Pandemics Through the Lens of Occupations

by Anand Chopra, Michael B. Devereux and Amartya Lahiri

This file provides the instructions to run the programs in the correct order to replicate the Figures in the paper.

1. Run “main.m” in MATLAB to generate Figures 4, 5, 6, 7, 13, 14, 15, 17 and 18 (Panel b).
2. Run “Occupation Specific Risk\main\_occ\_risk.m” in MATLAB to generate Figure 10.
3. Run “Heathcare Congestion\main\_congestion.m” in MATLAB to generate Figure 16.
4. Run “Figure\_1\_2\_3\_9.do” in STATA to generate Figure 1,2,3 and 9.
5. Run “Figure\_8.do” in STATA to generate Figure 8.
6. Run “Figure\_11.do” in STATA to generate Figure 11.
7. Run “Figure\_12.do” in STATA to generate Figure 12.

**Note 1**: Please adjust directories in each of the above MATLAB/STATA programs for the code to run correctly.

**Note 2:** The provided code will replicate all figures apart from Figure 18 Panel (a) which was taken from the paper by Chetty et al. (2020).

**Note 3:** Ado files needed for the STATA code are provided in the “Ado” folder.

**Data Availability Statement**

**Statement about Rights**

We certify that the authors of the manuscript have legitimate access to and permission to use the data used in this manuscript.

**Details on each Data Source**

* Annual data on median wages, salaried and commissions for one-, two- and three-digit NOC occupations is available from the 2016 Census Catalogue no. 98-400-X2016304 (Statistics Canada 2016). We filtered the two-digit NOC occupation data and divided median annual wages by 52 to get a weekly measure of wages. Click [here](https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/dt-td/Rp-eng.cfm?TABID=2&Lang=E&APATH=3&DETAIL=0&DIM=0&FL=A&FREE=0&GC=0&GID=1325190&GK=0&GRP=1&PID=110698&PRID=10&PTYPE=109445&S=0&SHOWALL=0&SUB=0&Temporal=2017&THEME=124&VID=0&VNAMEE=&VNAMEF=&D1=0&D2=0&D3=0&D4=0&D5=0&D6=0) to access the data and select “Total-Work activity during the reference year”. The data was constructed by Statistics Canada by taking a 25% sample of the Census and people 15 years and over working were considered.
* Data on number of workers, risk index and work from home shares for each two-digit occupation were provided by Baylis et al. (2020a,b).
* Employment data for February, March, April and June 2020 are available from the Table 14-10-0296-01 provided by Statistics Canada (2021). Click [here](https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1410029601&pickMembers%5B0%5D=1.11&pickMembers%5B1%5D=2.2&pickMembers%5B2%5D=4.1&cubeTimeFrame.startMonth=02&cubeTimeFrame.startYear=2020&cubeTimeFrame.endMonth=06&cubeTimeFrame.endYear=2020&referencePeriods=20200201%2C20200601) to access the data and select “British Columbia” for Geography and “February 2020” to “June 2020” as the relevant reference periods.
* Death data for Canada and British Columbia are available from the Canadian Covid-19

epidemiological data which is available at <https://health-infobase.canada.ca/covid-19/epidemiological-summary-covid-19-cases.html>.

* Death data for United States are available from The New York Times (2021). The url is <https://github.com/nytimes/covid-19-data>.
* Data on sectoral elasticities is 1 and weights are 0.1 over 10 sectors in “gdp\_sectors\_parameters.csv”. Data on elasticity of an occupation across sectors is 1 and weight is 1 in “sectors\_occ\_elasticities.csv” and “sectors\_occ\_elasticities.csv” respectively. This ensures that output is a linear sum of wages across all occupations and sectors.

**Dataset List**

|  |  |  |  |
| --- | --- | --- | --- |
| Data file | Source | Notes | Provided |
| Raw data/2-digit occupation aggregation.xlsx | Data on number of workers, risk index and work from home shares for 3-digit industry x 4-digit occupation were provided by Baylis et al. (2020b). | We aggregate them to two-digit occupation in the same Excel file, sheet “occ\_2\_aggregation”. |  |
| Raw data/ Occupation data on wages.xlsx | Data on median wages from Statistics Canada (2016), Census Catalogue no. 98-400-X2016304. | Combined data on no. of workers, risk index and share working from home from “2-digit occupation aggregation.xlsx”. | Yes |
| Raw data/ Occupation Employment.xlsx | Statistics Canada (2021), Table 14-10-0296-01 | Combined data on proportion working from home from “Occupation data on wages.xlsx” | Yes |
| Raw data/ Deaths data.xlsx | Data from US are from The New York Times (2021) and the Canadian Covid-19  epidemiological data. | Canadian Covid-19  epidemiological data is available at <https://health-infobase.canada.ca/covid-19/epidemiological-summary-covid-19-cases.html> | Yes |
| gdp\_sectors\_parameters.csv | Constructed by authors |  | Yes |
| sectors\_occ\_elasticities.csv | Constructed by authors |  | Yes |
| sectors\_occ\_elasticities.csv | Constructed by authors |  | Yes |

**Computational Requirements**

**Software Requirements**

* STATA (code was last run on version 14)
  + Labutil (Ado file provided, code by Cox (2000))
  + Splitvallabels (Ado file provided, code by Winter and Jann (2004))
* MATLAB (code was last run with MATLAB Release 2020a, previously also ran on MATLAB Release 2018b)

**Memory and Runtime Requirements**

The code was last run on a 4-core Intel-based laptop with i7 processor, 16 GB RAM and Windows 10.

The entire code took less than 10 minutes to execute.

**Description of programs/code**

Below, we provide a brief description of all the files in the replication package.

|  |  |  |
| --- | --- | --- |
| File Name | Sub-Directory | Type |
| **Data files** |  |  |
| gdp\_sectors\_parameters.csv | Also present in “Occupation Specific Risk” and “Healthcare Congestion” | CSV file |
| sectors\_occ\_elasticities.csv | Also present in “Occupation Specific Risk” and “Healthcare Congestion” | CSV file |
| sectors\_occ\_weights.csv | Also present in “Occupation Specific Risk” and “Healthcare Congestion” | CSV file |
| wage\_schedule.xlsx | Also present in “Occupation Specific Risk” and “Healthcare Congestion” | Excel worksheet |
| Deaths data.xlsx | Raw data | Excel worksheet |
| Occupation data on wages.xlsx | Raw data | Excel worksheet |
| Occupation Employment.xlsx | Raw data | Excel worksheet |
|  |  |  |
| **STATA programs** |  |  |
| Figure\_1\_2\_3\_9.do |  | STATA Do-file |
| Figure\_8.do |  | STATA Do-file |
| Figure\_11.do |  | STATA Do-file |
| Figure\_12.do |  | STATA Do-file |
| Labutil.ado | Ado | STATA Ado-file |
| Splivallabels.ado | Ado | STATA Ado-file |
|  |  |  |
| **MATLAB programs (by order of appearance in the code)** |  |  |
| main.m |  | MATLAB m-file |
| steady\_state\_cutoff.m |  | MATLAB m-file |
| go\_calibrate\_pis.m |  | MATLAB m-file |
| calibration\_pi.m |  | MATLAB m-file |
| policy\_susceptible.m |  | MATLAB m-file |
| sir\_dynamics.m |  | MATLAB m-file |
| policy\_susceptible\_linear.m |  | MATLAB m-file |
| sir\_dynamics\_linear.m |  | MATLAB m-file |
| policy\_susceptible\_rebate.m |  | MATLAB m-file |
| sir\_dynamics\_rebate.m |  | MATLAB m-file |
| policy\_susceptible\_vaccine.m |  | MATLAB m-file |
| sir\_dynamics\_vaccine.m |  | MATLAB m-file |
| main\_occ\_risk.m | Occupation Specific Risk | MATLAB m-file |
| steady\_state\_cutoff.m | Occupation Specific Risk | MATLAB m-file |
| go\_calibrate\_pis.m | Occupation Specific Risk | MATLAB m-file |
| calibration\_pi.m | Occupation Specific Risk | MATLAB m-file |
| policy\_susceptible.m | Occupation Specific Risk | MATLAB m-file |
| sir\_dynamics.m | Occupation Specific Risk | MATLAB m-file |
| main\_congestion.m | Healthcare Congestion | MATLAB m-file |
| steady\_state\_cutoff.m | Healthcare Congestion | MATLAB m-file |
| go\_calibrate\_pis.m | Healthcare Congestion | MATLAB m-file |
| calibration\_pi.m | Healthcare Congestion | MATLAB m-file |
| policy\_susceptible.m | Healthcare Congestion | MATLAB m-file |
| sir\_dynamics.m | Healthcare Congestion | MATLAB m-file |

**References**

Baylis, P., P.-L. Beauregard, M. Connolly, N. Fortin, D. A. Green, P. Gutierrez Cubillos, S. Gyetvay, C. Haeck, T. Laura Molnar, G. Simard-Duplain, H. E. Siu, M. TeNyenhuis, and C. Warman (2020a) “The distribution of Covid-19 related risks,” NBER Working Paper No. 27881

Baylis, P., P.-L. Beauregard, M. Connolly, N. Fortin, D. A. Green, P. Gutierrez Cubillos, S. Gyetvay, C. Haeck, T. Laura Molnar, G. Simard-Duplain, H. E. Siu, M. TeNyenhuis, and C. Warman (2020b) “Replication files for: The distribution of Covid-19 related risks, [https://github.com/pbaylis/vse-risk-tool](https://www.mail.ubc.ca/owa/redir.aspx?REF=YxoLVktsKOZ_5HI7FOwavsol8xdKRzSJ_UWtoJ-c72sdTNsKKCrZCAFodHRwczovL2dpdGh1Yi5jb20vcGJheWxpcy92c2Utcmlzay10b29s)”

Chetty, R., J. N. Friedman, N. Hendren, M. Stepner, and The Opportunity Insights Team (2020) “How did Covid-19 and stabilization policies affect spending and employment? A new real-time economic tracker based on private sector data,” NBER Working Paper No. 27431

Cox, N.J. (2000) “Labutil: Stata modules for managing value and variable labels”, Statistical Software Components S402501, Boston College Department of Economics

Statistics Canada (2016) “2016 Census of Population, Statistics Canada Catalogue no. 98-400-X2016304”

Statistics Canada (2021) “Table 14-10-0296-01: Labour force characteristics by occupation, monthly, unadjusted for seasonality (x 1,000)”

The New York Times (2021) “Coronavirus (Covid-19) Data in the United States”, Retrieved March 1 2021, from <https://github.com/nytimes/covid-19-data>

Winter, N. and B. Jann (2004) “Splitvallabels: Stata module to split up value labels for multi-line graph labelling”, Statistical Software Components S436401, Boston College Department of Economics