Readme File

This is the Readme file for the replication package for "Racial Disparities in the Auto Loan Market" by Alex Butler, Erik Mayer, and James Weston. See https://doi.org/10.1093/rfs/hhac029.

For any questions regarding the code, reach out to Erik Mayer at emayer@smu.edu.

The credit bureau data used in this paper are confidential and cannot be shared. However, in this replication package, we provide code to generate synthetic data sets with the same structure as the actual data. We then provide the code to conduct the empirical analysis and produce tables using the synthetic data. The code should be run in order, moving from the file prefixed with _0_ to _1_ and so on.

Code:
_0_GenerateSyntheticData.do creates the synthetic credit bureau data sets.
_1_AnalyzeMatching.do produces Tables 1 and 2.
_2_CreditAccessAnalysis.do produces Tables 3-6, and Figures 1 and 2.
_3_AutoLoanAnalysis.do produces Tables 7-9.
_4_CFPBDiD.do produces Table 10 and Figure 3.

Data:
Data_AutoLoans.dta is a synthetic data set of credit bureau auto loans.
Data_CBAllPersonYears.dta is a synthetic data set that stands in for the full 1% credit bureau panel.
Data_CBMortgages.dta is a synthetic data set of credit bureau mortgages.
Data_HmdaMortgages.dta is an actual data set of mortgages from HMDA (which is public).
Data_MainMatchedPanel.dta is a synthetic data set that stands in for the credit bureau/HMDA matched panel.
Data_StateEstimates.dta contains the actual state point estimates used to make Figures 1 and 2.
Data_YearEstimates_apr.dta and Data_YearEstimates_credit.dta contain the actual yearly point estimates used to make Figure 3.
Statemap50_coord.dta and Statemap50_fipslink.dta are tool data sets used to make maps in STATA.

Subfolders:
.\TablesAndFigures\ is where the table and figure output is stored.

Note that, because the credit bureau data sets used in this replication package must be synthetic (i.e., randomly generated), the results in the tables produced will obviously not match those in the paper. The only exception is Table 2, which uses the public HMDA data and will replicate directly. Because the state-level and year-level point estimates used in the figures are reported in the paper (and can hence be shared here as datasets), the figures will also replicate directly.