|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Definition** | **% of Missing’s** | **Missing years** | **Type** |
| indiv\_num | ID # |  | - | - |
| datdeb/ datdeb\_m/datdeb\_q | dates by day, moth, quarter |  | - | - |
| hplus | Desired hours in the reference week | 85 | - | Continuous |
| salred | Estimated monthly income | 8.86 | - | continuous |
| sexe | Gender | 0 | - | binary |
| married | Marital Status | 0 | - | binary |
| naf10 | Job Industry | 0 | - | categorical |
| empnbh | Hours worked in the week of interrogation | ~0 | - | continuous |
| ancentr | Job tenure (months) | 0.32 | - | continuous |
| nivet | Highest level of education | ~0 | - | categorical |
| urban | urban vs rural | 0 | - | binary |
| age | Age | 0 | - | continuous |
| etranger | Foreigner | ~0 | - | binary |
| dimanc | Works on Sundays | ~0 | - | categorical |
| samedc | Works on Saturdays | ~0 | - | categorical |
| horaic | Type of Work Schedule | ~0 | - | categorical |
| maisoc | WFH | ~0 | - | categorical |
| pnai28 | Birth country | ~0 | - | categorical |
| paimerc | Coutnry of birth (mother) | 17.11 | 2003-2004 | categorical |
| paiperc | Country of birth (father ) | 17.11 | 2003-2004 | categorical |
| zus | living in a “Zone urbaine sensible” | ~0 | - | binary |
| dep | department de residence | 0 | - | categorical |
| in\_tepa | Individual is in the survey during the shock | - | - | binary |

**Dictionary**

**Note:**

**df\_master.csv** holds the data for years 2003-2010 and has a total of **311,970 observations.**

* The above stats on missing values applies to df\_master.csv.

**df\_tepa.csv** holds data for individuals queried before and after the shock I am studying. It holds a total of **44,429 observations.**

* When moving to df\_tepa.csv, all stats on missing values become exact 0’s, apart from
  + ancentr: 0.41% of missing values
  + salred\_lowess: 7.95% of missing values
  + hplus: keeps 82.79% of missing values