

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 2018-2023 panel, `eligibility2023` reports status as of 2023 of all firms interviewed in 2023
- 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. `_2023_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
- 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**

```

. tab panel year

```

Panel: Firm interviewed in these years	Year of survey		Total
	2018	2023	
2018 only	376	0	376
2023 only	0	376	376
2018 and 2023	224	224	448
Total	600	600	1,200

The image below illustrates the structure of the dataset

	idstd	panelid	year	panel
1	649324	649324	2018	2018 and 2023
2	1193227	649324	2023	2018 and 2023
3	649327	649327	2018	2018 and 2023
4	1193228	649327	2023	2018 and 2023
5	649329	649329	2018	2018 and 2023
6	1193229	649329	2023	2018 and 2023