Question 1. The financial analyst considers monthly data on the short-term interest rate (the three month Treasury Bill rate) and on the AAA corporate bond yield in the USA. The data run from January 1950 to December 1999.

Let *DUS3MT* denote the changes in three-month Treasury Bill rate, and *DAAA* denote the changes in AAA bond rate. We consider the following results (with the standard errors given in parentheses)

$$\hat{DAAA}_t = \underset{(.007)}{0.006} + \underset{(.018)}{0.275} DUS3MT_t, \quad t = 1, \ldots, 600,$$

(1)

$$RSS = 17.486; \; DW = 1.447,$$

where *RSS* is the residual sum of squares and *DW* is the Durbin—Watson test. A researcher interpreting the residuals suggests that the errors show a positive correlation over time.

(a) **(5 marks)** □ Use the results above to test for the presence of first-order positive autocorrelation. Briefly describe the Durbin-Watson test: clearly specify the null and alternative hypothesis, test statistic, assumptions underlying the test, and the acceptance/rejection rule.

DW tables give two limiting values: lower d_L and upper d_U ones.

If $DW > d_U$ for given significance level the null hypothesis is not rejected, and if $DW < d_L$ it is rejected in fawor of one sided alternative $H_a : \rho > 0$. $DW = 1.447$, for 600 observations $1.447 < d_L(1\%) = 1.52$, so the null ($H_0 : \rho = 0$ where $u_t = \rho u_{t-1} + v_t$) is rejected, and we conclude on the presence of the positive autocorrelation of the first order.

□ Show that $DW \approx 2(1 - \hat{\rho})$ where $\hat{\rho}$ is autocorrelation coefficient.

The Durbin-Watson test is given by $DW = \frac{\sum_{t=2}^{T} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{T} \hat{u}_t^2}$, where $\hat{u}_t = Y_t - \hat{\beta}_1 - \hat{\beta}_2 X_t$ from the model $Y_t = \beta_1 + \beta_2 X_t + u_t$ and is used to test the null hypothesis $H_0 : \rho = 0$ against $H_a : \rho > 0$ where $u_t = \rho u_{t-1} + v_t$ and v_t satisfies standard Gauss-Markov conditions. Let us show that $DW \approx 2(1 - \rho)$:

$$DW = \frac{\sum_{t=2}^{n} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{n} \hat{u}_t^2} = \frac{\sum_{t=2}^{n} \hat{u}_t^2}{\sum_{t=1}^{n} \hat{u}_t^2} + \frac{\sum_{t=2}^{n} \hat{u}_{t-1}^2}{\sum_{t=1}^{n} \hat{u}_t^2} - 2 \frac{\sum_{t=2}^{n} \hat{u}_t \hat{u}_{t-1}}{\sum_{t=1}^{n} \hat{u}_t^2} \cong 2 - 2\hat{\rho} = 2(1 - \hat{\rho}).$$

□ What are the consequences of this correlation for the above regression results?

Autocorrelation makes OLS inefficient, meaning that another estimator with a lower population variance can be found, and it also invalidates significance tests.

(b) **(5 marks)** In an attempt to remove the autocorrelation you consider the following specification

$$\hat{DAAA}_t = \underset{(.007)}{0.005} + \underset{(.015)}{0.252} DUS3MT_t - \underset{(.018)}{0.080} DUS3MT_{t-1} + \underset{(.040)}{0.290} DAAA_{t-1},$$

(2)

$$RSS = 16.087; \; DW = 1.897.$$

□ Comment on the following statement 'The Durbin—Watson statistic is closer to 2, indicating that we have successfully removed the autocorrelation'. If you disagree with this statement, explain why, and suggest what test can be done based on the available data.

According the assumptions underlying DW test, no lagged dependent variables are allowed. So we disagree with the statement because the Durbin-Watson test is not valid in the presence of lagged dependent variables. Instead, we should use the Durbin *h* test: $d = \hat{\rho}\sqrt{\frac{n}{1 - ns_{Y_{t-1}}^2}} \sim N(0,1)$ under $H_0 : \rho = 0$, $\hat{\rho} \approx 1 - 0.5 \cdot DW$, $s_{Y_{t-1}}^2$ is the standard error of the coefficient estimate of lagged dependent variable, n is the number of observations adjusted so that to take into account lags used. Alternatively the Breusch-Godfrey test which is asymptotically valid with predetermined regressors, which are a weaker requirement than strict exogeneity.

□ What are the reasons to consider the second model as an improved version of the first?

F-test for restrictions

$$F = \frac{(RSS_R - RSS_U)/2}{RSS_U/(600-4)} = \frac{(17.486 - 16.087)/2}{16.087/(600-4)} = 25.91$$

while $F_{crit}^{5\%}(2,600) = 4.64$, so the null hypothesis $H_0 : \beta_3 = 0, \beta_4 = 0$ is not rejected.

In the absence of autocorrelation in both models, we would choose specification 1 as more efficient. But here specification **(1)** suffers from auticorrelation which invalidates efficiency considerations, while in specification **(2)** the autocorrelation is removed, so we choose specification **(2)**.

□ Can you estimate long term effect of $X_t = DUS3MT_t$ on $Y_t = DAAA_t$? Compare it with the short term effect.

The short term effect of $DUS3MT_t$ on $DAAA_t$ is 0.252. So the increase of $DUS3MT_t$ by one percentage point yealds the increase of $DAAA_t$ by 0.252 percentage points.

To evaluate long term marginal effect one should substitute in original ADL(1,1) model

$$DAAA_t = \beta_1 + \beta_2 DUS3MT_t + \beta_3 DUS3MT_{t-1} + \beta_4 DAAA_{t-1} + u_t$$

instead of current and lagged values their equilibrium values

$$\overline{DAAA} = \beta_1 + \beta_2 \overline{DUS3MT} + \beta_3 \overline{DUS3MT} + \beta_4 \overline{DAAA}$$

so

$$\overline{DAAA} = \frac{\beta_1}{(1-\beta_4)} + \frac{\beta_2 + \beta_3}{(1-\beta_4)} \overline{DUS3MT}$$

And so long term marginal effect can be evaluated as $\frac{\beta_2 + \beta_3}{(1-\beta_4)}$. Substituting the estimated values from equation (2) we get $\frac{0.252 - 0.080}{(1 - 0.290)} = 0.242$.

(c) (5 marks) Discuss the Common Factor Test as a model specification suitable for this model.

□ Derive Common Test specification from the ADL(1,1) model, corresponding to the equation (2).

Let $Y_t = DAAA_t$ and $X_t = DUS3MT_t$ we can remove the autocorrelation by rewriting the model using the following transformation: starting from specification of equation (1)

$$Y_t = \beta_1 + \beta_2 X_t + u_t \quad (*)$$

and substituting $u_t = \rho u_{t-1} + \varepsilon_t$ into it we get

$$Y_t = \beta_1 + \beta_2 X_t + \rho u_{t-1} + \varepsilon_t$$

Lagging equation (*) and multiplying it by ρ we get

$$\rho Y_{t-1} = \rho\beta_1 + \rho\beta_2 X_{t-1} + \rho u_{t-1} \quad (**)$$

So after subtracting equation (**) from (*) the lagged disturbance term u_{t-1} vanishes

$$Y_t = (1-\rho)\beta_1 + \rho Y_{t-1} + \beta_2 X_t - \beta_2 \rho X_{t-1} + \varepsilon_t.$$

with disturbance term ε_t what does not show the autocorrelation of the first order.

This is a restricted version of a more general $ADL(1,1)$ model:

$$Y_t = \lambda_1 + \lambda_2 Y_{t-1} + \lambda_3 X_t + \lambda_4 X_{t-1} + \varepsilon_t$$

under the restriction $\lambda_4 = -\lambda_2 \lambda_3$.

□ What extra information do you need to conduct this test? What assumptions are needed for this?

We can test this restriction using the Common Factor Test. Let RSS_R and RSS_U be the restricted sums of squares of the restricted and unrestricted model, respectively.

- Test statistic: $CF = n \times \log\left(\dfrac{RSS_R}{RSS_U}\right)$. Under H_0 of no autocorrelation of the first order, $CF \overset{asympt.}{\sim} \chi_1^2$.

Extra information needed: RSS_R (since RSS_U is given in (2)). Assumptions $AR(1)$ autocorrelation $u_t = \rho u_{t-1} + \varepsilon_t$ where $E(\varepsilon_t) = 0$, $E(\varepsilon_t^2) = \sigma_\varepsilon^2$ and $E(\varepsilon_t \varepsilon_s) = 0$ for $t \neq s$,

(d) (5 marks) □ After estimating unrestricted model corresponding to Common Test specification the value of $RSS_R = 17.4$ was obtained. Do the common factor test, and make the conclusion. What specificaion should be chosen on the base of the test, what are interpretations of its parameters?

It was found that the value of 17.4 was mistakenly obtained by rounding off the exact number 17.361. How this changes your conclusion and the interpretation of some parameters of the model?

$CF = 599 \times \ln\left(\dfrac{17.4}{16.087}\right) = 46.997 > 6.635 = \chi^2_{crit.,1\%}(1) \Rightarrow$ reject, and choose

$$Y_t = \lambda_1 + \lambda_2 Y_{t-1} + \lambda_3 X_t + \lambda_4 X_{t-1} + \varepsilon_t$$

The conclusion does not change if corrected value of $RSS_R = 17.361$ is used. In this case the coefficient 0.290 of $Y_{t-1} = DAAA_{t-1}$ can be interpreted as the effect of the former behaviour of dependent variable on the current one.

If Common Factor test does not reject null then the restricted version of the model should be chosen

$$Y_t = (1-\rho)\beta_1 + \rho Y_{t-1} + \beta_2 X_t - \beta_2 \rho X_{t-1} + \varepsilon_t.$$

Here where ρ can be interpeted as autocorrelation coefficient in AR(1) process $u_t = \rho u_{t-1} + \varepsilon_t$.

Question 2.
A researcher is considering the following alternative regression models for the expenditures on tobacco y_t (in billions of dollars) subject to the relative price index of tobacco x_t using annual data (1971-2015 - **45** observations) this time for USA:

$$y_t = \lambda + \mu x_t + w_t \qquad (1)$$

$$\Delta y_t = \gamma + \delta \Delta x_t + v_t \qquad (2)$$

where $\Delta y_t = y_t - y_{t-1}$, $\Delta x_t = x_t - x_{t-1}$, and u_t, v_t, and w_t are disturbance terms.
He got the following results of the estimation of these equations

$$\hat{y}_t = 114.1 - 0.38 x_t \qquad R^2 = 0.61$$
$$(2.59)\ (0.04)\quad DW = 0.16\quad RSS = 1781.3 \qquad (1')$$

$$\hat{\Delta y}_t = 0.29 - 0.36 \Delta x_t \qquad R^2 = 0.21$$
$$(0.43)\ (0.11)\quad DW = 2.27\quad RSS = 281.9 \qquad (2')$$

(a) (5 marks) You are given that the time series y_t and x_t are difference stationary, or they are integrated of order 1: $y_t \in I(1)$, $x_t \in I(1)$.

□ Explain what does it mean that time series is difference stationary and trend stationary.

Solution: If after removing the trend from a nonstationary series (detrending) the resulting variable becomes stationary, then the variable is called *trend stationary*.
If a nonstationary process can be transformed into a stationary process by differencing then the series is said to be *difference stationary*.

□ Give the examples of difference stationary and trend stationary time series. Show that the time series you have given posess the specified properties.

Solution: Let: $X_t = \alpha_0 + \alpha_1 t + u_t$, where $\mathrm{E}(u_t) = 0$, $\mathrm{var}(u_t) = \sigma^2$ and $\mathrm{E}(u_t u_{t-s}) = 0$ for all *s* and *t*.
Let:
$Z_t = X_t - \alpha_1 t = \alpha_0 + u_t$ (detrending).

$$\mathrm{E}(Z_t) = \mathrm{E}(\alpha_0 + u_t) = \alpha_0,$$

$$\mathrm{var}(Z_t) = \mathrm{Var}(\alpha_0 + u_t) = \sigma^2,$$

$$\mathrm{cov}(Z_t, Z_{t-s}) = \mathrm{E}[(Z_t - \mathrm{E}(Z_t))(Z_{t-s} - \mathrm{E}(Z_{t-s}))] = \mathrm{E}(u_t u_{t-s}) = 0.$$

This means that Z_t has constant mean and variance for all *t*, and covariance is zero for all $s > 0$. It implies that the series is trend stationary.

Solution: Let X_t be a random walk with a drift: $X_t = \beta_0 + X_{t-1} + \varepsilon_t$, where $\mathrm{E}(\varepsilon_t) = 0$, $\mathrm{var}(\varepsilon_t) = \sigma^2$ and $\mathrm{E}(\varepsilon_t \varepsilon_s) = 0$ for all *s* and *t*, $s \neq t$. Subtract X_{t-1} from both sides of equation to get: $\Delta X_t = X_t - X_{t-1} = \beta_0 + \varepsilon_t$.
Checking the same properties we conclude that ΔX_t is stationary. This implies that X_t is difference stationary.

Trend stationary time series instead of detrending can be differenced.
If $X_t = \alpha_0 + \alpha_1 t + u_t$ then $X_{t-1} = \alpha_0 + \alpha_1(t-1) + u_{t-1}$ so subtracting equations we get stationary time series $\Delta X_t = X_t - X_{t-1} = \alpha_1 + u_t - u_{t-1}$ Nevertheless it is called trend stationary time series. Its disturbance term $\varepsilon_t = u_t - u_{t-1}$ follows MA(1) process and so possibly posess some (negative) autocorrelation properties.

□ What is the main difference between two types of time series under consideration.

Solution: It is important to know whether a variable is difference stationary or trend stationary because for difference stationary variables shocks have a permanent effect whereas for trend stationary variables shocks are transitory.

(b) (5 marks) □ What is cointegration? How to test whether time series y_t and x_t are cointegrated (you are given again that $y_t \in I(1)$, $x_t \in I(1)$.

Solution: Definition: Several time series are called cointegrated if
1) They all are of the same order of integration (say I(1))
2) There is a linear combination of them such that it is stationary
The second condition is equivalent to the following: the residuals of regression of one of them on others are stationary. So to test time series y_t and x_t for cointegration is is sufficient to do the following:
1) Check whether y_t and x_t are of the same order of integration (we are given $y_t \in I(1)$, $x_t \in I(1)$).
2) Run regression estimated by OLS: $y_t = \alpha_0 + \alpha_1 x_t + u_t$ to get its residuals e_t
3) Run new regression of the type $\Delta\hat{e}_t = \theta e_{t-1} + \sum \theta_i \Delta e_{t-i}$ and test θ using the ADF test.
$H_0 : \theta = 1 \quad \Rightarrow \quad e_t$ is non-stationary. This means that y_t and x_t are not cointegrated. For y_t and x_t to be cointegrated the null has to be rejected.

□ What are relative advantages and disadvantages of the models (1) and (2)?

Let z_t are residuals of equation (1’). To test them for stationarity the researcher runs the regression

$$\hat{\Delta z_t} = \underset{(0.058)}{-0.118} z_{t-1} \qquad R^2 = 0.07$$

□ What is the meaning of this equation? What is your conclusion from the test (**use the following critical values for ADF t-statistic** $t_{ADF}^{1\%} = -2.62$, $t_{ADF}^{5\%} = -1.95$, $t_{ADF}^{10\%} = -1.61$)?

Solution: This is Dickey-Fuller equation of the type $\Delta z_t = (\beta - 1) z_{t-1} + u_t$ to test $H_0 : \beta = 1$ against $H_a : \beta < 1$ Evaluate $t = \dfrac{b}{s.e.(b)} = \dfrac{-1.118}{0.058} = -2.03$. As $-2.62 < t = -2.03 < -1.94$ null hypothesis of non-stationarity is rejected at 5%. It means that y_t and x_t are cointegrated.

(c) (5 marks) Consider a simple $ADL(1,1)$ model

$$y_t = \alpha_1 + \alpha_2 y_{t-1} + \alpha_3 x_t + \alpha_4 x_{t-1} + u_t \qquad (3)$$

where y_t and x_t are $I(1)$.

□ Derive error correction model based on (3).

Solution: One of the possible solutions is following. Rewrite (3) as

$$y_t - y_{t-1} = \alpha_1 + \alpha_2 y_{t-1} - y_{t-1} + \alpha_3 x_t - \alpha_3 x_{t-1} + \alpha_3 x_{t-1} + \alpha_4 x_{t-1} + u_t$$

$$\Delta y_t = \alpha_1 - (1-\alpha_2) y_{t-1} + \alpha_3 \Delta x_t + (\alpha_3 + \alpha_4) x_{t-1} + u_t$$

$$\Delta y_t = \alpha_3 \Delta x_t - (1-\alpha_2)\left[y_{t-1} - \frac{\alpha_1}{(1-\alpha_2)} - \frac{(\alpha_3+\alpha_4)}{(1-\alpha_2)} x_{t-1} \right] + u_t$$

$$\Delta y_t = \alpha_3 \Delta x_t - (1-\alpha_2)[y_{t-1} - \beta_1 - \beta_2 x_{t-1}] + u_t$$

or,

$$\Delta y_t = \alpha_3 \Delta x_t - \pi[y_{t-1} - \beta_1 - \beta_2 x_{t-1}] + u_t$$

where

$$\pi = (1-\alpha_2);\ \beta_1 = \frac{\alpha_1}{(1-\alpha_2)} \text{ and } \beta_2 = \frac{(\alpha_3+\alpha_4)}{(1-\alpha_2)}.$$

This equation is the Error Correction Model.

The point of rearranging the ADL(1,1) model in this way is that, although Y_t and X_t are both I(1), all of the terms in the regression equation are I(0) and hence the model may be fitted using least squares in the standard way.

Bonus answer: Of course, the β parameters are not known and the cointegrating term is unobservable. One way of overcoming this problem, known as the Engle–Granger two-step procedure, is to use the values of the parameters estimated in the cointegrating regression to compute the cointegrating term. It can be demonstrated that the estimators of the coefficients of the fitted equation will have the same properties asymptotically as if the true values had been used.

(d) (5 marks) Let model (4') is the result of estimation error correction model, obtained in (c) for years 1960-1996

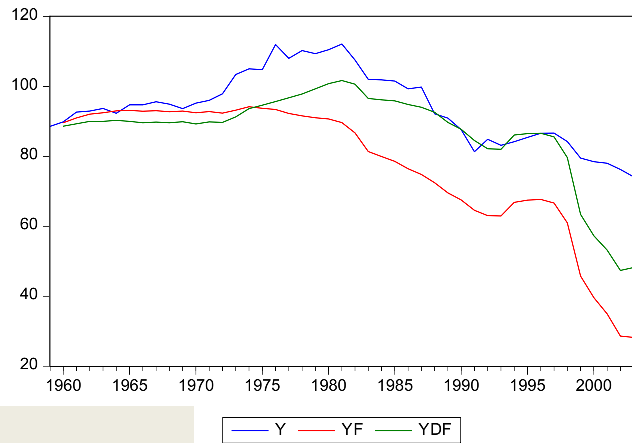
$$\hat{\Delta y_t} = \underset{(0.217)}{-0.795}\Delta x_t - \underset{(0.061)}{0.128} Z_{t-1} \qquad R^2 = 0.32 \qquad (4')$$

□ Give interpretation to the estimated model. Explain how error correction mechanism works.

Solution: The first term corresponds to the difference model (short run behavior), the second term is the error-correction mechanism (describing long run behavior), both coefficients are significant. The size of the adjustment is proportional to the discrepancy. The effect of this term is to reduce the discrepancy between y_{t-1} and its cointegrating level. The error-correction mechanism allows you to overcome the short-sightedness of the model. At the same time, the correction value is relatively small (-0.128). It means that each year only 12.8 of discrepancy is covered by an adjustment using a cointegration relationship.

Let variable yf_t be forecast obtained using by error correction model and ydf_t be forecast obtained using difference model (2).

□ Comment the following graph, demonstrating predictive power of different models



Solution: Tobacco expenditure has a complicated dynamics, with a rise at the beginning of the period followed by a decline in subsequent years. However, the difference model generally provides a more accurate forecast, while the error correction model overestimates the impact of relative prices on tobacco use decline. This may be due to the poor quality of the cointegration relationship: it uses a linear model, while the dynamics of tobacco expenditure is non-linear.

Question 3.

The student is interested in the influence of the price level and income on the consumption of soft drinks SD_{it}. Suppose that the time series C_t (Coca-cola), P_t (Pepsi), S_t (Sprite), U_t (7-up), M_t (Myrinda), F_t (Fanta) for 1988-2017 form the panel (as observations for particular units of some general type of soft drinks (SD_{it})). DPI_t is the disposable personal income and $PRSD_{it}$ is the relative price index for corresponding SD_{it}.Let the model under investigation be

$$\log(SD_{it}) = \beta_1 + \beta_2 \log(DPI_t) + \beta_3 \log(PRSD_{it}) + u_{it}.$$

(a) (5 points) What are the advantages of panel data analysis comparing to cross-section regression and time series analysis? What are typical problems associated with the panel data models? The student decided to use the approach based on the fixed effects panel data model. Help the student to understand what is LSDV method and how to use it for estimation of panel data model under consideration. Explain clearly what are fixed effects.

Solution

What are advantages of panel data analysis comparing to cross-section regression and time series analysis?

Larger data set. If in time series data T observations are available, with panel data T*n observations are available, where n is the number of units and T is the number of periods.

Unobserved heterogeneity problem (more about it later) is eliminated or mitigated.

Dynamics can be explored better (compared to cross-section). Although with cross-section one could investigate dynamics by asking retrospective questions, it is not very reliable, as people forget details over time.

Panel data are often of higher quality. For example, national surveys are usually rather well designed and well organized.

What are typical problems associated with the panel data models?

1. Cost of data collection
2. Issue of modeling unobserved heterogeneity

The main question in the panel data analysis is the problem of the origin of unobserved heterogeneity $\sum_{p=1}^{s}\gamma_p Z_{pi} = \alpha_i$ in the model of the type

$$Y_{it} = \beta_0 + \sum_{j=1}^{k}\beta_j X_{jit} + \sum_{p=1}^{s}\gamma_p Z_{pi} + \varepsilon_{it} \qquad (1).$$

Fixed effect approach assigns it to the fixed characteristics of individual elements (here to the types of soft drinks), while random effect approach assigns it to the random factors.

In the **Least Squares Dummy Variable (LSDV) fixed effect method** the unobserved effect is brought explicitly into the model. A set of dummy variables D_i is defined, where D_i is equal to 1 in the case of an observation relating to an individual *i* and 0 otherwise. The model can be written as

$$Y_{it} = \sum_{j=2}^{K}\beta_j X_{jit} + \sum_{i=1}^{n}\alpha_i D_i + u_{it}.$$

(vi)

The unobserved effect is now being treated as the coefficient of the specific individual *i*.

The term $\alpha_i D_i$ represents a fixed effect on the dependent variable Y_i for individual *i*. We use the model without intercept otherwise we will fall into the dummy variable trap. Entering the details, we note that the method calculates its own intercept as an average of all coefficients for dummies, and represents fixed effects as the deviations of these coefficients from the calculated mean. The method is based on the assumption that the elasticities in income and prices are equal for all soft drinks, and heterogeneity is reflected only in the intercept.

It should be noted that inclusion of a set of dummies leads to losing of additional degrees of freedom, so estimators become less efficient.

(b) (5 points) How to test the presence of unobserved heterogeneity if fixed effect model or to put it another way, that there are considerable differences between the pooled model and the model with fixed effects? Give some details: indicate corresponding test and the data needed for it, null hypothesis, distribution of the test statistic, the number of degrees of freedom and the decision rule.

Solution We start from LSDV fixed effects model of the type

$$\log(SD_{it}) = \beta_2\log(DPI_t) + \beta_3\log(PRSD_{it}) + \sum_{i=1}^{6}\gamma_i D_i + u_{it}. \qquad t = 1, 2, ..., 30 \quad (1)$$

where D_i are dummies corresponding different drinks under consideration. To take into account that the elasticities of DPI_t and $PRSD_{it}$ could be different, it is possible to run instead 6 different regressions for 6 drinks ($i = 1, 2,..., 6$),

$$\log(SD_{it}) = \gamma_1 + \gamma_2\log(DPI_t) + \gamma_3\log(PRSD_{it}) + u_{it}$$

and then perform an F-test using the values of RSS' of evaluated models:

$$F = \frac{(RSS_{POOLED} - RSS_{LSDV})/(n-1)}{RSS_{LSDV}/(n\cdot T - n - K)} = \frac{(RSS_{POOLED} - RSS_{LSDV})/(6-1)}{RSS_{LSDV}/(6\cdot 30 - 6 - 2)} = \frac{(RSS_{POOLED} - RSS_{LSDV})/5}{RSS_{LSDV}/172}$$

or $F = \dfrac{(R^2_{LSDV} - R^2_{POOLED})/(n-1)}{(1-R^2_{LSDV})/(n\cdot T - n - K)} = \dfrac{(R^2_{LSDV} - R^2_{POOLED})/(6-1)}{(1-R^2_{LSDV})/(6\cdot 30 - 6 - 2)} = \dfrac{(R^2_{LSDV} - R^2_{POOLED})/5}{(1-R^2_{LSDV})/172}$

Here RSS_{LSDV} and R^2_{LSDV} - the values of RSS and R^2 for LSDV fixed effect model, RSS_{POOLED} and R^2_{POOLED}- the values of RSS and R^2 for pooled model, $n = 6$ - number of crosssection objects, $T = 30$ - number of time periods, $K = 2$ - number of variables.

The rule for degrees of freedom follows general principles of F-test and is clear from the formula above. If H_0 is not rejected there is no significant differences between elasticities.

(c) (5 points) What other methods within the fixed effects panel data model can the student apply? What are their comparative advantages and disadvantages? Could the researcher have used a random effects regression in the present case? How to choose between random effects and fixed effects panel model?

Solution

What other methods within the fixed effects panel data model can the student apply?

1. Within groups method
2. First differences method

What are their comparative advantages and disadvantages?

1. All three methods (LSDV, Within groups and First difference) leads to losing $n-1=4-1=3$ degrees of freedom.
2. Autocorrelation problem may appear if researcher uses First difference method (and in a weakened form in the Within groups method), which may even be useful, since the resulting autocorrelation of the MA (1) type is negative, which, superimposed on the positive autocorrelation in the original data, can solve or alleviate the problem of autocorrelation as a whole.
3. Variation of new explanatory variables is smaller, thus precision of estimates decreases
4. All variables are constant in time

Could the researcher have used a random effects regression in the present case?

The student is recommended to use Darbin-Wu-Hausman (DWH) test to choose between fixed and random effects. It is standard for majority of econometric computer programs and is based on using chi-square statistics with degrees of freedom equal to the number of variables in the equation under consideration (2 in our case) (as it compares estimates of coefficients obtained by two alternative models). Under H_0 that there is no difference between coefficients obtained by two alternative models – fixed and random panel models (which means that unobserved heterogeneity α_i as a part of disturbance term, is not correlated with DPI_t and $PRSD_{it}$) both fixed effect and random effect models provide us with consistent estimates. We choose in this case random effect models as it retains in disturbance term all unobserved heterogeneity, there is no reduction of degrees of freedom typical for fixed effects models.

If H_0 is rejected, so there are essential differences between coefficients obtained using fixed and random effects models, we choose fixed effects model, because rejecting of H_0 means that main assumption of independence of the disturbance term from regressors is violated so using random effects model we are under risk of getting inconsistent estimates of parameters. So we have to suffice the fixed effects model that always gives consistent estimates.

It should be noted that random effect model could not be applied in this case from economic point of view as for this model first condition of sample being drawn randomly from the population, but Fanta belongs to Coca Cola company, while Myrinda, 7up and Sprite belong to PepsiCo, thus these soft drinks could not be considered separately, because their pricing policies are not independent.

Question 4.

A researcher has annual time series data on labour employed, L_t, measured in thousands of employees, output, Y_t, measured in billion dollars at constant prices, and average annual wages, w_t, measured in thousands of dollars at constant prices, for a certain industry for 25 years and fits the following regression (standard errors in parentheses; d is the Durbin-Watson statistic):

$$\widehat{L_t} = \underset{(1.91)}{10.11} + \underset{(1.20)}{6.13}Y_t - \underset{(0.14)}{0.21}w_t + \underset{(0.10)}{0.80}L_{t-1} \qquad d = 1.88$$

(a) (5 points) Explain how the model may be interpreted as a partial adjustment model

Suppose that the target labour force in year t, L_t^*, is given by

$$L_t^* = \alpha + \beta_1 Y_t + \beta_2 w_t + u_t$$

and partial adjustment towards the target is achieved each year:

$$L_t - L_{t-1} = \lambda(L_t^* - L_{t-1})$$

Substituting the first equation into the second, one obtains the regression model used by the researcher:

$$L_t = \alpha\lambda + \beta_1\lambda Y_t + \beta_2\lambda w_t + (1-\lambda)L_{t-1} + u_t$$

(b) (5 points) Analyse dynamic behaviour of the model.

Short-run (current period) behaviour: A \$ billion increase in output will lead to the employment of 6.13 thousand additional workers; a one-thousand dollar increase in annual wages causes employment to fall by 0.21 thousand workers.

Adjustment behaviour: The coefficient of L_{t-1} implies that only 0.20 of the discrepancy between target and previous actual employment is eliminated in any year

Long-run behaviour: Looking at the theoretical coefficients of the model, if one divides the coefficient of Y_t by the estimate of λ one obtains an estimate of β_1: $b_1 = 6.13/0.20 = 30.65$. This implies that a \$ billion increase in output will increase target employment by 30.65 thousand workers. Similarly, if one divides the coefficient of w_t by the estimate of λ one obtains an estimate of β_2: $b_2 = 0.21/0.20 = 1.05$. This implies that a \$ thousand increase in annual wages will decrease target employment by 1.05 thousand workers.

The researcher fits a logarithmic version of the same model, using the same data:

$$\widehat{\log L_t} = 0.18 + 0.14 \log Y_t - 0.12 \log w_t + 0.82 \log L_{t-1}$$
$$(0.03)\ (0.03)\qquad (0.06)\qquad (0.10)\qquad d = 1.90$$

(c) (5 points) Again, briefly provide an interpretation of the short run and the long run dynamics of the regression and perform appropriate statistical tests.

Short-run (current period) behaviour: The short-run elasticity of employment with respect to output is 0.14; the short-run elasticity with respect to annual wages is -0.12.

Adjustment behaviour: The coefficient of L_{t-1} implies that only 0.18 of the discrepancy between target and previous actual employment is eliminated in any year.

Long-run behaviour: Looking at the theoretical coefficients of the model, if one divides the coefficient of $log\, Y_t$ by the estimate of λ one obtains an estimate of the long-run elasticity with respect to output: $0.14/0.18 = 0.78$. Similarly, if one divides the coefficient of $log\, w_t$ by the estimate of λ one obtains an estimate of the long-run elasticity with respect to wages: $-0.12/0.18 = -0.67$.