**Price Elasticity of Demand for Air Travel**

Use the data in AIRFARE.DTA for this exercise. We are interested in estimating the price elasticity of air travel demand. The main variables in the data set are:

*passen*: average passengers per day

*fare*: average airfare.

*dist*: the route distance (in miles)

*concen*: market share of the largest firm (how monopolistic, or competitive is the market)

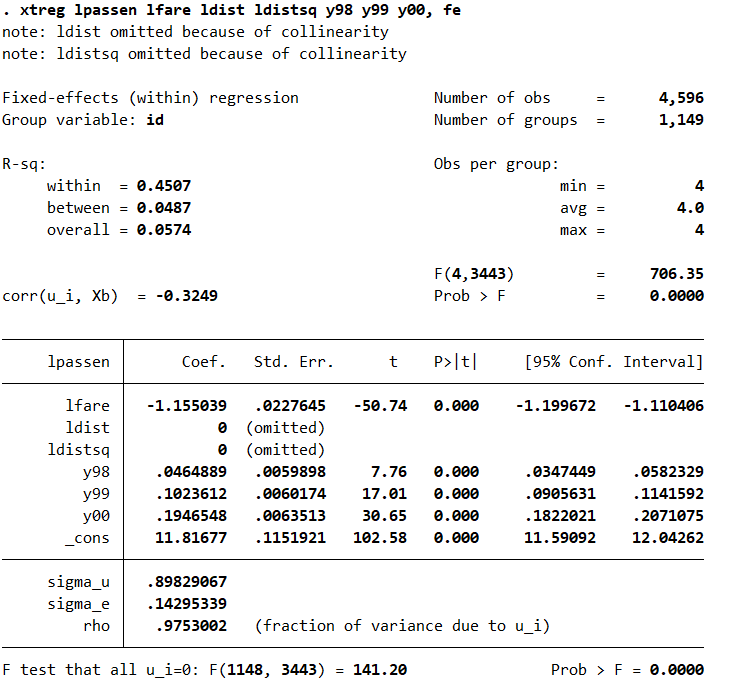
*y98, y99, y00*: yearly dummy variables

***Question 1***

The price elasticity of air travel demand can be measured by using the OLS regression model above, controlling for the year specific effects (y98, y99, y00) and the route distance (ldist and ldistsq). Estimated price elasticity of air travel demand can be interpreted as, one percent increase of average one way fare (in dollars - price) decreases the number of average passengers per day (demand) by 0.565 percent. Since we took the year 1997 as our reference point we can find the average change in demand for air travel by looking the coefficients of the yearly dummy variables.

The average change in demand for air travel from year 1997 to 2000 is 13.8 percent increase of “average passengers per day” which is the 100 times coefficient of the variable *y00* (log-level relation)*,* since the year 1997 is the reference and *y00* represents the average change in demand from 1997 to 2000.

***Question 2***



Estimated Price Elasticity of Air Travel Demand:

* According to the fixed effects model, when average one way fare (price) increases one percent, the number of average passengers per day (demand) decreases by 1.155 percent.

The fixed effects model above shows that ldist and ldistsq (route distance variables) are omitted. We are not able to estimate the effect of route distance on air travel demand with fixed effects model since route distance is a time-invariant variable. Route distance between two locations does not change by time and a fixed effects method uses time variation to estimate the model.

**Fixed Effects or OLS**

The aim of the fixed effects model is estimating causal effects in the presence of time-invariant endogeneity. When T > 2 (number of time periods) which is 4 in our case (1997, 1998, 1999, 2000), using fixed effects model is more efficient than using OLS differencing. In general, using fixed effect estimation on panel data should give better results.

However in our model, included distance variables (ldist and ldistsq) are time-invariant variables (not changing time-dependently) which can be problematic. Route distance variables appear as significant variables in the first model which makes their effects on demand critical. Thus, not including them might lower the accuracy of our model.

According to Wooldridge (2010), pooled OLS should be used when you select a different sample for each year/month/period of the panel data and Fixed effects should preferred when you are going to observe the same sample of individuals/countries/states/cities/etc. In our data, routes are being used as samples (each id represents) and we observe each same routes for four different years. Using fixed effect model is a better choice (In FE model, standard error of *lfare* is smaller and R-squared of the model is greater than pooled OLS model of the first question).

***Question 3***

Simultaneity bias or reverse causation appears when the dependent variable and explanatory variables are jointly determined. The classical theory of supply and demand suggests that when price increases, demand decreases. On the other hand, prices tend to rise when demand exceeds supply. The systems of supply demand equations are typical cases for the simultaneity bias (There might be a simultaneity bias problem in our price elasticity model, too.). These systems can be estimated by two-stage least squares method and instrumental variables.

According to the definition of instrumental variable,

* it does not appear in regression (no direct effect on y),
* it is highly correlated with the endogenous variable,
* it is uncorrelated with the error term.

The conditions listed must be met (or assumed) to use a variable as an instrumental variable. Shortly, an instrumental variable has an effect on a dependent variable only through the channel of endogenous variable. Thus, to be treated as a valid instrumental variable we should assume that concen is not correlated with error term (exogeneity of instrumental variable), and it is correlated with the endogeneous variable (lfare) with an undirect effect on lpassen. The assumption where Concen is highly correlated with lfare is called “relevance assumption”.

Exogeneity Assumption & Relevance Assumption

Assume that exogeneity is not satisfied for lfare:

Cov(lfare, u) 0

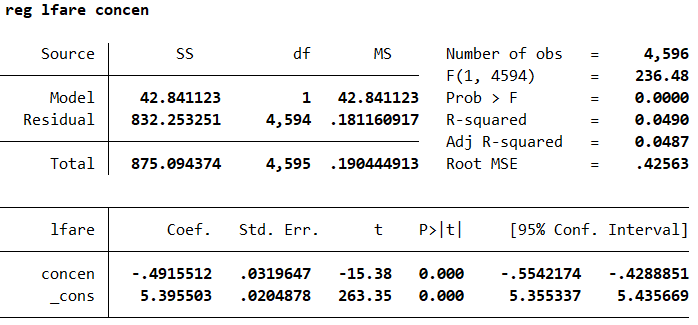
and there is an IV (concen) for which:

Cov(concen, u) = 0

⇒Cov(concen, y − β0 − β1x) = 0

Cov(concen, y) − 0 − β1Cov(concen, lfare) = 0

⇒β1 = Cov(concen, y)/Cov(concen, lfare)



Is *concen* is highly correlated with lfare?

Yes, the p-value is almost 0 which makes the effect significant. The relevance assumption is satisfied. The market concentration of the largest firm clearly affects the prices (airfares).

*Validity of the Instrumental Variable:*

It doesn’t appear as regressor. ✓

It is correlated with fare. ✓

It may not be uncorrelated with the error term (?) 🡪 concen might correlated with the country, economic stability, many extra conditions like pandemic (Covid-19) etc.

***Question 4***

Using the concen as an instrumental variable, we are going to estimate the demand function by Two Stage Least Squares (2SLS) estimator.

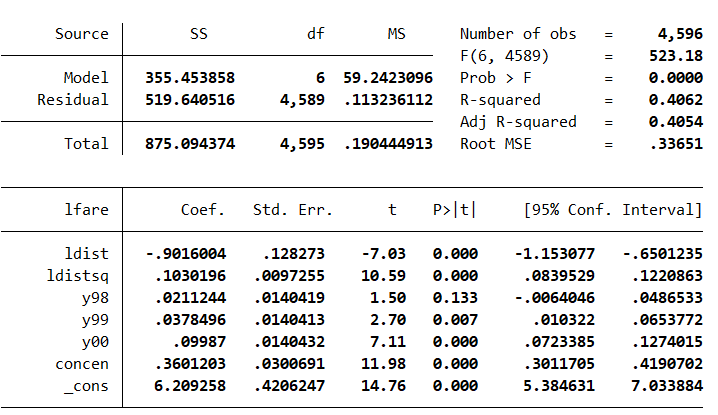
Two step estimation of the 2SLS method has the following procedure;

Model

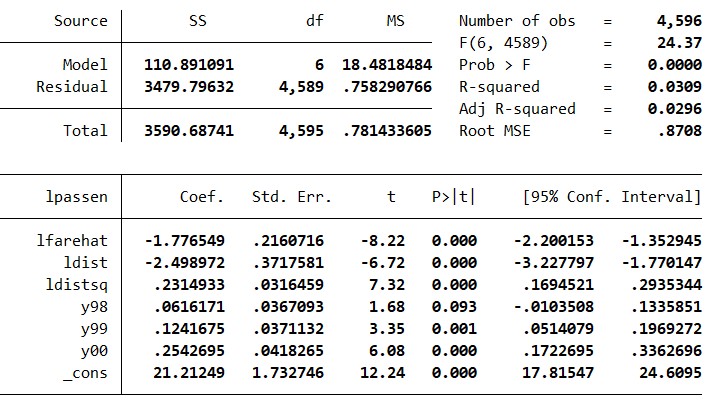
* First, the endogeneous explanatory variable “x” is predicted using only the exogeneous information and the instrumental variable.
* Second, the main OLS model is estimated with “x” replaced by its prediction from the previous stage.

Now, we are going to estimate demand function using Two Stage Least Squares estimater manually (with two steps version as we discussed in the class).

First Step – Prediction model of lfare (by exogeneous information and IV)

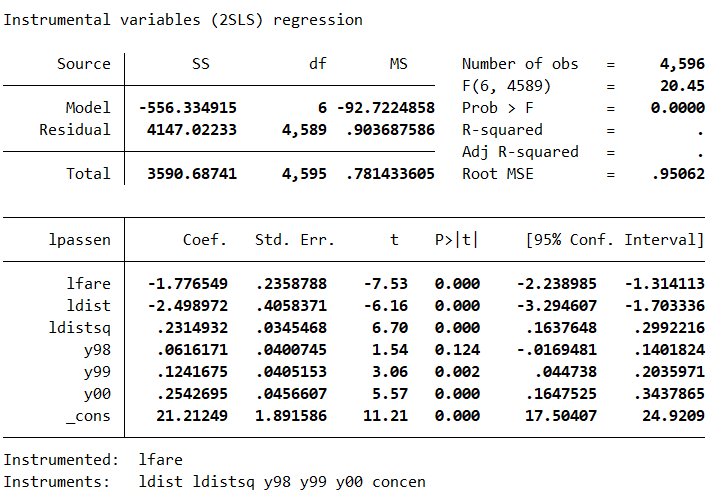


Second Step – OLS model with lfare replaced lfarehat from the first step



The price elasticity of demand goes as follows; if average one way airfare (price) increases one percent, the number of average passengers per day (demand) decreases by 1.777 percent.

The estimation of the same model using “*ivreg”*:



When we estimate the same model with “ivreg” command in STATA, we see that the coefficients appear to be the same for each variable which shows that our 2-step 2SLS model is correct. However, there is a main difference that can be detected easily: Standard errors of the variables are different. The standard errors in the second stage of 2-step 2SLS model are wrong, which are corrected in the model with *“ivreg”* command of STATA. The standard errors are larger in the correct “*ivreg*” version.

Mert Ekici 26772

ekicimert@sabanciuniv.edu