# PER CAPITA INCOME, CONSUMPTION PATTERNS, AND $CO_2$ EMISSIONS - README FOR REPLICATION PACKAGE -

Justin Caron and Thibault Fally.*

July 2021

**Abstract**

This document describes the content of the replication package attached to "Per Capita Income, Consumption Patterns, and $CO_2$ Emissions" to be published in the *Journal of the Association of Environmental and Resource Economists*. The estimation of model parameters, as well as model calibration and simulation are all implemented using the GAMS modeling language. Compilation of results and graph-making are done in STATA. This note describes how to replicate results.

*Justin Caron: Department of Applied Economics, HEC Montreal. Thibault Fally: Department of Agricultural and Resource Economics, University of California Berkeley.

## Introduction

Estimation and simulation are implemented in the GAMS (General Algebraic Modeling System) modeling language. The GAMS programs and STATA do-files included in the package allow the reader to replicate all of paper's results. Result generation is mostly automated except for manually copying results from GAMS into STATA. The user can use the `runall.bat` batch file to follow the workflow.

In the following, we describe the list of GAMS programs, do-files and data files contained in the replication folder. The most important files are `demand.gms`, which describes the estimation of demand parameters, and `CF_CO2_NH.gms`, which includes the general equilibrium model used for counterfactual simulation.

## Data

All of the data used to produce the paper's main results are freely available online.

**GTAP:** The main estimation and simulation exercise is based on the GTAP (Global trade analysis project) version 8 dataset, which is freely available from the GTAP website at www.gtap.agecon.purdue.edu. Robustness checks use GTAP9, which as of early 2021 still requires a licence.

Extraction of the raw GTAP8 data and translation into GAMS format relies on the GTAPinGAMS code developed by Thomas Rutherford (University of Wisconsin). The following describes the steps required for the creation of the GAMS-ready GDX file (`GAMS/data/gtap8gas.gdx`) used for analysis.

- Download the GTAP data from the GTAP website

  at https://www.gtap.agecon.purdue.edu/databases/archives.asp. Chose "Download GTAPAgg 8.1". Also download the appropriate GTAP licence file at "GTAPAgg License" and copy the `gtapagg.lic` file to the `GTAPagg` folder.

- Run `gtapagg.exe` to extract the whole dataset. We don't aggregate any of the dimensions here. Thus click on "View/change regional aggregation" and chose the "1 to 1" option. Do the same for "sectoral aggregation" and "factor aggregation". Create the aggregated database.

- `gtapagg.exe` creates a ZIP folder including a number of datafiles in the *HAR* format. Copy this ZIP file to the `GAMS/data` folder

- Then run the `make-GTAPAGG.bat` batch file from the `/GAMS/data/GTAP data extraction` folder. This successively runs the `ZIP2GDX.gms` file (extracts data from the ZIP folder into a GAMS-readable *GDX* file) and the `FILTER.GMS` file which filters out extremely small values in the GTAP data (this step isn't strictly necessary).

- Run `gtapaggr.gms`, using the sectoral mapping in `gtap8gas.map` to aggregate the two sectors pertaining to natural gas, GDT and GAS, into a single GAS sector.

- Output is the GAMS-useable `GTAP8gas.gdx` file provided in the `/GAMS/Data` folder.

**Other data sources:**

- Bilateral trade cost determinants. From CEPII. Imported into the `GAMS/data/importdist_gtap8.gdx` file.
- Non-CO2 emissions data. Made available by GTAP as "Non-CO2 Emissions 8 Data Base" from https://www.gtap.agecon.purdue.edu/resources/free_resources.asp.

  Copied to the `GAMS/data/GTP_NCO2_MMTCEQ_v8.gdx` file.

As the above data sources are freely available we have left them in the package to facilitate replication.

## GAMS files for estimations and simulation

**Running GAMS** All GAMS programs have been used and tested with GAMS version 24.2.3 on a PC computer. The code should work on MAC as well. We solve the estimation and simulation models with the following solvers:

- CONOPT non-linear solver for Maximum Likelihood estimation of gravity and demand equations. Other non-linear solvers should work. Alternatively NLLS estimation of gravity could be done using CPLEX or any solver accepting quadratic optimization.
- PATH non-linear solver. Used to solve the simulation model, which is a formulated as a mixed-complementarity (MCP) model. Other non-linear solvers may work.

You will need the appropriate licences.

The following batch file can be used to execute all GAMS files in the correct order:

- **Runall.bat** - a windows batch file (text editable) which runs all GAMS files necessary for estimation and simulation. The content of this file describes the whole workflow, up to the creation of GDX files to be copied to STATA. Executing this batch file in one go should replicate all estimation and simulation results but will take a long time. Alternatively, it can be used to follow the workflow piece-by-piece.

To run the batch file you will need to add GAMS to your computer's PATH file.
Alternatively each file can be ran independently from GAMS (specifying the appropriate options).

### Data preparation and estimation

The following describes the files used for the estimation of gravity and demand parameters as well as the preparation of the parameters subsequently used to calibrate the simulation model.

These programs should be executed in the following order:

- **gravity.gms** - Estimates the gravity parameters and implied $\Phi$ parameters. The file estimates gravity Poisson ML but can also be configured to estimate an OLS model.
  - Output: `gravityestimates_gtap8_all.gdx`
- **demand.gms** - Main estimation file for demand parameters. Estimates the CRIE demand system with different specifications for $\theta$. The file must be run with the following options for the *SPEC* (specification) global variable (see batch file):
  - *tc*: the theta parameter is estimated from the $\Phi$ parameter (derived from trade costs). Benchmark specification.
  - *theta4*: $\theta$ is fixed at 4 in all sectors
  - *homonotc*: homothetic version with identical prices across countries (no trade costs)
  - Output: `estimates/estimates_gtap8gas_logweighted_tc_rall.gdx`
- **demand_clm_flex_quadratic_t4.gms** - estimates quadratic non-homothetic CES preferences as in Comin et al. 2016 (CLM). The file assumes $\theta = 4$ for all sectors as default.
  - Output: `estimates/estimates_gtap8gas_logweighted_theta4_CLM_flexquadratic.gdx`
- **demand_clm_flex_quadratic_t4.gms** - estimates quadratic non-homothetic CES preferences as in Comin et al. 2016 (CLM). The file assumes $\theta = 4$ for all sectors as default.
  - For standard log-linear NH CES: **demand_clm_t4.gms**
  - For shifter flexible NH CES: **demand_clm_flex_shifter_t4.gms**
  - Output: `estimates/estimates_gtap8gas_logweighted_theta4_CLM_flex_quadratic.gdx`
- **datapreparation.gms** - prepares IO coefficients, trade shares, and other parameters to be subsequently passed on to the simulation simulation programs below.

  Also inverts the IO matrix to compute total $CO_2$ intensities and prepares all sector-level reporting parameters to be exported to STATA. F Loads non-$CO_2$ other GHG data as well.

  The file must be run for the following demand specifications, i.e. options for the *DEMANDEST* variable: *tc*, *theta4* and *tccsbeta*. The first two are with average input shares, the later with country-specific input shares.
  - Output: `estimates/co2_data_TC.gdx`
- **CF_compute_fittedprod_shares.gms** - Computes fitted production shares (consistent with fitted consumption shares) for use in simulation model.

The file must be run for the following demand specifications, i.e. options for the *DEMANDEST* variable: *tc* and *theta4*, with SPEC=NH, for CRIE; and *theta4* with SPEC=NH_CLM, NH_CLMquad, NH_CLMshift for NH CES preferences.

  - Output: `fitted_prod_shares_SPEC_DEMANDEST.gdx`

### General equilibrium simulations

- **cf_co2_nh.gms** - This is the main file containing the general equilibrium GE model which is used to simulate counterfactual changes in emissions based on CRIE preferences.

  The following global variables must be defined:

  - *SPEC*: defines demand specifications: *NH* (non-homothetic demand) or *H* (homothetic demand)
  - *DEMANDEST*: defines the demand estimation of $\theta$: *TC* (default) or *THETA4*
  - *PRODSPEC*: define fossil fuel supply elasticity: *resprod_06* (elasticity of 0.75, default), *resprod_15* (1.5), *resprod_50* (5), *resprod_06_gas15* (0.75 for coal and oil, 1.5 for gas); *noresprod* (no productivity growth in the resource sectors).

For the robustness checks, the *DEMANDEST* global variable can be switched to *THETA4* and *THETA8*.

All results and parameters are exported to the `results` folder.

**CF_co2_NH_CLM_flex.gms** - file including the same GE model as above, but with NH CES preferences from Comin et al. (2016) (CLM). Options for DEMANDEST are *NH_CLMstd* (standard loglinear NH CES); *NH_CLMquad* (flexible quadratic, default); *NH_CLMshift* (flexible shifter).

**CO2_estimations.gms** loads demand estimates, coefficients from datapreparation and simulation results, computes statistics and prepares output for graphs/tables. For the "content" computations/graphs, set fitdem=no; to compile "simulation" results, set fitdem= yes

## STATA do-files for graph and table generation

The following describes do-files required to compile the tables and graphs included in the main text. Code to generate tables and graphs in the appendix is available upon request.

The do-files can be found in the `STATA` subfolder. The generated graphs are saved in the `Figures` sub-folder.

- Table 1: All data is in the `co2_data_tc.gdx` file, COEFFS parameter.

- Figure 1: Use the `fdanalysis.do` STATA do-file importing data from `estimates_merged.gdx` (a gdx merge of the `estimates_XX.gdx` files for all 4 demand specifications, use gdxmerge.exe).
- Figure 2: Use the `figures_sectorcorrelations.do` do-file importing data from `CO2_DATA_TC.GDX`, COEFFS parameter.
- Figures 3, 4, 5: Use the `figures_co2_content.do` do-file importing data from `results/simulationresults_tomerge/merged.gdx` (a gdx merge of files in the `results/simulationresults_tomerge.gdx` folder). CONTENTS parameter.
- Table 2: results are in the MODELFIT parameter in the `co2stats_tc_NH_respro06_fittedem_no.gdx` file.
- Figure 6: Use the `figures_simulations.do` do-file importing data from `results/simulationresults_tomerge/merged.gdx` (a gdx merge of files in the `results_simulationresults_tomerge.gdx` folder). CHG_ENERGY_CO2 parameter.