1. The date on which the author first downloaded, or obtained in some other way, the original data file (e.g. when was the data collected, when was it documented, and what version of data is used for the final analysis).
2. Explanation for where each of the data files can be downloaded from (in the case of this paper, writing the link to freely available data).
3. Additional information that might help the reader understand data better. Things that could be mentioned here are: file structures (e.g., the delimiters used to separate variables, or, in rectangular files without delimiters, the columns in which the variables are stored) or how to import the data files to be compatible with different softwares (apart from STATA).
4. Create a codebook for each data file. For example, here is a codebook format for the file “JDE HM data -- individual level.dta”. I do not know more details about the variables so I have only filled in the columns I can understand from the data. Such a codebook should be included for all the nine data files (and values should also be inputted into STATA such that the variable descriptions are visible while running “describe datafile.dta”.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable name (one word) | value label | Variable label (description) | Valid range | Value indicating missing data | Value indicating inapplicable data |
| vill |  | village name |  |  |  |
| Bari |  | bari name |  |  |  |
| hhid |  | household ID |  |  |  |
| age |  | age of respondent |  |  |  |
| mothereduc |  | education of respondents mother |  |  |  |
| educ |  | education of respondent (in years) |  |  |  |
| garmentvill |  | whether garment village |  |  |  |
| cementfloor |  | whether house has cement flooring |  |  |  |
| work |  | whether respondent is working |  |  |  |
| garment |  | whether garment village |  |  |  |
| wage |  | wage earned by respondent |  |  |  |
| exper |  | yrs of work for respondent |  |  |  |
| agefirstbirth |  | respondents age at first birth |  |  |  |
| female |  | whether respondent is female |  |  |  |
| marriageage |  | respondents age at marriage |  |  |  |
| married |  | whether respondent is married |  |  |  |
| haschild |  | whether respondent has any children |  |  |  |