This document describes the STATA programs used to generate the results in Tables 2 and 3 of Section 3.6 in Güvenen, Karahan, Ozkan and Song 2021, Econometrica. As explained in the manuscript, this section uses data from the Panel Study of Income Dynamics. We provide documentation for all of the programs that take the PSID raw files and produce the numbers in Tables 2 and 3. Two output files are produced. Table2.xlsx contains the numbers in Table 2 and psid_stats_all_6groups.xlsx contains those in Table 3.

Raw data, consisting of the family files for years 1999, 2001, 2003, 2005, 2007, 2009, 2011, and 2013 and the individual file, are provided in the folder "Intermediate." They have been downloaded from the PSID website (https://simba.isr.umich.edu/data/data.aspx).

Below is a list of the programs that are provided with the replication package and a short description for them:
1) master_PSID.do
   - sets up the main directory and the relevant paths for the code to run
   - runs the other programs in the correct order and produces the results in Tables 2 and 3.
   - the folder structure should mimic the one provided in the replication package.
2) process_family_files.do
   - processes the family files downloaded from the PSID website by selecting only the variables relevant to the analysis
3) topcoding.do
   - deals with top coded observations
4) process_individual_files.do
   - processes the individual file by keeping the family number, sequence number, and relationship to household head
   - reshapes the data to long format
5) merge_family_individual.do
   - merges the family and individual files
6) make_annual.do
   - annualizes some monthly and quarterly variables
7) clean_data.do
   - cleans the dataset by removing those that are household head intermittently, by making the race, age and education variables consistent over time, removing households with a female head
8) higher_order_moments_6groups.do
   - computes the higher-order moments of earnings and wage changes for the entire sample (ages 25-55) and separately for ages 25-39 and 40-55. These numbers are outputted to Table2.xlsx.
   - reproduces Table 3 in psid_stats_all_6groups.xlsx.