This package contains two files designed to help read individual level DHS data into Stata. The first file addresses the problem that versions of Stata before Version 7/SE will read in only up to 2047 variables and most of the individual files have more variables than that. The file will read in the .do, .dct and .dat file and output new .do and .dct files with only a subset of the variables specified by the user. The second file deals with earlier DHS surveys in which .do and .dct file do not exist and only .sps and .sas files are provided. The file will read in the .sas and .sps files and output a .dct and .do file. If necessary the first file can then be run again to select a subset of variables.

1 Program READDHS

In order to run this file the .do, .dct and .dat file should be saved in the same folder as the program, and they should all be re-named so the name is only three letters long, and all capitals. For example, for Tanzania they might be named TAN.DO, TAN.DCT AND TAN.DAT. The program will initially run as follows (with the bolded text the user entries).

ORIGINAL DO? **TAN.DO**

ORIGINAL DCT? **TAN.DCT**

OUTPUT DO? **TAN2.DO**

OUTPUT DCT? **TAN2.DCT**

DATA FILE? **TAN.DAT**

Having entered this information, the program will then say "ENTER BEGINNING LETTER OR D IF DONE". In this part of the program the user specifies the variables they want outputted. The program takes advantage of the fact that there is a set of variables that begin with the letter 'v', a set that begin with 'b', 'h', etc. When a letter is specified every variable beginning with that letter is outputted (note that if one desired the entire file outputted into two separate files you could run the program twice, and output half of the letters each time.) The program will take up to 10 letters, and once you have entered the last letter type 'D' and it will finish. Note that the program is case-sensitive. The letters entered should be lowercase, with the exception of the final 'D'.

Once finished, the file TAN.DO can be run normally from within Stata and should require no additional modifications.

2 Program READDHS2

In order to run this file, the .sas and .sps file should be saved in the same folder as the program. The files should both be re-named so the name is only three letters long, and all capitals (for example, TAN.SAS and TAN.SPS). The file will prompt for the original SAS file (TAN.SAS) and original SPS file (TAN.SPS), and then for the output DO file (TAN.DO) and output DCT file (TAN.DCT). It will output these files.

Important! Note that if you then run the READDHS file the variable names will all begin with capitals. Therefore, when you get to the step where you tell it which subset of variables you would like to extract, use capital letters rather than lowercase.