ReadMe for Albouy, Graf, Kellogg, Wolff, *JAERE* forthcoming

The file ClimateRegressionData_150327.dta contains all of the data used in the analysis in Albouy et al. Below is a description of the variables in the file, in the order that they appear. For more detail on these variables, please consult the paper and especially the data appendix. The file has 2057 observations, one for each Puma in the contiguous 48 states.

With the exception of the summary statistics table, all of the analysis for the paper is carried out in Matlab. The do-file that generates the summary statistics is SummaryStats_150327.do. The do-file DataToMatlab_Puma_150327.do exports data from ClimateRegressionData_150327.dta to a large number of Matlab-ready csv files. A folder named Inputs_PUMA is required to hold these files.

The main analysis file that is capable of running nearly every model in the paper is HetBin_150402.m. Beginning on row 228, there are a series of "toggles" that allow for different dependent variables, controls, functional forms, etc. This file calculates analytic standard errors for the homogenous preference parameter estimates but does not calculate any boostrapped standard errors.

Q1_D2_FE0_S7_99_150402.m calculates point estimates and bootstrapped standard errors for the reference case model in the paper (column I in tables 2 – 5). Similar files can be constructed for the other specifications.

VARIABLES IN ClimateRegressionData_150327.dta

- statefip: FIP code for the state
- PumaID: Puma code. This is the unique identifier.
- msa, msaname: metropolitan statistical area code and name
- Wage_orig: wage differential based on place of residence
- Wage: wage differential based on place of wage, commuting cost corrected
- Price: housing cost differential
- QOL_orig: QOL based on Wage_orig
- QOL_25_1: "traditional" QOL that over-weights wage differentials relative to price differentials
- QOL_GM: QOL that reduces the influence of wage differentials by 1/3
- QOL_Dahl: QOL with a Dahl migration adjustment
- QOL_fulladj: QOL adjusted for migration using Notowidigdo elasticities
- QOL_doubleadj: QOL adjusted for migration using doubled Notowidigdo elasticities
- QOL_noncol: QOL calculated only for individuals without a 4 year degree
- QOL_col: QOL calculated only for individuals with at least a 4 year degree
- QOL: the main QOL used in the paper
- Population: total population
- Population_Col: 4 year college and above population
- Population_NonCol: Less than 4 year college population

- Area: area in square miles
- PopDens: Population / Area
- PopDens_Weighted: population density weighted by population within each census tract in the puma
- Income: mean wage income
- age: mean age
- sch_*: percent of population that is a high school dropout, high school graduate, some college, 4 year college degree, graduate degree
- min_*: percent of population that is Hispanic, black, native American, Asian, or other
- puma_name_*: variants of the puma name
- Pop1990_Imputed: 1990 population imputed to 2000 Pumas using census tract data
- Pop2000_Imputed: 2000 population imputed to 2000 Pumas using census tract data
- PopChg19902000: percentage change in population from 1990 to 2000 using imputed data
- qol_popadj: adjustment to QOL based on 1990 to 2000 population change and Notowidigdo elasticities (QOL_fulladj = QOL + qol_popadj)
- elev_m: average elevation in meters
- slope_pct: average slope (percent)
- min_sea: minimum distance to sea
- mean_sea: average distance to sea
- on_sea: dummy for on the sea
- min_lake: minimum distance to Great Lake
- mean_ lake: average distance to Great Lake
- on_ lake: dummy for on the Great Lake
- HDD_Pres: HDD, 1970-1999
- CDD_Pres: CDD, 1970-1999
- bin1-bin222: 1970-1999 temperature bins based on average daily temperatures. bin1 is t<-50C. bin2 is t in [-50,-49.5). bin102 is [0,0.5). bin222 is t>=60C. 65F is equal to 18.3333C. bin138 is [18,18.5). Average is 18.25. This bin contains the 65F "bliss point"
- Tbin1 – Tbin222: 1970-1999 temperature bins based on daily high and low temperatures. We assume a within-day sinusoidal curve between the daily high and low.
- Pbin1-Pbin102: 1970-1999 precipitation bins based on daily precipitation. Pbin1 is p=0cm. Pbin2 is p in (0,0.1]. Pbin 102 is p>10cm.
- Precip1-Precip12: 1970-1999 average annualized precipitation by month of year, in inches
- DewPt1-DewPt12: 1970-1999 average dew point by month of year (degrees Fahrenheit)
- Temp1-Temp12: 1970-1999 average temperature (F) by month of year
- RelHum1-RelHum12: 1970-1999 average relative humidity (%) by month of year
- RHbin1-RHbin10: 1970-1999 average number of days in relative humidity bins. Bins are deciles of the overall relative humidity distribution
- DPbin1-DPbin10: 1970-1999 average number of days in dew point bins. Bins are deciles of the overall dew point distribution.
- Sun1-Sun12: 1970-1999 average relative humidity (%) by month of year. Data interpolated using all weather stations reporting sunshine
- Sun4S1-Sun4S12: 1970-1999 average relative humidity (%) by month of year. Data interpolated using the four nearest weather stations reporting sunshine

- DeltaA2Ensemble30Year_bin1 – bin222: A2 ensemble temperature bin changes from 1970-1999 to 2070 – 2099.
- DeltaA2Ensemble30YearCloudFrac1 – 12: A2 ensemble cloud percentage changes from 1970-1999 to 2070 – 2099. By month of year.
- DeltaA1F1Ensemble30Year_bin1 – bin222: A1FI temperature bin changes from 1970-1999 to 2070 – 2099.
- DeltaA1F1Ensemble30YearCloudFrac1 – 12: A1FI cloud percentage changes from 1970-1999 to 2070 – 2099. By month of year.
- Delta_HDD_A2Ensemble_30Year - Delta_CDD_A1F1_30Year: Changes in HDD and CDD from 1970-1999 to 2070 – 2099.
- DeltaA2Ensemble30YearPrecipT1 - DeltaA1F130YearTempAvg12: Changes in precipitation, relative humidity, dew point, and average temperature from 1970-1999 to 2070 – 2099 for both the A2 ensemble and A1FI. By month of year.
- BootstrapMSA: sequential codes that group msas
- BootstrapState: sequential codes that group states