

## GENERAL STRUCTURE OF THE ENTERPRISE SURVEYS PANEL DATASETS

The Enterprise Surveys panel datasets have the following common format:

- Variable `panel` allows easy identification of panel observations
- Variable `panelid` is the same across the waves for the same firm
- Variable `eligibility<year>` reports eligibility status of all firms interviewed in the previous wave as of the <year> of the latest wave
  - e.g. in 2013-2023 panel, `eligibility2023` reports status as of 2023 of all firms interviewed in 2013
- Variable `status<year>` reports the status of eligible firms as to whether they have been selected in the sample to be interviewed or the interview was a refusal.
- Wherever possible variables are matched across waves. If needed, matches are made by converting variable names in older waves to variable names in the most recent wave
- Due to methodological changes and evolution of the survey instrument it is not possible to match all variables in the datasets
- Variables that are not matched across waves are named as `_year_variable`, with the year in which the variable was collected (e.g. `_2013_date`)
- It is recommended that users thoroughly familiarize themselves with the questionnaires from each of the years contained in the dataset before proceeding with analysis
- Some monetary unit variables in 2002 and 2005 surveys (in US currency) are converted into the local currency units (LCU) using the market, period average, exchange rates. The sources of the exchange rates are the International Financial Statistics (IFS - IMF) websites.
- Weights are representative of the universe for the year that the firm was interviewed. They are **not** panel weights.

To understand more about the panel structure, please type in Stata: `tab panel year`

Panel: Firm interviewed in these years	Year of survey			Total
	2009	2013	2023	
2009 only	137	0	0	137
2013 only	0	168	0	168
2023 only	0	0	380	380
2009 and 2013 only	112	112	0	224
2013 and 2023 only	0	83	83	166
2009, 2013 and 2023	119	119	119	357
Total	368	482	582	1,432

The image below illustrates the structure of the dataset

	<code>idstd</code>	<code>panelid</code>	<code>year</code>	<code>panel</code>
1	453951	453951	2009	2009, 2013 and 2023
2	527762	453951	2013	2009, 2013 and 2023
3	1158526	453951	2023	2009, 2013 and 2023
4	453952	453952	2009	2009, 2013 and 2023