

README for the Reproducibility Package for Human Capital Expenditure and Its Effectiveness on Multi-Programme Coverage: A Policy Prioritization Investigation

Authors: Omar A. Guerrero¹, Daniele Guariso², Gonzalo Castañeda³, and Michael Weber⁴

¹ University of Helsinki ² Euro-Mediterranean Center on Climate Change ³ Centro de Investigación y Docencia Económicas (CIDE), Mexico City ⁴ The World Bank

Overview

This repository contains the code and data to replicate the analysis performed in the World Bank Working Paper titled: *Human Capital Expenditure and Its Effectiveness on Multi-Programme Coverage: A Policy Prioritization Investigation*. The analysis makes use of the Policy Priority Inference (PPI) framework to model connections between government expenditure and target indicator coverage in Mexico between 2016 and 2022. The paper presents various types of analysis that aim at identifying potential opportunities and challenges in coordinating expenditure strategies that aim at improving Mexico's stock of human capital.

Paper abstract

This paper uses AI-enhanced agent computing to determine how to allocate budgetary resources within a large set of heterogeneous government programs targeting human capital. Our approach considers essential features of the budget allocation process: multidimensionality, interdependencies between policy issues, and the political economy of public officials' collective action. We use highly disaggregated Mexican data covering the 2016-2022 period across 49 human capital programs of the federal government and focus on how expenditure affects program coverage (the proportion of the population with a public problem and who has access to various government benefits to mitigate those problems) in the short run. We answer the following research questions: how sensitive is program coverage to changes in public expenditure?; what are the structural bottlenecks behind poor coverage response?; and what are the optimal budgetary allocations that could boost the performance of a multidimensional objective function?

The structure

The repository is organized into three folders: - **code**: contains all the scripts needed to process the raw datasets and perform the simulations for the analysis - **data**: provides all the necessary data - **figures**: provides high resolution files of all the figures in the report - **tables**: contains the tables from the report

The code

The code is organized into sequential Python scripts. They should be run in the order indicated by the number in the filenames. Files from 1 to 9 are for processing data. Files from 11 to 14 calibrate the PPI model and run all the experiments. Scripts 21 onward produce all the figures in the report and saves them in the **figures** folder. Once all files have been run, the user can take the output data files and replicate the figures presented in the report.

Policy Priority Inference (PPI)

The analysis requires the PPI toolkit, which can be installed for Python through pypi. Further information on PPI can be found in the book: Complexity Economics and Sustainable Development. The open source-code for the PPI framework can be found in its official repository.

Data Availability Statement

1. Inventario CONEVAL de Programas y Acciones Federales de Desarrollo Social Source: CONEVAL Year: 2023 URL: <https://www.coneval.org.mx/evaluacion/ipfe> Access Date (Month-Year): 10-2024 Note: - Access Type: Open License: - License URL: - Citation: CONEVAL. 2023. “Inventario CONEVAL de Programas y Acciones Federales de Desarrollo Social” [dataset]. <https://www.coneval.org.mx/evaluacion/ipfe>. Accessed 10-2024.
2. Worldwide Governance Indicators Source: World Bank Year: 2023 URL: <https://www.worldbank.org/en/publication/worldwide-governance-indicators> Access Date (Month-Year): 10-2024 Note: - Access Type: Open License: - License URL: - Citation: World Bank. 2023. “Worldwide Governance Indicators” [dataset]. <https://www.worldbank.org/en/publication/worldwide-governance-indicators>. Accessed 10-2024.
3. Consumer Price Indices Source: World Bank Year: 2023 URL: https://data360.worldbank.org/en/dataset/FAO_CP Access Date (Month-Year): 10-2024 Note: - Access Type: Open License: - License URL: - Citation: World Bank. 2023. “Consumer Price Indices” [dataset]. https://data360.worldbank.org/en/dataset/FAO_CP. Accessed 10-2024.
4. World Population Prospects Source: United Nations Department of Economic and Social Affairs Year: 2023 URL: <https://population.un.org/wpp/> Access Date (Month-Year): 10-2024 Note: - Access Type: Open License: - License URL: - Citation: United Nations Department of Economic and Social Affairs. 2023. “World Population Prospects” [dataset]. <https://population.un.org/wpp/>. Accessed 10-2024.
5. Human Capital Index Source: World Bank Year: 2023 URL: <https://humancapital.worldbank.org> Access Date (Month-Year): 10-2024 Note: - Access Type: Open License: - License URL: - Citation: World Bank. 2023. “Human Capital Index” [dataset]. <https://humancapital.worldbank.org>. Accessed 10-2024.

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. I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package. Appropriate permissions are documented in the LICENSE.txt file.

Instructions for Replicators

If all the necessary libraries have been properly installed, all needed is cloning this repository (preserving the folder structure) and run each script sequentially. The repository already provides all the intermediate data files, so it is not necessary to run every single script. For example, if you want to modify the experiments from section 4.1 (sensitivity analysis), you can go straight to script 12, modify it, and run it.

Software Requirements

- Python 3.12+ and R 4.0.5
- Python dependencies: pandas, numpy, matplotlib, scipy, scikit-learn, joblib, rpy2, policy-priority-inference
- R dependencies: sparsebn, sparsebnUtils, ccdrAlgorithm