## Data and Methods

### Data

Our data set is the product of four individual data sources. To locate street barriers, we identified an initial list of street barriers from a Washington University in St. Louis) term paper that examined the barriers phenomenon [(Waldron 2010)](https://www.zotero.org/google-docs/?oCk1q7). From this list, we used the City’s orthoimagery, Google Street View, and fieldwork visits to identify which barriers are still in place. Checking against these additional sources resulted in a publicly available data set of point data describing the locations of barriers,[[1]](#endnote-1) which in turn are divided into two groups: barriers known to remain in existence (n=280) and those that were removed by the City at some past date (n=55).

Second, publicly available address-level crime data were obtained from the St. Louis Police Department. Third, census tract-level data from the 2016 American Community Survey 5-year estimates were procured. These data were used to construct explanatory variables for modeling the locations of both barriers and violent crimes. Fourth, to measure population change over time, we utilized census tract data from previous decennial censuses. Census tract boundaries often shift from census to census, making it difficult to measure temporal variation at the census tract level. In order to achieve a stable unit of analysis over time, we used the R-based package ‘areal’ (Prener and Revord 2019) to spatially interpolate tract-level data into one-square kilometer grid squares, estimating grid square-level data based on the census tracts that intersected with each grid square and the size of those intersections. The creation of the grid "fishnet" yielded a dataset with 205 grid squares with corresponding demographic attributes as well as a count of barriers known to still exist and violent crimes that occurred within each grid. This analytical data set is also available for public download as part of this manuscript's Open Science Framework repository.[[2]](#endnote-2)

### Methods

We designed this study to address two primary goals. This first goal was to determine whether barriers are contemporarily effective as a crime prevention tool. To this end, we employed regression techniques designed for modeling count variables, in this case the occurrence of violent crime per grid square in 2016, the most recent year for which we have accurate barrier location data. However, spatial autocorrelation, a process that refers to the spatial clustering of a given variable, is a known issue in using generalized linear models to model the location of violent crime incidents (Helbich & Arsanjani 2015). Following Helbich and Arsanjani (2015), we adopt the spatial eigenvector filtering technique to control for spatial autocorrelation.

An important factor when deploying spatial eigenvector filtering is choosing the best method to conceptualize spatial relationships. Following similar studies (Helbich and Arsanjani 2015), we chose the C-coding or globally standardized spatial weight, which places greater emphasis over areal units with higher spatial linkages, as opposed to the W-coding or row standardized spatial weight, which is popular due to its intuitive spillover-effect interpretation but gives more emphasis to areal units on the edge of the study area, with fewer linkages (Patuelli et al. 2012).

The second goal was to ascertain the social correlates of barrier location. We accomplish this by estimating models with social data from two periods in time. First, to assess the social correlates of barrier location in a period the preceded the vast majority of barrier installations, we regressed barrier counts on 1980 census data. Second, to assess contemporary social correlates of barrier location, we regressed barrier counts on 2016 ACS data. Because a majority of the grid squares have zero values (N=138), we employed zero-inflated models for this set of analyses, assuming a fundamentally different relationship between social variables and barrier location between grid squares with zero barriers and grid squares with one or more barriers. Accordingly, zero-inflated models produce both a model to estimate the correlates of grid squares with count data and a model to estimate the correlates of grid squares with zero values.

Independent variables

1. *Link to repository removed for peer review* [↑](#endnote-ref-1)
2. *Link to repository removed for peer review* [↑](#endnote-ref-2)