README and Replication File Guide for

"Sectoral Digital Intensity and GDP Growth After a Large Employment Shock: A Simple Extrapolation Exercise"

June 18, 2021

Overview

The codes in this replication package construct the analysis file from the following data sources: StatsCanada; IPUMS Canada; Gallipoli and Makridis, 2018; Vancouver School of Economics, 2020; Dingel and Neiman, 2020. The do file *covid_GDPanalysis* generates the most of the final data and produce the most results in the paper. The others do files in the Codes folder are essentially for industry index computing.

Data Availability and Provenance Statements

Statement about Rights

☑ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

Summary of Availability

☑ All data **are** publicly available.

☐ Some data **cannot be made** publicly available.

☐ **No data can be made** publicly available.

Data Availability Statement

- Employment data: Statistics Canada, `https://doi.org/10.25318/1410035501-eng`
- GDP data: Statistics Canada, `https://doi.org/10.25318/3610043401-eng` and `https://doi.org/10.25318/3610040201-eng`
- IPUMS Canada 2011 sample data: `https://doi.org/10.18128/D020.V7.2`
- O*NET data
  - Gallipoli, G. and Makridis, C. (2018). Structural Transformation and the Rise of Information Technology. Journal of Monetary Economics.
  - Dingel, J.I. and Neiman, B. (2020). How many jobs can be done at home. Journal of Public Economics.
- Occupational Characteristic data: Baylis, P., Beauregard, P.-L., Connolly, M., Fortin, N., Green, D., Gutierrez-Cubillos, P., Haeck, C., Gyetvay, S., Molnar, T., Simard-Duplain, G., Siu, H., Tenyenhuis, M., and Warman, C. (2020). The Distribution of COVID-19 Related Risks. NBER working paper.

Accessing the Data

- Raw Data
  - Employment Data: *RawData/StatsCan/emp_Data*, the latest employment data available upon submission is in March 2021.
  - GDP Data: *RawData/StatsCan/GDP_Data*, the latest GDP data available upon submission is in February 2021 (i.e. the results in our paper are under the ${version} 2102)
  - IPUMS Data: *RawData/IPUMS/canada_ipums2011.dta*
  - O*NET Data: *RawData/ONET*, includes O*NET data from Dingel and Neimann (2020) and Gallipoli and Makridis (2018) for 3 digits and 5 digits soc codes [1].
  - Occupational Characteristic Data: *RawData/VSE*, the database inside folder is provided by Vancouver School of Economics (2020)
- Preliminary Data
  - *Generated_data/allyears_final_update_data${version}.dta*: this dataset includes GDP, employment and industry indices related, at industry level
  - *Generated_data/gdp predictions_new_agri.xlsx*: this dataset includes the predicted growth rate of GDP and levels of GDP by using projected employment, at resilience level
  - *Generated_data/gdp predictions_new_realized_agri.xlsx*: this dataset includes the predicted growth rate of GDP and levels of GDP by using realized employment, at resilience level
  - *Generated_data/industry_index*: IT Intensity, WFH Index and Resilience Data based on IT intensity index computed from Giovanni and Makridis (2018), Dingel and Neimam (2020), and Vancouver School of Economics (2020)
- Final Data
  - *Generated_data/update_new_merged_data_model_noagri${version}.dta*: this dataset includes the growth rate of real GDP, levels of real GDP, and merges with our predictions (Baseline, Benign Outlook) using projected employment, at resilience level
  - *Generated_data/update_new_realized_merged_data_model_noagri${version}.dta*: this dataset includes the growth rate of real GDP, levels of GDP, and merges with our predictions using realized employment, at resilience level
  - *Generated_data/gdp_predictions_ind_prov*: datasets in this folder includes the predicted growth rate of GDP by industry and by province respectively

Replication File Details

All related ado files are under *data/Ados*.

- *Codes/crosswallk_ind_occ.do*
  - This code constructs crosswalk from IPUMS Canada (2011) occupations to 3-digit soc 2010 codes and labels the industry code provided from IPUMS. This codes needs to be run in *ipums_occu_it_dingel_wfh.do*.
- *Codes/ipums_occu_it_dingel_wfh.do*
  - This codes cleans IPUMS data and then generates the IT intensity index and Dingel WFH index at industry level and province level, also computes the WFH index based on VSE database at industry and province level.

[1] For the computation, refer to Gallipoli, G. and Makridis, C. (2018). Structural Transformation and the Rise of Information Technology. Journal of Monetary Economics

    - Input data (all located in RawData):
        - "${data}/canada_ipums2011.dta"
        - "${data}/onet_avg_2002_2016_3dig.dta"
        - "${data}/DN_index2.dta"
        - "${data}/gdp_indus_prov_canada_growthrates_2019.dta"
        - "${data}/employ_industry.dta"
        - "${wfh}/`province'/work_from_home_score.xlsx"
    - Output data (all located in Generated_data/industry_index):
        - "${data}/it_intensity_industry.dta"
        - "${data}/dingel_index_industry.dta"
        - "${data}/it_index_province.dta"
        - "${data}/dingel_index_province.dta"
        - "${wfh}/vse_index_province.dta"
        - "${wfh}/vse_index_industry.dta"

- *Generated_data/industry_index/Indices by province.xlsx*
    - Use this Excel file to replicate industry indices by province figures: Figure 1 (main paper)
        - Input data source:
            - "${data}/it_intensity_industry.dta"
            - "${data}/dingel_index_industry.dta"
            - "${wfh}/vse_index_industry.dta"
        - Output figures: indices_province.pdf

- *Generated_data/industry_index/Indices by industry.xlsx*
    - Use this Excel file to replicate industry indices by industry figures: Figure 2 (main paper)
        - Input data source:
            - "${data}/it_intensity_province.dta"
            - "${data}/dingel_index_province.dta"
            - "${wfh}/vse_index_province.dta"
        - Output figures: indices_industry.pdf

All the following subprograms to generate our results are contained in the code *Codes/covid_GDPanalysis.do*.

- **dataprep**
    - This subprogram imports the raw GDP and employment data and merges with the industry indices.
        - Input data (all located in RawData):
            - "${empData}/emp_databaseLoadingData_`latestEmpData'.csv"
            - "${gdpData}/ `latestGDPdata'.csv"
            - "${index}/matchind_it_ci_wfh_merged.dta"
        - Output data: "${revision}/allyears_final_update_data${version}.dta"

- **emp_gdp_stats**
    - This subprogram computes the growth rate of employment, growth rate of GDP and level of GDP at resilience level.
        - Input data:
            - "${revision}/allyears_final_update_data${version}.dta"
            - "${revision}/gdp predictions_new_agri.xlsx"

        - "${revision}/gdp predictions_new_realized_agri.xlsx"
    - Output data: "${revision}/update_new `realized'_merged_data_model_noagri${version}.dta"

- **gdp_plot**
  - Run this subprogram to replicate GDP level/growth rates at resilience level figures: Figure 4, Figure 5, Figure 6 (main paper); Figure A1 (Appendix A.2)
    - Input data: "${revision}/update_new `realized'_merged_data_model_noagri${version}.dta"
    - Output figures:
      - "${GDP_plot${version}}/quarter_noagri.png"
      - "${GDP_plot${version}}/gdp_levels_`name'_noagri.png" (`name' for baseline or Benign Outlook recovery)
      - "${GDP_plot${version}}/gdp_levels_`name'_real_emp_noagri.png"

- **industry_plot**
  - Run this subprogram to replicate GDP growth rates at industry level figures/tables: Figure 7, Figure 9 (main paper); Table A3, Table A5 (Appendix A.2)
    - Input data:
      - "${revision}/allyears_final_update_data${version}.dta"
      - "${revision}/gdp_prediction_new.xlsx"
    - Output figures/tables:
      - "${GDP_plot${version}}/industry_no_agri.png"
      - "${GDP_table}/gdp_industry_table${version}.tex"
      - "${GDP_plot${version}}/dingel_industry_no_agri.png"
      - "${GDP_table}/gdp_industry_table_dingel${version}.tex"

- **province_plot**
  - Run this subprogram to replicate GDP growth rates at province level figures/tables: Figure 8 (main paper); Table A4 (Appendix A.2)
    - Input data: "${revision}/ind_prov_estimation_new.xlsx"
    - Output figures/tables:
      - "${GDP_plot${version}}/province.png"
      - "${GDP_table}/gdp_province_table.tex"
      - "${GDP_table}/gdp_province_table_dingel.tex"

- **regression**
  - This subprogram runs the regression of elasticity of GDP growth to Employment growth, generates Table 1 and Table 3 (main paper)
    - Input data:
      - "${revision}/allyears_final_update_data${version}.dta"
      - "${dingel}/matchind_it_ci_dingel.dta"
    - Output tables:
      - "${GDP_table}/new_table1${version}.tex"
      - "${GDP_table}/dingel_table${version}.tex"

- **bubble**
  - Run this subprogram to replicate employment drop level after onset of the pandemic across sectors by industry indices figures: Figure 3 (main paper)
    - Input data:

        - "${empDataProvince}/`province'.csv"
        - "${revision}/emply_indus_latest2.xlsx"
    * Output figures: "${GDP_plot}/bubble_`name'.png" (`name' for IT share, home-shorability and resilience)

- **GDP_level_table**
  - Run this subprogram to replicate GDP levels at resilience level for predicted recoveries and real GDP tables: Table A1 (Appendix A.2)
    * Input data: "${revision}/update_new `realized'_merged_data_model_noagri${version}.dta"
    * Output tables:
        - "${GDP_table}/gdp_level_table_pannelA${version}.tex"
        - "${GDP_table}/gdp_level_table_pannelB${version}.tex"

- **GDP_growth_table**
  - Run this subprogram to replicate GDP growth rates at resilience level for predicted recoveries and real GDP tables: Table A2 (Appendix A.2)
    * Input data: "${revision}/update_new `realized'_merged_data_model_noagri${version}.dta"
    * Output tables:
        - "${GDP_table}/gdp_growth_table_pannelA${version}.tex"
        - "${GDP_table}/gdp_growth_table_pannelB${version}.tex"

Data Citation

- Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 7.2 [dataset]. Minneapolis, MN: IPUMS, 2019. `https://doi.org/10.18128/D020.V7.2`
- Statistics Canada. (2019). Table: 36-10-0434-01 Gross Domestic Product by industry accounts [Data table]. `https://doi.org/10.25318/3610043401-eng`
- Statistics Canada. (2020). Table: 36-10-0402-02 Gross Domestic Product by industry accounts by industry by provinces [Data table]. `https://doi.org/10.25318/3610040201-eng`
- Statistics Canada. (2020). Table: 14-10-0355-02 Employment by occupation, industry or sector [Data table]. `https://doi.org/10.25318/1410035501-eng`
- Vancouver School of Economics, V. (2020). Measures of occupational heterogeneity. In VSE COVID Risk/Reward Assessment Tool.
- Dingel, J.I. and Neiman, B. (2020). How many jobs can be done at home. Journal of Public Economics.
- Gallipoli, G. and Makridis, C. (2018). Structural Transformation and the Rise of Information Technology. Journal of Monetary Economics.