

# Replication package for “Business disruptions from social distancing”

Please cite as

Koren, Miklós and Rita Pető. 2020. “Replication package for «Business disruptions from social distancing»” (Version 1.3) [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.4016325>

## License and copyright

All text and graphs (\*.md, \*.txt, \*.tex, \*.eps) are [CC-BY-4.0](#). All code (\*.do, Makefile, \*.py) are subject to the [3-clause BSD license](#). All derived data (data/derived/\*) are subject to [Open Database License](#). Please respect the copyright and license terms of original data vendors (data/raw/\*).

## Data Availability Statements

The mobility data used in this paper (SafeGraph 2020) is proprietary, but may be obtained free of charge for COVID-19-related research from the COVID-19 Consortium. The authors are not affiliated with this consortium. Researchers interested in access to the data can apply at <https://www.safegraph.com/covid-19-data-consortium> (data manager: Ross Epstein, [ross@safegraph.com](mailto:ross@safegraph.com)). After signing a Data Agreement, access is granted within a few days. The Consortium does not require coauthorship and does not review or approve research results before publication. Datafiles used: /monthly-patterns/patterns\_backfill/2020/05/07/12/2020/02/patterns-part[1-4].csv.gz (Monthly Places Patterns for February 2020, released May 7, 2020), /monthly-patterns/patterns/2020/06/05/06/patterns-part[1-4].csv.gz (Monthly Places Patterns for February 2020, released June 5, 2020) and /core/2020/06/Core-USA-June2020-Release-CORE\_POI-2020\_05-2020-06-06.zip (Core Places for June 2020, released June 6, 2020). The COVID-19 Consortium will keep these datafiles accessible for researchers. The authors will assist with any reasonable replication attempts for two years following publication.

All other data used in the analysis, including raw data, are available for reuse with permissive licenses. Raw data are saved in the folder data/raw/. The Makefile in each folder shows the URLs used to download the data.

## SafeGraph

### Citation

SafeGraph. “Patterns [dataset]”; 2020. Downloaded 2020-06-20.

### License

Proprietary, see <https://shop.safegraph.com/> or <https://www.safegraph.com/covid-19-data-consortium> (data manager: Ross Epstein, [ross@safegraph.com](mailto:ross@safegraph.com))

## O\*NET

### Citation

U.S. Department of Labor/Employment and Training Administration, 2020. “O\*NET Online.” Downloaded 2020-03-12.

### License

CC-BY-4.0 <https://www.onetonline.org/help/license>

## **Current Employment Statistics**

### **Citation**

U.S. Bureau of Labor Statistics. 2020. "Current Employment Statistics."  
<https://www.bls.gov/ces/> Downloaded 2020-03-15.

### **License**

Public domain: <https://www.bls.gov/bls/linksite.htm>

## **National Employment Matrix**

### **Citation**

U.S. Bureau of Labor Statistics. 2018. "National Employment Matrix."  
<https://www.bls.gov/emp/data/occupational-data.htm> Downloaded 2020-03-15.

### **License**

Public domain: <https://www.bls.gov/bls/linksite.htm>

## **Crosswalk**

### **Citation**

U.S. Bureau of Labor Statistics. 2019. "O\* NET-SOC to Occupational Outlook Handbook Crosswalk." <https://www.bls.gov/emp/classifications-crosswalks/nem-onet-to-soc-crosswalk.xlsx> Downloaded 2020-03-15.

### **License**

Public domain: <https://www.bls.gov/bls/linksite.htm>

## **American Time Use Survey**

### **Citation**

U.S. Bureau of Labor Statistics. 2018. "American Time Use Survey."  
<https://www.bls.gov/tus/>.

We are using the following files:

- Respondent File
- Replicate Weights
- Leave Module 2017-18

### **License**

Data is in [public domain](#).

## **County Business Patterns**

### **Citation**

U.S. Bureau of the Census. 2017. “County Business Patterns.” Available at <https://www.census.gov/programs-surveys/cbp.html>

## License

<https://www.census.gov/data/developers/about/terms-of-service.html>

# Dataset list

## Raw data

Data file	Source	Notes	Provided
data/raw/bls/industry-employment/ces.txt	BLS Current Employment Statistics	Public domain	Yes
data/raw/bls/atus/*.dat	BLS Time Use Survey	Public domain	Yes
data/raw/bls/employment-matrix/matrix.xlsx	BLS National Employment Matrix	Public domain	Yes
data/raw/bls/crosswalk/*.xlsx	O*NET-SOC to Occupational Outlook Handbook Crosswalk	Public domain	Yes
data/raw/onet/*.csv	O*NET Online	Creative Commons 4.0	Yes
data/raw/census/cbp/*.txt	County Business Patterns	Public domain	Yes
not-included/safegraph/02/*.csv	SafeGraph	Available with Data Agreement with SafeGraph	No
not-included/safegraph/05/*.csv	SafeGraph	Available with Data Agreement with SafeGraph	No

## Clean data

Data file	Source	Notes	Provided
data/clean/industry-employment/industry-employment.dta	BLS Current Employment Statistics	Public domain	Yes
data/clean/time-use/atus.dta	BLS Time Use Survey	Public domain	Yes
data/clean/employment-matrix/matrix.dta	BLS National Employment Matrix	Public domain	Yes
data/clean/onet/risks.csv	O*NET Online	Creative Commons 4.0	Yes
data/clean/cbp/zip_code_business_patterns.dta	County Business Patterns	Public domain	Yes

## Derived data

Data file	Source	Notes	Provided
-----------	--------	-------	----------

Data file	Source	Notes	Provided
data/derived/occupation/*	Various sources	Public domain	Yes
data/derived/time-use/atus_working_at_home_occupationlevel.dta	BLS Time Use Survey	Public domain	Yes
data/derived/crosswalk/*	Various sources	Public domain	Yes
not-included/safegraph/naics-zip-*.csv	SafeGraph	Available with Data Agreement with SafeGraph	Yes, with permission of SafeGraph
data/derived/visit/visit-change.dta	SafeGraph	Aggregated to 3-digit NAICS industries	Yes, with permission of SafeGraph

## Computational requirements

### Software Requirements

- Stata (code was last run with version 16.1)
  - `estout` (from <http://www.stata-journal.com/software/sj14-2/>)
  - running `make install` from the root of the folder will install `estout` locally, and should be run once.

Portions of the code use bash scripting (`make`, `wget`, `head`, `tail`), which may require Linux or Mac OS X.

The entry point for analysis is `analysis/Makefile`, which can be run by [GNU Make](#) on any Unix-like system by

```
cd analysis
make
```

The dependence of outputs on code and input data is captured in the respective Makefiles.

We have used Mac OS X, but all the code should run on Linux and Windows platforms, too.

### Hardware

The analysis takes a few minutes on a standard laptop.

### Description of programs

1. Raw data are in `data/raw/<vendor>/<dataset>`. This data is saved as it has been received from the data publisher, downloaded by the respective Makefiles. Each folder has a `README.md` with data citation and license terms.
2. Clean data are in `data/clean/<dataset>`. Each folder has a `Makefile` that specifies the steps of data cleaning.
3. Analysis data are in `data/derived/<topic>`. Each folder has a `Makefile` that specifies the steps of sample specification, variable selection and modification.
4. Analysis codes are in `analysis/`. A `Makefile` specifies how the `.do` files have to be run to reproduce the analysis.

## Instructions

After uncompressing the replication package, change the working directory to `analysis/` and run the Makefile there.

```
mkdir koren-peto
```

```
cd koren-peto
unzip ../koren-peto.zip
cd analysis/
make
```

The Makefile assumes that Stata is accessible from your shell via the command `stata`. If you use any other command to call Stata, introduce an alias.

## List of tables and programs

Exhibit	File	Code
headcount number in abstract	analysis/headcount.log line 66	analysis/headcount.do
Table 1	NA	NA
Table 2	exhibit/table2.csv	analysis/toplist.do
Table 3	exhibit/table3.tex	analysis/employment-growth.do
Table 4	exhibit/table4.csv	analysis/counterfactual.do
Figure 1	exhibit/fig1.eps	NA
Figure 2	exhibit/fig2.eps	analysis/scatter_customers_teamwork_byoccupation.do
Figure 3	exhibit/fig3.eps	analysis/scatterplot_validation_customer_atus.do
Figure 4	exhibit/fig4.eps	analysis/scatterplot_validation_teamwork_atus.do
Supplementary Information 1	exhibit/SI1.csv	analysis/create_SI1.do

## References

1. Koren, Miklós and Rita Pető. Business disruptions from social distancing; PLoS ONE 2020. <https://doi.org/10.1371/journal.pone.0239113>
2. National Center for O\*NET Development. O\*NET OnLine [dataset]; 2020. <https://www.onetonline.org/>
3. U S Bureau of the Census. County Business Patterns [dataset]; 2017. <https://www.census.gov/data/datasets/2017/econ/cbp/2017-cbp.html>
4. SafeGraph. Patterns [dataset]; 2020. <https://www.safegraph.com/covid-19-data-consortium>
5. U S Bureau of Labor Statistics. American Time Use Survey [dataset]; 2018. <https://www.bls.gov/tus/#data>
6. U S Bureau of Labor Statistics. Current Employment Statistics [dataset]; 2020. <https://www.bls.gov/ces/>
7. U S Bureau of Labor Statistics. National Employment Matrix [dataset]; 2020. <https://www.bls.gov/emp/tables/industry-occupation-matrix-industry.htm>