Raw Data

FAOSTAT_data_2017.csv	Area, production and yield by country. Downloaded from FAOstat (http://www.fao.org/faostat/en/#data/QC)
FAOSTAT_stocks_data_2017.csv	This is the data that Roberts and Schlenker used to generate quantity demanded. *I have not processed this dataset yet because I am not exactly sure what it measures*. Downloaded from FAOstat (http://www.fao.org/faostat/en/#data/BC)
CPI.csv	Consumer price index (https://fred.stlouisfed.org/series/CPIAUCSL)
hemisphere.dta	denotes the hemisphere that each country lies in (1=north; 0=south)

Code to create Stata datasets for analysis

setup_quantities.do	Reads "FAOSTAT_data_2017.csv" and "hemisphere.dta" and creates two datasets: 1. caloric_panel.dta 2. global_quantities.dta
setup_prices.do	Downloads futures data from quandl and creates four datasets of daily futures prices (Cfut.dta, RRfut.dta, Sfut.dta, Wfut.dta), one for each commodity, and one dataset of annual prices for use in estimation (global_prices.dta).

Stata datasets

caloric_panel.dta — sample: 1961-2014. Contains area, production and yield by country

Variables:

Year = year
country = country code
country_str = country name
northern = dummy variable for northern hemisphere
maize_area = hectares of land planted to maize
maize_prod = maize production in metric tonnes
maize_yield = maize production in tonnes per hectare
ln_maize_yield = natural log of maize yield
yhat_maize_cntry = trend yield for maize
kappa_maize = calorie weight (tonnes per calorie*2000*365)
same variables for rice, soybeans and wheat
trendsp1, trendsp2 = variables used to estimate cubic spline to create trend yield
area = total area planted to the four crops
prod = calorie-weighted production
yield_trend_sum = calorie-weighted sum of trend yield

(Note: maize and corn are different names for the same thing.)

global_quantities.dta — sample: 1961-2014. Contains area, production and yield for the world

Variables:

Same as on caloric_panel.dta plus

yield_trend = calorie-weighted average trend yield (Y in Hendricks, Janzen and Smith)

yield_shock = calorie-weighted yield shocks. ***You can use the log of this variable as the ω in your supply models.***

Cfut.dta, RRfut.dta, Sfut.dta, Wfut.dta

sample: 1959-2016. Contains daily futures prices on all traded contracts in the sample period; one column for each contract.

Variables:

date = date (ddmmmyyyy)

pXXXXA = price on date of contract that expires in month A of year XXXX (months: F=Jan, G=Feb, H=Mar, J=Apr, K=May, M=Jun, N=Jul, Q=Aug, U=Sep, V=Oct, X=Nov, Z=Dec).

contXXXXA = expiration date of contract that expires in month A of year XXXX.

global_prices.dta — sample: 1959-2016. Contains annual prices for estimation

Variables:

year = year

C_spot_cont = expiration date of futures contract used to define spot price (will be Nov 1 or Dec 1 of current year)

C_spot_price = spot price

C_fut_cont = expiration date of futures contract used to define futures price relevant for next year's supply (will be Nov 1 or Dec 1 of next year)

C_spot_price = futures price

same variables for rice, soybeans and wheat

cpi = consumer price index

Code to estimate supply elasticities

analysis.do — reads in global_quantities.dta and global_prices.dta and generates estimates of supply elasticities