Author: Kien Hoang-Le

University of Nottingham

Created date: 20 July 2023

Description: Documents record all the variables which are used in my dissertation, and can be found in the data file *reg\_panel.dta.*

--------------------------------------------------------

The panel data includes 3 main categories of variables: (1) mortgage variables from the HMDA dataset, (2) toxic releases variables from the TRI dataset and (3) census and macroeconomic variables on the county-level from multiple sources.

The following table provides details:

|  |  |
| --- | --- |
| Variables from Toxic Release Inventory dataset | |
| **Variables name** | **Sources and description** |
| FRS ID | Facility ID made by TRI |
| State, County, FIPS | Geographical Identifiers |
| Latitude, Longitude | Geographical Identifiers |
| Carcinogen | Binary variable, indicating the type of chemical substance observed at a specific data point is carcinogen or not |
| Classification | 3-level variables. Indicates if the chemical is classified as a dioxin or dioxin-like compound, a Persistent Bioaccumulative and Toxic chemical, or a general EPCRA Section 313 chemical.  Values: {TRI, PBT, DIOXIN} where:   * TRI = General EPCRA Section 313 Chemical * PBT = Persistent Bioaccumulative and Toxic * DIOXIN = Dioxin or Dioxin-like compound |
| Fugitive Air | An estimate of the total quantity of the toxic chemical released as fugitive air emissions at the reporting facility. |
| Stack Air | An estimate of the total quantity of the toxic chemical released as fugitive air emissions at the reporting facility. |
| Total Release | Total onsite release and total offsite release |

|  |  |
| --- | --- |
| Variables from HMDA dataset | |
| **Variables name** | **Sources and description** |
| LEI | Lender’s Identification |
| Rate Spread | The difference between the covered loan’s annual percentage rate (APR) and the average prime offer rate (APOR) for a comparable transaction as of the date the interest rate is set |
| U.S. 30-year Treasury yield spread | The difference between the covered loan’s annual percentage rate (APR) and average 30-Year US treasury bond yield of that year |
| Applicant Race | Race of the first applicant or borrower. Aggregated based on the HMDA data:  1: White  2: Asian  3: Black and African American  4: Native Americans |
| Applicant Age | Age of the first applicant or borrower:  7 age bins, from under 25 to over 74 |
| Loan-to-Value ratio | The ratio of the total amount of debt secured by the property to the value of the property relied on in making the credit decision.  Divided into decile bins |
| Property Value  Property | The value of the property securing the covered loan or, in the case of an application, proposed to secure the covered loan, relied on in making the credit decision. |
| Income | The gross annual income, in thousands of dollars, relied on in making the credit decision, or if a credit decision was not made, the gross annual income relied on in processing the application. Divided into decile bins |
| Loan Purpose | Purchase a new home or refinance  0 for new purchase  1 for refinance |
| Regulatory Agency | The integer code corresponding to an institution's regulatory agency  1 OCC Office of the Comptroller of the Currency  2 FRB Federal Reserve System  3 FDIC Federal Deposit Insurance Corporation  5 NCUA National Credit Union Administration  7 HUD Department of Housing and Urban Development  9 CFPB Consumer Financial Protection Bureau |

|  |  |
| --- | --- |
| Variables from US Census data Source: **US Census Bureau** and **U.S Department of Labour, Bureau of Labour Statistics** | |
| **Variables name** | **Sources and description** |
| County Land Area | 2020 land area of the County (square meters) |
| County Housing Density | 2020 housing unit density of the County (square miles) |
| County Unemployment Rate | County’s unemployment rate in each year |
| Census Urban/Rural | Define whether a census tract is urban or rural |