This document explains the data and program files contained in BPS_data&codes.ZIP. These were the files used to generate the results contained in "The Quantity and Quality of Life and the Evolution of World Inequality," by Gary S. Becker, Tomas J. Philipson, and Rodrigo R. Soares (*American Economic Review*, v95, n1, March 2005).

Files:

- section3.dta

This is a Stata data file containing the data used to generate the aggregate results from Section 3. It contains life expectancy, income, and population numbers for all the countries included in the sample, for the years 1960, 1990 and 2000. The variables are named as *rgdpttYEAR* (per capita income adjusted for terms of trade, from the Penn World Tables version 6.1), *lifeYEAR* and *popYEAR* (life expectancy at birth and population, from the World Bank's World Development Indicators).
The file also contains the variables calculated in the paper, generated by the file compensating.do.

- compensating.do

This is a Stata "do" file, containing the commands used to generate the variables and summary statistics discussed in Section 3. The variable *w1990* indicates the value of the life expectancy gains between 1960 and 1990 measured in yearly income. The variables *w2000_60* and *w2000_90* indicate the same variable for 2000, the first one measuring the value of the gains between 1960 and 2000, and the second one the gains between 1990 and 2000. An analogous variables, defined as *fullYEAR*, denotes the "full-income" (value of longevity gains in yearly income plus gdp) corresponding to these same periods.
The variable *NPV_w2000_60* denotes the lifetime net present value of *w2000_60*.
The other commands contained in the file are the ones used to calculate the statistics discussed in Sections 1 and 3.

- section4.dta

This is a Stata data file containing the data used in the calculations from Section 4. The file contains variables related to life expectancy in 1965 and 1995 (*life65* and *life95*) and population in 1965 (*pop65*). The other variables were constructed directly from the World Health Organization Mortality Database using age and cause specific number of deaths and population, according to the strategy described in Section 4.2 (the original World Health Organization data are not included because of the enormous size of the dataset). The variables are defined as follows: *dlife_rCC* denotes the change in life expectancy attributable to cause of death "*CC*," according to the groups of causes of death described in Section 4.1; *dlife020_CC*, *dlife2050_CC* and *dlife50_CC* denote the change in life expectancy attributable to cause of death "*CC*" and, respectively, age groups 0-20, 20-50, and above 50.

- full_decomposition.do

This is a Sata "do" file containing the commands used to run the regression results presented in Section 4.3, as part of the decomposition of life expectancy changes exercise.