README for Jobless Development

Franziska Ohnsorge, Richard Rogerson, and Zoe Leiyu Xie

DATA AVAILABILITY

Most of the data used for producing results in the paper are publicly available. The public data sources used include ILOSTAT; WDI; UN; Penn World Table; GGDC/UNU-WIDER; Bento and Restuccia (2021); World Bank Enterprise Survey; World Bank Women, Business and the Law; Economic Freedom of the World.

Real GDP data (\IntermData\gdp.dta) obtained from World Bank Prospects Group's real GDP database. The underlying data used to produced this file was provided by DECPG. The raw data is not to be shared, and it was provided for verification purposes to the PRWP verification team.

REQUIREMENTS

Stata code requires estout to run. It is recommended to run the program in local drive to avoid syncing issues with OneDrive or Network drive.

SECTION I: DATASETS

Raw data files in folder "InputData"

- Files "0_ILOSTAT …csv". Source: ILOSTAT. URL: https://rshiny.ilo.org/index.html. Access year: September-October 2023. Specific download instruction: Copy provided URL into preferred browse. Choose "Data Explorer". In browser that opens choose Variable by searching the name in "indicator.label" column of the .csv file (also listed in detail below). Then set filter according to details below. Set time slider to choose all available time period (NOTE: THE AVAILABLE TIME PERIOD DOWNLOADED AFTER SEPTEMBER-OCTOBER 2023 WILL BE DIFFERENT FROM THE DATASET INCLUDED IN THE INPUTDATA FOLDER.). After all filters are set, then choose "Export", "filtered", "both", "csv" to download.

- "0_ILOSTAT Employment Model.csv": indicator.label ("Employment by sex and age – ILO modelled estimates, Nov. 2022 (thousands)"). Select all Regions, all Reference areas, all Sexes, and all Age groups.

- "0_ILOSTAT Employment Survey.csv": indicator.label ("Employment by sex and age (thousands)"). Select all Regions, all Reference areas, all Sexes, and all Age groups.

- "0_ILOSTAT Informal Rate Survey.csv": indicator.label ("Informal employment rate by sex and age (%)"). Select all Regions, all Reference areas, all Sexes, and all Age groups.

- “0_ILOSTAT Informal Sector Rate Survey.csv”: indicator.label (“Informal employment rate by sex and economic activity (%)”). Select all Regions, all Reference area, all Sexes, and the top 16 Economic activities (“Agriculture, Non-Agriculture: Total”, “Agriculture, Non-Agriculture: Agriculture”, “Agriculture, Non-Agriculture: Non-agriculture”, “Broad sector: Total”, “Broad sector: Agriculture”, “Broad sector: Industry”, “Broad sector: Services”, “Broad sector: Not classified”, “Aggregate: Total”, “Aggregate: Agriculture”, “Aggregate: Manufacturing”, “Aggregate: Construction”, “Aggregate: Mining and quarrying, Electricity, gas and water supply”, “Aggregate: Trade, Transportation, Accommodation and Food, and Business and Administrative Services”, “Aggregate: Public Administration, Community, Social and other Services and Activities”, “Aggregate: Not classified”).
- “0_ILOSTAT Labor Force Survey.csv”: indicator.label (“Labour force by sex and age (thousands)”). Select all Regions, all Reference areas, all Sexes, and all Age groups.
- “0_ILOSTAT Non-Formal Rate Survey.csv”: indicator.label (“Share of employment outside the formal sector by sex and age (%)”). Select all Regions, all Reference areas, all Sexes, and all Age groups.
- “0_ILOSTAT Non-Formal Sector Rate Survey.csv”: indicator.label (“Share of employment outside the formal sector by sex and economic activity (%)”). Select all Regions, all Reference area, all Sexes, and the top 16 Economic activities (“Agriculture, Non-Agriculture: Total”, “Agriculture, Non-Agriculture: Agriculture”, “Agriculture, Non-Agriculture: Non-agriculture”, “Broad sector: Total”, “Broad sector: Agriculture”, “Broad sector: Industry”, “Broad sector: Services”, “Broad sector: Not classified”, “Aggregate: Total”, “Aggregate: Agriculture”, “Aggregate: Manufacturing”, “Aggregate: Construction”, “Aggregate: Mining and quarrying, Electricity, gas and water supply”, “Aggregate: Trade, Transportation, Accommodation and Food, and Business and Administrative Services”, “Aggregate: Public Administration, Community, Social and other Services and Activities”, “Aggregate: Not classified”).
- “0_ILOSTAT Population by Age UN.csv”: indicator.label (“Population by sex and age – UN estimates and projections, July 2022 (thousands)”, NOTE: THIS VARIABLE IS NO LONGER AVAILABLE ON ILOSTAT AND HAS BEEN REPLACED BY THE UPDATED VERSION “Population by sex and age – UN estimates and projections, July 2024 (thousands)”). Select all Regions, all Reference areas, all Sexes, and 6 aggregate band Age groups (“Aggregate bands: Total”, “Aggregate bands: <15”, “Aggregate bands: 15-24”, “Aggregate bands: 25-54”, “Aggregate bands: 55-64”, “Aggregate bands: 65+”).
- “0_ILOSTAT Sector Age Employment Survey.csv”: indicator.label (“Employment by sex, age and economic activity (thousands)”). Select all Regions, all Reference areas, all Sexes, 4 aggregate band Age groups (“Aggregate bands: Total”, “Aggregate bands: <15”, “Aggregate bands: 15-24”, “Aggregate bands: 25-54”, “Aggregate bands: 55-64”), and the top 14 Economic activities (“Broad sector: Total”, “Broad sector: Agriculture”, “Broad sector: Non-agriculture”, “Broad sector: Industry”, “Broad

sector: Services", "Broad sector: Not classified", "Aggregate: Total", "Aggregate: Agriculture", "Aggregate: Manufacturing", "Aggregate: Construction", "Aggregate: Mining and quarrying, Electricity, gas and water supply", "Aggregate: Trade, Transportation, Accommodation and Food, and Business and Administrative Services", "Aggregate: Public Administration, Community, Social and other Services and Activities", "Aggregate: Not classified").

- "0_ILOSTAT Sector Educ Employment Survey.csv": indicator.label ("Employment by sex, economic activity and education (thousands)"). Select all Regions, all Reference areas, all Sexes, the 6 broad sector Economic activities ("Broad sector: Total", "Broad sector: Agriculture", "Broad sector: Non-agriculture", "Broad sector: Industry", "Broad sector: Services", "Broad sector: Not classified"), and 5 Education groups ("Aggregate levels: Total", "Aggregate levels: Less than basic", "Aggregate levels: Basic", "Aggregate levels: Intermediate", "Aggregate levels: Advanced").

- "0_ILOSTAT Sector Employment Model.csv": indicator.label ("Employment by sex and economic activity – ILO modelled estimates, Nov. 2022 (thousands)", NOTE: THIS VARIABLE IS NO LONGER AVAILABLE ON ILOSTAT AND HAS BEEN REPLACED BY THE UPDATED VERSION "Employment by sex and economic activity – ILO modelled estimates, Nov. 2024(thousands)"). Select all Regions, all Reference areas, all Sexes, and all Economic activities.

- "0_ILOSTAT Sector Employment Survey.csv": indicator.label ("Employment by sex and economic activity (thousands)"). Select all Regions, all Reference areas, all Sexes, and the top 18 Economic activities ("Agriculture, Non-Agriculture: Total", "Agriculture, Non-Agriculture: Agriculture", "Agriculture, Non-Agriculture: Non-agriculture", "Broad sector: Total", "Broad sector: Agriculture", "Broad sector: Non-agriculture", "Broad sector: Industry", "Broad sector: Services", "Broad sector: Not classified", "Aggregate: Total", "Aggregate: Agriculture", "Aggregate: Manufacturing", "Aggregate: Manufacturing, Electricity, gas and water supply", "Aggregate: Construction", "Aggregate: Mining and quarrying, Electricity, gas and water supply", "Aggregate: Trade, Transportation, Accommodation and Food, and Business and Administrative Services", "Aggregate: Public Administration, Community, Social and other Services and Activities", "Aggregate: Not classified").

- "0_ILOSTAT WAP Survey.csv": indicator.label ("Working-age population by sex and age (thousands)"). Select all Regions, all Reference areas, all Sexes, and 4 aggregate band Age groups ("Aggregate bands: Total", "Aggregate bands: 15-24", "Aggregate bands: 25-54", "Aggregate bands: 55-64")

- Files "1_WDI_…xls". Source: WDI. URL: https://databank.worldbank.org/source/world-development-indicators. Access year: September-November 2023. Specific download instruction: Copy provided URL into preferred browse. Database is pre-selected as "World Development Indicators". Country select all available, Time select all available (up to 2022 at the time of download), Variable name in "Metadata - Indicators" tab of each Excel file.

- "LFS_WDI.xlsx". Source: WDI. URL: https://databank.worldbank.org/source/world-development-indicators. Access year: September 2023. Specific download instruction: Copy provided URL into preferred browse, Database is pre-selected as "World Development Indicators", Country select all available, Time select all available (up to 2022 at the time of download), Variable name in "Series - Metadata" tab of the Excel file.

- "GGDC10Sector_jan15.xlsx". Source: GGDC 10-sector database. URL: https://www.rug.nl/ggdc/structuralchange/previous-sector-database/10-sector-2014. Access year: May 2024.

- "informal-economy-database.xlsx". Source: World Bank Informal Economy Database. URL: https://www.worldbank.org/en/research/brief/informal-economy-database. Access year: September 2023.

- "PennTablepwt1001.xlsx". Source: Penn World Table. URL: https://www.rug.nl/ggdc/productivity/pwt/?lang=en. Access year: November 2023.

- "UNPopulationByAgeCompletedataset.xlsx". Source: UN World Population Prospects 2022. URL: https://population.un.org/wpp/downloads?folder=Archive&group=Standard%20Projections. Access year: March 2024. Specific download instruction: Copy URL into preferred browser. Choose "Standard Projections (…)" as File Type. Choose "Population" as Major Topic/Special Groups. Download file "Population by Select Age Groups - Both Sexes (XLSX)". The downloaded data file is very large. Remove all tabs except "Medium variant" and "NOTES" tabs.

- "UNUWider etd-release2021.xlsx". Source: UNU-WIDER Economic Transformation Database. URL: https://www.wider.unu.edu/database/etd-economic-transformation-database. Access year: May 2024.

Raw data files in "InputData"

- Files "ES-Indicators-Database-Global-Methodology_November_20_2023.dta" and "New_Comprehensive_November_20_2023.dta". Source: World Bank Enterprise Survey. URL: https://login.enterprisesurveys.org/content/sites/financeandprivatesector/en/library/combineddata.html. Access year: December 2023. Specific download instruction: Copy URL into preferred browser. Log in either as WB staff or create account and log in.

- "IMF_FAS.xlsx". Source: IMF Financial Access Survey. URL: https://data.imf.org/?sk=e5dcab7e-a5ca-4892-a6ea-598b5463a34c&sid=1460043522778. Access year: January 2025. Specific download instruction: Copy URL into preferred browser. Log into IMF Data (if no login credential, then create one before logging in). Once logged into the webpage, choose "Download data in Excel" in the sidebar.

- "economic-freedom-of-the-world-2023-master-index-data-for-researchers-iso.xlsx". Source: Economic Freedom of the World. URL: https://efotw.org/economic-freedom/dataset?geozone=world&page=dataset&min-year=2&max-year=0&filter=0. Access year: December 2023. Specific download instruction: Copy URL into preferred browser. Choose "Download Entire Dataset" (green button as of March 2024).

- "WBL-1971-2023-Dataset.xlsx". Source: Women, Business and the Law. URL: https://wbl.worldbank.org/en/wbl-data. Access year: January 2024.

- "Data_BRAES_Final.dta". Source: Bento and Restuccia (2017). URL: Diego Bento's personal webpage https://drive.google.com/file/d/1Ohmp-ZN94mG98kPvrhdqgoMhuOAaESSE/view. Access year: December 2023.

Processed intermediary .dta data files are under directory "InputData".

- WAP_wdi. Source: WDI. Variables: SP.POP.TOTL; SP.POP.TOTL.MA.IN; SP.POP.TOTL.FE.IN; SP.POP.1564.TO.ZS; SP.POP.1564.FE.ZS; SP.POP.1564.MA.ZS. Year created/downloaded: September 2023.

- classification.dta. Sources: Variable "ifscode": IMF World Economic Outlook (https://www.imf.org/external/pubs/ft/weo/2022/02/weodata/co.pdf). Variables "Market" and "Commodityexport": Table 1.2 of Global Economic Prospects (e.g., the Jan 2024 issue) for EMDE and commodity exporter classifications. Variable "Region": World Bank's classification of geographic region (https://datahelpdesk.worldbank.org/knowledgebase/articles/378834-how-does-the-world-bank-classify-countries). Variable "Smallstates": World Bank's list of small states (e.g., https://www.worldbank.org/en/news/statement/2024/06/26/small-states-and-small-states-forum-members). Variable "wbcode": IBAN Alpha-3 code (https://www.iban.com/country-codes).

- gdp. Source: Internal database provided by DECPG, based on World Bank Global Economic Prospects; IMF World Economic Outlook; World Development Indicators; Haver Analytics.

- controls_WDI. Source: WDI. Created by running controls_WDI.do at the time when the program was created. "control_WDI_new.dta" is created when running controls_WDI.do now. Because WDI database is updated over time, the dataset is saved under a different name to ensure the same results are generated consistently.

Note: Datasets may change over time, and fresh downloads may not be the same as the datasets downloaded at the time when the authors created the database. Such datasets include (but not limited to) WDI, IMF Financial Access Survey, ILO, GGDC, Penn World Tables, Economic Freedom of the World, Women, Business and the Law, World Bank Enterprise Survey, World Bank Informal Economy Database, and UN World Population Prospects.

Final data used for regressions are under directory “OutputData”. The main dataset is “world_extended.dta”

Main results file “FiguresTables.xlsx” is in the root directory, containing tabs for each of the figure (Figures 1 to 4), table (Tables 1 to 6). Numbers discussed in Section 3.2 and not in tables are computed in tab “Table2” column. Numbers discussed in Section 43 and not in tables are computed in tab “Corr_Margins” column AH.

SECTION II: CODES

The program has two parts. First part uses Stata to process raw data and run regressions. Second part uses Excel to make figures and tables and compute numbers mentioned in the text. The main Stata code is “Main.do” under the root directory. It calls all Stata files, including:

- **“Import_ILO.do”: imports ILOSTAT data and produces intermediate data file “InputData_ilostat.dta”.**
- **“controls_ES.do”: imports Enterprise Survey data and produces intermediate data file “InputData_ES.dta”**
- **“controls_WDI.do”:imports WDI data and produces intermediate data file “InputData_WDI_new.dta”. NOTE RUNNING Controls_WDI.do WILL CREATE DATASET BASED ON LATEST WDI DATABASE. A BACK UP COPY OF THE ORIGINAL WDI DATASET USED IS SAVED IN THE FOLDER “controls_WDI.dta”.**
- **“controls_others.do”: imports various other datasets and produces intermediate data file “InputData_others.dta”**
- “Compile_input_data.do”: imports raw data under “InputData”, produces intermediate data files under “InputData”, and finally combines all intermediate data files into final data file under “OutputData”.
- “Regression-yearly.do”: runs regressions at yearly frequency, exports regression result as .csv files under “RegResults”, calls “Regression-yearly_Margins.do” to compute margins. The latter exports computed margins to “FiguresTables.xlsx” file tab “Corr_Margins”, which is used to compute numbers mentioned in Section 4.3 of the paper (column AH). “Regression-yearly.do” also exports country female FE and average GDP per capita data to “FiguresTables.xlsx” tab “Figure4”.
- “Regression-decade.do”: runs regressions at decade frequency, exports regression result as .csv files under “RegResults”.
- In addition, the “Main.do” code also exports data to others tabs of “FiguresTables.xlsx”, which is used to produce figures and tables.

The Excel file "FiguresTables.xlsx" contains data exported from Stata code and produces figures and tables. It includes the following tabs: Main tabs:

- "Figure1" tab: makes Figure 1 using data exported by "Main.do".
- "Figure2" tab: makes Figure 2 using converted fixed effects data in "converted_FE" tab.
- "Figure3" tab: makes Figure 3 using converted fixed effects data in "converted_FE" tab.
- "Figure4" tab: makes Figure 4 using country fixed effects and average GDP per capita data exported by "Regression-yearly.do". It converts the country fixed effects using regression coefficients from the "Table_…" tabs.
- "Table1" tab: makes Table 1 using regression results from the "Table…" tabs.
- "Table2" tab: makes Table 2 using converted fixed effects data in "converted_FE" tab.
- "Table3" tab: makes Table 3 using regression results from "Corr" tab.
- "Table4" tab: makes Table 4 using regression results from "Table_…" tabs.
- "Table5" tab: makes Table 5 using country fixed effects in "raw_FE2" tab.
- "Table6" tab: makes Table 6 using regression results from "Table_…" tabs.
- "Corr_Margins" tab: computes numbers mentioned in Section 4.3 of the text using margins numbers exported by "Regression-yearly_Margins.do".

Intermediate calculation tab:

- "converted_FE" tab: computes converted country fixed effects using raw country fixed effects in "raw_FE" tab and regression coefficients in "Table_…" tabs. Data in this tab is also used by "Main.do" to produce estimates for the "Corr" tab.

Source data tabs: tab:

- "Exportclassification" tab: country's commodity exporter status, exported by "Main.do".
- "raw_FE2" tab: raw country fixed effects variables, exported by "Main.do".
- "raw_FE" tab: raw country fixed effects variables, exported by "Main.do".
- "Corr" tab: correlation result produced by "Main.do" using data converted country fixed effects data from "converted_FE" tab.
- "Table_…" tabs: regression results produced by "Regression-yearly.do" and "Regression-decade.do" and exported into the Excel file by "Main.do"

SECTION III: FOLDERS

The folder structure should be self-explanatory.

Root

```
├── All codes: Stata .do files and FiguresTables.xlsx
└── InputData ├── RawData
                  └── RawControlVars
                  └── IntermData
└── OutputData
└── RegResults
```

Root contains the Excel file and all Stata .do files. “InputData” contains raw and intermediate data. Raw data are under “InputData” or “InputData” folders. Intermediate data are under “InputData”. “OutputData” contains final .dta data file used for regression, and .dta data files produced by Stata regressions. “RegResults” contains .csv regression results files exports from Stata codes.