

# The Persistent Effects of Peru's Mining *Mita*: Instructions for reproducing the tables and figures

Melissa Dell

July, 2010.

Included are the data and programs to reproduce all tables and figures in “The Persistent Effects of Peru's Mining *Mita*.” The tables may be reproduced using only Stata (version 10 or higher). Reproducing the figures requires R, Python, and ArcGIS. Details are provided below.

## 1 Geospatial variables

The program `prepare_gis.do` creates the Stata files `gis_dist.dta` and `gis_grid.dta`, which contain the geographic variables required for reproducing Tables 1 through 9. The data files used as inputs by this program are `dbnd.csv`, `elevdis.dbf`, `sldist.dbf`, `d2_potd.csv`, `d2_huand.csv`, `dis_xy.dta`, `inmita.dbf`, `outmita.dbf`, `gr_coor.dbf`, `e3.dbf`, and `s3.dbf`. These data files were, in turn, produced by GIS programs that are also included and that are described at the bottom of this document.

## 2 Summary Statistics (Table 1)

The Stata program for creating Table 1 is `maketable1.do`. The data used to make this table are: `gis_grid.dta`, `consumption.dta`, `spec_check1572.dta`, and `budget1572.dta`. The file `consumption.dta` is produced by the program `maketable2.do`, described below, and the files `spec_check1572.dta` and `budget1572.dta` are produced by the program `maketable5.do`, also described below. Hence, these programs must be run before `maketable1.do` in order for it to function.

## 3 Living Standards (Tables 2 through 4)

The Stata programs for creating Tables 2 through 4 are `maketable2.do`, `maketable3.do`, and `maketable4.do`, respectively. To produce Tables 3 and 4, you must first run the program `maketable2`, as Tables 3 and 4 use data files created by this program.

The household data needed to reproduce the results in columns (1) through (3) of Tables 2 and 3 and columns (1) through (7) of Table 4 are contained in `enaho01_2001iv_200.dta`, `sumaria_2001iv.dta`, and `enaho01b_2001iv.dta`. The height data to reproduce columns (4) through (7) of Tables 2 and 3 and columns (8) through (12) of Table 4 are contained in `ctalla.dta`. The GIS data need to reproduce Tables 2 through 4 are in `gis_dist.dta` and `nrivers.csv` (the latter is required for Table 4 only).

## 4 1572 Tribute and Population (Table 5)

The Stata program for creating Table 5 is `maketable5.do`. The data files on 1572 tribute and population used to make this table are `tribute1572.dta`, `1572budget.csv`, and `1572demo-graphic.csv`. The program also uses the data file `gis_dist.dta`.

## 5 Haciendas (Table 6)

The Stata program for creating Table 6 is `maketable6.do`. The data files on *haciendas* and land distribution used to make this table are `hacienda1689.csv`, `hacienda1845.csv`, `hacienda1940.csv`, `apurimac-d2.dta`, `arequipa-d2.dta`, `ayacucho-d2.dta`, `cusco-d2.dta`, and `puno-d2.dta`. The program also uses the data file `gis_dist.dta`.

## 6 Education (Table 7)

The Stata program for creating Table 7 is `maketable7.do`. The data files on education used to make this table are `educ1876.csv`, `educ1940.csv`, and `enaho01a_2001iv_300.dta`. The program also uses the data file `gis_dist.dta`.

## 7 Roads (Table 8)

The Stata program for creating Table 8 is `maketable8.do`. The data files on roads used to make this table are: `distr.dbf`, `vecd'X'.dbf` and `depd'X'.dbf`, where 'X'=1/5, and 9. These `.dbf` files can, in turn, be reproduced by the program file `roads.py`, included and described at the bottom of this document. The program also uses the data file `gis_dist.dta`.

## 8 Consumption Channels (Table 9)

The Stata program for creating Table 9 is `maketable9.do`. The data files on occupation, agricultural market participation, and the agricultural labor force used to make this table are: `occupation1993.csv`, `apurimac-d2.dta`, `arequipa-d2.dta`, `ayacucho-d2.dta`, `cusco-d2.dta`, `puno-d2.dta`, `apurimac-d4.dta`, `arequipa-d4.dta`, `ayacucho-d4.dta`, `cusco-d4.dta`, and `puno-d4.dta`. The program also uses the data file `gis_dist.dta`.

## 9 Appendix (Table A3)

The Stata program for creating Table A3 is maketableA3.do. The data files used by this program are produced by the programs maketable2 through maketable9, and are hacienda1689.dta, hacienda1845.dta, hacienda1940.dta, educ1876.dta, educ1940.dta, educ2001.dta, roads.dta, and ag\_mkts1994.dta.

## 10 Maps (Figures 1 and A1)

The mapping (.mxd), shape, and raster files necessary for reproducing the maps in Figures 1 and A1 are in the folder Maps1\_A1\_Arc. Load the files Figure1.mxd and FigureA1.mxd into ArcGIS and make sure that the settings on your machine are such that relative directory paths - as opposed to the full paths - are used. Each map will display once its .mxd file is loaded.

## 11 RD Figures

The program, shape, and data files necessary for reproducing the RD figures (Figures 2, 3, and A2) are contained in the folder RDfigs. To reproduce the figures, run the file mita\_run.R, using the statistical package R.

## 12 GIS data

This section describes the python programs for producing the geospatial data files used as inputs by the program prepare\_gis.do, described above. The final outputs from these python programs are also included, so that those without access to the specialized geospatial software required to run them can still reproduce the tables. These programs require ArcGIS desktop as well as the ArcGIS 3D and spatial analyst extensions. Directory paths in the python files must be changed accordingly, including the directory paths of the Arc toolboxes that are loaded at the top of each file.

The distances to Potosí, Huancavelica, and the *mita* boundary, as well as mean elevation and slope by district, are calculated by the file gis\_prep.py. The raw data files used by this program are in the geodatabase mita.gdb and are listed at the top of the python file. To construct the data necessary for comparing elevation and slope in Table 1, the program geogrid.py can be used. The raw data files are contained in mita.gdb and described at the top of the python file. Finally, the road network variables used in Table 8 are constructed by roads.py, with the raw data again located in mita.gdb.