Eviews for Panel Data

George Chobanov
Eviews for Panel Data

• The Genesis of Econometric Relationships
• Types of Data
• Advantages of Using Panel Data
• Limitations and Challenges in Applying Panel Data
• Panel Data Sources
• Panel Data Studies
The Genesis of Econometric Relationships

- The **econometric relationship** of one economic variable $Y$ on one or more economic variables $X_1, X_2, ..., X_k$ is said to be regression, of historical reasons.

- Indeed the economic variable $Y$ is **explained** by the economic variables $X_1, X_2, ..., X_k$ by a relationship of $Y = f(X_1, X_2, ..., X_k) + u$, with $f$ staying for the deterministic and with $u$ – incorporating all the uncertainty (randomness, unpredictability)

- Specifications on $f$ and $u$; **classical conditions**:
  - $f$ – linear
    
    $Y = \beta_0 + \beta_1 X_1, \beta_2 X_2, ..., \beta_k X_k + u$
  - $u$ – normally distributed
Types of Data

- Data for specification or estimation of the parameters of $f$ and $u$ are based on observation of the variables $Y$ and $X_1, X_2, \ldots, X_k$.
- The observations on the variables are denoted by the same letters, but adding an index $j$: $Y_j$ and $X_{1j}, X_{2j}, \ldots, X_{kj}$, for $j=1,2,\ldots,n$.
- Three types of data: time series, cross-sectional, and pooled (panel) data
  - Time series data: Observations on values of a variable taken at different times; daily, weekly, monthly, quarterly, annually. Stationary?!
  - Cross-sectional data: Observations on values of a variable taken at the same point in time from different populations (sample spaces). Heterogeneity?!
  - Pooled data: Observations on values of a variable taken at different times and from different populations (sample spaces).
  - Panel (longitudinal) data: Observations on values of a variable taken at different times and from different populations (sample spaces), but particularly for the same statistical units of the population.
Advantages of Using Panel Data

- Panel data have the theoretical ability for deriving general economic relationships isolating the effects of specific actions, treatments, or more general polices.
- Panel data give a richer source of variation, which allows more efficient estimation of parameters.
- In panel data one can control individual heterogeneity.
- Panel data are better than cross-sectional data for studying complex issues of dynamic behavior.
Limitations and Challenges in Applying Panel Data

• Heterogeneity Bias

The theoretical ability of panel data for deriving general economic relationships isolating the effects of specific actions, treatments, or more general polices is based on the assumption that economic data are generated from controlled experiments in which the outcomes are random variables with a probability distribution that is a smooth function of the various variables describing the conditions of the experiment. Indeed homogeneity of data is assumed. From probabilistic point of view:

The economic variable $y$ is supposed to be generated by a parametric probability distribution function $P(y|\Theta)$, where $\Theta$ is an m-dimensional real vector, identical for all statistical units at all times. If it is not the case, there is a heterogeneity bias. Example:
Limitations and Challenges in Applying Panel Data

- Selectivity Bias:
  appears, if the sample may not be randomly drawn from the population
Panel Data Sources

- National Longitudinal Survey of Labor Market Experience (NLS) in US
- Panel Study of Income Dynamics (PSID) at University of Michigan
- Socio-Economic Panel (SEP) in the Netherlands
- German Social Economics Panel (GSOEP)
- British Household Panel Study (BHPS)
- National Data Collection Units (NDUs) of the Statistical Office of the European Communities (Eurostat)
- Mannheim Innovation Panel (MIP)
- Mannheim Innovation Panel-Service Sector (MIP-S)
- European Community Innovation Survey (CISI and CISII)
- European Community Household Panel (ECHP)
Panel Data Studies

- **Notations:** Panel data have two dimensions: time $t=1,2,\ldots,T$ and population $i=1,2,\ldots,N$ and therefore could be put in a matrix or a panel. For the classical linear model:

\[ Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \cdots + \beta_k X_{kit} + u_{it} \]

- **Microeconomic Relationships**
  - Relationships in Households Behavior
    \[ Y_{it} = \beta_0 + \beta_1 X_{1it} + u_{it} \]
    $X$ – income, $Y$ - consumption
  - Relationships in Enterprises Behavior

- **Macroeconomic Relationships**
  - Inflation-Growth Relationship
  - Governance Performance – Transaction Costs Relationship for Measuring Transaction Cost Efficiency