

27/10/2005

**Read Me File describing the data sets and computer programmes used in
Piketty/Postel-Vinay/Rosenthal “Wealth Concentration in a Developing
Economy: Paris and France 1807-1994”**

A. Files supplied

A.1. Data Files

Demographic files: this includes three Excel files offering reference demographic data on Parisian deaths, population totals and age distributions (published demographic sources used to construct these tables are described on-line)

Wealth-at-death 19th century micro-files: this includes ten Stata files (1807.ind.dta, 1817.ind.dta,..., 1887.ind.dta, 1902.ind.dta). These Stata files were constructed by extracting selected variables from the raw, individual-level Excel files on positive estates (1807.indiv.xls, 1817.ind.xls,..., 1887.ind.xls, 1902.ind.xls) and by adding the relevant number of observations corresponding to adult decedents with zero estate (using demographic files). For instance, the 1807 raw Excel file includes 3 647 decedents with positive estates, while the 1807 Stata file includes the total number of 20-year-old-and-over decedents in Paris in 1807, i.e. 11 622 decedents (including 3 647 with positive estate and $11\ 622 - 3\ 647 = 7975$ decedents with zero estate). The raw Excel files on positive estates were constructed using individual estate tax records stored at the Archives de Paris (all technical details on data collection are given on-line). The collection of this new material is a central contribution of the present research project.

Wealth-at-death 20th century published tabulations : this includes one Excel file containing estate tabulations compiled and published by the French Finance Ministry for France and Paris for a selected number of years since 1902 (no such tabulation was ever compiled prior to 1902). The references of the official publications where these raw tabulations were originally published are reported on the “ReadMe” sheet of the Excel file. These published 20th century tabulations were already used by Piketty (2001, 2003).

A.2. Program files

This includes ten do-files in text format (do1807.txt, do1817.txt,..., do1887.txt, do1902.txt) that can be readily applied to the Stata files to replicate the results.

A.3. Output files

This includes two Excel files containing all tables and graphs generated by the data files and program files.

The file WPVersion27102005.xls contains the extensive set of tables and graphs generated by the data files and program files. This file offers an updated version of the statistical appendix of the working paper (Piketty, Postel-Vinay, Rosenthal, CEPR Discussion Paper 4631, September 2004).

The file AERVersion27102005.xls contains solely the tables and graphs that were included in the published AER version. This file is an extract of the previous file and does not include all appendix tables (all details are provided on the readme sheet).

B. How to use the files to replicate the results

Most results included in the (extensive) output file WPVersion27102005.xls can be replicated by applying the program files do1807.txt, do1817.txt,..., do1887.txt, do1902.txt to the data files 1807.ind.dta, 1817.ind.dta,..., 1887.ind.dta, 1902.ind.dta.

More specifically, applying the program files do1807.txt, etc. to the data files 1807.ind.dta, etc. automatically generates the results reported for the 19th century period (i.e. years 1807, 1817, 1827, 1837, 1847, 1857, 1867, 1877, 1887, 1902) on appendix tables A1, A2, A3, A4(I) and A4(II) (as well as on tables 5 and 6).

All other tables and figures were constructed using the results reported on these appendix tables (the exact appendix table used is reported in footnotes to each table and figure).

The results reported on appendix tables A1, A2 and A3 for the 20th century period (i.e. years 1913, 1929, 1938, 1947 and 1994) were computed by applying Pareto interpolation techniques to 20th century published tabulations. These computations are described at length in Piketty (2001, 2003), to whom we borrow the methodology and results used in the present paper for the 20th century period. We also reproduce in appendix tables A5, A6 and A7 the complete 20th century series.