DESCRIPTION OF DATA AND DO FILES FOR THE PAPER

"CASTES AND LABOR MOBILITY"

BY HNATKOVSKA, LAHIRI, PAUL -- FORTHCOMING AEJ: APPLIED ECONOMICS

_____

This folder contains the following data files:

1. **"quota.dta"** contains reservation quotas for SC/STs in all Indian states, for each year of survey round. The sources for this data are:
   - For 1983:
     Dynamics of Reservation Policy, 1985, Edited by Haroobhai Mehta, Hasmukh Patel, Patriot Publishers, Link House, Bahadur Shah Zafar Marg, New Delhi 10002
   - For 1987-88:
     Mechanics of reservation for scheduled castes & scheduled tribes in services : with explanations & illustrations, including updated lists of SCs & STs / by B.M. Masand ; historical background by P.S. Krishnan, New Delhi : Ankur Publishers, 1991
   - For all other rounds:
     Report of the Commissioner for Scheduled Castes and Scheduled Tribes, Government of India, New Delhi (reports for various years)

2. **"deflators2_rXX.dta", where XX=38,43,50,55,61** indicating the survey round, contains state-level poverty lines that we used to convert wages into real terms. Poverty line information is available separately for rural and urban sectors. In our analysis we express all wages in 1983 rural Maharashtra prices. The source of this data is Planning Commission of India, Government of India, New Delhi.

3. The NSS data is not provided due to confidentiality issues, but can be obtained from:
   Deputy Director General
   Computer Centre
   M/O Statistics and PI
   East Block No. 10 R.K. Puram
   New Delhi-110066

   Enquiries relating to procurement of data can be made at :
   Email nssdat@gmail.com
   (+91) 11 - 26107267 (Mr. M.L.Phillip, Deputy Director)
   Tel/Fax :   (+91) 11 – 26107267, (+91) 11 – 26160652,(+91) 11 - 26107649

The folder also contains the following do files to extract, clean the data, including the NSS data, and to merge all datasets to create our working dataset. We also include detailed do files to replicate all the results in the paper.

TO REPLICATE THE RESULTS PLEASE USE THE FOLLOWING STEPS:

1.   Extract all variables by NSS round as described in **README_VARIABLES.txt**. Please note, these variables come from different raw data files within each NSS round. Therefore, you need to merge variables to get round specific data set. Also, keep if special_code == 1 | special_code == 3 for round 38, 41 and 50 (Read NSS documentation). The merging keys(variables) for each round are given below.

Round 38: sector state region stratum sub_sample sub_round fsu hh_no person_no
Round 43: sector state region stratum sub_sample sub_round fsu sss hh_no person_no
Round 50: sector state region stratum sub_sample sub_round fsu sss hh_no person_no
Round 55: sector state region stratum sub_sample sub_round fsu hg_sb sss hh_no person_no
Round 61: sector state region stratum sub_sample sub_round fsu hg_sb sss hh_no person_no

2.   Use round specific do files to clean variables and generate consistent variables over the rounds:
Round 38: cleaning_38.do
Round 43: cleaning_43.do
Round 50: cleaning_50.do
Round 55: cleaning_55.do
Round 61: cleaning_61.do

3.   Use price deflators files to convert nominal quantities to real quantities. Merge by "state_old" and "sector" (already coded in "cleaning_XX.do"):
Round 38: deflators2_r38.dta
Round 43: deflators2_r43.dta
Round 50: deflators2_r50.dta
Round 55: deflators2_r55.dta
Round 61: deflators2_r61.dta

4.   Once you have all "dataXX_clean.dta" files from step 2, run **"merge_data.do"** file to get the final dataset for the analysis.

5.   Add the reservation quota information which is given in the data file **"quota.dta"**. Merge by "round" and "state_old".

6.   Use **"main.do"** to replicate the results in all tables and figures in the paper, except Tables 6 and 8. Tables 6 and 8 in the can be replicated using **"tables6and8.do"** file.