**Data Appendix: Construction of ethnic enclave measures**

We use data from the 1990 and 2000 5% file of the U.S. Censuses and the 2005 to 2011 American Community Surveys (ACS) to construct our ethnic enclave measures. Both data sets were extracted from the Integrated Public Use Microdata Series ([Ruggles et al., 2010](#_ENREF_67)). The Census is fielded every ten years to enumerate and collect demographic information on the population. The 1990 and 2000 5% files of the Census include approximately 12.5 million and 14 million individuals respectively. The ACS is a large-scale nationally representative survey which is designed to replace the decennial long-form Census. Each ACS provides data for over three million individuals.

Both the Census and ACS contain detailed information on country of origin. We limit the sample to the top 50 sending countries in the 1990 Census as sample sizes for particular countries become small outside this range. These countries comprise over 86% of immigrants in the TUS. We delete observations in which the specific birth country cannot be determined. Examples include Africa (not specified) and Americas (not specified).

We construct our proxy for ethnic enclaves as the proportion of individuals from the same birth country in the metropolitan statistical area (MSA) of residence. MSA information is available for roughly 70% of the TUS, and we exclude respondents with missing MSA information. For example, if a respondent reports that he was born in Ireland and resides in the New York City MSA, we assign him the following value: .

Because we do not have full overlap between the TUS data, and the Census and ACS surveys we must impute our enclave measure for non-Census and non-ACS years (i.e., 1995, 1996, 1998, and 1999). Moreover, not all years of our Census (recall that we rely on the 5% sample) or ACS surveys contain the universe of MSAs in the U.S. Thus, we must impute enclave measures for missing MSAs. Since MSAs are at a much finer geographic level than the state, many of the birth country-by-MSA-by-year cells are small or empty, however. In total, we must impute information for 57% of our MSA/birth country/year cells. To provide enclave values for all MSAs in all years we utilize prediction equations that regress the percent of residents from each of the 50 included birth countries on birth country, MSA, and year fixed effects, and birth country by state fixed effects using least squares. The adjusted from this regression is 0.98, suggesting that our prediction model has reasonably good fit.

The MSA coding in the TUS (and the CPS more broadly) changed substantially between the 2003 and 2006/07 supplements (specifically, the CPS switched from using the Office of Management and Budget’s June 1993 to June 2003 geographic area identification system). We match MSAs between the two classification systems, and achieve a match rate of 95%. More details on our matching procedure are available on request.