.0256	0.0222
	(0.0271)	(0.0268)
R^2	0.00126	0.000232
RSS	1678.2	1586.1
n	1511	1511
F-stat	1.906	0.368
Standard Errors	Regular	Robust

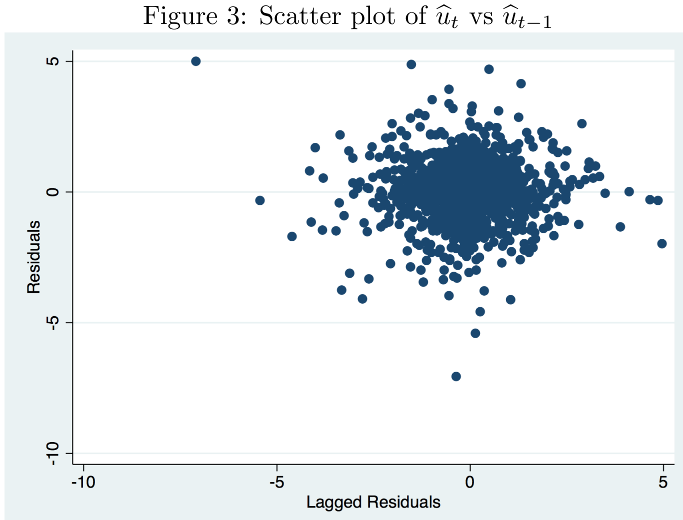
Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The magnitude of the coefficient is pretty different than our OLS estimate (it's now -0.018). Again though, our slope coefficient is not statistically significant ($p = 0.55$ for our two-tailed test that the slope is 0), so we fail to reject the Efficient Market Hypothesis.

(k) ...Probably not? We can't formally tell here. Here's a graph that plots residuals in time t vs the residuals in time $t+1$:

Figure 3: Scatter plot of $\widehat{u}_t$ vs $\widehat{u}_{t-1}$



It doesn't look like there's much of a trend in the previous plot, indicating that the errors are probably not correlated.

Unfortunately, we can't formally test for autocorrelation here. *We cannot use the Durbin-Watson test when including lagged dependent variables as regressors*, which we did in our regressions.

Problem 2

For all of these, recall that our model in matrix form is as follows:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}$$

Thus, we can rewrite our estimator $\widetilde{\boldsymbol{\beta}}$ in terms of the true coefficients $\boldsymbol{\beta}$ and the true errors, as follows:

$$
\begin{aligned}
\widetilde{\boldsymbol{\beta}} &= \left(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{y} \\
\text{(subbing in for } \boldsymbol{y}\text{)} \quad &= \left(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}(\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{u}) \\
&= \boldsymbol{\beta} + \left(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{u}
\end{aligned}
$$

(a) Yes, it is unbiased. Recall that we assume that $\mathbb{E}[u_i] = 0 \Rightarrow \mathbb{E}[\boldsymbol{u}] = \mathbf{0}_n$ where $\mathbf{0}_n$ is an $(n \times 1)$ matrix.

$$
\begin{aligned}
\mathbb{E}[\widetilde{\boldsymbol{\beta}}] &= \mathbb{E}\left[\boldsymbol{\beta} + \left(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{u}\right] \\
\text{(fixed regressors assumption)} \quad &= \boldsymbol{\beta} + \left(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\underbrace{\mathbb{E}[\boldsymbol{u}]}_{=\mathbf{0}_n} \\
&= \boldsymbol{\beta}
\end{aligned}
$$

(b) Recall that by assumption, $\mathbb{V}[\boldsymbol{u}] = \boldsymbol{\Omega}$. Thus,

$$
\begin{aligned}
\mathbb{V}[\widetilde{\boldsymbol{\beta}}] &= \mathbb{V}\left[\boldsymbol{\beta} + \left(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{u}\right] \\
\text{(fixed regressors assumption)} \quad &= \left(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\underbrace{\mathbb{V}[\boldsymbol{u}]}_{=\boldsymbol{\Omega}}\left(\left(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\right)' \\
\text{(properties of transposes)} \quad &= \left(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\underbrace{\boldsymbol{\Omega}^{-1}\boldsymbol{\Omega}}_{=\boldsymbol{I}_n}\boldsymbol{\Omega}^{-1}\boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1} \\
&= \left(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}\underbrace{\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}}_{=\boldsymbol{I}_{k+1}} \\
&= \left(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}
\end{aligned}
$$

(c) Yes. There are two ways to approach this.

(i) Note that

$$
\begin{aligned}
\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X} &= \begin{bmatrix}\boldsymbol{x}_1 & \cdots & \boldsymbol{x}_n\end{bmatrix}\begin{bmatrix}\frac{1}{\sigma_1^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2^2} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \frac{1}{\sigma_n^2}\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_1' \\ \vdots \\ \boldsymbol{x}_n'\end{bmatrix} \\
&= \sum_{i=1}^{n}\frac{1}{\sigma_i^2}\boldsymbol{x}_i\boldsymbol{x}_i' \\
&= n \times \left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\boldsymbol{x}_i\boldsymbol{x}_i'\right)
\end{aligned}
$$

If using the lecture assumptions, you can note that as n grows, this $((k+1)\times(k+1))$ matrix is growing infinitely large (as in, all of its elements are approaching infinity) - this is because you'd expect the second term, which is a "sample average", to stay relatively constant, except it's being blown up by n. Thus, $\mathbb{V}[\widetilde{\boldsymbol{\beta}}] = \left(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}$ is shrinking to $\mathbf{0}_{(k+1)\times(k+1)}$, a $\mathbf{0}_{(k+1)\times(k+1)}$ matrix of zeros. Thus,

- $\mathbb{E}[\widetilde{\boldsymbol{\beta}}] = \boldsymbol{\beta}$
- $\mathbb{V}[\widetilde{\boldsymbol{\beta}}] \xrightarrow{p} \mathbf{0}_{(k+1)\times(k+1)}$

Thus, $\widetilde{\boldsymbol{\beta}}$ is consistent.

(ii) If you use the textbook assumptions, by i.i.d. data, you can use LLN and LIE to prove consistency. Recall what we showed about $\widetilde{\boldsymbol{\beta}}$ earlier:

$$\begin{aligned}\widetilde{\boldsymbol{\beta}} &= \boldsymbol{\beta} + \left(\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{u} \\ \text{(multiplying by 1)} \quad &= \boldsymbol{\beta} + \left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}\left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{u}\right)\end{aligned}$$

Consider the second term in the product for a second. If you did the matrix math out (similar to what we did above):

$$\begin{aligned}\frac{1}{n}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{u} &= \frac{1}{n}\begin{bmatrix}\boldsymbol{x}_1 & \cdots & \boldsymbol{x}_n\end{bmatrix}\begin{bmatrix}\frac{1}{\sigma_1^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2^2} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \frac{1}{\sigma_n^2}\end{bmatrix}\begin{bmatrix}u_1 \\ \vdots \\ u_n\end{bmatrix} \\ &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\boldsymbol{x}_i u_i \\ \text{(Law of Large Numbers)} \quad &\xrightarrow{p} \mathbb{E}\left[\frac{1}{\sigma_i^2}\boldsymbol{x}_i u_i\right] \\ \text{(Law of Iterated Expectations)} \quad &= \mathbb{E}\left[\frac{1}{\sigma_i^2}\boldsymbol{x}_i \underbrace{\mathbb{E}[u_i|\boldsymbol{x}_i]}_{=0}\right] \\ &= \mathbb{E}\left[\frac{1}{\sigma_i^2}\boldsymbol{x}_i \cdot 0\right] \\ &= \mathbf{0}_{k+1}\end{aligned}$$

The last part is true because in the textbook, we assume $\mathbb{E}[u_i|\boldsymbol{x}_i] = 0$.

Additionally, similar to part (i), we see that

$$\begin{aligned}\left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right) &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\boldsymbol{x}_i\boldsymbol{x}_i' \\ \text{(by LLN)} \quad &\xrightarrow{p} E\left[\frac{1}{\sigma_i^2}\boldsymbol{x}_i\boldsymbol{x}_i'\right]\end{aligned}$$

This matrix is of full rank (assuming the $\sigma_i^2 \neq 0$ and none of the elements in $\boldsymbol{x}_i$ are 0 for all i), so

$$\begin{aligned}\widetilde{\boldsymbol{\beta}} &= \boldsymbol{\beta} + \left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}\left(\frac{1}{n}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{u}\right) \\ &\xrightarrow{p} \boldsymbol{\beta} + \mathbb{E}\left[\frac{1}{\sigma_i^2}\boldsymbol{x}_i\boldsymbol{x}_i'\right]^{-1}\mathbf{0}_{k+1} \\ &= \boldsymbol{\beta}\end{aligned}$$

Again, like we showed earlier, $\widetilde{\boldsymbol{\beta}}$ is consistent.

Problem 3

(a) I'm assuming all of my OLS assumptions hold from my original regression except for zero correlation across my errors. Therefore, I'm still assuming that $\mathbb{E}[u_t] = 0$ for all t. That's all there is to this problem - I was assuming it originally, and it's still true here.

(b) I'm assuming $Var(u_t) = \sigma_u^2$, $Cov(u_{t-1}, \epsilon_t) = 0$, and $Var(\epsilon_t) = \sigma_\epsilon^2$ for all t.

$$\begin{aligned} Var(u_t) &= Var\left(\rho u_{t-1} + \epsilon_t\right) \\ &= \rho^2 \underbrace{Var(u_{t-1})}_{=\sigma_u^2} + \underbrace{Var(\epsilon_t)}_{=\sigma_\epsilon^2} + 2\rho \underbrace{Cov\left(u_{t-1}, \epsilon_t\right)}_{=0} \\ &= \rho^2 Var(u_t) + \sigma_\epsilon^2 \end{aligned}$$

Moving $Var(u_t)$ to the left hand side of the equation, we get

$$\begin{aligned} (1-\rho^2)Var(u_t) &= \sigma_\epsilon^2 \\ \Rightarrow Var(u_t) &= \frac{\sigma_\epsilon^2}{1-\rho^2} \end{aligned}$$

(c) I'm assuming $Cov(\epsilon_t, \epsilon_s) = 0$ for all $t \neq s$. I'm also going to assume that $Cov(u_{t-s}, \epsilon_t) = 0$ for all $s \geq 1$ (i.e., all current and future ϵs are uncorrelated with past us). I'm also still assuming my u_ts are homoskedastic, so $Var(u_t) = \sigma_u^2$ for all t.

Note if $s = 0$, then we just have $Cov(u_t, u_t) = Var(u_t)$, which we solved for in the previous problem. I'm going to consider $s \geq 1$.

$$\begin{aligned} Cov(u_t, u_{t-s}) &= Cov(\rho u_{t-1} + \epsilon_t, u_{t-s}) \\ &= \rho Cov(u_{t-1}, u_{t-s}) + \underbrace{Cov(\epsilon_t, u_{t-s})}_{=0} \\ \text{(Stop here if } s=1\text{)} \quad &= \rho Cov(u_{t-1}, u_{t-s}) \\ &= \rho Cov(\rho u_{t-2} + \epsilon_{t-1}, u_{t-s}) \\ &= \rho^2 Cov(u_{t-2}, u_{t-s}) + \underbrace{Cov(\epsilon_{t-2}, u_{t-s})}_{=0} \\ \text{(Stop here if } s=2\text{)} \quad &= \rho^2 Cov(u_{t-2}, u_{t-s}) \\ &\vdots \\ &= \rho^s \underbrace{Cov(u_{t-s}, u_{t-s})}_{=Var(u_t)} \\ &= \rho^s Var(u_t) \\ \text{(from part (b))} \quad &= \rho^s \left(\frac{\sigma_\epsilon^2}{1-\rho^2}\right) \end{aligned}$$

Notice the above expression can be generalized for all s where $0 \leq s \leq t$.

This is like an extension to what we did with AR(1) processes.

- Recall that when we're testing for autocorrelation in the Durbin-Watson test, our *alternative* hypothesis is that $Corr(u_t, u_{t-1}) = \rho \neq 0$.
- The above process is a slightly different model of autocorrelation - we're saying that consecutive errors predict one another with some random error. This modeling assumption allows us to derive the entire variance/covariance structure of our errors, which is what we just did.