Documentation for Dubois, Griffith and O'Connell (2020) "How well targeted are soda taxes?" (openicpsr-120232)

The purchase data are from Kantar UK and cover June 2009-Dec 2014.

We use data from;

(i) their Worldpanel (now called the FMCG purchase panel): see https://www.kantarworldpanel.com/en/Consumer-Panels-/FMCG

(ii) their Food on-the-go survey (now called the Out-of-home survey): see https://www.kantarworldpanel.com/global/Consumer-Panels/Out-of-Home

The data are available commercially from Kantar. Up to date contact information for Kantar UK are provided here: https://www.kantarworldpanel.com/global/Coverage/worldpanel/United-Kingdom

The Kantar data that is hosted at the Institute for Fiscal Studies (IFS) are accessible to academic visitors to the institute. IFS hosts several academic visitors each year. Researchers interested in visiting the IFS should get in touch with Emma Hyman (emma_h@ifs.org.uk). The ability for any individual to visit the IFS will depend on capacity at IFS and available research funds.

The advertising data are from AC Nielsen and cover calendar years 2009-2014.

AC Nielsen collect information on TV advertising expenditure at the brand-slot level: see https://www.nielsen.com/eu/en/solutions/measurement/advertising-expenditure/

The data are available commercially from AC Nielsen. Researchers interested should contact Kyle Lee (kylie.lee@nielsen.com).

The AC Nielsen data that is hosted at the Institute for Fiscal Studies (IFS) are accessible to academic visitors to the institute. IFS hosts several academic visitors each year. Researchers interested in visiting the IFS should get in touch with Emma Hyman (emma_h@ifs.org.uk). The ability for any individual to visit the IFS will depend on capacity at IFS and available research funds.

The data on UK weather can be downloaded directly from the UK Met Office using the link https://www.metoffice.gov.uk/research/climate/maps-and-data/historic-station-data. Series for all stations (with the exception of cwmystwyth, which closed from 2011 April) should be downloaded. We use the period June 2009-Dec 2014.

Program structure

The majority of the programs are in Stata format (*.do). Some analysis is done in MatLab (*.m). The file structure is as follows: Files with -

- **0** prefix are master files that call other do files and set up file paths
- **1** prefix prepare data for estimation of the demand model
- **2** prefix estimate demand model
- **3** prefix prepare data for Tables and Figures in Section I
- **4** prefix prepare data for running counterfactuals
- **5** prefix prepare data for Tables and Figures in Section II
- **6** prefix prepare data for Tables and Figures in Section III
- **7** prefix prepare data for Monte Carlo confidence bands

- **8** prefix prepare data for Tables and Figures in Section IV
- **9** prefix combines data with confidence brands and produces Tables and Figures in paper
- **A** prefix prepares data and produces tables and figures in Online Appendix

Description of individual programs

0.a.Run.do master file which calls all other programs

0.b.paths.do sets up file paths (for folders that must be created before running)

0.c.globals.do defines globals used throughout

1.a.createdata.do sets up data for estimation of demand model. Calls **1.a.i.advertising.do** and **1.a.ii.athominventories.do**

2.a.estimation.do estimates demand model by maximum likelihood

3.a.decriptives.do creates data for tables in Sections I.A, I.B, and I.D

3.b.demandlinkages.do creates data for analysis in Section I.C

4.a.filesetup.do sets up estimates for counterfactuals

5.a.passthroughdesc.do sets up data for analysis in Section II.C

5.b.estimates.do extracts coefficient estimates. Need to run **estimate.m** first. (This files computes standard errors for moments of the coefficient distribution and take draws from the estimate vcov matrix for the Monte Carlo confidence bands)

5.c.priceeffects.do computes price elasticities

6.a.counterfactual.do conducts baseline counterfactual tax analysis

7.a.coefdraws.do sets up files with draws from estimate vcov matrix

7.b.prepare_ci.do combines draws with data

7.c.priceeffects_ci.do computes price elasticities with draws

7.d.counterfactual_ci.do conducts baseline counterfactual tax analysis with draws

8.a.jackknife.do estimates jackknife version of on-the-go demand

8.b.jackknife.do computes bias correction

8.c.foodindata.do sets up data for estimation of the at-home demand model. Calls **8.c.i.foodinweather.do**

8.d.estimation.in.do estimates at-home demand model by maximum likelihood

8.e.filesetupin.do sets up at-home estimates for counterfactuals in robustness section

8.f.foodin.do computes counterfactual for food in segment

8.g.passthrough1.do prepares data for pass-through simulation

8.h.passthrough2.do combines simulated prices with data. Need to run **Run.m** first (which is in the Passthrough folder). This calls a set of .m files contained in the folder and simulates tax-passthrough.

8.i.eqcounterfactual.do computes counterfactual with simulated pass-through

9.a.combines main results, robustness and confidence bands

9.b.tables.do produces all tables in the paper

9.c.graphs.do produces figures in paper (with the exception of Figures 5(d)-(f) and 6(d)-(f) which are produced by running **graphs.m**)

A.a0.Run.do master file for appendix

A.0b.paths.do sets file paths for appendix

A.1a.Kantar.do sets up Kantar data for appendix

A.1b.NDNS.do sets up NDNS data for appendix

A.1c.NHANES.do sets up NHANES date for appendix

A.2.Output.do produces tables and figures for appendix

Data created from data preparation programs

1.a.createdata.do	allprodcodes.dta
1.a.createdata.do	fotg_rawattributes.dta
1.a.createdata.do	prodcode_purchases.dta
1.a.createdata.do	attributes.dta
1.a.createdata.do	week.dta
1.a.createdata.do	hhno_WPdates.dta
1.a.createdata.do	FOTGpurchases.dta
1.a.createdata.do	indvno_panelstats.dta
1.a.createdata.do	indvno_sample.dta
1.a.createdata.do	hhno_sample.dta
1.a.createdata.do	snackpurchases_raw.dta
1.a.createdata.do	rmproduct_avail.dta
1.a.createdata.do	all_purchases.dta
1.a.createdata.do	transaction_prices.dta
1.a.createdata.do	prices.dta
1.a.createdata.do	index_vars.dta
1.a.createdata.do	options_vars.dta
1.a.createdata.do	Estimation_data.dta
1.a.i.advertising.do	week_ad.dta
1.a.i.advertising.do	alladv.dta
1.a.i.advertising.do	wk_n.dta
1.a.i.advertising.do	nationalad.dta
1.a.i.advertising.do	addata.dta
1.a.i.advertising.do	hhym_weather_string.dta
1.a.ii.athominventories.do	day_stock.dta
1.a.ii.athominventories.do	raw_athomedrinks.dta
1.a.ii.athominventories.do	athomedrinktypes.dta
1.a.ii.athominventories.do	hhno_adequiv.dta
1.a.ii.athominventories.do	athome_inventories.dta
3.a.decriptives.do	sugars_prod.dta
3.a.decriptives.do	active.dta
3.a.decriptives.do	demogs.dta
3.a.decriptives.do	weekindata.dta
3.a.decriptives.do	aggdiet_a.dta
3.a.decriptives.do	centiles.dta
3.a.decriptives.do	age.dta
3.a.decriptives.do	size.dta

3.a.decriptives.do	purchyearfull.dta
3.a.decriptives.do	hhpurch_base.dta
3.a.decriptives.do	T_out.dta
3.a.decriptives.do	storestats.dta
3.a.decriptives.do	mean_price.dta
3.a.decriptives.do	productstats.dta
3.a.decriptives.do	agestats.dta
3.a.decriptives.do	expstats.dta
3.a.decriptives.do	sugstats.dta
3.b.demandlinkages.do	volume_out.dta
3.b.demandlinkages.do	link_days.dta
3.b.demandlinkages.do	volume_in.dta
3.b.demandlinkages.do	rf_demogs.dta
3.b.demandlinkages.do	demandlinkage.dta
5.a.passthroughdesc.do	passthroughdesc.dta
8.c.foodindata.do	rawdata.dta
8.c.foodindata.do	drinkpurchases_raw_allhh.dta
8.c.foodindata.do	drinkpurchases_raw.dta
8.c.foodindata.do	rmprodsizeyear_avail.dta
8.c.foodindata.do	rmprodsize_avail.dta
8.c.foodindata.do	rmproduct_avail.dta
8.c.foodindata.do	priceweights.dta
8.c.foodindata.do	prrmmapping.dta
8.c.foodindata.do	prices.dta
8.c.foodindata.do	Estimation_data_cs.dta
8.c.i.foodinweather.do	hhym_weather_string_in.dta