



**United Nations Economic Commission for Europe
Statistical Division**

Integrating data from different sources for the bigger picture

**Steven Vale
UNECE**

steven.vale@unece.org





Why Integrate?

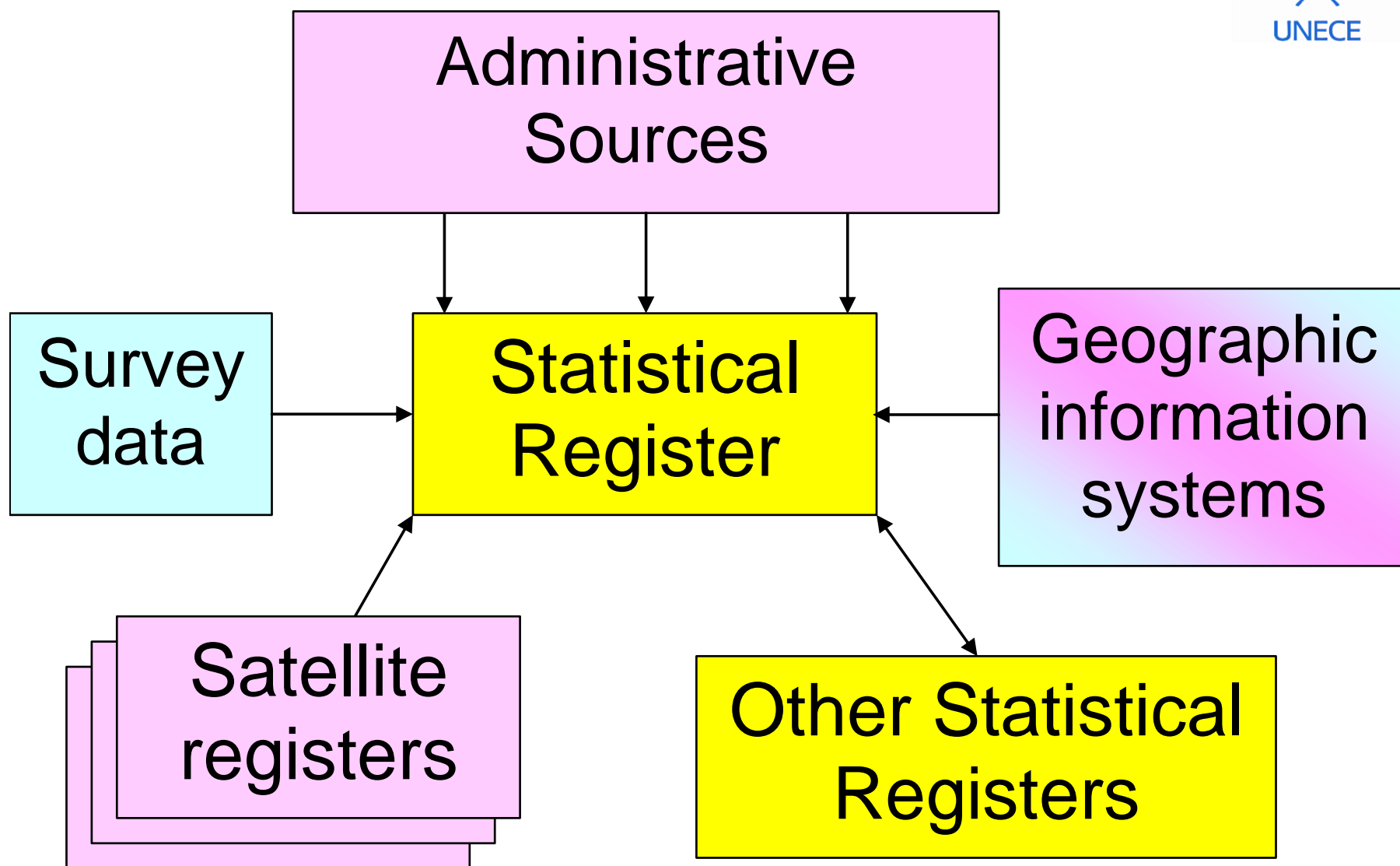
- Reduce response burden
- Reduce costs
- Estimate missing data
- Build efficient sampling frames
- Produce more detailed statistics



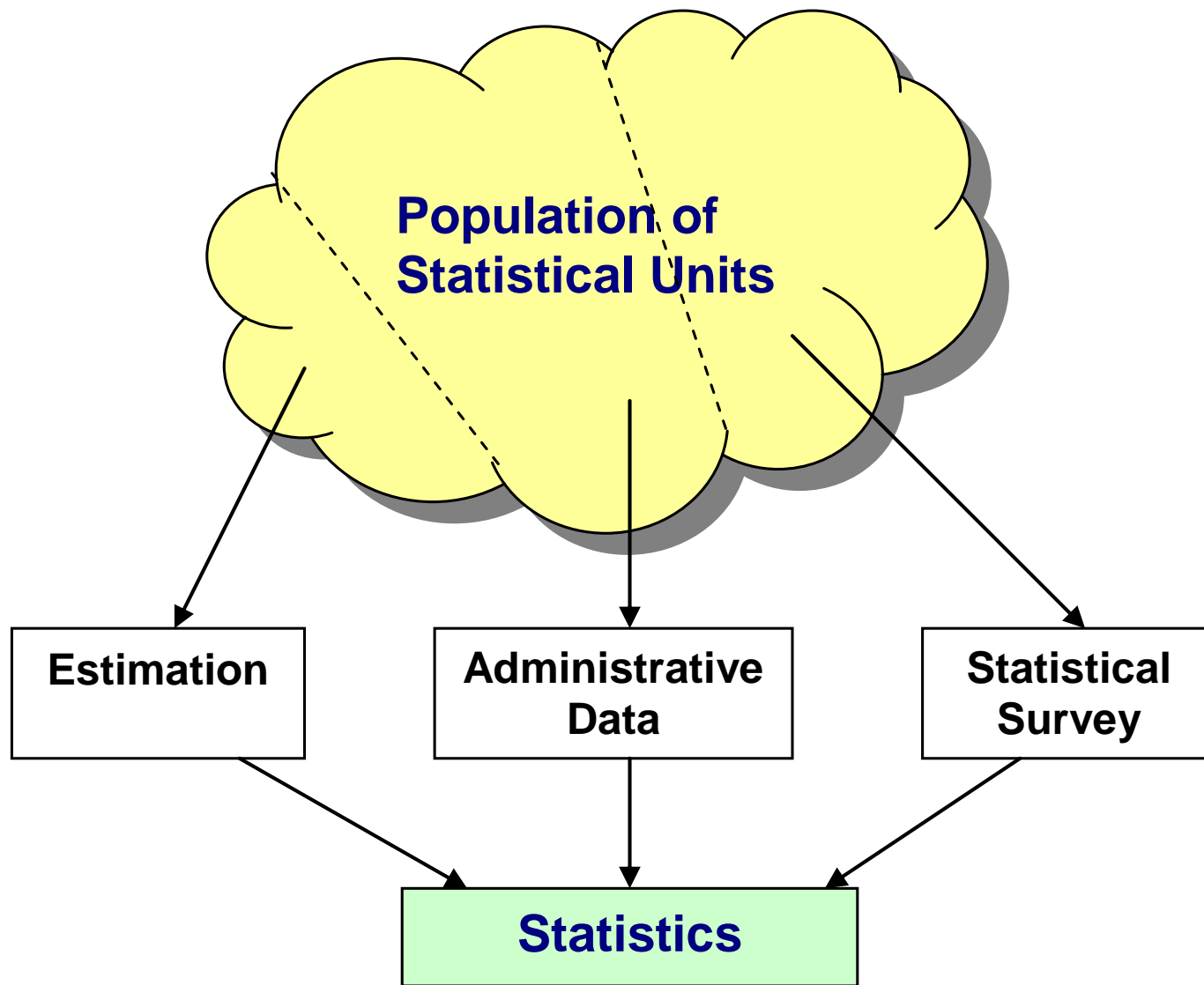
Models for Data Integration

- ❖ Statistical registers
- ❖ Statistics from mixed source models
 - Split population model
 - Split data approach
- ❖ Pre-filled questionnaires
- ❖ Using additional sources for non-responders
- ❖ Using additional sources for validation

Statistical Registers

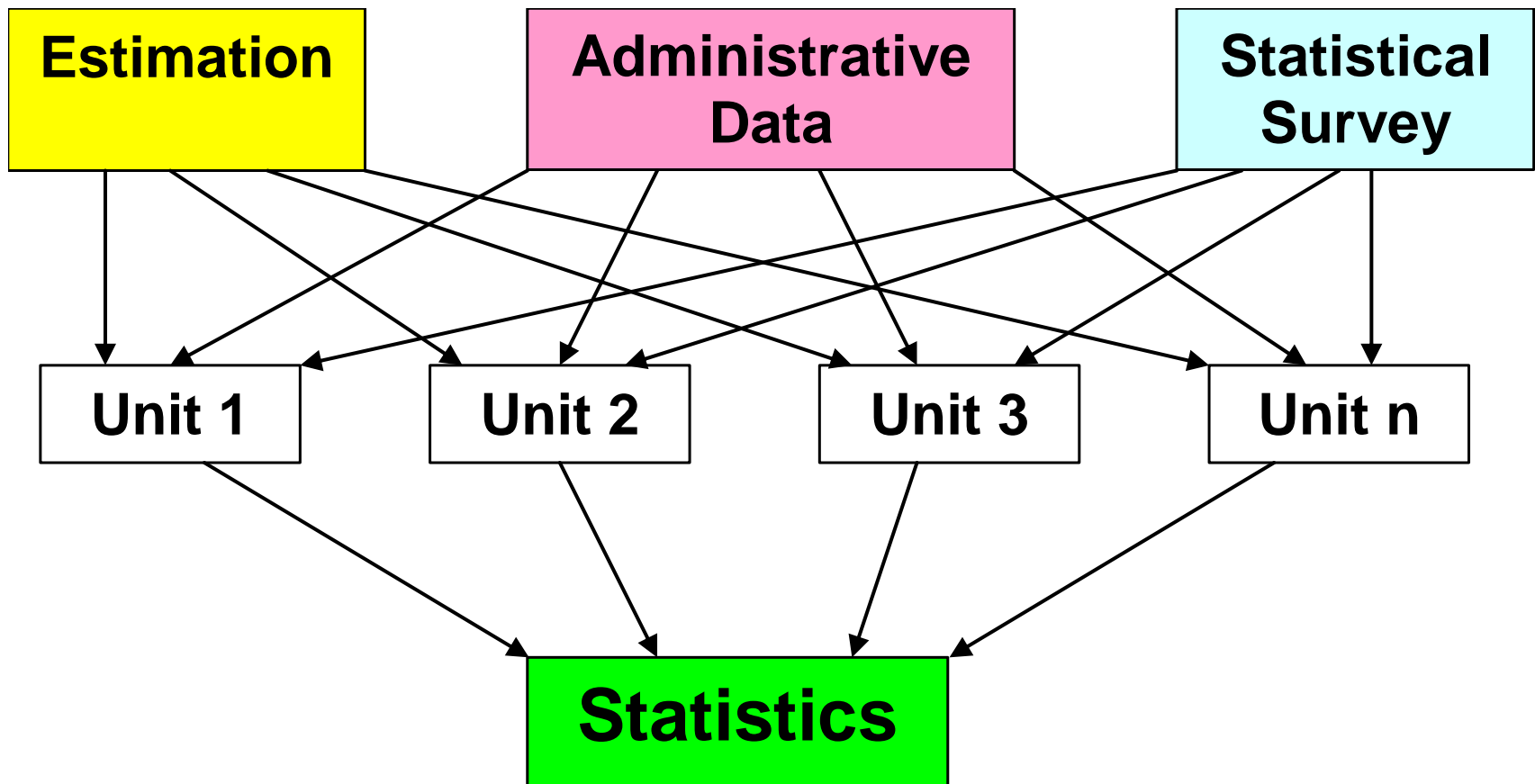


Split Population Model



Split Data Approach

- ❖ Several sources of data for each unit





Pre-filled Questionnaires

- ❖ Survey questionnaires are pre-filled with data from other sources where possible
- ❖ Respondents check the data are correct
- ❖ Reduces response burden
..... but may introduce bias!

2. Your workplace business activity (see guidance note 5)

Our records show your business activity as:

Manufacture of wooden furniture

2a. If your current business activity is different, please describe it fully in the boxes below (one letter per box with a space between each word).



Using additional sources for non-responders

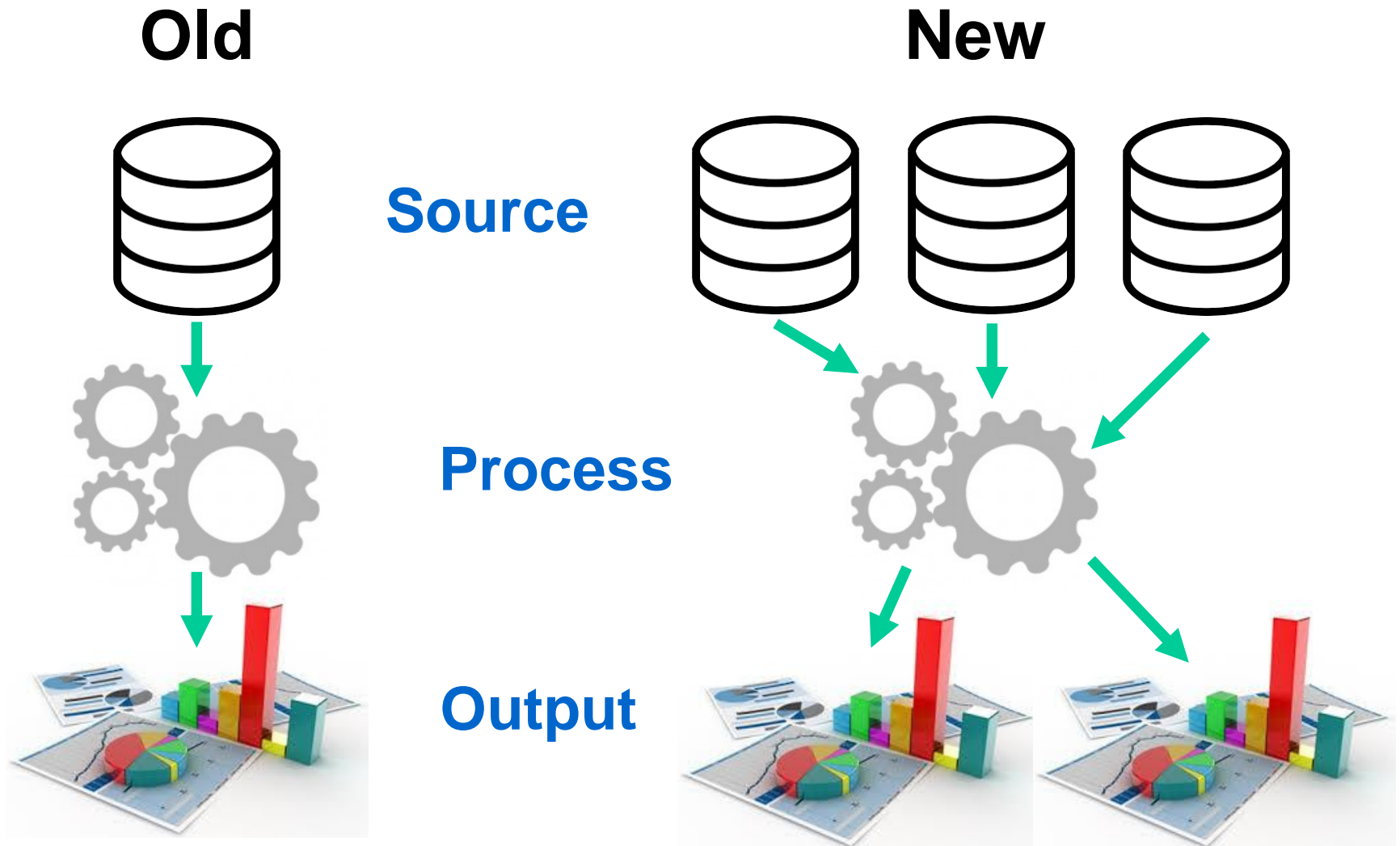
- ❖ Other sources are used directly to supply variables for units that do not respond to statistical surveys
- ❖ Often used for less important units, so that response-chasing resources can be focused on key units



Using additional sources for validation

- ❖ Plausibility checks – are levels and trends consistent?
- ❖ “Data confrontation” to understand differences between sources

Changing the data paradigm





Thank-you!

steven.vale@un.org