# Data Appendix

Devynck Tom, Druilhe Théo, Fouquet Damien

February 2024

Each year thousands of financial institutions report data about mortgages to the public, under the Home Mortgage Disclosure Act (HMDA). We use here the data collected under HMDA in Michigan in 2022. For mortgage applications, the data will tell you if the consumer got the mortgage—look for applications that were "originated"—or if the consumer was denied, didn't complete the application, or something else happened. The data also have information about pre-approvals and loans sold from one institution to another. For each record, you can learn about the loan, the property characteristics, the applicant demographics, and the lender. The data are modified to protect applicant and borrower privacy.

The initial database presented 470k observations, after data-cleaning, we kept 243k observations. The representivity of the final dataset has been tested.

Table 1: Specifications of thedata collected under HMDA, Michigan 2022.

| Variable | Label | Type | Modality |
|---|---|---|---|
| derived_race | Race of the applicant | Categorical | White Black, Asian and Native |
| loan_purpose | Purpose of the loan | Categorical | 1 for Home purchase |
| | | | 2 for Home improvement |
| | | | 31 for Refinancing |
| | | | 32 for cash-out Refinancing |
| | | | 4 for Other Purpose |
| loan_amount | Amount of the loan | Numerical | |
| loan_to_value_ratio | ratio of the total amount of debt secured by the property to the value of the property | Numerical | |
| interest_rate | interest rate for the covered | Numerical | |
| loan_term | number of months after which the legal loan or application obligation will mature | Numerical | |
| property_value | value of the property in dollars | Numerical | |
| income | Income of the applicant in dollars | Numerical | |
| applicant_credit_score_type | name and version of the credit scoring model used to generate the credit score | | |
| Categorical | 1 for Equifax Beacon 5.0 | | |
| | | | 2 for Experian Fair Isaac |
| | | | 3 for FICO Risk Score Classic 04 |
| | | | 4 for FICO Risk Score Classic 98 |
| | | | 5 for VantageScore 2.0 |
| | | | 6 for VantageScore 3.0 |
| | | | 7 for More than one credit scoring model |
| | | | 8 for Other credit scoring model |
| | | | 9 for Not applicable |
| | | | 11 for Exempt |
| applicant_sex | Sex of the applicant | Categorical | 1 for Male |
| | | | 2 for Female |
| applicant_age | Age of the applicant | Categorical | <25 |
| | | | 25-34 |
| | | | 35-44 |
| | | | 45-54 |
| | | | 55-64 |
| | | | 65-74 |
| | | | >74 |
| deny | Has the application been deny? | Dummy | |
| same_sex | Is the co-applicant of the same sex of the applicant? | Dummy | |