Local Pollution as a Determinant of Residential Electricity Demand

By Alberto Salvo

Prepared in April 2020

The computer code described below prepares and combines open-access environmental/ economic data (made available) and proprietary household-level utility usage microdata (available with restrictions). The code generates all the results included in the article and in the online appendix. All code runs in Stata version 15 or higher.

Data made available replicates the analysis of: electricity demand at the level of identifying variation (Table B.3's robustness test after collapsing the data to area by usage period); the home energy behaviors survey (Table B.5, Figs. 3 and B.4); and aggregate health (Table C.1).

The following files containing source data and code are described in what follows:

(i) source_pollution_(area).csv
(ii) upload_nea_pollution.do
(iii) source_noaa_igra2.dta
(iv) upload_noaa_igra2.do
(v) source_weather_(station number)_yyyymm.csv
(vi) upload_mss_weather.do
(vii) source_weather_nus_geography.xlsx
(viii) upload_nusg_weather.do
(ix) source_fire_archive_M6_5540.csv
(x) upload_nasa_firms.do
(xi) source_usdos_pollution_beijing_2013_2014.dta
(xii) source_usdos_pollution_kolkata_2016.csv
(xiii) source_usdos_pollution_dhaka_2017.csv
(xiv) upload_pollution_beijing_kolkata_dhaka.do
(xv) source_prices_monthly.csv
(xvi) upload_prices.do
(xvii) source_ema_industrial_electricity.csv
(xviii) upload_ema_industrial_electricity.do
(xix) source_hits_income.dta
(xx) upload_hits_income.do
(xxi) source_holidays.dta
(xxii) aggregated_usage_data_to_replicate_table_b_3.dta
(xxiii) aggregated_usage_data.dta
(xxiv) Survey of Home Energy Behaviors.pdf
(xxv) Survey+of+Home+Energy+Behaviors_November+21%2C+2018_00.37.csv
(xxvi) home_energy_behaviors_survey.do
(xxvii) source_moh_polyclinic_visits.csv

(xxviii) analysis_health_polyclinic_visits.do
(xxix) run_non_proprietary.do
(xxx) source_RESTRICTED_microdata_consumption_201209to201312.csv
(xxxi) source_RESTRICTED_microdata_consumption_201401to201412.csv
(xxxii) source_RESTRICTED_microdata_consumption_201501to201512.csv
(xxxiii) source_RESTRICTED_microdata_actual_or_estimated.csv
(xxxiv) source_RESTRICTED_microdata_billing_period.csv
(xxxv) (will run only with proprietary data) upload_spservices_microdata.do
(xxxvi) (will run only with proprietary data) combine_proprietary_data.do
(xxxvii) (will run only with proprietary data) run_proprietary.do
(xxxviii) source data for Fig. 2.xlsx
(xxxix) source data for Fig. A.3.xlsx
(xl) source data for Fig. A.4.xlsx
(xli) source data for Fig. B.5.xlsx

Proprietary household-level utility usage microdata can be purchased from SP Services Ltd. These source data files include the term "RESTRICTED" in their label. Please see "Data Extraction Request Form (endorsed).pdf." This is the data extraction request form that I signed in June 2015 with SP Services Ltd, and was later endorsed by the Vice-Dean of NUS' Faculty of Arts & Social Sciences. Upon reasonable request, the microdata are available on an NUS Department of Economics (or equivalent institutional) computer to replicate all published results from the deposited computer code.

1. Stata do file "upload_nea_pollution.do" reads the National Environment Agency (NEA) source data files "source_pollution_(district).csv," by district of Singapore, and prepares the following variables:
   a. PM2.5, North, 24-hour mean ($\mu g\ m^{-3}$)
   b. PM2.5, East, 24-hour mean ($\mu g\ m^{-3}$)
   c. PM2.5, South, 24-hour mean ($\mu g\ m^{-3}$)
   d. PM2.5, West, 24-hour mean ($\mu g\ m^{-3}$)
   e. PM2.5, Center, 24-hour mean ($\mu g\ m^{-3}$)

   Figs. A.7 (partial) and A.8 are generated.

2. Stata do file "upload_noaa_igra2.do" reads the NOAA Integrated Global Radiosonde Archive, IGRA Version 2 source data file "source_ noaa_igra2.dta," and prepares the following variables:
   a. (dpdpG) Dew point depression at the surface, 8 am (°C)
   b. (pos_tpGrad_0to78m) Positive atmospheric thermal gradient surface to 1000 mbar, 8 am (yes=1)

   Figs. A.12, A.13, and A.14 are generated.

3. Stata do file "upload_mss_weather.do" reads the Meteorological Service Singapore (MSS) source data files "source_weather_(station number)_yyyymm.csv," by weather station and calendar month, and prepares the following variables, among others:

a. (tp) Temperature at the surface, source MSS, daily mean (°C)
b. (dv_pp) Precipitation, daily total > 0 mm (yes=1)
c. (ws) Wind speed at the surface, daily mean (km h^{-1})

Figs. A.7, A.9, A.10, and A.11 are generated.

4. Stata do file "upload_nusg_weather.do" reads the NUS Geography (NUSG) weather station source data file "source_weather_nus201201to201512.xlsx," and prepares the following variables:
   a. (hm) Relative humidity, daily mean (%)
   b. (rd) Solar radiation (average of 7am to 6pm mean, W/m^2)

5. Stata do file "upload_nasa_firms.do" reads the NASA Fire Information for Resource Management System (FIRMS) source data file "source_fire_archive_M6_5540.csv", combines this with wind direction data from the temporary data file "(temporary) noaa_igra2_daily," and prepares the following variables:
   a. (idw_frp) Inverse-distance-weighted radiative power (MW km^{-1})
   b. (ddw_frp) Direction-difference-weight. radiative power (MW×1000)
   c. (idwddw_frp) Inv.-dist. × direction-diff. weigh. radiative power (MW km^{-1})

   Figs. A.17 and A.18 are generated.

6. Stata do file "upload_pollution_beijing_kolkata_dhaka.do" reads the US Department of State's source data files "source_usdos_pollution_beijing_2013_2014.dta," "source_usdos_pollution_kolkata_2016.csv," and "source_usdos_pollution_dhaka_2017.csv," and computes daily mean PM2.5 concentrations for Beijing, Kolkata, and Dhaka in the years indicated by the file titles. Fig. B.2 is generated.

7. Stata do file "upload_prices.do" reads the source data file "source_prices_monthly.csv," which contains prices in the economy, and prepares the following variables:
   a. (pe) Household electricity price, inclusive of a 7% General Services Tax and adjusted for changes in the Consumer Price Index (CPI base June 2014). The source for the price series (electricity tariff, low tension supplies, domestic) is the Energy Market Authority (EMA). The source for the CPI series is "M212191 - Consumer Price Index (CPI)," Singapore Department of Statistics. As shown in Fig. A.4, the price series is expressed in US$/kWh at 1.26 SG$ per 1 US$.
   b. (pg) Household natural gas price, inclusive of a 7% General Services Tax and adjusted for changes in the CPI. The source for the natural gas price series is City Gas Pte Ltd. As shown in Fig. A.4, the price series is expressed in US$/kWh at 1.26 SG$ per 1 US$.
   c. (laborcost) Inflation-adjusted overall unit labor cost in the economy. The source for the nominal labor cost index is "M183171 - Unit Labour Cost Index (Base Year 2010 = 100), Quarterly," Singapore Department of Statistics. I use the above CPI series to adjust for inflation.

d. (wages) Inflation-adjusted average wage in the economy. The source for the nominal wage series is "M182931 - Average Monthly Nominal Earnings Per Employee, Quarterly," Ministry of Manpower. I use the above CPI series to adjust for inflation.
e. (pg_us_impexp) As a proxy for world natural gas prices, I take the average between the US natural gas import price and the US natural gas export price. I convert this average price from US$ to SG$ using the nominal exchange rate, and then adjust for inflation in Singapore. The source for the SG$ to US$ exchange rate series is the Monetary Authority of Singapore's Financial Database. I use the above CPI series to adjust for inflation. The source for the US natural gas import price is the "Price of U.S. Natural Gas LNG Imports (Dollars per Thousand Cubic Feet)" from the US Energy Information Administration (EIA). The source for the US natural gas export price is the "Price of Liquefied U.S. Natural Gas Exports (Dollars per Thousand Cubic Feet)" from EIA.

8. Stata do file "upload_ema_industrial_electricity" reads the Energy Market Authority's (EMA) monthly industrial electricity consumption series, available in the source data file "source_ema_industrial_electricity.csv."

9. Stata do file "upload_hits_income.do" reads the 2008 Household Interview Travel Survey (HITS) source data file "source_hits_income.dta." It prepares an average income measure by residential local market, i.e., postal code by dwelling type. The market-level income measure is the proportion of households where at least one member states earning over a specified cutoff, computed by postal code by dwelling type. Fig. A.5 is generated.

10. Stata do file "home_energy_behaviors_survey.do" reads "Survey+of+Home+Energy+Behaviors_November+21%2C+2018_00.37.csv." The latter is the source data file containing responses to the survey that I implemented in fall 2018 on the Qualtrics online platform. Please see "Survey of Home Energy Behaviors.pdf." The do file replicates the findings reported in Table B.5, and generates Figs. 3 and B.4.

11. Stata do file "analysis_health_polyclinic_visits.do" reads "source_moh_polyclinic_visits.csv." The latter is the key source data file required to replicate the health responses documented in Appendix C. Specifically, the do file replicates the health regressions reported in Table C.1.

12. Stata do file "run_non_proprietary.do" sequentially runs the do files 1 to 11 above, which call only on data that are made available. The do file also calls on the available Stata data file "aggregated_usage_data_to_replicate_table_b_3.dta" and replicates the results reported in Table B.3, Estimation on data aggregated to the level of identifying variation.

13. (**will run only with proprietary data**) Stata do file "upload_spservices_microdata.do" prepares the household-level utility usage microdata and stores it in the temporary data file "(temporary) utilities_130k_monthly.dta," from the proprietary text files supplied by SP Services Ltd:
    a. source_RESTRICTED_microdata_consumption_201209to201312.csv

b. source_RESTRICTED_microdata_consumption_201401to201412.csv
c. source_RESTRICTED_microdata_consumption_201501to201512.csv
d. source_RESTRICTED_microdata_actual_or_estimated.csv
e. source_RESTRICTED_microdata_billing_period.csv

Upon reasonable request, these files are available on an NUS Department of Economics (or equivalent institutional) computer in order to run the do file. The following functions are uploaded to memory and then run in sequence (i) prepareUsage, (ii) prepareAorEreading, (iii) prepareBillingPeriod, and (iv) mergeData.

14. (**will run only with proprietary data**) Stata do file "combine_proprietary_data.do" combines the above household-level utility usage microdata, environmental variables, and other economic variables, preparing the electricity and natural gas estimation samples. The following functions are uploaded to memory (i) assignRegion, (ii) assignHolidays, (iii) actualUse, (iv) covariatesAvgUpToPast366d, (v) estimationSample, (vi) pollutionNextBillCycle, and (vii) aggregatedData. Functions (iii) to (vi) are then run in sequence; these functions call on functions (i) and (ii). See the descriptions to each function that are provided in the do file. The function assignHolidays calls on the Stata source data file "source_holidays.dta." Figs. A.1 and A.2 are generated in function actualUse. The function covariatesAvgUpToPast366d takes a while to run as it averages covariates over the same days that are concurrent to each usage observation, for all observations in the sample. For example, the 61 days closing on May 1, 2013 for some households, the 61 days up to May 2, 2013 for others, and so forth. The function aggregatedData prepares two aggregated samples, both of which are made available. One aggregated sample is used in Stata do file "run_non_proprietary.do" (point 12 above) to replicate the results reported in Table B.3, Estimation on data aggregated to the level of identifying variation. The other sample consists of electricity and natural gas use data aggregated to usage period by two-digit zip code by dwelling type triple, to allow approximate verification.

15. (**will run only with proprietary data**) Stata do file "run_proprietary.do" sequentially runs the do files 13 and 14 above. As such, it should be run only after running the do file "run_non_proprietary.do," described in point 12. The code opens the electricity or natural gas estimation samples and replicates the following results, in order:
    a. Table 2's electricity regressions
    b. Table 2's natural gas regressions
    c. Table 3's electricity regressions
    d. Figs. 4 and 5's heterogeneous electricity and natural gas responses by apartment type
    e. Table B.1's robustness tests and other robustness mentioned in the response document to reviewers (additional clustering strategies, permutation tests)
    f. Table B.4's robustness test, adding pollution in the next billing cycle as a regressor
    g. Fig. B.1 (also see file titled "Fig. B.1 showing 95 percent confidence intervals.xlsx" for 95% CI)
    h. Table 4 and Fig. B.3, household electricity demand under counterfactual PM2.5 pollution and heat
    i. Figs. 1 and A.6

j. Aggregate household consumption series for Fig. A.3, panel C