Codebook

This codebook contains the definition of variables in all datasets used in Liao, Steven, Neil Malhotra, and Benjamin J. Newman. Forthcoming. "Local Economic Benefits Increase Positivity toward Foreigners." *Nature Human Behaviour*.

Datasets

- `cces-merge-rep.RData`: A respondent-year dataset that merges Cooperative Congressional Election Study (CCES, https://cces.gov.harvard.edu/) panel data and local characteristics. The data contains 108 variables and 28,500 observations.
  - `case_id`: CCES ID for each respondent
  - `weight`: CCES weight for each respondent
  - `zcta_10`: Respondent's residential ZIP Code Tabulation Area (ZCTA) in 2010
  - `zip_10`: Respondent's residential ZIP code in 2010
  - `zip_type`: Type of ZIP code
  - `year`: Year
  - `anticorrupt`: Year = 2013 or after?
    1 = Yes
    0 = No
  - `time`: Time period?
    1 = 2010
    2 = 2012
    3 = 2014
  - `time1`: Year = 2010?
    1 = Yes
    0 = No
  - `time2`: Year = 2012?
    1 = Yes
    0 = No
  - `time3`: Year = 2014?
    1 = Yes
    0 = No
  - `city`: Respondent's city in 2010
  - `county_fips`: Respondent's county FIPS in 2010
  - `county_name`: Respondent's county name in 2010
  - `cbsa`: Respondent's CBSA code in 2010
  - `cbsa_title`: Respondent's CBSA name in 2010
  - `state_abb`: Respondent's state abbreviation in 2010
  - `state_name`: Respondent's state name in 2010

- `path`: Grant legal status to all illegal immigrants who have held jobs and paid taxes for at least 3 years, and not been convicted of any felony crimes?
  1 = No
  0 = Yes
- `border`: Increase the number of border patrols on the US-Mexican border?
  1 = Yes
  0 = No
- `police`: Allow police to question anyone they think may be in the country illegally?
  1 = Yes
  0 = No
- `scale`: Anti-immigration attitudes, $((\texttt{path} + \texttt{border} + \texttt{police})/3) \times 100$
- `stay_10_12`: Respondent has the same ZIP code in 2010 and 2012?
  1 = Yes
  0 = No
- `stay_12_14`: Respondent has the same ZIP code in 2012 and 2014?
  1 = Yes
  0 = No
- `stay_10_14`: Respondent has the same ZIP code in 2010 and 2014?
  1 = Yes
  0 = No
- `birthyr`: Respondent's age, $(\text{Year of birth} - 1992)/(-73)$
- `educ`: What is the highest level of education you have completed?
  1 = No HS
  2 = HS graduate
  3 = Some college
  4 = 2-year college
  5 = 4-year college
  6 = Post-grad
- `employ`: Which of the following best describes your current employment status?
  1 = Full-time
  2 = Part-time
  3 = Temporarily laid off
  4 = Unemployed
  5 = Retired
  6 = Permanently disabled
  7 = Homemaker
  8 = Student
  9 = Other
- `faminc`: Thinking back over the last year, what was your family's annual income?
  1 = Less than $10,000
  2 = $10,000 – $19,999
  3 = $20,000 – $29,999

4 = $30,000 – $39,999
5 = $40,000 – $49,999
6 = $50,000 – $59,999
7 = $60,000 – $69,999
8 = $70,000 – $79,999
9 = $80,000 – $99,999
10 = $100,000 – $119,999
11 = $120,000 – $149,999
12 = $150,000 – $199,999
13 = $200,000 – $249,999
14 = $250,000 – $349,999
15 = $350,000 – $499,999
16 = $500,000 or more
17 = $150,000 or more
18 = $250,000 or more
97 = Prefer not to say
98 = Skipped
99 = Not Asked

- `income`: `faminc` with 97, 98, and 99 recoded as NAs.
- `immstat`: Which of these statements best describes you?
  1 = Immigrant Citizen
  2 = Immigrant non-citizen
  3 = First generation
  4 = Second generation
  5 = Third generation
- `race`: What racial or ethnic group best describes you?
  1 = White
  2 = Black
  3 = Hispanic
  4 = Asian
  5 = Native American
  6 = Mixed
  7 = Other
  8 = Middle Eastern
- `ideo`: Thinking about politics these days, how would you describe your own political viewpoint?
  1 = Very liberal
  2 = Liberal
  3 = Moderate
  4 = Conservative
  5 = Very Conservative
- `pid3`: Generally speaking, do you think of yourself as a...?
  1 = Democrat

2 = Independent
3 = Republican

- `white`: `race` = 1?
  1 = Yes
  0 = No
- `ownhome_10`: Home Ownership status?
  1 = Own
  2 = rent
  3 = Other
- `ownhome_12`: Home Ownership status?
  1 = Own
  2 = rent
  3 = Other
- `ownhome_14`: Home Ownership status?
  1 = Own
  2 = rent
  3 = Other
- `stable_homeowner`: `ownhome_12` = 1 & `ownhome_14` = 1?
  1 = Yes
  0 = No
- `value_per_year`: Average Zillow median value per square feet (all homes)
- `china_undergraduate`: Weighted number of Chinese international undergraduates in ZCTA
- `log_china_undergraduate`: Log Chinese international undergraduates in ZCTA, log(`china_undergraduate` + 1)
- `log_china_undergraduate_2012`: 2012 log Chinese international undergraduates in ZCTA
- `log_china_undergraduate_2010`: 2010 log Chinese international undergraduates in ZCTA
- `log_china_undergraduate_lead_1`: Log Chinese international undergraduates in ZCTA, $t + 1$
- `china_graduate`: Weighted number of Chinese international graduate students in ZCTA
- `log_china_graduate`: Log Chinese international graduate students in ZCTA, log(`china_graduate` + 1)
- `log_china_graduate_2012`: 2012 log Chinese international graduate students in ZCTA
- `log_china_graduate_2010`: 2010 log Chinese international graduate students in ZCTA
- `log_china_graduate_lead_1`: Log Chinese international graduate students in ZCTA, $t + 1$
- `india_undergraduate`: Weighted number of Indian international undergraduate students in ZCTA

- log_india_undergraduate: Log Indian international undergraduates in ZCTA, log(india_undergraduate + 1)
- log_india_undergraduate_2012: 2012 log Indian international undergraduates in ZCTA
- log_india_undergraduate_2010: 2010 log Indian international undergraduates in ZCTA
- log_india_undergraduate_lead_1: Log Indian international undergraduates in ZCTA, $t+1$
- enroll_zcta: Total individuals enrolled in college or above
- ln_enroll_zcta: Log total individuals enrolled in college or above, log(enroll_zcta + 1)
- share_enroll_zcta: The share of individuals in a ZCTA enrolled in college or above
- share_edu_ba_above: The share of individuals in a ZCTA who have a BA degree or above
- pop_zcta_tot: Total population in ZCTA
- pop_zcta_tot_thousand: Total population in ZCTA (thousands)
- aland_sqmi_zcta: Total land area (square miles) in ZCTA
- pop_density_zcta: Population density, pop_zcta_tot/aland_sqmi_zcta
- ln_pop_density_zcta: Log population density, log(pop_density_zcta + 0.001)
- pop_adult_zcta: Total adult population in ZCTA
- ln_pop_adult_zcta: Log total adult population in ZCTA, log(pop_adult_zcta + 1)
- share_pop_adult_zcta: The share of adult population in ZCTA, (pop_adult_zcta/pop_zcta_tot)
- pop_cn_zcta: Total Chinese population in ZCTA
- ln_pop_cn_zcta: Log total Chinese population in ZCTA, log(pop_cn_zcta + 1)
- share_pop_cn_zcta: The share of Chinese population in ZCTA, (pop_cn_zcta/pop_zcta_tot)
- pop_white_zcta: Total White population in ZCTA
- ln_pop_white_zcta: Log total White population in ZCTA, log(pop_white_zcta + 1)
- share_pop_white_zcta: The share of White population in ZCTA, (pop_white_zcta/pop_zcta_tot)
- employ_zcta: The share of employed in ZCTA
- employ_zcta_100: The percentage of employed in ZCTA

- `unemploy_zcta`: The share of unemployed in ZCTA
- `median_household_income_zcta`: The median household income ($10,000)
- `median_tax_zcta`: The median real estate tax in ZCTA
- `median_home_value_zcta`: The median home value in ZCTA (U.S. Census, self-reported)
- `effective_tax_zcta`: The effective real estate tax rate (%), (`median_tax_zcta` / `median_home_value_zcta` $\times$ 100)
- `vacancy_zcta`: The share of houses vacant in ZCTA
- `n_establish_zcta`: Total number of establishments (U.S. Census)
- `suitabilitynext`: Log Chinese Undergraduates in ZCTA (Lead)
- `china_inter`: Log Chinese Undergraduates in ZCTA $\times$ Post-anti-corruption Campaign (2014)
- `china_inter_placebo`: Log Chinese Undergraduates in ZCTA $\times$ Pre-anti-corruption Campaign (2012)
- `india_inter`: Log Indian Undergraduates in ZCTA $\times$ Post-anti-corruption Campaign (2014)
- `chinagraduate_inter`: Log Chinese Graduate Students in ZCTA $\times$ Post-anti-corruption Campaign (2014)
- `suitabilitynext_inter`: Log Chinese Undergraduates in ZCTA (Lead) $\times$ Pre-anti-corruption Campaign (2012)
- `suitability2012_inter`: Log Chinese Undergraduates in ZCTA (2012) $\times$ Post-anti-corruption Campaign (2014)
- `suitability2012_inter_placebo`: Log Chinese Undergraduates in ZCTA (2012) $\times$ Pre-anti-corruption Campaign (2012)
- `suitability2010_inter`: Log Chinese Undergraduates in ZCTA (2010) $\times$ Pre-anti-corruption Campaign (2012)
- `suitability2012_india_inter`: Log Indian Undergraduates in ZCTA (2012) $\times$ Post-anti-corruption Campaign (2014)
- `suitability2012_china_grad_inter`: Log Chinese Graduate Students in ZCTA (2012) $\times$ Post-anti-corruption Campaign (2014)
- `log_china_undergraduate_stable`: Log Chinese Undergraduates in ZCTA $\times$ Stable Homeowner
- `time3_stable_homeowner`: Post-anti-corruption Campaign (2014) $\times$ Stable Homeowner
- `homeowner_threeway`: Log Chinese Undergraduates in ZCTA $\times$ Post-anti-corruption Campaign (2014) $\times$ Stable Homeowner
- `log_china_undergraduate_birthyr`: Log Chinese Undergraduates in ZCTA $\times$ Age
- `time3_birthyr`: Post-anti-corruption Campaign (2014) $\times$ Age

  - `birthyr_threeway`: Log Chinese Undergraduates in ZCTA × Post-anti-corruption Campaign (2014) × Age
  - `log_china_undergraduate_educ`: Log Chinese Undergraduates in ZCTA × Education
  - `time3_educ`: Post-anti-corruption Campaign (2014) × Education
  - `educ_threeway`: Log Chinese Undergraduates in ZCTA × Post-anti-corruption Campaign (2014) × Education
  - `log_china_undergraduate_white`: Log Chinese Undergraduates in ZCTA × White
  - `time3_white`: Post-anti-corruption Campaign (2014) × White
  - `white_threeway`: Log Chinese Undergraduates in ZCTA × Post-anti-corruption Campaign (2014) × White
  - `log_china_undergraduate_income`: Log Chinese Undergraduates in ZCTA × Income
  - `time3_income`: Post-anti-corruption Campaign (2014) × Income
  - `income_threeway`: Log Chinese Undergraduates in ZCTA × Post-anti-corruption Campaign (2014) × Income

- `pew-merge.RData`: A respondent-year dataset that merges Pew (https://www.pewresearch.org/download-datasets/) cross-sectional data and local characteristics. The data contains 61 variables and 12,492 observations.

  - `surveyid`: Pew survey ID
  - `year`: Year of survey
  - `month`: Month of survey
  - `psraid`: Respondent's ID in cross-section survey
  - `weight`: Respondent's weight in cross-section survey
  - `zipcode`: Respondent's ZIP code
  - `zcta`: Respondent's ZIP Code Tabulation Area (ZCTA)
  - `city`: Respondent's city
  - `county_fips`: Respondent's county FIPS
  - `county_name`: Respondent's county name
  - `cbsa`: Respondent's CBSA code
  - `cbsa_title`: Respondent's CBSA name
  - `state_abb`: Respondent's state abbreviation
  - `state_name`: Respondent's state name
  - `cnthreat`: Respondent's perception of a "China Threat"
    3 = Major threat
    2 = Minor threat
    1 = Not a threat

- `russiathreat`: Respondent's perception of a "Russia Threat"
  3 = Major threat
  2 = Minor threat
  1 = Not a threat
- `iranthreat`: Respondent's perception of an "Iran Threat"
  3 = Major threat
  2 = Minor threat
  1 = Not a threat
- `koreathreat`: Respondent's perception of a "North Korea Threat"
  3 = Major threat
  2 = Minor threat
  1 = Not a threat
- `male`: Male Respondent?
  1 = Yes,
  0 = No
- `education`: What is the highest level of school you have completed or the highest degree you have received?
  1 = Less than high school (Grades 1–8 or no formal schooling)
  2 = High school incomplete (Grades 9–11 or Grade 12 with NO diploma)
  3 = High school graduate (Grade 12 with diploma or GED certificate)
  4 = Some college, no degree (includes community college)
  5 = Two year associate degree from a college or university
  6 = Four year college or university degree/Bachelor's degree (e.g., BS, BA, AB)
  7 = Some postgraduate or professional schooling, no postgraduate degree (e.g. some graduate school)
  8 = Postgraduate or professional degree, including master's, doctorate, medical or law degree (e.g., MA, MS, PhD, MD, JD, graduate school)
- `income`: Last year, that is in..., what was your total family income from all sources, before taxes?
  1 = Less than $10,000
  2 = 10 to under $20,000
  3 = 20 to under $30,000
  4 = 30 to under $40,000
  5 = 40 to under $50,000
  6 = 50 to under $75,000
  7 = 75 to under $100,000
  8 = 100 to under $150,000
  9 = $150,000 or more
- `age`: Respondent's age
- `latino`: Respondent's race is Latino
  1 = Yes
  0 = No

- `white`: Respondent's race is White
  1 = Yes
  0 = No
- `black`: Respondent's race is Black
  1 = Yes
  0 = No
- `asian`: Respondent's race is Asian
  1 = Yes
  0 = No
- `otherdk`: Respondent's race is Other/Don't Know
  1 = Yes
  0 = No
- `pid5`: Respondent's party identification?
  1 = Strong Democrat
  2 = Moderate Democrat
  3 = Independent
  4 = Moderate Republican
  5 = Strong Republican
- `ideology`: Respondent's ideology?
  1 = Very Conservative
  2 = Conservative
  3 = Moderate
  4 = Liberal
  5 = Very liberal
- `value_per_year`: Average Zillow median value per square feet (all homes)
- `anticorrupt`: Year = 2013 or after?
  1 = Yes
  0 = No
- `china_undergraduate`: Weighted number of Chinese international undergraduates in ZCTA
- `log_china_undergraduate`: Log Chinese international undergraduates in ZCTA, log(`china_undergraduate` + 1)
- `china_graduate`: Weighted number of Chinese international graduate students in ZCTA
- `log_china_graduate`: Log Chinese international graduate students in ZCTA, log(`china_graduate` + 1)
- `india_undergraduate`: Weighted number of Indian international undergraduate students in ZCTA
- `log_india_undergraduate`: Log Indian international undergraduates in ZCTA, log(`india_undergraduate` + 1)
- `enroll_zcta`: Total individuals enrolled in college or above

  - `ln_enroll_zcta`: Log total individuals enrolled in college or above, log(`enroll_zcta` + 1)
  - `share_enroll_zcta`: The share of individuals in a ZCTA enrolled in college or above
  - `share_edu_ba_above`: The share of individuals in a ZCTA who have a BA degree or above
  - `pop_zcta_tot`: Total population in ZCTA
  - `aland_sqmi_zcta`: Total land area (square miles) in ZCTA
  - `pop_density_zcta`: Population density, `pop_zcta_tot`/`aland_sqmi_zcta`
  - `ln_pop_density_zcta`: Log population density, log(`pop_density_zcta` + 0.001)
  - `pop_adult_zcta`: Total adult population in ZCTA
  - `ln_pop_adult_zcta`: Log total adult population in ZCTA, log(`pop_adult_zcta` + 1)
  - `share_pop_adult_zcta`: The share of adult population in ZCTA, (`pop_adult_zcta`/`pop_zcta_tot`)
  - `pop_cn_zcta`: Total Chinese population in ZCTA
  - `ln_pop_cn_zcta`: Log total Chinese population in ZCTA, log(`pop_cn_zcta` + 1)
  - `share_pop_cn_zcta`: The share of Chinese population in ZCTA, (`pop_cn_zcta`/`pop_zcta_tot`)
  - `pop_white_zcta`: Total White population in ZCTA
  - `ln_pop_white_zcta`: Log total White population in ZCTA, log(`pop_white_zcta` + 1)
  - `share_pop_white_zcta`: The share of White population in ZCTA, (`pop_white_zcta`/`pop_zcta_tot`)
  - `employ_zcta`: The share of employed in ZCTA
  - `unemploy_zcta`: The share of unemployed in ZCTA
  - `median_household_income_zcta`: The median household income ($10,000)
  - `median_tax_zcta`: The median real estate tax in ZCTA
  - `median_home_value_zcta`: The median home value in ZCTA (U.S. Census, self-reported)
  - `effective_tax_zcta`: The effective real estate tax rate (%), (`median_tax_zcta` / `median_home_value_zcta` $\times$ 100)
  - `vacancy_zcta`: The share of houses vacant in ZCTA

- `price-merge.RData`: A ZCTA-year dataset that merges Zillow data (https://www.zillow.com/research/data/) on home prices and local characteristics. The dataset contains 48 variables and 163,285 observations.

- `zcta`: ZIP Code Tabulation Area (ZCTA)
- `year`: Year
- `anticorrupt`: Year = 2013 or after?
  1 = Yes
  0 = No
- `anticorrupt_lead_1`: Year = 2012 or after?
  1 = Yes
  0 = No
- `county_fips`: County FIPS
- `county_name`: County name
- `cbsa`: CBSA code
- `cbsa_title`: CBSA name
- `state_abb`: State abbreviation
- `state_name`: State name
- `value_per_year`: Average Zillow median value per square feet in ZCTA
- `ln_value_per_year`: Log average Zillow median value per square feet in ZCTA
- `china_undergraduate`: Weighted number of Chinese international undergraduates in ZCTA
- `china_undergraduate_2012`: 2012 weighted number of Chinese international undergraduates in ZCTA
- `ln_china_undergraduate`: Log Chinese international undergraduates in ZCTA, log(`china_undergraduate` + 1)
- `ln_china_ug_2012`: 2012 log Chinese international undergraduates in ZCTA
- `china_graduate`: Weighted number of Chinese international graduate students in ZCTA
- `china_graduate_2012`: 2012 weighted number of Chinese international graduate students in ZCTA
- `ln_china_graduate`: Log Chinese international graduate students in ZCTA, log(`china_graduate` + 1)
- `ln_china_g_2012`: 2012 log Chinese international graduate students in ZCTA
- `india_undergraduate`: Weighted number of Indian international undergraduate students in ZCTA
- `india_undergraduate_2012`: 2012 weighted number of Indian international undergraduate students in ZCTA
- `ln_india_undergraduate`: Log Indian international undergraduates in ZCTA, log(`india_undergraduate` + 1)
- `ln_india_ug_2012`: 2012 log Indian international undergraduates in ZCTA
- `enroll_zcta`: Total individuals enrolled in college or above

- `ln_enroll_zcta`: Log total individuals enrolled in college or above, log(`enroll_zcta` + 1)
- `share_enroll_zcta`: The share of individuals in a ZCTA enrolled in college or above
- `share_edu_ba_above`: The share of individuals in a ZCTA who have a BA degree or above
- `pop_zcta_tot`: Total population in ZCTA
- `aland_sqmi_zcta`: Total land area (square miles) in ZCTA
- `pop_density_zcta`: Population density, `pop_zcta_tot`/`aland_sqmi_zcta`
- `ln_pop_density_zcta`: Log population density, log(`pop_density_zcta` + 0.001)
- `pop_adult_zcta`: Total adult population in ZCTA
- `ln_pop_adult_zcta`: Log total adult population in ZCTA, log(`pop_adult_zcta` + 1)
- `share_pop_adult_zcta`: The share of adult population in ZCTA, (`pop_adult_zcta`/`pop_zcta_tot`)
- `pop_cn_zcta`: Total Chinese population in ZCTA
- `ln_pop_cn_zcta`: Log total Chinese population in ZCTA, log(`pop_cn_zcta` + 1)
- `share_pop_cn_zcta`: The share of Chinese population in ZCTA, (`pop_cn_zcta`/`pop_zcta_tot`)
- `pop_white_zcta`: Total White population in ZCTA
- `ln_pop_white_zcta`: Log total White population in ZCTA, log(`pop_white_zcta` + 1)
- `share_pop_white_zcta`: The share of White population in ZCTA, (`pop_white_zcta`/`pop_zcta_tot`)
- `employ_zcta`: The share of employed in ZCTA
- `unemploy_zcta`: The share of unemployed in ZCTA
- `median_household_income_zcta`: The median household income ($10,000)
- `median_tax_zcta`: The median real estate tax in ZCTA
- `median_home_value_zcta`: The median home value in ZCTA (U.S. Census, self-reported)
- `effective_tax_zcta`: The effective real estate tax rate (%), (`median_tax_zcta` / `median_home_value_zcta` $\times$ 100)
- `vacancy_zcta`: The share of houses vacant in ZCTA

- `cmps-merge.RData`: A respondent-year dataset that merges Collaborative Multiracial Post-Election Surveys (CMPS, https://cmpsurvey.org/) cross-sectional data and local characteristics. The dataset contains 12 variables and 7,154 observations.

- `year`: Year
- `zcta`: Respondent's ZIP Code Tabulation Area (ZCTA)
- `asian_neighborhood_tri`: Respondent's perceived neighborhood prevalence of Asians?
  2 = Entirely
  1 = Mostly
  0 = Partially
- `log_china_undergraduate`: Log Chinese international undergraduates in ZCTA
- `homeowner`: Own home?
  1 = Yes
  0 = No
- `age`: Respondent's age
- `edu`: Respondent's highest level of education?
  1 = Grades 1–8
  2 = Some High School
  3 = High School graduate
  4 = Some College/ Technical School
  5 = College graduate
  6 = Post-graduate education
- `white`: Race = White?
  1 = Yes
  0 = No
- `income`: Respondent's household income, categorical:
  1 = Less than $20,000
  2 = $20,000 to $39,999
  3 = $40,000 to $59,999
  4 = $60,000 to $79,999
  5 = $80,000 to $99,999
  6 = $100,000 to $150,000
  7 = More than $150,000
- `median_household_income_zcta`: The median household income ($10,000) in a ZCTA
- `ln_pop_density_zcta`: Log population density in a ZCTA
- `share_edu_ba_above`: The share of individuals in a ZCTA who have a BA degree or above

- `foia-aggregate.RData`: Aggregated Freedom of Information Act (FOIA) data on international students in the United States by origin country, level of education, and year. The dataset contains 5 variables and 16,252 observations
  - `iso3n`: Origin country ISO numeric code
  - `year`: Year
  - `student_level`: Student's academic level? Secondary, Undergraduate, Graduate, Total

  - `n`: Total number of students
  - `cty`: Origin country name

- `nar-df.Rdata`: A country-year dataset on foreign acquisitions of U.S. residential real estate between 2010 and 2015 from the National Association of Realtors (NAR, https://www.nar.realtor/sites/default/files/reports/2016/2016-profile-of-international-home-buying-activity-06-06-2016.pdf). The dataset contains 5 variables and 35 observations.
  - `cty`: Origin country
  - `year`: Year
  - `sale`: Total dollar volume of residential real estate sales to foreign buyer (billions)
  - `sale_total`: Total dollar volume of residential real estate sales to top five buyers (billions)
  - `share`: Share of sales among top five buyers

- `did-vehicle.RData`: A list object that contains difference-in-differences regression results analyzed and saved by the authors for producing Table S11. The list includes 4 elements.
  - $ `did.vehicle.a`: plm regression output for Column 1 of Table S11
  - $ `did.vehicle.b`: plm regression output for Column 2 of Table S11
  - $ `did.vehicle.c`: plm regression output for Column 3 of Table S11
  - $ `did.vehicle.c.lfe`: lfe regression output for simulations