Replication instructions: Non-Pecuniary Benefits: Evidence from the Location of Private Company Sales

Mark Jansen* Adam Winegar†

February 12, 2024

This document details the replication procedure for *Non-Pecuniary Benefits: Evidence from the Location of Private Company Sales*, published in the Review of Corporate Finance Studies. The code used to create the graphs and tables in the paper is written in Stata 18. Conducting the entire replication procedure should take no more than 20 minutes on a standard machine (using actual real data).

Statement about Rights

The authors of the manuscript have legitimate access to and permission to use the data used in this manuscript. Some data cannot be made publicly available.

Data sources

We use several sources of data to conduct our empirical analysis, which are summarized in Table 1. We do not provide the data for replication, but note when the data is publicly available.

*University of Utah, Eccles School of Business. E-mail: mark.jansen@eccles.utah.edu
†BI Norwegian Business School. E-mail: adam.w.winegar@bi.no

File	Source	Notes	
BLS_Super_input_concentration.dta	BEA IO Tables	Public	Provided
BLS_Super_output_concentration	BEA IO Tables	Public	Provided
BP_CBSA.dta	Hand Collected data	Public	Provided
CBSA_Geographic_Data	Multiple Sources (see Internet Appendix A for Description)	Public	Provided
distances.dta	John Gardner Website	Public	Provided
IPUMS_migration_1990_cbsa.dta	IPUMS Data Library	Public	Provided
IPUMS_migration_2000_cbsa.dta	IPUMS Data Library	Public	Provided
New_QOL_CBSA_1980_2000.dta	Calculations from Albouy (2016)	Public	Provided
QOL_Predictors_BP.dta	Multiple Sources (see Internet Appendix A for Description)	Public	Provided
Regression_Analysis_pseudo.dta	Pratts Stats	Proprietary	Pseudo
scaled_transactions_cbsa.dta	Pratts Stats (Aggregated)	Proprietary	Provided
seegert_collapsed.dta	Calculations from Nathan Seegert	Public	Provided

Table 1: **Data sets used in the empirical analysis**

1. `BLS_Super_input_concentration.dta` is compiled from the Bureau of Economic Analysis (BEA) Input-Output Matrix Tables. Values are mapped to Bureau of Labor Statistics Super Sectors and then calculated for the input from each super sector for the given super sector. These are then used as weights for the concentration of input industries. The underlying files and computations are available upon request.

2. `BLS_Super_output_concentration.dta` is compiled from the Bureau of Economic Analysis (BEA) Input-Output Matrix Tables. Values are mapped to Bureau of Labor Statistics Super Sectors and then calculated for the ouput from each super sector for the given super sector. These are then used as weights for the concentration of output industries. The underlying files and computations are available upon request.

3. `BP_CBSA.dta` indicator for whether the given CBSA is a Best Place. See Internet Appendix Table B.1. for a list of Best Places. See Internet Appendix A for a description of the variable.

4. `CBSA_Geographic_Data.dta` aggregated geographic variables to the CBSA. From various sources including the U.S. Census, BEA, Bureau of Labor Statistics (BLS), United States Department of Education National Center for Education Statistics Common Core of Data, the Tax Foundation, and the Federal Housing Finance Agency (FHFA). Data was aggregated to CBSA using the National Bureau of Economic Research (NBER) county-CBSA crosswalk. The underlying files and computations are available upon request.

5. `IPUMS_migration_1990_cbsa.dta` is from IPUMS https://usa.ipums.org/usa/ using the cleaning file provided by IPUMS. Additional aggregation to CBSA uses the National Bureau of Economic Research (NBER) county-CBSA crosswalk. Filters follow Albouy (2016). The underlying files and computations are available upon request.

6. `IPUMS_migration_2000_cbsa.dta` is from IPUMS https://usa.ipums.org/usa/ using the cleaning file provided by IPUMS. Additional aggregation to CBSA uses the National Bureau of Economic Research (NBER) county-CBSA crosswalk. Filters follow Albouy (2016). The underlying files and computations are available upon request.

7. `New_QOL_CBSA_1980_2000.dta` based on data from IPUMS https://usa.ipums.org/usa/ using the cleaning file provided by IPUMS. Additional calculations follow Albouy (2016), described in Internet Appendix A. The underlying files and computations are available upon request.

8. `QOL_Predictors_BP.dta` aggregated geographic variables to the CBSA. From various sources including the U.S. Census, BEA, Bureau of Labor Statistics (BLS), U.S. Department of Agriculture, United States Department of Education National Center for Education Statistics Common Core of Data, National Oceanic and Atmospheric Administration (NOAA), Environmental Protection Agency (EPA), National Archive of Criminal Justice Data (NACJD), and the from Florida Department of Law Enforcement. Data was aggregated to CBSA using the National Bureau of Economic Research (NBER) county-CBSA crosswalk. The underlying files and computations are available upon request.

9. `Regression_Analysis_pseudo.dta` is transaction data from Pratt Stats merged with the `CBSA_Geographic_Data.dta` based on transaction CBSA and year. A sample of the transactions are provided. The data related to transactions have been randomized due to its proprietary nature.

10. `scaled_transactions_cbsa.dta` is transaction counts from Pratt Stats and headquarter counts from Reference USA. This data has been aggregated to cbsa and year to preserve its proprietary nature.

11. `seegert_collapsed.dta` calculations from Nathan Seegert.

I Computational requirements

The following libraries are requires to run the STATA code: `estout`, `egenmore`, `rangestat`, `mmerge` and `reghdfe`. The libraries are installed when running `RunAll.do`.

II Replication procedure

The steps below outline the procedure which needs to be followed to replicate the graphs and tables in the paper:

1. Prepare the raw data files as explained in the section "Data sources" under "`Data`" and set in "`RunAll.do`" Create a folder to store the results, "`Tables`" and set in "`RunAll.do`". Create a folder for the mapping files under "`Data/Mapping.do`".

2. Generate main regression tables for transactions data using `Main_Regressions.do`. This will generate Tables 2-6, Internet Appendix Tables D.1-D.7, Figure 1, and Internet Appendix Figures A.1, D.1.

3. Generate the Best Places Decomposition Tables using `BP_Decomposition.do`. This will generate Table 7, and Internet Appendix Table D.8 and D.9.

4. Generate the Entrepreneur Migration and Rate Tables using `IPUM_Migration.do`. This will generate Table 8 and 9 and Internet Appendix Table D.10.

References

Albouy, D., 2016. What are cities worth? Land rents, local productivity, and the total value of amenities. Review of Economics and Statistics 98, 477–487.