Econometrics

Types of economic data

- Econometric models use economic data

- Types of economic data
  - Cross-sectional data
  - Time series data
  - Pooled cross sections
  - Panel/longitudinal data

- Which econometric models we apply depends on the type of data used

Cross-sectional data

- Data for people, households, businesses, countries, cities, etc.
- Data are at a given point of time/during a given period – no time dimension
- Typically denote the individual by i
- Widely used in microeconomics

Example: cross-sectional data on wages

Wage y_i Hourly wage	Educ x_{1i}	Exper x_{2i}	Tenure x_{3i}	Female x_{4i} Female=1	Married x_{5i}	lwage (log wage)	expersq (experience squared)
3.1	11	2	0	1	0	1.131402	4
3.24	12	22	2	1	1	1.175573	484
3	11	2	0	0	0	1.098612	4
6	8	44	28	0	1	1.791759	1936
5.3	12	7	2	0	1	1.667707	49
8.75	16	9	8	0	1	2.169054	81
11.25	18	15	7	0	0	2.420368	225
5	12	5	3	1	0	1.609438	25

Note i, y_i, x_{1i}, etc. Every observation/row is for person i.

Time series data

- Data can be macroeconomic, financial, etc.
- Examples: stock and bond prices, GDP, growth rates
- Observations are over time
- The time dimension can be annual, monthly, daily, etc.
- Time series may have trend (e.g. rising values over time), seasonality (e.g. higher values in a given month), and cycles (e.g. every 3-5 years).
- Observations may be serially correlated (errors are correlated from one period to the next).

Time series data on minimum wages

Year t	Avgmin y_t Avg min wage	Avgcov x_{1t} Avg coverage	Prgnp x_{2t} GDP	Prunemp x_{3t} Umemployme nt rate
1950	0.198	0.201	878.7	15.4
1951	0.209	0.207	925	16
1952	0.225	0.226	1015.9	14.8
1953	0.311	0.231	1081.3	14.5
1954	0.313	0.224	1104.4	15.3
1955	0.369	0.236	1138.5	13.2
1956	0.447	0.245	1185.1	13.3
1957	0.488	0.244	1221.8	12.8

Note t, y_t, x_{1t}, etc. Every observation/row is at time t.

Pooled cross sections

- Example: house prices in two periods but not the same houses are sold.
- Two or more cross sections
- Cross sections are drawn independently of each other
- Used to estimate effect of new policy, for example effect of new tax on house prices

Pooled cross-sectional data on house prices

Year t	Price y_t	Rooms x_{1t}	Baths x_{2t}	lprice	y81
1978	60000	7	1	11.0021	0
1978	40000	6	2	10.59663	0
1978	34000	6	1	10.43412	0
1978	63900	5	1	11.06507	0
1981	49000	6	1	10.79958	1
1981	52000	5	1	10.859	1
1981	68000	6	2	11.12726	1
1981	54000	6	1	10.89674	1

Note t, y_t, x_{1t}, etc. Every observation is at time t. Before and after period.

Panel data or longitudinal data

- Example: employment data across individuals and over time
- Same cross-sectional units over time
- Have both cross-sectional i and time series t dimensions

Panel data on wages

Person id i	Year t	lwage y_{it}	Exper x_{1it}	Educ x_{2it}	Hours x_{3it}
13	1980	1.19754	1	14	2672
13	1981	1.85306	2	14	2320
13	1982	1.344462	3	14	2940
17	1980	1.675962	4	13	2484
17	1981	1.518398	5	13	2804
17	1982	1.559191	6	13	2530
18	1980	1.515963	4	12	2332
18	1981	1.735379	5	12	2116
18	1982	1.631744	6	12	2500

Note i, t, y_{it}, x_{1it}, etc. Every observation is for unit i at time t.