

Readme for

“Discrimination in hiring: Evidence from Retail”

Alan Benson
Univ. of Minnesota

Simon Board
UCLA

Mortiz Meyer-ter-vehn
UCLA

1 Data description

This section provides a data description, data and code availability statement, and computational requirements for the three sources of data: (1) the proprietary company data, (2) American Community Survey data, and (3) simulated theory data.

1.1 Proprietary company data

Data from the company are covered by a Data Use Agreement between the company and the University of Minnesota. Dissemination of the raw data are prohibited due sensitivity, the IRB, and proprietary clauses in the data use agreement. Affiliates of the University of Minnesota may access the raw data subject to the terms of the data use agreement and IRB by contacting bensona@umn.edu.

Company data covered by the agreement include sales per hour, the sales per hour targets, race of worker, race of manager, department, location, date of hire, and internal employment history.

The full raw data are 21 GB in CSV and Excel format and include a panel of over 20 million observations at the person-month level. The *de minimus* Stata DTA with only the observations and variables required for the study is 178 MB. Analyses were performed on Stata 16.1 SE.

1.2 ACS data

The ACS data come from usa.ipums.org ([Ruggles et al., 2020](#)) and are available on this replication package. The ACS data provided in this replication package may be reproduce using the accompanying extract, `ipumsdo.do`.

ACS data are used to calculate metropolitan and state racial composition. The raw data are approximately 3 GB. It was run and merged from Stata 16.1 SE.

1.3 Simulation

The Matlab program used to create the simulation data are included in this package. There are no special computing requirements required to run it.

2 Statement about rights

I certify that the authors of the manuscript have legitimate access to and permission to use the data used in this manuscript. I certify the authors of the manuscript have documented permission to redistribute/publish the data contained within this replication package.

3 Instructions for data preparation and analysis

- step1-create.do is a Stata do-file that creates the files used in the analysis
- step2-analysis.do produces the tables in the paper and robustness tables
- Fig1_RES.m reproduces Figure 1, showing the theory simulation The
- cbsadata subfolder has the cbsa race controls from the ACS

References

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2020. "IPUMS USA: version 10.0 [dataset]." *Minneapolis, Mn: Ipums*, 10: D010.