Overview

The code in this replication package prepares the data sets used in "Salience and Taxation with Imperfect Competition" and runs all the analyses used in the data using Stata. All dofiles required to prepare the analytical files are called by BUILD.do. All dofiles required to produce the results are called by ANALYSES.do.

Data Availability and Provenance Statements

This paper uses the following three datasets:

1. The Nielsen Consumer Panel (HMS)
2. The Nielsen Retail Scanner Dataset (RMS)
3. Tax rates and exemptions on consumer goods, hand-collected by the authors.

The two Nielsen databases (HMS and RMS) are confidential. They are provided through Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. Information pertaining to data access can be found at http://research.chicagobooth.edu/nielsen/.

The tax data used in this study have been deposited in the Zenodo repository (DOI: 10.5281/zenodo.7526306). The data were collected by the authors (all sources are listed in Table OA.2), and are publicly available under a Creative Commons Non-commercial license. These data are included in the replication archive.

Data Citations

Kroft, Kory, Jean-William P. Laliberté, René Leal-Vizcaíno, and Matthew J. Notowidigdo (2022). Replication data for: Salience and Taxation with Imperfect Competition. DOI: 10.5281/zenodo.7526306

NielsenIQ (2016). Data provided by the Kilts Center for Marketing Data Center. Accessed on 05-05-2016.

Dataset list

Data file	Source	Notes	Provided
`Numerous raw files (one per year)` purchases_`yr'.tsv trips_`yr'.tsv panelists_`yr'.tsv	Nielsen HMS	Confidential	No
`Numerous raw files (one per year and per module)` `m'_`yr'.tsv	Nielsen RMS	Confidential	No
`KLLN_modules_hierarchy.dta`	Nielsen	List of modules used in paper. File created by authors.	Yes
`stores_panel.dta`	Nielsen	List of stores in Nielsen data (Confidential). File created by authors.	No
`products.dta`	Nielsen	List of products in Nielsen data (Confidential). File created by authors.	No
`countyrates_quarterly.dta`	Authors	List of tax rates by county	Yes
`taxability_quarterly.dta`	Authors	List of tax exemptions	Yes
`tau.dta`	Authors	Tax rate by product-country-quarter	Yes
`border.dta`	Authors	List of border counties	Yes

`statecoord.dta`
`statedb.dta`
`uscoord.dta`
`usdb.dta`
	US Census Bureau – converted to dta format by authors	GIS coordinates or states and counties.	Yes

Computational requirements

Software Requirements

- Stata (BUILD.do will install user-written packages)
- Python
- Excel (for calibrations)

Memory and Runtime Requirements

The codes were last run on a desktop running Windows 7 Professional with an Intel® Xeon® Processor E5-1660 v4 (8C, 3.2GHz, 3.8GHz Turbo, 2400MHz, 20MB, 140W) with 256GB if RAM. Importing the RMS data takes about 14 days. The remaining codes together take over 20 days to run.

Description of programs

- Programs in `programs/data prep` will prepare all of the necessary datasets.
- Programs in `programs/analysis` will prepare all of the results.

Instructions to Replicators

- Edit `programs/BUILD.do` and `programs/ANALYSES.do` to adjust the default path.
- Also edit `programs/data prep/RSM_TSV_reading.py` to adjust paths. This file takes the raw RMS data files (in tsv format) and saves them as Stata data files (in dta format).
- Run, in order:
  - `programs/data prep/RSM_TSV_reading.py`
  - `programs/BUILD.do`
  - `programs/ANALYSES.do`
- Calibrations are conducted directly in Excel, in the file `output/Tables_scai.xlsx`
- All Tables in the paper are stored in `output/Tables_scai.xlsx`

List of tables and programs

Figure/Table #	Programs	Output files
Table 1	n.a. (no data)	
Table 2	RMS_fullsample.do RMS_border.do	tax_statetrends.csv price_statetrends.csv border_main.csv statetrends_HMSweighted.csv
Table 3	HMS_fullsample.do variances.do	HMS_baseline.csv HMS_Xs.scml variances.scml
Table 4	HMS_mixed.do variances.do	variances.scml
Table 5	n.a. (no data / calibrations)	Tables_scai.xlsx
Table 6	n.a. (no data / calibrations)	Tables_scai.xlsx
Table OA.1	n.a. (no data)	
Table OA.2	n.a. (no data)	
Table OA.3	RMS_fullsample.do RMS_border.do	loneway_col1.scml loneway_col2.scml

		loneway_col3.scml
Table OA.4	RMS_border.do	border_stateIV.csv border_fs_log.scml
Table OA.5	RMS_fullsample.do	HHI_heterogeneity.csv
Table OA.6	RMS_fullsample.do RMS_border.do	reducedform_full.csv reducedform_border.csv
Table OA.7	RMS_fullsample.do	hiloP_statetrends.csv
Table OA.8	RMS_fullsample.do RMS_border.do	tax_statetrends.csv price_statetrends.csv border_main.csv
Table OA.9	RMS_fullsample.do	robustness_trends.csv
Table OA.10	RMS_fullsample.do	noALCI_statetrends.csv
Table OA.11	HMS_fullsample.do	HMS_qRMS.csv HMS_qRMS_levels.csv
Table OA.12	HMS_mixed.do variances.do	variances.scml
Table OA.13	n.a. (no data / calibrations)	Tables_scai.xlsx
Table OA.14	n.a. (no data / calibrations)	Tables_scai.xlsx
Table OA.15	n.a. (no data / calibrations)	Tables_scai.xlsx
Table OA.16	n.a. (no data / calibrations)	Tables_scai.xlsx
Figure OA.1	n.a. (no data)	
Figure OA.2	maps.do	RMS_sample.pdf HMS_sample.pdf
Figure OA.3	maps.do	county_rates_q32008.pdf
Figure OA.4	maps.do	food_taxability.pdf
Figure OA.5	maps.do	county_rates_change.pdf
Figure OA.6	maps.do	county_rates_nbchange.pdf