

de la Croix, David, and Matthias Doepke
"To Segregate or To Integrate: Education Politics and Democracy"
July 2008

The empirical analysis is based on data from the 2000 US Census.

The dataset used in the paper is a 1% sample from the 2000 U.S. Census. The data is made available by Steven Ruggles, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander. Integrated Public Use Microdata Series: Version 4.0 [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor], 2008. The data is available online at: <http://usa.ipums.org/usa>.

The extract must contain the following variables:

YEAR
SERIAL
NUMPREC
HHWT
STATEFIP
GQ
METRO
HHINCOME
PERNUM
NCHILD
RELATE
AGE
HISPAND
RACAMIND
RACASIAN
RACBLK
RACPACIS
RACWHT
RACOTHER
SCHOOL
SCHLTYPE

The size of the dataset is approximately 118 MB. To run the do-file "SegregateOrIntegrate.do" STATA 7 or higher is recommended (due to restrictions on the length of variable names in STATA 6 and earlier). The authors used STATA-SE 9.

The hardware must also have sufficient memory to run the ordered logit regressions.

For the computations presented in the paper, the authors used a UNIX-machine with 5 GB of memory allocated to STATA.

"SegregateOrIntegrate.do" is the command file for all the results presented in Tables 1, 2, 3, as well as in footnote 24.

Some results presented in Table 1 are based on computations in an Excel spreadsheet ("Table 1.xls"), which the authors also made available electronically.

(Los Angeles, August 17, 2008)