**Codebook for Excel workbook of inequality measures**

**Each sheet pertains to a continent or group of countries and each row corresponds to a single household expenditure survey. The following variables are given**

**Country country name**

**Survey\_id Survey id number that corresponds to the bibliographical reference**

**BL\_gini\_income Income Gini coefficient using Kakwani’s Beta-Lorenz functional form**

**BL\_p5010 50/10 percentage ratio using Kakwani’s Beta-Lorenz functional form**

**BL\_p9050 90/50 percentage ratio using Kakwani’s Beta-Lorenz functional form**

**BL\_p9010 90/10 percentage ratio using Kakwani’s Beta-Lorenz functional form**

**naive\_gini\_income Income Gini coefficient using linear interpolation**

**naive\_p5010 50/10 percentage ratio using linear interpolation**

**naive\_p9050 90/50 percentage ratio using linear interpolation**

**naive \_p9010 90/10 percentage ratio using linear interpolation**

**regln\_gini\_income Income Gini coefficent using lognormal functional form**

**regln \_p5010 50/10 percentage ratio using lognormal functional form**

**regln \_p9050 90/50 percentage ratio using lognormal functional form**

**regln \_p9010 90/10 percentage ratio using lognormal functional form**

**A note on the inequality measures we estimate**

**Estimating inequality measures from surveys with individual data is straightforward, but many historical studies report tables of data on household income and expenditures where the households are grouped by income, or sometimes total expenditures. These grouped data allow us to make estimates of the degree of income/expenditure inequality among the same of households. With such grouped data we can calculate the total income of all households and also the total income shares of each income group. To estimate inequality we then need to use these shares to estimate how income shares would rise across all households, if we had data on every household. There are two ways to do this: first, a curve could be fitted to the group shares, or secondly we could simply join the group shares by drawing a line between them. To fit a curve, we need a formula for the curve. Here we offer two possibilities: Lognormal (LN) and Beta-Lorenz (BL). We also offer the simplest join the dots method (Naïve).**

**We can then use these methods to estimate inequality. We offer a Gini index, which is on a scale between 0 (perfect equality and 1 (perfect inequality)., as well as some popular percentile ratios. An example of a percentile ratio is the 90/10 ratio, which is the ratio of the income of the household for which 90% of households have lower income, to the income of the house for which only 10% of households have lower income. The 50/10 and 90/50 ratios are similarly defined.**

**By the standards of inequality measures from modern (post-1960) household expenditure surveys., many of those presented here are low. This is primarily the outcome of the survey target group being drawn from a relatively narrow social or economic class, rather than being a sample of the entire population (or nearly so) as is generally case with modern surveys. After World War Two many of the surveys were samples of the entire population and we would expect that these would show higher estimates of inequality than the earlier surveys solely for this reason. It is advisable, therefore, to examine our inequality estimates in conjunction with the survey description given in the bibliography.**