

Inter-city wage differentials and intra-city workplace centralization

Jim Dewey

Bureau of Economic and Business Research
University of Florida

Gabriel Montes-Rojas

Bureau of Economic and Business Research
University of Florida
and
City University London

1. Information required to replicate the results

Creating the Data Sets

Download the 10% US 2000 Census from www.ipums.org

Run the do-file **Centrality_allMSAs_W.do** (this will access and run **variables_US.do**). This will generate the following STATA databases:

Index_allMSAs.dta : this dataset contains the centrality index by the Standard Occupation Categories used in the US 2000 Census for each of the seven cities used in the study.

US_allvars.dta : this dataset is used for the regression results in Table 4. (Since this is a random sample, the results will not match those of the paper exactly, but, since the sample is very large, the results will be close. To replicate our results exactly, you would need to use **US_allvars_0.dta**, which contains the sample used for the reported results.)

Obtaining Results

Tables 1 2 3.do : Tables 1, 2 and 3 can be replicated using this do-file (requires **Index_allMSAs.dta**).

Table 4.do : Table 4 regression results can be replicated using this do-file (requires **US_allvars.dta**).

Teachers index.do : This do-file compared predicted wages for elementary, and middle school, teachers, across cities (see Section 4; requires **US_allvars.dta**).

2. Occupational Centrality Index Data

Occupational Centrality Indices.xls: Contains the centrality index values for 475 occupations as calculated and used in the paper (for use or reference or for by other researchers).

Additional information on Occupation Categories

Census Occupation Categories: <http://www.census.gov/hhes/www/ioindex/occ2000t.pdf>

Standard Occupation Categories (SOC) http://www.bls.gov/oes/2001/oes_stru.htm#00-0000

Census Occupation Codes and SOC in the PWPUMA Sample:

The Census 2000 codes are matched to equivalent 2000 SOC codes and use the following notational rules for SOC codes. In general, when a Census code matches a 2000 SOC detailed occupation, the detailed SOC code is cited, e.g., 11-1011, 11-9199, and 13-1022. However, if a Census code combines all of several 2000 SOC detailed occupations within a broad occupation or minor group, the SOC broad occupation or minor group code is cited, e.g., 11-2020 and 25-1000. Furthermore, if the census code aggregates two or more SOC categories in a way that does not have a single SOC equivalent code, the SOC code as shown on the EEO File contains alpha characters, e.g., 13-11XX and 47-50YY. More info: http://www.eeoc.gov/stats/census/occupational_data.html