Appendix E

Computer Output of EViews, MINITAB, Excel, and STATA

In this appendix we show the computer output of *EViews*, MINITAB, Excel, and STATA, which are some of the popularly used statistical packages for regression and related statistical routines. We use the data given in Table E.1 from the textbook website to illustrate the output of these packages. Table E.1 gives data on the civilian labor force participation rate (CLFPR), the civilian unemployment rate (CUNR), and real average hourly earnings in 1982 dollars (AHE82) for the U.S. economy for the period 1980 to 2002.

Although in many respects the basic regression output is similar in all these packages, there are differences in how they present their results. Some packages give results to several digits, whereas some others approximate them to four or five digits. Some packages give analysis of variance (ANOVA) tables directly, whereas for some other packages they need to be derived. There are also differences in some of the summary statistics presented by the various packages. It is beyond the scope of this appendix to enumerate all the differences in these statistical packages. You can consult the websites of these packages for further information.

E.1 *EViews*

Using Version 6 of *EViews,* we regressed CLFPR on CUNR and AHE82 and obtained the results shown in Figure E.1.

This is the standard format in which *EViews* results are presented. The first part of this figure gives the regression coefficients, their estimated standard errors, the t values under the null hypothesis that the corresponding population values of these coefficients are zero, and the p values of these t values. This is followed by R^2 and adjusted R^2. The other summary output in the first part relates to the standard error of the regression, residual sum of squares (RSS), and the F value to test the hypothesis that the (true) values of all the slope coefficients are simultaneously equal to zero. Akaike information and Schwartz criteria are often used to choose between competing models. The lower the value of these criteria, the better the model is. The method of maximum likelihood (ML) is an alternative to the method of least squares. Just as in OLS we find those estimators that minimize

FIGURE E.1
EViews output of civilian labor force participation regression.

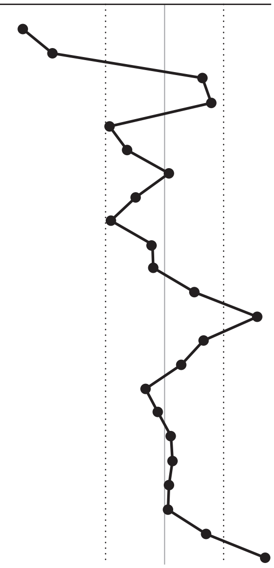
Dependent Variable: CLFPR
Method: Least Squares
Sample: 1980–2002
Included observations: 23

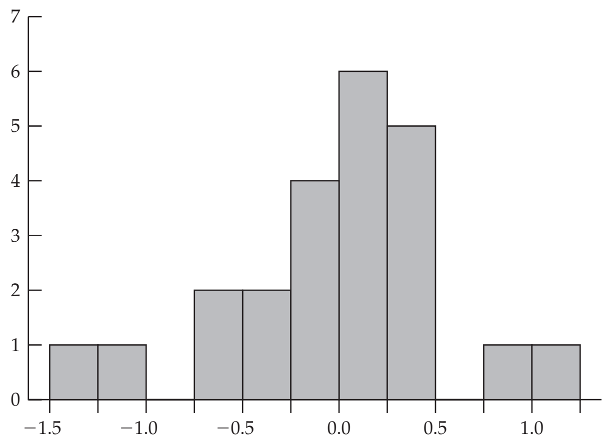
Variable	Coefficient	Std. Error	*t*-Statistic	Prob.
C	80.90133	4.756195	17.00967	0.0000
CUNR	−0.671348	0.082720	−8.115928	0.0000
AHE82	−1.404244	0.608615	−2.307278	0.0319

R-squared	0.772765	Mean dependent var	65.89565
Adjusted R-squared	0.750042	S.D. dependent var	1.168713
S.E. of regression	0.584308	Akaike info criterion	1.884330
Sum squared resid	6.828312	Schwarz criterion	2.032438
Log likelihood	−18.66979	*F*-statistic	34.00731
Durbin–Watson stat	0.787625	Prob(*F*-statistic)	0.000000

Obs	Actual	Fitted	Residual
1980	63.8000	65.2097	−1.40974
1981	63.9000	65.0004	−1.10044
1982	64.0000	63.6047	0.39535
1983	64.0000	63.5173	0.48268
1984	64.4000	64.9131	−0.51311
1985	64.8000	65.1566	−0.35664
1986	65.3000	65.2347	0.06526
1987	65.6000	65.8842	−0.28416
1988	65.9000	66.4103	−0.51027
1989	66.5000	66.6148	−0.11476
1990	66.5000	66.5819	−0.08186
1991	66.2000	65.8745	0.32546
1992	66.4000	65.4608	0.93923
1993	66.3000	65.8917	0.40834
1994	66.6000	66.4147	0.18530
1995	66.6000	66.7644	−0.16441
1996	66.8000	66.8425	−0.04251
1997	67.1000	67.0097	0.09032
1998	67.1000	66.9974	0.10263
1999	67.1000	67.0443	0.05569
2000	67.2000	67.1364	0.06355
2001	56.9000	66.4589	0.44105
2002	66.6000	65.5770	1.02304

Residual Plot





Series: Residuals
Sample 1980–2002
Observations 23

Mean	−1.39e-14
Median	0.063552
Maximum	1.023040
Minimum	−1.409735
Std. Dev.	0.557116
Skewness	−0.593013
Kurtosis	3.752631
Jarque–Bera	1.890898
Probability	0.388505

the error sum of squares, in ML we try to find those estimators that maximize the possibility of observing the sample at hand. Under the normality assumption of the error term, OLS and ML give identical estimates of the regression coefficients. The Durbin–Watson statistic is used to find out if there is first-order serial correlation in the error terms.

The second part of the *EViews* output gives the actual and fitted values of the dependent variable and the difference between the two, which represent the residuals. These residuals are plotted alongside this output with a vertical line denoting zero. Points to the right of the vertical line are positive residuals and those to the left represent negative residuals.

The third part of the output gives the histogram of the residuals along with their summary statistics. It gives the Jarque–Bera (JB) statistic to test for the normality of the error terms and also gives the probability of obtaining the stated statistics. The higher the probability of obtaining the observed JB statistic, the greater is the evidence in favor of the null hypothesis that the error terms are normally distributed.

Note that *EViews* does not give directly the analysis-of-variance (ANOVA) table, but it can be constructed easily from the data on the residual sum of squares, the total sum of squares (which will have to be derived from the standard deviation of the dependent variable), and their associated degrees of freedom. The F value given from this exercise should be equal to the F value reported in the first part of the table.

E.2 MINITAB

Using Version 15 of MINITAB, and using the same data, we obtained the regression results shown in Figure E.2.

MINITAB first reports the estimated multiple regression. This is followed by a list of predictor (i.e., explanatory) variables, the estimated regression coefficients, their standard errors, the T (= t) values, and the p values. In this output S represents the standard error of the estimate, and R^2 and adjusted R^2 values are given in percent form.

This is followed by the usual ANOVA table. One characteristic feature of the ANOVA table is that it breaks down the regression, or explained, sum of squares among predictors. Thus of the total regression, sum of squares of 23.226, the share of CUNR is 21.404 and that of AHE82 is 1.822, suggesting that relatively, CUNR has more impact on CLFPR than AHE82.

A unique feature of the MINITAB regression output is that it reports "unusual" observations; that is, observations that are somehow different from the rest of the observations in the sample. We have a hint of this in the residual graph given in the *EViews* output, for it shows that the observations 1 and 23 are substantially away from the zero line shown there. MINITAB also produces a residual graph similar to the *EViews* residual graph. The St Resid in this output is the standardized residuals; that is, residuals divided by S, the standard error of the estimate.

Like *EViews,* MINITAB also reports the Durbin–Watson statistic and gives the histogram of residuals. The histogram is a visual picture. If its shape resembles the normal distribution, the residuals are perhaps normally distributed. The normal probability plot accomplishes the same purpose. If the estimated residuals lie approximately on a straight line, we can say that they are normally distributed. The Anderson–Darling (AD) statistic, an adjunct of the normal probability plot, tests the hypothesis that the variable under consideration (here residuals) is normally distributed. If the p value of the calculated AD statistic is reasonably high, say in excess of 0.10, we can conclude that the variable is normally distributed. In our example the AD statistic has a value of 0.481 with a p value of about 0.21 or 21 percent. So we can conclude that the residuals obtained from the regression model are normally distributed.

FIGURE E.2 MINITAB output of civilian labor force participation rate.

Regression Analysis: CLFPR versus CUNR, AHE82

The regression equation is
CLFPR = 81.0 − 0.672 CUNR − 1.41 AHE82

Predictor	Coef	SE Coef	T	P
Constant	80.951	4.770	16.97	0.000
CUNR	−0.67163	0.08270	−8.12	0.000
AHE82	−1.4104	0.6103	−2.31	0.032

S = 0.584117   *R*-Sq = 77.3%   *R*-Sq(adj) = 75.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	23.226	11.613	34.04	0.000
Residual Error	20	6.824	0.341		
Total	22	30.050			

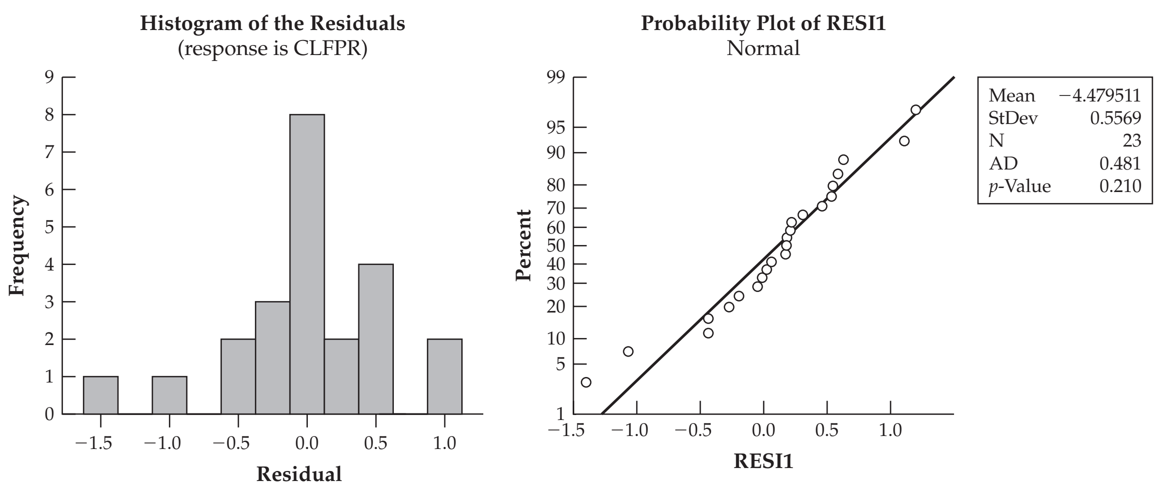
Source	DF	Seq SS
CUNR	1	21.404
AHE82	1	1.822

Unusual Observations

Obs	CUNR	CLFPR	Fit	SE Fit	Residual	St Resid
1	7.10	63.800	65.209	0.155	−1.409	−2.50*R*
23	5.80	66.600	65.575	0.307	1.025	2.06*R*

R denotes an observation with a large standardized residual.

Durbin–Watson statistic = 0.787065



E.3 Excel

Using Microsoft Excel we obtained the regression output shown in Table E.2.

Excel first presents summary statistics, such as R^2, multiple R, which is the (positive) square root of R, adjusted R^2, and the standard error of the estimate. Then it presents the ANOVA table. After that it presents the estimated coefficients, their standard errors, the t values of the estimated coefficients and their p values. It also gives the actual and estimated

TABLE E.2
Excel Output of Civilian Labor Force Participation Rate

Summary Output	
Regression Statistics	
Multiple *R*	0.879155
R Square	0.772914
Adjusted *R*	0.750205
Standard E	0.584117
Observation	23

ANOVA

	df	SS	MS	*F*	Significance *F*
Regression	2	23.22572	11.61286	34.03611	3.65E-07
Residual	20	6.823846	0.341192		
Total	22	30.04957			

	Coefficient	Standard Err	*t* Stat	*p*-value	Lower 95%	Upper 95%
Intercept	80.95122	4.770337	16.96971	2.42E-13	71.00047	90.90196
CUNR	−0.671631	0.082705	−8.120845	9.24E-08	−0.84415	−0.499112
AHE82	−1.410432	0.610348	−2.310867	0.031626	−2.683594	−0.13727

values of the dependent variable and the residual graph as well as the normal probability plot.

A unique feature of Excel is that it gives the 95 percent (or any specified percent) confidence interval for the true values of the estimated coefficients. Thus, the estimated value of the coefficient of CUNR is −0.671631 and the confidence interval for the true value of CUNR coefficient is (−0.84415 to −0.499112). This information is very valuable for hypothesis testing.

E.4 STATA

Using STATA we obtained the regression results shown in Table E.3.

Stata first presents the analysis of variance table along with the summary statistics such as R^2, adjusted R^2, and the root mean-squared-error (MSE), which is just the standard error of the regression.

Then it gives the values of the estimated coefficients, their standard errors, their *t* values, the *p* values of the *t* statistics, and the 95 percent confidence interval for each of the regression coefficients, which is similar to the Excel output.

E.5 Concluding Comments

We have given just the basic output of these packages for our example. But it may be noted that packages such as *EViews* and STATA are very comprehensive and contain many of the econometric techniques discussed in this text. Once you know how to access these packages, running various subroutines is a matter of practice. If you wish to pursue econometrics further, you may want to buy one or more of these packages.

TABLE E.3
STATA Output of Civilian Labor Force Participation Rate

tm

Statistics/Data Analysis

Project: Data of Table E.1

tm

8.0

Statistics/Data Analysis

gress clfpr cunr ahe82

Source	SS	df	MS
Model	23.2256929	2	11.6128465
Residual	6.82384072	20	.341192036
Total	30.0495337	22	1.36588789

Number of obs	=	23
F(2, 20)	=	34.04
Prob > F	=	0.0000
R-squared	=	0.7729
Adj R-squared	=	0.7502
Root MSE	=	.58412

clfpr	Coef.	Std. Err.	*t*	*p* > |*t*|	[95% Conf. Interval]	
cunr	−.6716305	.0827045	−8.12	0.000	−.8441491	−.4991119
ahe82	−1.410433	.6103473	−2.31	0.032	−2.683595	−.1372707
_cons	80.95122	4.770334	16.97	0.000	71.00048	90.90197

References

www.eviews.com
www.stata.com
www.minitab.com
Microsoft Excel

R. Carter Hill, William E. Griffiths, George G. Judge, *Using Excel for Undergraduate Econometrics,* John Wiley & Sons, New York, 2001.