HOMEWORK 1 REPLICATION OF TABLES 1 AND 2 FROM DGR (2011) USING THE 2007 SURVEY OF CONSUMER FINANCES DATA

HUAYAN GENG (HGENG1@BINGHAMTON.EDU)

Files Descriptions

HW1.do is the STATA command file applying to obtain all values of replications and graphs.

hw1_2007scf.log is the log file recording commands in HW1.do and results generated by it.

HW1_SUMMARY.pdf it the summary file of this homework including problems met, solutions, replication tables and Lorenz Curve graphs.

Key Points in Replication

In this section, I list worthy problems I met during replication and how I solved it.

1. I apply summary dataset of the 2007 Survey of Consumer Finances (SCF). On the website SCF website, it states that all dollar variables have been adjusted to 2013 dollars. Therefore, I have to figure out the inflation-adjustment factor of this data. Fortunately, on SCF website, I found *Changes in U.S. Family Finances from 2010 to 2013: Evidence from the Survey of Consumer Finances* which is given by Board of Governors of the Federal Reserve System, with https://www.federalreserve.gov/pubs/bulletin/2014/pdf/scf14.pdf. Hereby, the adjustment factor in 2007 is 1.1228. Thus, on the first line of my command file, I generate a variable named INF = 1.1228 as inflation-adjustment factor for my replication.

2. This dataset is generated from a survey, so there must be a weight for us to find population distribution and statistics. In the dataset, the variable, *wtg*, is the sampling weight variable for all individual. Before realizing it, I simply delete all duplicate observations and treat rest of observations equally weighted, causing huge differences between my replication and paper does. Thus, all distributions, statistics and graphs are weighted by variable *wtg*.

3. Definitions of variables in 2007 summary dataset are unknown due to no codebook for summary data. However, we can find variable definition file of summary dataset in 2013 and 2007 survey dataset codebook. Therefore, I search variable names in 2013 definition file to find the formula presented with survey dataset variables and then search those variables (survey variable in formula) in 2007 survey codebook. Finally, I find all the variable definitions in summary dataset to generate my own variables.

4. For Gini Coefficient, I try both methods including non-negative values only of the variables (i.e. earning, income and wealth) and all real value of the variables. I found that the results generated by containing all values are closer to original paper. However, when I draw Lorenz Curve, it only contains non-negarive data.

Table 1 Replication

Red letter means slightly higher than that in original paper, while blue letter means slightly lower.

Income 0 quantile is -506.0 thousand 2007 USD in DGR (2011).

Wealth 0, 40, 60, 90 and 95 quantiles are respectively -474.0, 64.7, 197.7, 908.4 and 1890 thousand 2007 USD in DGR (2011).

Date: 09/07/2017.

Replication results are very similar to DGR (2011).

TABLE 1. Replication of Quantiles of the 2007 Earnings, Income, and Wealth Distribution ($\times 10^3$ 2007 USD)

Income	-505.7	4.2	8.9	12.3	20.1	36.3	58.8	98.7	142.0	207.2	680.7	187,000
Quantiles	0	1	5	10	20	40	60	80	90	95	99	100
Earnings	-1,547	0.0	0.0	0.0	0.0	25.7	50.4	87.5	126.1	180.2	497.0	162,000
Wealth	-473.7	-31.3	-4.6	0.0	7.3	64.9	197.9	496.9	910.3	1,900	8,374	141,000

TABLE 2 REPLICATION

Red letter means slightly higher than that in original paper, while blue letter means slightly lower.
Income Gini index is 0.58 in DGR (2011).
Wealth coefficient of variation is 6.02, location of mean is 80% and mean/median is 4.61 in DGR (2011).
Replication results are extremely similar to DGR (2011).

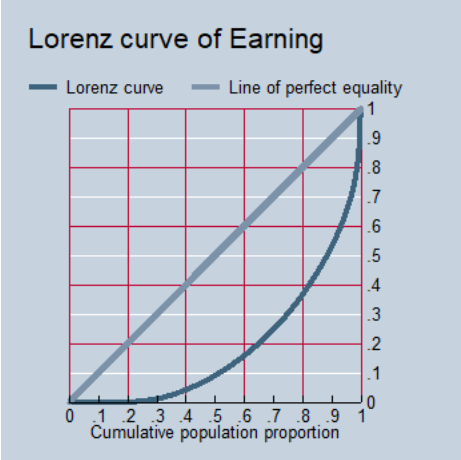
TABLE 3. Replication of Concentration and Skewness of the Distributions

	Earnings	Income	Wealth
Coefficient of variation	3.60	4.32	6.01
Variance of the logs	1.29	0.99	4.53
Gini index	0.64	0.57	0.82
Top 1% / lowest 40%	183	88	1,526
Location of mean (%)	69	74	81
Mean / median	1.72	1.77	4.60

LORENZ CURVE

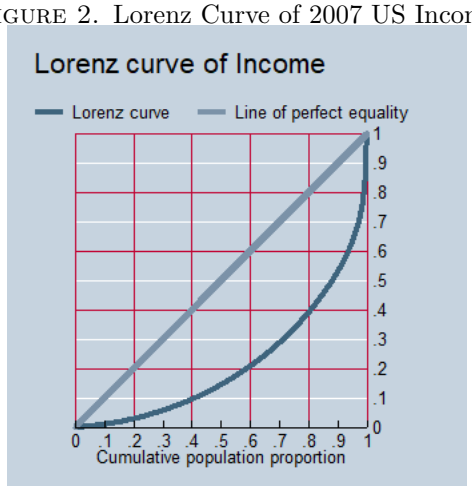
Earning

FIGURE 1. Lorenz Curve of 2007 US Earning



Income

FIGURE 2. Lorenz Curve of 2007 US Income



Wealth

FIGURE 3. Lorenz Curve of 2007 US Wealth