Introductory Econometrics

Tutorial 10

<u>PART A:</u> To be done before you attend the tutorial. The solutions will be made available at the end of the week.

1. b. If FDI increases by 1%, gdp increases by approximately 0.22%, the amount of bank credit remaining constant. [*correct answer*]

2.

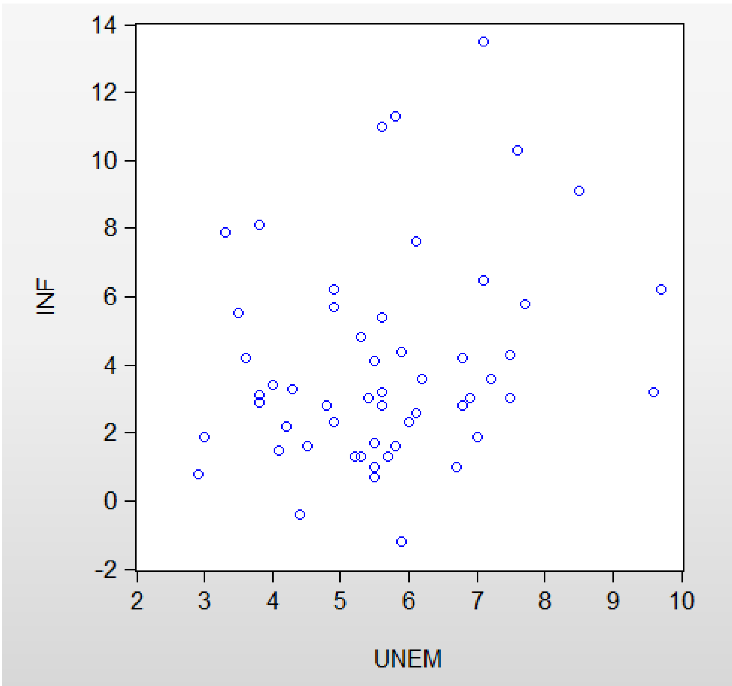
Y4 is white noise

Y1 is AR(1)

Y2 is AR(2)

Y3 is AR(3).

(a)



No there isn't. If anything, there seems to be a vague positive relationship.

(b)

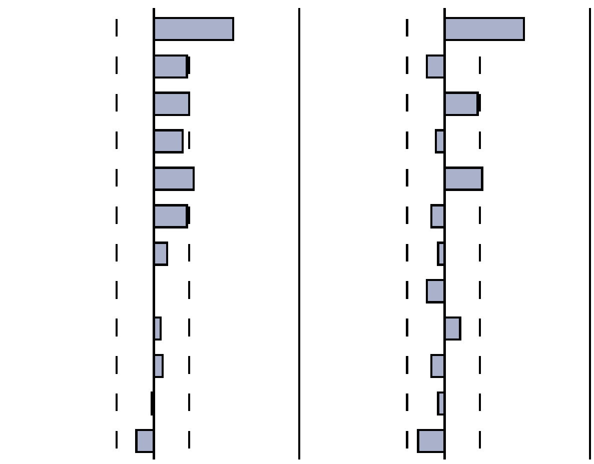
Dependent Variable: INF
Method: Least Squares
Sample: 1948 2003
Included observations: 56

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.053565	1.547957	0.680617	0.4990
UNEM	0.502378	0.265562	1.891752	0.0639

R-squared	0.062154	Mean dependent var	3.883929
Adjusted R-squared	0.044786	S.D. dependent var	3.040381
S.E. of regression	2.971518	Akaike info criterion	5.051084
Sum squared resid	476.8157	Schwarz criterion	5.123418
Log likelihood	-139.4304	Hannan-Quinn criter.	5.079128
F-statistic	3.578726	Durbin-Watson stat	0.801482
Prob(F-statistic)	0.063892		

Sample: 1948 2003
Included observations: 56

Autocorrelation Partial Correlation



	AC	PAC
1	0.568	0.568
2	0.247	-0.112
3	0.262	0.253
4	0.218	-0.050
5	0.285	0.273
6	0.254	-0.095
7	0.096	-0.038
8	0.012	-0.109
9	0.064	0.114
10	0.069	-0.090
11	-0.006	-0.040
12	-0.117	-0.175

$$
\begin{aligned}
inf_t &= \beta_0 + \beta_1 unem_t + u_t \\
u_t &= \rho u_{t-1} + e_t \\
H_0 &: \rho = 0 \\
H_1 &: \rho \neq 0 \\
&\text{Estimate of the aux. regression} \\
\hat{u}_t &= 2.046 - 0.381 unem_t + 0.643 \hat{u}_{t-1} + \hat{e}_t \\
R^2_{\hat{u}} &= 0.379, \quad n = 55 \\
BG &= nR^2_{\hat{u}} \sim \chi^2_1 \quad \text{under } H_0 \\
BG_{calc} &= 55 \times 0.379 = 20.845 > BG_{crit} = 3.84 \\
&\text{We reject the null and conclude the error are serially correlated}
\end{aligned}
$$

The partial correlogram of residuals shows a very significant AR(1), but borderline at lags

3 and 5 as well. We estimate the model with AR(1) errors.

$$\begin{aligned} inf_t &= \underset{(2.288)}{7.287} - \underset{(0.330)}{0.664} unem_t + \hat{u}_t \\ \hat{u}_t &= \underset{(0.095)}{0.782} \hat{u}_{t-1} + \hat{e}_t \end{aligned}$$

(c)

The FGLS estimates in the model with AR(1) error support the Phillips curve hypothesis.

Do not forget to bring your answers to PART A and a copy of the tutorial questions to your tutorial.

<u>Part B:</u> **This part will be covered in the tutorial. It is still a good idea to attempt these questions before the tutorial.**
The purpose of this tutorial is to use dummy variables when using time series data.

1. (a) *For this tutorial, I recommend that you ask assignment groups to sit together and do this together, and you just walk around and answer their questions. By now, in particular after the first assignment, they should be able to be pretty self-sufficient in transforming variables and running regressions in Eviews. If you feel that many are having similar issues, then interject and present a general explanation to the tutorial class. Just a suggestion, please feel free to ignore if you assess that baby-birds are not ready for unassisted flights yet.*

   In Eviews, "`series sat = (dow=7)`" *and* "`series sun = (dow=1)`" *will create dummy variables for Saturday and Sunday. Then* "`series wknd = sat + sun`" *creates a dummy for weekends. Of course there are several different ways to generate this.*

   (b) *As the question recommends, students are free to use any software that they know to get the best look at their data. Here is what I have done: I used vlookup to change the dow to names for each day, and then used pivot chart and pivot table to get a bar chart for aveload in each day of the week in Excel. I generated the scatter plot in Eviews.*

Column Labels						
mon	tue	wed	thu	fri	sat	sun
2393.923077	2412.076923	2450.076923	2475.884615	2447.307692	2286.192308	2236.576923

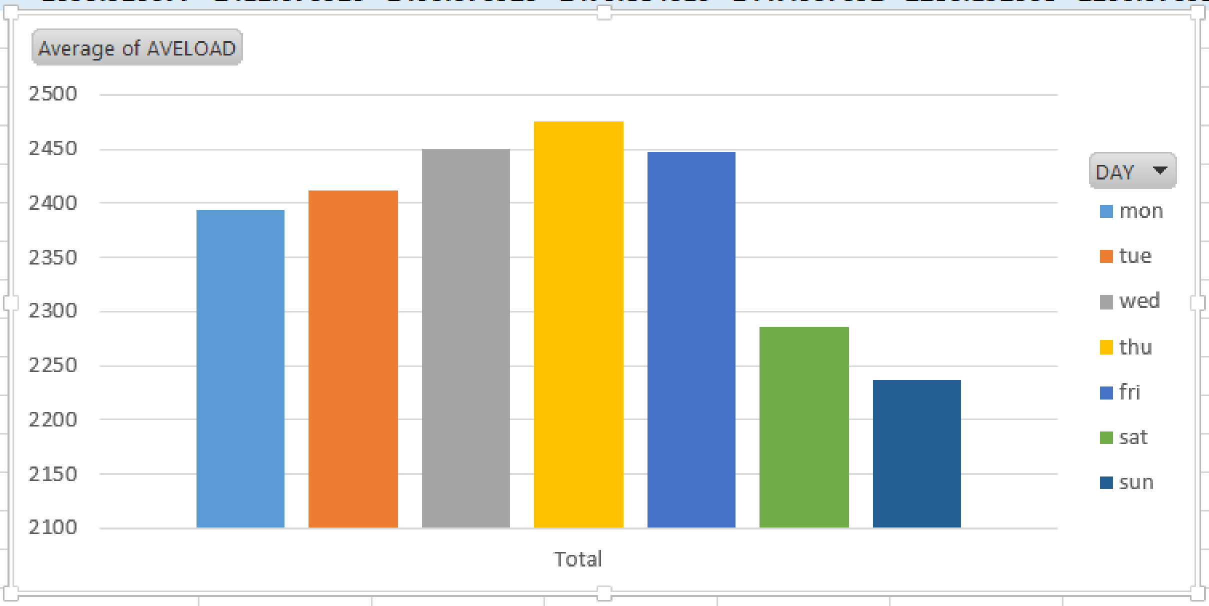


Figure 1:

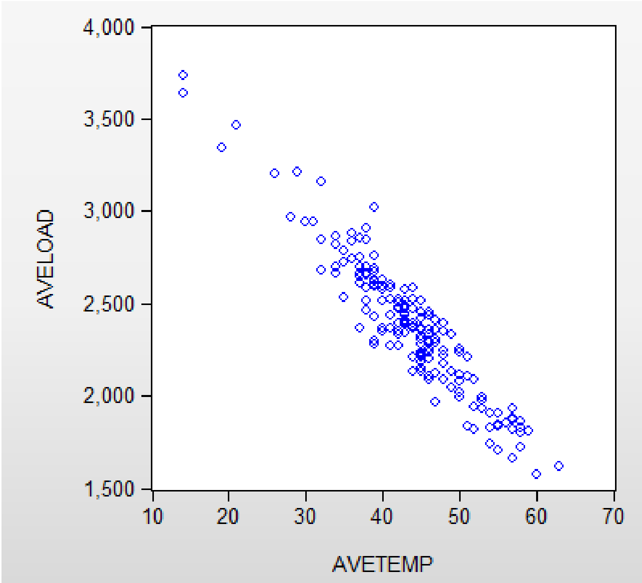
(c)

Dependent Variable: AVELOAD
Method: Least Squares
Sample: 10/01/2005 3/31/2006
Included observations: 182

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4332.783	41.79892	103.6578	0.0000
AVETEMP	-43.58037	0.940259	-46.34931	0.0000
WKND	-143.6717	16.35973	-8.782032	0.0000
PUBHOL	-121.5734	31.00930	-3.920546	0.0001

R-squared	0.927222	Mean dependent var	2386.005
Adjusted R-squared	0.925995	S.D. dependent var	365.8870
S.E. of regression	99.53527	Akaike info criterion	12.06063
Sum squared resid	1763494.	Schwarz criterion	12.13105
Log likelihood	-1093.518	Hannan-Quinn criter.	12.08918
F-statistic	755.9290	Durbin-Watson stat	0.942970
Prob(F-statistic)	0.000000		

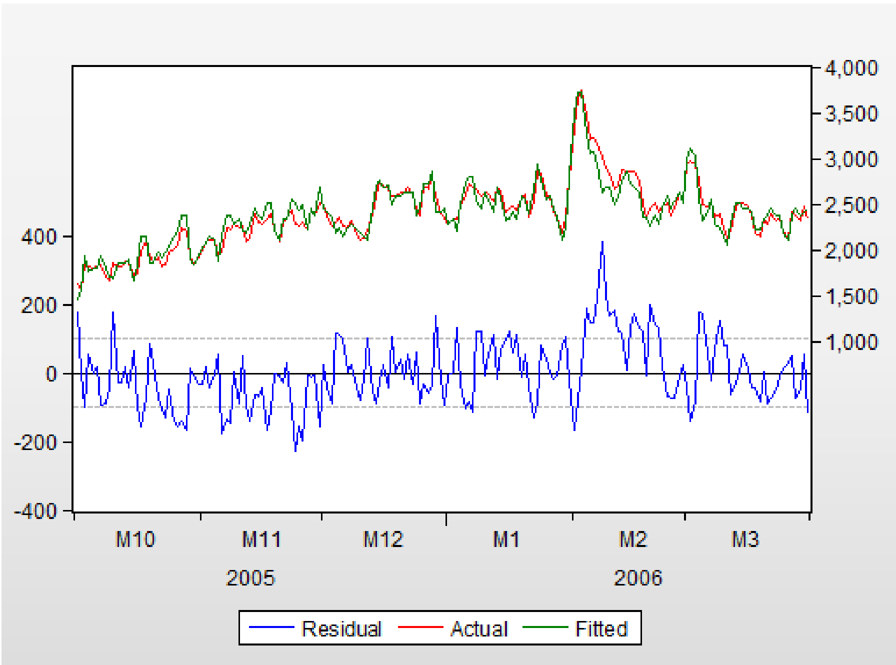
OLS will still be unbiased but no longer BLUE. Also, the OLS standard errors will not be correct and cannot be used for inference. So, the above model cannot be used



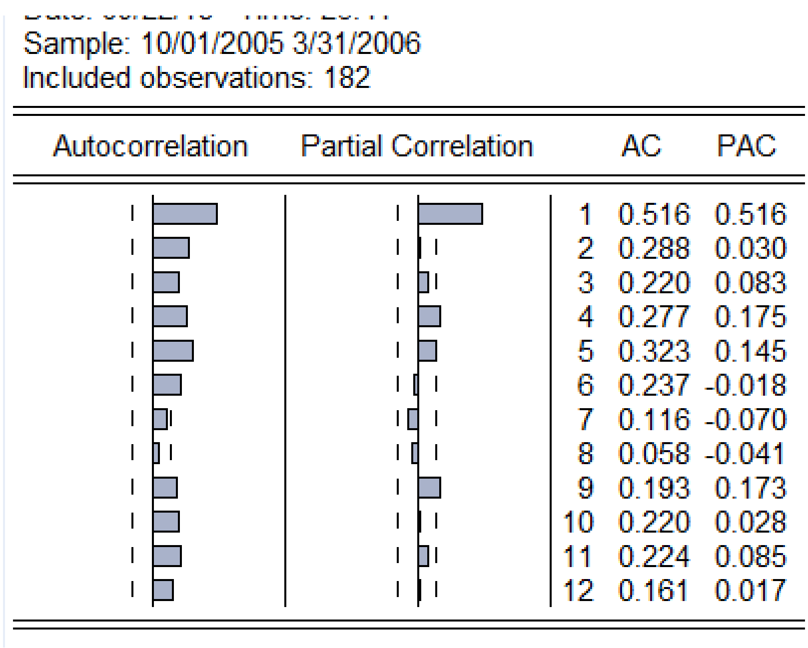
4,000
3,500
3,000
2,500
2,000
1,500
AVELOAD
10
20
30
40
50
60
70
AVETEMP

Figure 2:

(d)



400
200
0
-200
-400
4,000
3,500
3,000
2,500
2,000
1,500
1,000
M10
M11
M12
M1
M2
M3
2005
2006
Residual
Actual
Fitted



Sample: 10/01/2005 3/31/2006
Included observations: 182

Autocorrelation	Partial Correlation		AC	PAC
		1	0.516	0.516
		2	0.288	0.030
		3	0.220	0.083
		4	0.277	0.175
		5	0.323	0.145
		6	0.237	-0.018
		7	0.116	-0.070
		8	0.058	-0.041
		9	0.193	0.173
		10	0.220	0.028
		11	0.224	0.085
		12	0.161	0.017

$$
\begin{aligned}
aveload_t &= \beta_0 + \beta_1 avetemp_t + \beta_2 wknd_t + \beta_3 pubhol_t + u_t \\
u_t &= \rho_1 u_{t-1} + \rho_2 u_{t-2} + \rho_3 u_{t-3} + \rho_4 u_{t-4} + \rho_5 u_{t-5} + \rho_6 u_{t-6} + \rho_7 u_{t-7} + e_t \\
H_0 &: \rho_1 = \rho_2 = \rho_3 = \rho_4 = \rho_5 = \rho_6 = \rho_7 = 0 \\
H_1 &: \text{at least one of the AR parameters is not zero} \\
& \quad \text{Estimate of the aux. regression} \\
& \quad \texttt{uhat c avetemp wknd pubhol uhat(-1 to -7)} \\
R^2_{\hat{u}} &= 0.340, \qquad n_{\hat{u}} = n - 7 = 182 - 7 = 175 \\
BG &= n_{\hat{u}} \times R^2_{\hat{u}} \sim \chi^2_7 \qquad \text{under } H_0 \\
BG_{calc} &= 175 \times 0.340 = 59.5 > BG_{crit} = 14.07 \\
& \quad \text{We reject the null and conclude the error are serially correlated}
\end{aligned}
$$

The BG auxiliary regression only the first AR term is significant. The correlogram also shows a significant first order autocorrelation, with the 4th partial autocorrelation also close to being significant. We start with adding AR(1) errors and check the residual correlogram again to see if there is a need to consider a longer AR model for the errors. Note that unlike OLS which has an exact analytical formula ($(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$), feasible GLS does not have a fixed formula and relies on minimising sum of squared residuals numerically, and there are numerous algorithms for minimising a function numerically, each of which works best in some situations and fails in others. Therefore, although you get the same answer for OLS estimates no matter which software you choose, you may get different answers from different software when using FGLS. Even different versions of the same software may give you different answers if the default estimation method is different in different versions. The following two tables give the results of the same estimation command in EViews 8 and

EViews 10:

Dependent Variable: AVELOAD
Method: Least Squares
Date: 09/23/18 Time: 00:18
Sample (adjusted): 10/02/2005 3/31/2006
Included observations: 181 after adjustments
Convergence achieved after 13 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3582.163	85.98445	41.66059	0.0000
AVETEMP	-26.34896	1.466071	-17.97250	0.0000
WKND	-146.2891	10.02355	-14.59454	0.0000
PUBHOL	-48.19952	19.12181	-2.520656	0.0126
AR(1)	0.914579	0.030275	30.20876	0.0000

R-squared	0.965152	Mean dependent var	2390.249
Adjusted R-squared	0.964360	S.D. dependent var	362.3839
S.E. of regression	68.41268	Akaike info criterion	11.31623
Sum squared resid	823732.0	Schwarz criterion	11.40459
Log likelihood	-1019.119	Hannan-Quinn criter.	11.35205
F-statistic	1218.633	Durbin-Watson stat	2.183140
Prob(F-statistic)	0.000000		

Inverted AR Roots	.91

Dependent Variable: AVELOAD
Method: ARMA Maximum Likelihood (OPG - BHHH)
Sample: 10/01/2005 3/31/2006
Included observations: 182
Convergence achieved after 17 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3576.414	71.92891	49.72151	0.0000
AVETEMP	-26.59983	1.099957	-24.18261	0.0000
WKND	-146.6367	11.46094	-12.79447	0.0000
PUBHOL	-48.23138	15.66494	-3.078938	0.0024
AR(1)	0.912171	0.029491	30.93090	0.0000
SIGMASQ	4545.861	520.2919	8.737135	0.0000

R-squared	0.965856	Mean dependent var	2386.005
Adjusted R-squared	0.964886	S.D. dependent var	365.8870
S.E. of regression	68.56262	Akaike info criterion	11.33559
Sum squared resid	827346.7	Schwarz criterion	11.44121
Log likelihood	-1025.538	Hannan-Quinn criter.	11.37841
F-statistic	995.7266	Durbin-Watson stat	2.186760
Prob(F-statistic)	0.000000		

Inverted AR Roots	.91

Note that the results are close, but are not the same. In fact version 10 does not use FGLS as its default, but it uses the maximum likelihood estimation method (you will be introduced to maximum likelihood in ETC3400 and ETC3410). Note that maximum likelihood uses all 182 observations. Since inference with time series is all based on large samples, whether a method uses all observations or all observations minus 1 does not matter at all.

Sample: 10/02/2005 3/31/2006
Included observations: 181
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC
		1	-0.117	-0.117
		2	-0.006	-0.020
		3	-0.025	-0.028
		4	-0.018	-0.025
		5	0.086	0.082
		6	-0.013	0.006
		7	-0.056	-0.057
		8	-0.082	-0.093
		9	0.060	0.042
		10	-0.000	0.000
		11	0.105	0.105
		12	0.061	0.100

Residuals are now white noise. So the regression with AR(1) errors has taken care of serial correlation in errors. Compare these results with OLS results that we started with.

(e)

We add $wknd \times avetemp$ to the model and test if it is significant

Dependent Variable: AVELOAD
Method: ARMA Maximum Likelihood (OPG - BHHH)
Sample: 10/01/2005 3/31/2006
Included observations: 182
Convergence achieved after 17 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3569.919	71.31876	50.05582	0.0000
AVETEMP	-26.43281	1.099069	-24.05019	0.0000
WKND	-113.9042	54.95640	-2.072629	0.0397
WKND*AVETEMP	-0.750741	1.300690	-0.577187	0.5646
PUBHOL	-48.87897	16.01936	-3.051244	0.0026
AR(1)	0.911599	0.029970	30.41691	0.0000
SIGMASQ	4536.934	518.6744	8.747172	0.0000

R-squared	0.965923	Mean dependent var	2386.005
Adjusted R-squared	0.964755	S.D. dependent var	365.8870
S.E. of regression	68.69070	Akaike info criterion	11.34458
Sum squared resid	825722.0	Schwarz criterion	11.46781
Log likelihood	-1025.356	Hannan-Quinn criter.	11.39453
F-statistic	826.7383	Durbin-Watson stat	2.191117
Prob(F-statistic)	0.000000		

Inverted AR Roots	.91

$$
\begin{aligned}
H_0 &: \beta_{wknd*avetemp} = 0 \\
H_1 &: \beta_{wknd*avetemp} \neq 0 \\
t_{\hat{\beta}_{wknd*avetemp}} &= \frac{\hat{\beta}_{wknd*avetemp}}{se\left(\hat{\beta}_{wknd*avetemp}\right)} \sim t_{176} \text{ whether you use } n = 182 \text{ or } n = 181 \text{ does not matter} \\
t_{calc} &= -0.577 \text{ based on EViews 10 results, } t_{crit} = 1.98 \\
-t_{crit} &< t_{calc} < t_{crit} \implies \text{ cannot reject the null} \\
&\quad \text{There is no evidence that sensitivity to temperature is different in weekends}
\end{aligned}
$$

(f)

Dependent Variable: AVELOAD
Method: ARMA Maximum Likelihood (OPG - BHHH)
Included observations: 182
Convergence achieved after 18 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3565.053	74.47101	47.87169	0.0000
AVETEMP	-26.38951	1.059103	-24.91685	0.0000
SAT	-131.2799	13.34733	-9.835670	0.0000
SUN	-161.1398	12.32969	-13.06925	0.0000
PUBHOL	-43.81095	15.85584	-2.763080	0.0063
AR(1)	0.917351	0.029483	31.11492	0.0000
SIGMASQ	4371.970	487.5982	8.966338	0.0000

R-squared	0.967162	Mean dependent var	2386.005
Adjusted R-squared	0.966036	S.D. dependent var	365.8870
S.E. of regression	67.43033	Akaike info criterion	11.30789
Sum squared resid	795698.6	Schwarz criterion	11.43112
Log likelihood	-1022.018	Hannan-Quinn criter.	11.35785
F-statistic	859.0334	Durbin-Watson stat	2.113989
Prob(F-statistic)	0.000000		

Inverted AR Roots	.92

$$
\begin{aligned}
H_0 &: \beta_{SAT} = \beta_{SUN} \\
H_1 &: \beta_{SAT} \neq \beta_{SUN} \\
F &= \frac{(SSR_r - SSR_{ur})/1}{SSR_{ur}/(182-6)} \sim F_{1,176} \qquad \text{under } H_0 \\
F_{calc} &= \frac{823732.0 - 795698.6}{795698.6/176} = 6.2 \\
F_{calc} &> F_{crit} \approx 3.92 \implies \text{ we reject the null}
\end{aligned}
$$

There is significant evidence that the intercept of the relationship between average load and average temperature for Saturday is different from the intercept for Sunday.

EViews 8 output for the last two parts is includes here just in case. Results are qualitatively similar, although numerically slightly different.

Dependent Variable: AVELOAD
Method: Least Squares
Sample (adjusted): 10/02/2005 3/31/2006
Included observations: 181 after adjustments
Convergence achieved after 13 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3575.797	87.10410	41.05199	0.0000
AVETEMP	-26.21442	1.494041	-17.54598	0.0000
WKND	-118.7534	56.30834	-2.108984	0.0364
WKND*AVETEMP	-0.632062	1.271754	-0.497000	0.6198
PUBHOL	-48.73931	19.19382	-2.539324	0.0120
AR(1)	0.914560	0.030416	30.06789	0.0000

R-squared	0.965201	Mean dependent var	2390.249
Adjusted R-squared	0.964207	S.D. dependent var	362.3839
S.E. of regression	68.55950	Akaike info criterion	11.32587
Sum squared resid	822570.9	Schwarz criterion	11.43190
Log likelihood	-1018.991	Hannan-Quinn criter.	11.36885
F-statistic	970.7848	Durbin-Watson stat	2.186015
Prob(F-statistic)	0.000000		

Inverted AR Roots	.91

Dependent Variable: AVELOAD
Method: Least Squares
Sample (adjusted): 10/02/2005 3/31/2006
Included observations: 181 after adjustments
Convergence achieved after 12 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3574.143	86.52460	41.30783	0.0000
AVETEMP	-26.13974	1.441347	-18.13563	0.0000
SAT	-130.6782	11.43111	-11.43180	0.0000
SUN	-160.9822	11.25601	-14.30188	0.0000
PUBHOL	-43.72755	18.83197	-2.321984	0.0214
AR(1)	0.918624	0.029536	31.10218	0.0000

R-squared	0.966522	Mean dependent var	2390.249
Adjusted R-squared	0.965565	S.D. dependent var	362.3839
S.E. of regression	67.24636	Akaike info criterion	11.28719
Sum squared resid	791362.9	Schwarz criterion	11.39322
Log likelihood	-1015.491	Hannan-Quinn criter.	11.33018
F-statistic	1010.449	Durbin-Watson stat	2.114395
Prob(F-statistic)	0.000000		

Inverted AR Roots	.92