Read-Me file for creating data used in Kreisman/Rangel 2014.

I. Folder structure and path directories.

The structure of the raw data, cleaned data, dofiles and resulting tables and figures is shown in the appendix to this document.

All computation was run using Stata version 11. To run any/all programs, the user should first run the dofile *setup.do*. This program sets the path directories to the user's local or computer or remote server.[1]

The only modifications required of the user is to modify line 2 of *setup.do*, which is:

2. global directory ''ENTER YOUR DIRECTORY PATH HERE"

Between the quotes, the user should put the directory path to where the *skincolor/* folder is on their machine.

II. Extracting the data.

The data were extracted from the NLSY's web investigator, which can be found at https://www.nlsinfo.org/investigator/pages/login.jsp.

Contained within the folder skincolor/ are the NLSY taglist, the codebook, dictionary, and logfile. These can be imported directly to stata, or the taglist (skincolor.NLSY97) can be opened in the web investigator.

III. Preparing the data for analysis.

The data are prepared for analysis in a sequential set of dofiles. Each of these dofiles is called in the main dofile (2_create_skin_color_data.do). This calls:

- 1.2_create_interviewer.do, which creates interviewer characteristics, saved to 1.2_interviewer.dta
- 1.3_newentry.do, which defines labor market entry, saved as newentry.dta.
- 1.4_create_weekly_lfs.do, which creates the share of weeks respondents were employed in each survey year, saved as 1.4_weekly_lfs.dta. (This is computationally intensive and should be run on a MP server).
- 1.5_create_labels.do, which adds variable labels to the data.

The dofile 2_create_skin_color_data.do, which calls each of the above sub-dofiles, results each of the sub-datasets described above, and two main datasets used in the analysis. These are:

- skin_color_wide.dta, where the unit of observation is the respondent.
- skin_color.dta, which is a balanced panel where the unit of observation is person*survey year.

A log file, 2_create_skin_color.log, is created.

IV. Tables and figures.

All tables and figures are creates in the dofile 3_tables_and_figures.do. A log file ''3_tables_and_figures.log'' is defined at the top.

---

[1] Executing program 1.4_create_weekly_lfs.do is computationally intensive requiring a large amount of space, therefore we suggest running this on a server with at least 4g of available memory.

Appendix: Folder structure

skincolor/			
	read_me.pdf		
	data/		
		1.2_interviewer.dta	
		1.3_newentry.dta	
		1.4_weekly_lfs.dta	
		skincolor_wide.dta	
		skincolor.dta	
	do/		
		1_setup.do	
		2_create_skincolor_data.do	
		3_tables_and_figures.do	
		create/	1.2_create_interviewer.do
			1.3_create_newentry.do
			1.4_create_weekly_lfs.do
			1.5_create_labels.do
	extracts/		
		skincolor/	
			extract.log
			skincolor-value-labels.do
			skincolor.cdb
			skincolor.csv
			skincolor.dct
			skincolor.NLSY97