The file “Assessment Gap_Main Paper Run” produces estimates, figures, and tables for “The Assessment Gap: Racial Inequalities in Property Taxation.” This code calls several ancillary functions for efficiency; these are provided in the subfolder “Ancillary Functions.”

This code runs over the core dataset, compiled as described in the paper and the Online Appendix. The dataset uses proprietary, commercially available data from both ATTOM (property records) and Atlas Investment Research (government shapefiles), as well as administrative data originating from the Cook County Assessor’s Office.

A large amount of additional detail is available in both the paper and the Online Appendix. At a high level, replication of the core dataset would involve the following steps:

1) Obtain property-level records of assessment and transactions from ATTOM or another source
2) Obtain shapefiles for government boundaries from Atlas Investment Research or another source
3) At the property-level, match assessments for tax year *t* to transactions occurring in tax year *t*
4) Using geolocation detail, associate each property with the set of governments encompassing it; and then group properties taxed and served by the same unique set of overlapping governments together as a taxing jurisdiction
5) Associate each observed transaction with an observation in HMDA records, to connect transactions to race and ethnicity of home buyer (or refinancing homeowner)
6) Exploit the time-series structure of the data to associate observed assessment ratios with race and ethnicity of the home *seller*
7) Using property geolocation detail, merge in desired demographic and geographical variables (from ACS/Census or other) by state-year, county-year, ZIP-code-year, tract-year or block-group-year as appropriate

The resulting dataset would contain property-level observations of assessment ratios, each associated with an identifier for taxing jurisdiction, and the race and ethnicity of the home seller, plus a range of geographic covariates.

The “Main Paper Run” code applied to such a dataset, will produce the analyses used in our paper. Column (1) of Table 1 tests balance on the extensive margin of whether a transaction is observed. This requires sampling unsold properties from the raw ATTOM files. Thereafter, the balance test parallels exactly the approach reflected in the “HMDA Merge Analysis” block. The analysis of appeals outcomes in Cook County further imports property-level records of appeals from the Cook County Assessor’s Office, which are merged to ATTOM records based on a property level PIN that appears in both datasets.