2010

Using School-Age Populations to Identify Hard-to-Count Populations: A Report to the Secretary of the Commonwealth

Henry C. Renski
University of Massachusetts - Amherst, hrenski@larp.umass.edu

Susan Strate
Donahue Institute, sstrate@donahue.umassp.edu

John Gaviglio
Donahue Institute, jgaviglio@donahue.umassp.edu

Sonya Smith

Bill Proulx

UMASS DONAHUE INSTITUTE

UMass Donahue Institute Population Estimates Program

Using School-Age Populations to Identify Hard-to-Count Populations

A Report to the Secretary of the Commonwealth

January 2010

Report Contributors
Henry Renski, PhD, Principle Investigator
Susan Strate, Program Manager
John Gaviglio, State Data Center Manager
Sonya Smith, Research Assistant
Bill Proulx, Research Analyst

Contents

Background

At the request of the Secretary of the Commonwealth, the UMASS Donahue Institute conducted an investigation of alternate data sources and methods for identifying areas with significant concentrations of historically Hard-to-Count (HTC) populations so that they may be fully counted in the upcoming Decennial Census of 2010. Historically hard-to-count populations typically include minority groups, such as African-Americans and Hispanics, low-income households, and recent international immigrant populations.

Identifying HTC populations can be challenging, namely because of the lack of detailed demographic data for small areas. At present, the Decennial Census of Population of 2000 is the only source for detailed demographic population characteristics for small areas, such as census tracts, block groups, or towns with less than 20,000 persons.[1] Much has changed in the Commonwealth over the past ten years, particularly through the influx of immigrant and refugee populations. So while the 2000 Census can help identify areas of past HTC concentrations, it cannot identify areas of recent change.

Although the Commonwealth does not conduct its own census or large-scale geographic survey, the Massachusetts Department of Education does collect annual data on the demographic characteristic of students enrolled in all public, private and charter schools in the state. To the extent that school-age children mirror the demographic and ethnic composition of their surrounding neighborhoods, enrollment data provides an alternative, albeit imperfect, indicator of neighborhood demographic change.

This study uses data on public elementary school enrollments to identify areas where there has been a relative increase in hard-to-count populations since 2000. School enrollment statistics from the Massachusetts Department of Education include several measures associated with hard-to-count and undercounted populations: race (White, Hispanic, African American, Asian), likely recent immigrants (measured by students whose first language is not english and/or have limited english proficiency), and low income students. We then built a Geographic Information System (GIS) database to show where these enrollment changes are occurring, with particular interest in identifying changes occurring outside of past concentrations of Hard-To-Count areas, as identified in the Census Bureau's Planning Database—based on 2000 counts. Included with this report is an excel spreadsheet listing the individual schools included in our study, as well as a Geographic Information Systems database showing their locations.

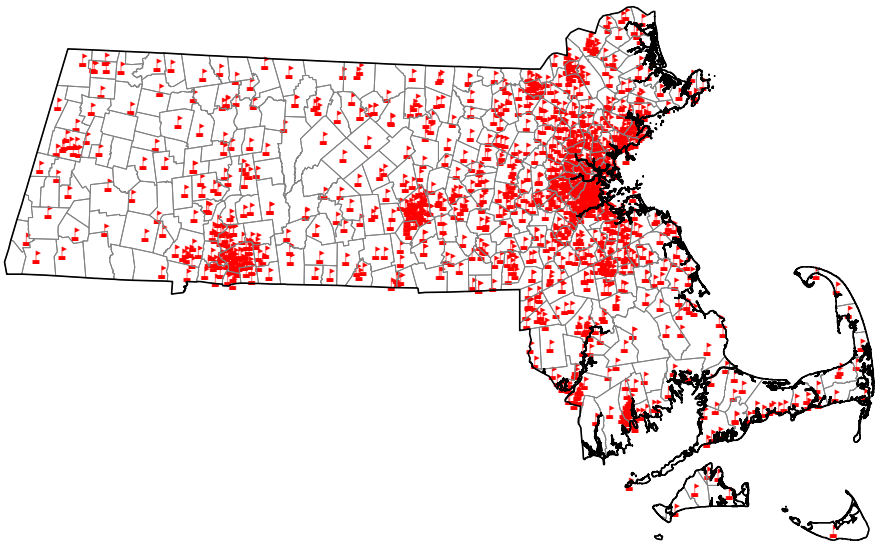
[1] The recently initiated American Community Survey (ACS) promises annual updates for key demographic characteristics down to a census tract level, but not until after the 2010 census.

Data and Methods

This section provides a brief summary of the data and methods used to estimate student enrollment changes in HTC populations. There are three primary data sources used in this study. The first is a GIS layer showing the locations of all primary and secondary private, public and charter schools in the state of Massachusetts. This layer was produced by MASS GIS and is current as of 2009. In addition to each school's location, this layer contains information on the name of the school, the school ID number and contact information, whether it is a public, private, charter or other type of school, and the grade levels taught.

We decided to focus only on public schools because private and charter schools typically draw students from a wider area and thus their student body may not necessarily reflect the community where the school is located. Among the public schools, we restrict our analysis to elementary schools. Schools that only offer classes for grades sixth grade or higher are not include because middle and high schools often have regional (or district-wide) jurisdictions that draw students from a broad area. Elementary schools are generally smaller and more localized to a specific town or neighborhood. Because the jurisdictions of middle and high schools typically overlap (or subsume) elementary school jurisdictions, there is little to no loss of geographic coverage from their exclusion. We also exclude schools that only offer kindergarten and/or pre-kindergarten programs. The final database includes 1,089 public elementary schools spread throughout the state (Figure 1).[2]

Figure 1
Public elementary schools in Massachusetts, 2009



[2] Because rural schools often pull students from larger jurisdictions, the density of schools is higher in dense population centers. Our school-based indicators will more accurately reflect the community in the area immediately surrounding the school in these areas.

The second primary data source is enrollment counts from the Massachusetts Department of Elementary and Secondary Education (DoED). In October of each year, every school in the state conducts a census of its students and reports this information to the DoED. The DoED, in turn, makes this information available to the public through its online data portal. Each school reports its total enrollment as well as student characteristics, such as race (White, African American, Hispanic, Asian, American Indian or Tribal), whether the student's first language is not english (FLNE), whether student is of limited english proficiency (LEP), and whether the student's family qualifies as low income (LI).

The third data source is a GIS layer showing census tract-level HTC scores as provided by the Census Bureau's Tract Level Planning Database. The HTC scores are a composite index of 12 variables associated with historically HTC population characteristics taken from the 2000 Census. The Census identifies tracts with HTC index scores higher than 70 as likely areas of high non-return rates for the upcoming census count. We use this layer to identify schools that saw considerable demographic change outside of previously identified HTC neighborhoods.

While offered as a partial solution to the dilemma posed by the lack of post-2000 demographic data, our study is still subject to several important limitations. In most cases, we expect changes in school enrollments to reflect changing in community characteristics. But there are circumstances where enrollments change due to administrative reasons, such as redistricting, school closings, openings, or expansions. Busing or school choice may also alter the composition of the student body in a way that does not necessarily represent the community. Because of these limitations, we emphasize that this data should be used in conjunction with other indicators of neighborhood change.

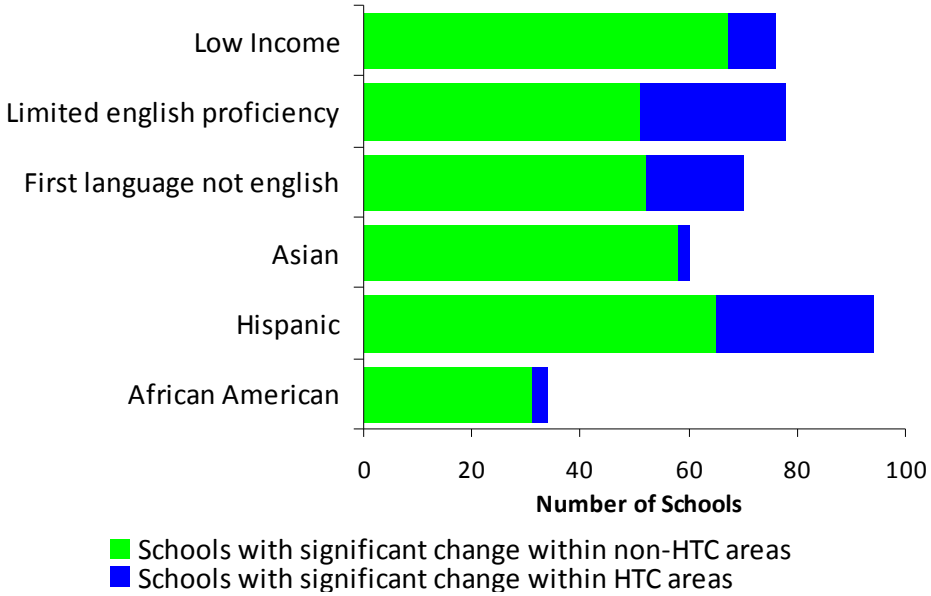
Indentifying School Enrollment Change

The main purpose of our study is to identify areas where there likely has been a large demographic shift in HTC populations since 2000. There are two general types of change indicators that could be used to measure changing populations: (1) absolute measures of change, such as the difference between enrollments in two time periods; and (2) relative measures of change that control for size differences between schools. Although they may help identify areas with a large increase in the number of HTC persons, absolute measures of change heavily favor big schools with large enrollments and may obscure demographic shifts in smaller communities.[3] In this study, we use a relative measure of change. For each of the measured characteristics (race, income, English proficiency, etc.) we identify schools where there was considerable demographic change according to the percentage point difference in the enrollment share between 2000 and 2009.[4] For example, if African Americans comprised 10% of a schools total enrollment in 2000 and 20% of the enrollment in 2009, the percentage point change in the African American population was .10. We measure these percentage point shifts in standard deviations from their mean to identify schools with the greatest *relative* increases or decreases in their enrollment shares. As a rule of thumb, we designated schools with significant demographic shifts as those with a percentage point increase 1.5 standard deviations or higher.

The majority of all schools with significant changes in their demographic composition are outside of traditional HTC areas. Out of the 1,089 schools included in the study, 262 scored as having significant demographic change on at least one of the six criteria. Of these 262, 213 (81%) were outside of traditional HTC census tracts. Figure 2 breaks this down by our six demographic groups, showing the number of schools identified as having significant demographic change, and, of those, how many are in HTC areas. A listing of these specific schools is provided in an excel workbook as a companion to this report (see spreadsheet Raw Data). We identified 34 schools with significant increases in their share of African American students, of which 31 (91%) are outside of traditional HTC areas. The Hispanic population showed the greatest number of schools (94 schools) with significant increases in their enrollment share, 69% of which are outside of traditional HTC areas. Asian populations show the highest share of schools with significant demographic shifts outside of traditional HTC areas (97%). There are slightly more schools with large increases in their share of students with limited english proficiency (78) compared to schools with large increases in students for whom english is not their first language. However, a higher percentage of schools with high concentrations of first language not english students are located outside of traditional HTC areas. Low income populations (88%) also have a fairly large percentage of relative enrollment growth occurring outside of HTC areas.

[3] The accompanying spreadsheets and GIS databases include the annual enrollment counts for different targeted populations which can be used to calculate absolute change.
[4] In cases where 2000 data was not available, we use the earliest year provided.

Figure 2: Selected demographic groups change in share of enrollment by number of schools within Hard-to-Count and outside of Hard-to-Count Census Tracts



Figures 3 through 8 show where these demographic shifts are occurring. These maps were created using ArcGIS v. 9.3 and the data and files used to create these maps are included in the report package. African American enrollments show a spatial shift from historic concentrations southwest of Boston and Springfield to increasing shares to just north of Boston, Worcester and the Randolph/Brockton area (Figure 3).[5] Hispanic enrollments have increased in many parts of the state, including the Boston area, Springfield, Worcester, Framingham, Fitchburg, Lawrence and New Bedford (Figure 4). The Asian population appears to be shifting from inner-cities enclaves in places like Boston, Fitchburg, Lowell, Worcester and Lawrence toward the outer reaches of the Boston metropolitan area (Figure 5).

[5] Keep in mind that because we are looking at changes in enrollment shares, not counts, and that these upward and downward shifts may be due to the influx or outflow of other races.

Figure 3
African American (AA) populations
Change in the share of enrollment, 2000 to 2009

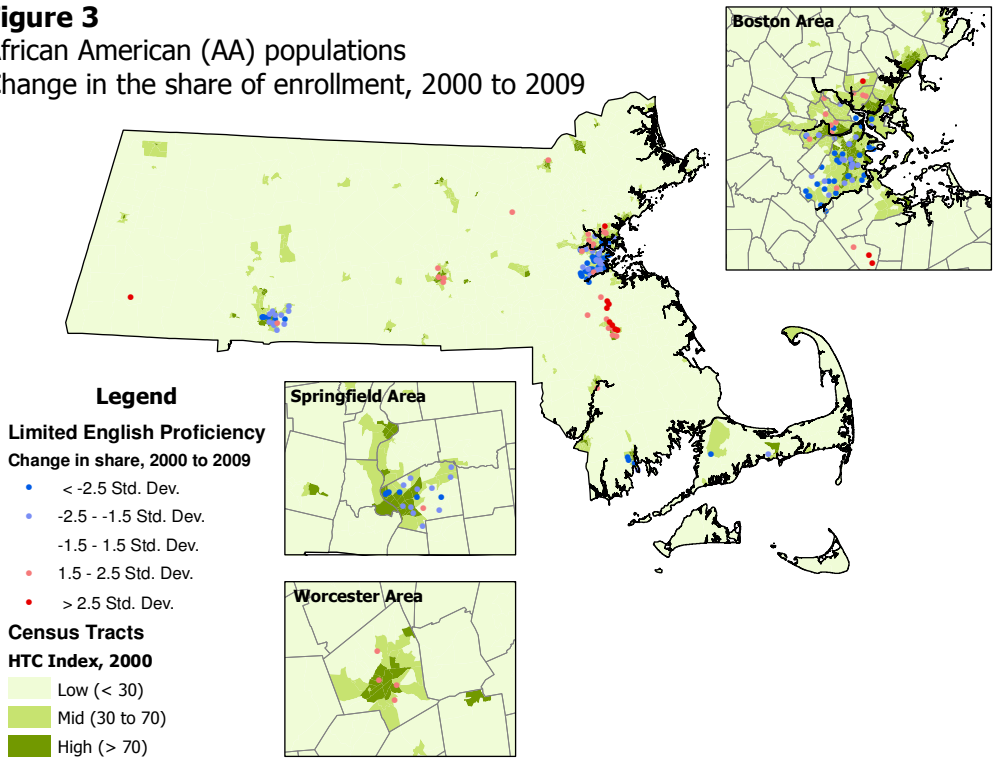


Figure 4
Hispanic (HISP) populations
Change in the share of enrollment, 2000 to 2009

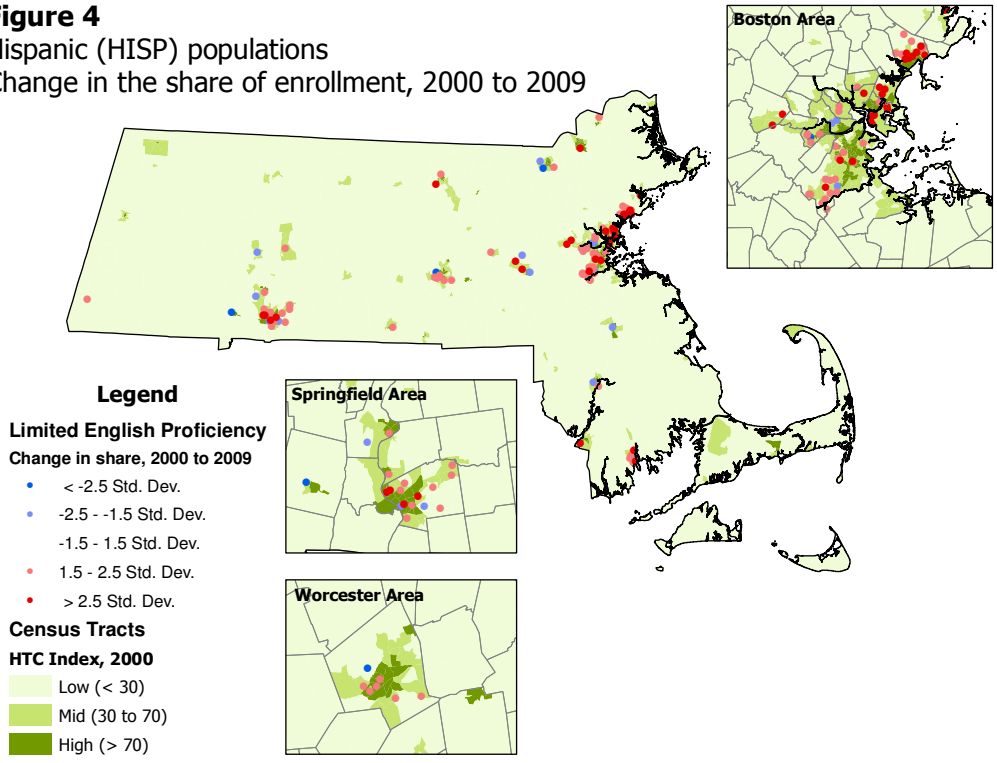
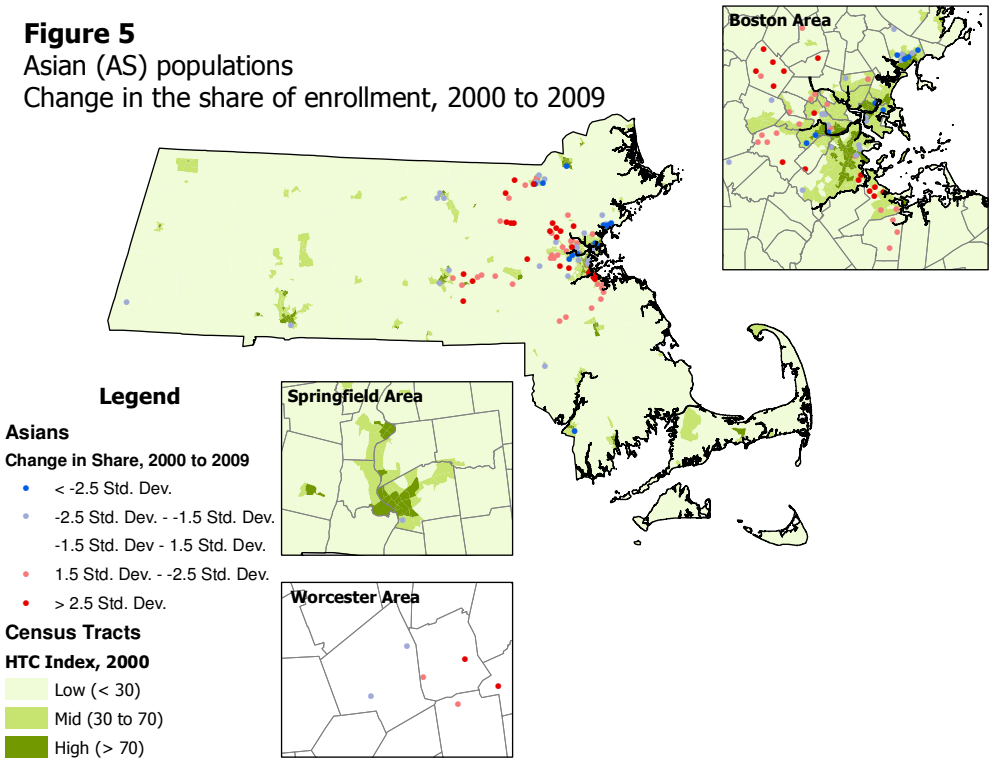


Figure 5
Asian (AS) populations
Change in the share of enrollment, 2000 to 2009



As expected, there is considerable overlap in the measures of Limited English Proficiency and First Language not English (Figures 6 and 7). Both measures are likely to include many of the same students. The share of students whose First Language is not English is typically higher than the share of limited English proficiency for most schools. The difference may relate to how recently the household moved to the US—second generation immigrants will likely not have limited English proficiency but English may not be their first language. The differences may also be due to the prevalence of English in a recent immigrant's county of origin or their level of education. In either case, Limited English Proficiency appears to be the preferred measure for identifying HTC populations. As shown in Figure 7, the largest increases in the enrollment of students with limited English proficiency occurred in traditional manufacturing centers, such as Worcester, Lawrence, and Lowell. Noticeable changes also occurred in Lynn, Brockton, Quincy, and in other pockets both in and surrounding the city of Boston.

Figure 6
First Language Not English (FLNE) populations
Change in the share of enrollment, 2000 to 2009

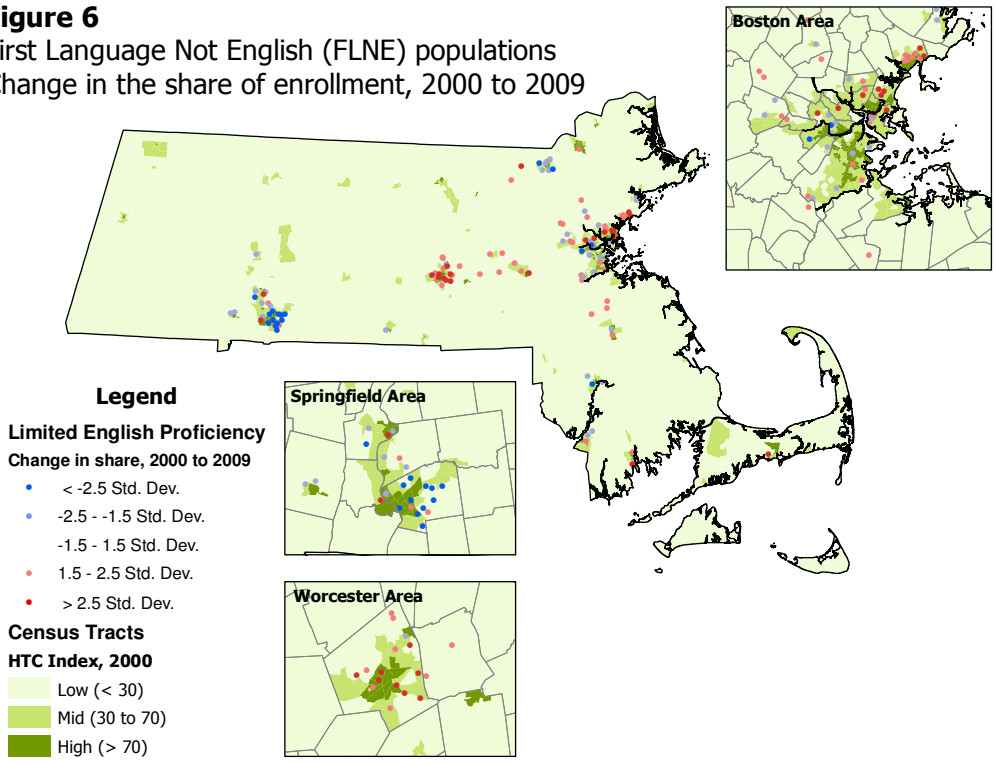
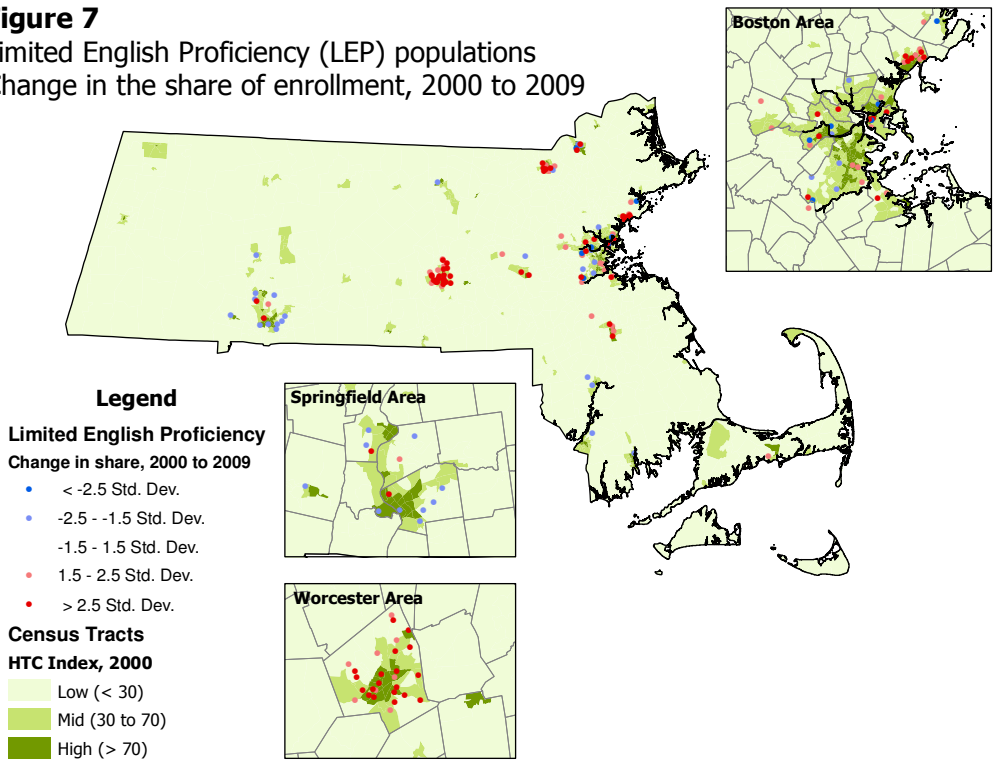
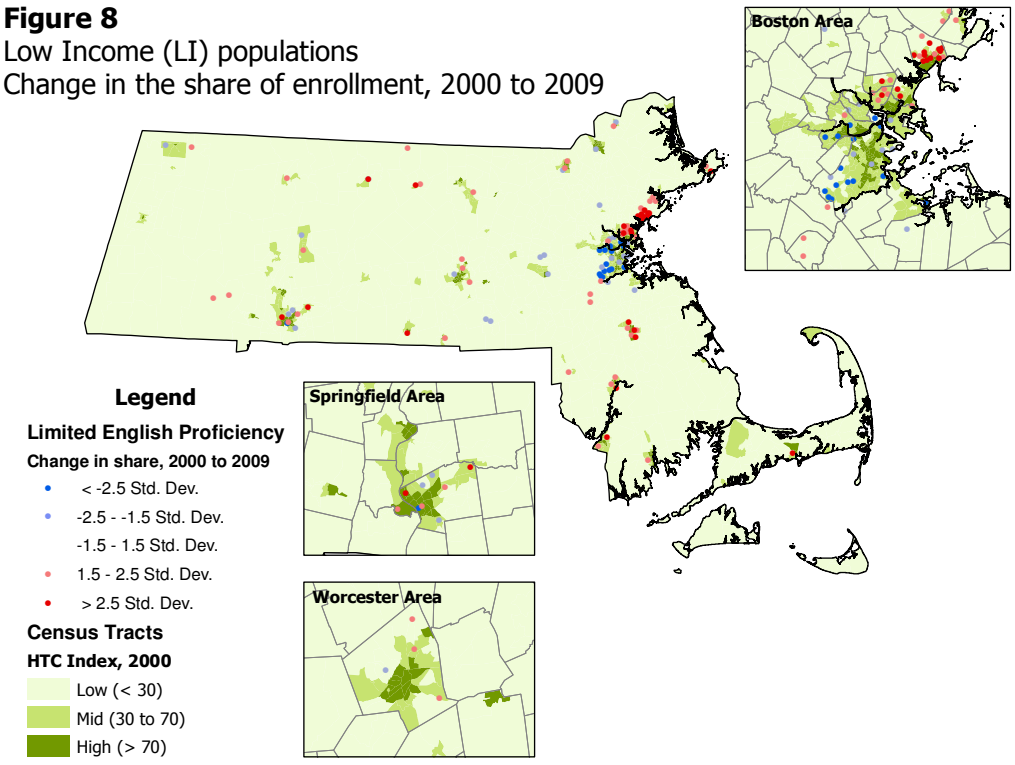


Figure 7
Limited English Proficiency (LEP) populations
Change in the share of enrollment, 2000 to 2009



Our final measure is enrollment shares of students from low-income households. Within the Boston area we see a spatial shift occurring from Boston to communities just to the north, Everett, Malden, Revere, Lynn, and Salem (Figure 8). There have also been increases scattered throughout the rest of the state, most notably in Brockton, Taunton and Fall River (to the Southeast), Gloucester (to the Northeast), Worcester, Athol, Gardiner and Southbridge (Central MA), and Springfield.

Figure 8
Low Income (LI) populations
Change in the share of enrollment, 2000 to 2009



Investigating the Components of School Enrollment Change

A secondary objective of this study was to construct an index of the "hard-to-count" populations by combining the six indicators of demographic enrollment change in counts into a single number. To do this, we used a statistical technique known as principle components analysis. Principle components analysis looks for overlap among the different measures and combines the shared portion of the different measures into common indices (or factors). For example, there is considerable overlap among the First Language not English and Limited English Proficiency populations. These may also coincide with at least some of the change found among the Hispanic and low-income populations, but some might separately relate to recent Asian immigrants.

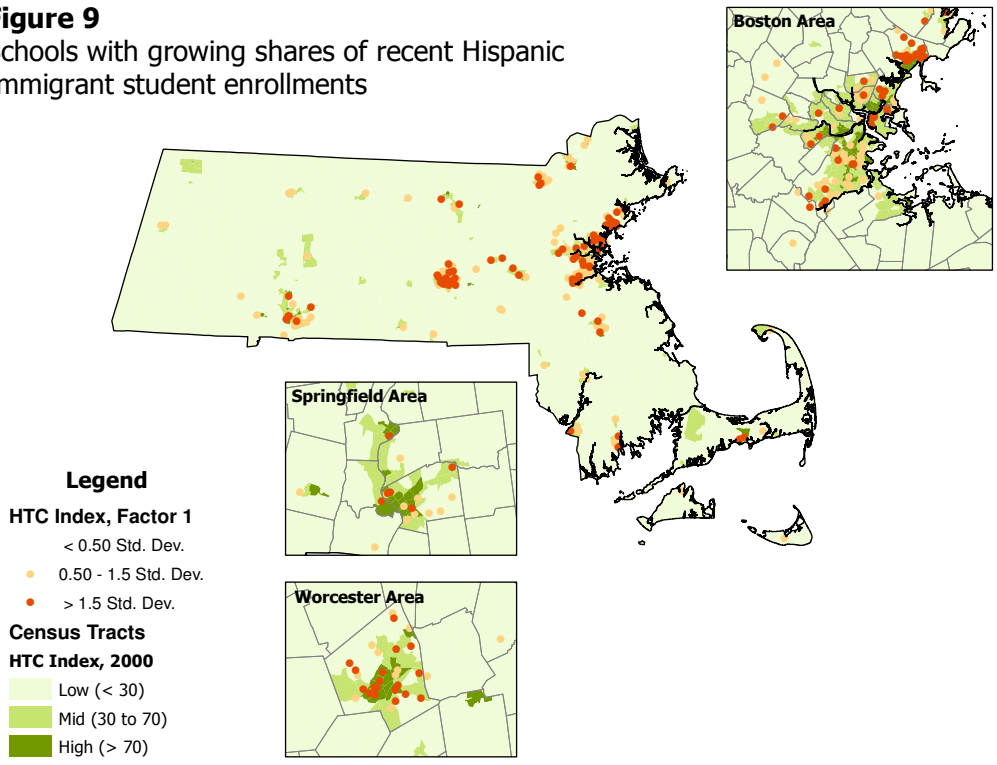
While our original objective was to develop a single index to summarize the combined strength of all our demographic measures, the principle components analysis actually revealed three quite distinct processes. This can be seen in Table 1, where standardized scores show the relative contribution of each demographic measure to each factor. Higher coefficient scores indicate measures that are more closely related to the overall factor. Negative scores identify demographic that have an inverse association with the overall factor.

Table 1: Principle Components Analysis, Standardized Scoring Coefficients

	Factor 1	Factor 2	Factor 3
Hispanics	0.378	-0.251	-0.150
African Americans	-0.072	0.640	-0.152
Asians	-0.124	0.206	0.692
First Language Not English	0.345	0.139	0.392
Limited English Proficiency	0.372	0.074	0.196
Low Income	0.214	0.426	-0.358

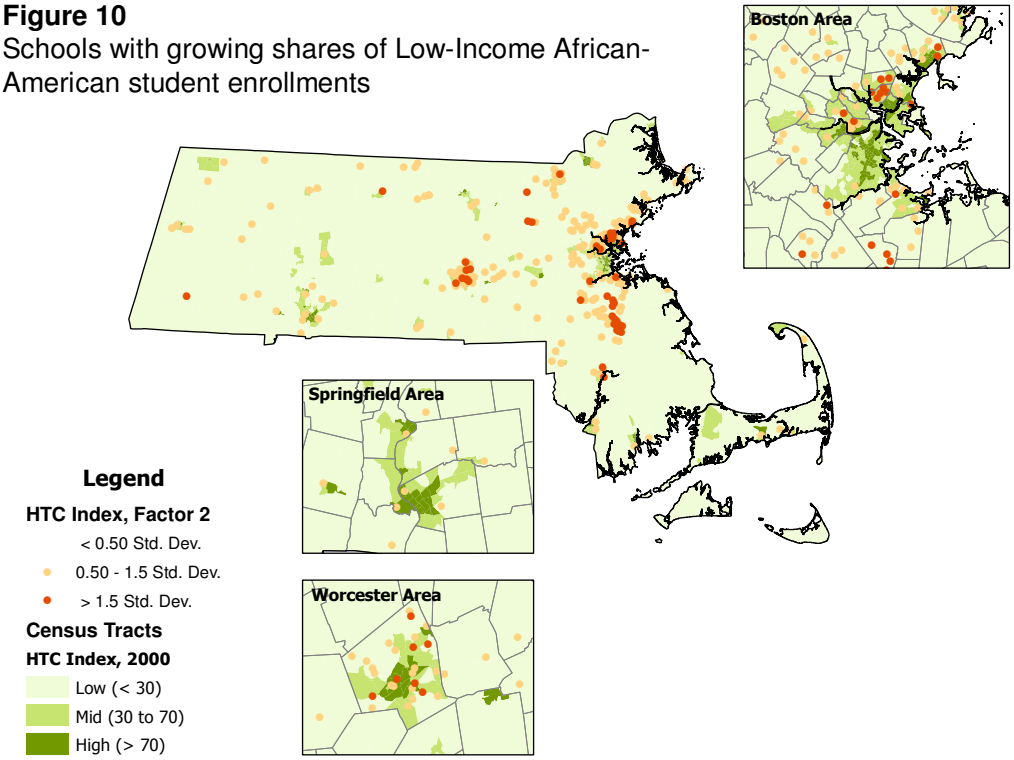
The first factor is dominated by recent growth of the Hispanic population, which also coincides with changes in FLNE and LEP enrollments and, to a lesser extent, with increasing shares of low-income students. Considering combination of these four characteristics, Factor 1 is most likely identifying areas with growth in their population of recent Hispanic immigrants. Figure 9, shows that these changes are predominantly occurring in the 'gateway' cities of Lowell, Lawrence, Springfield, Worcester and Brockton, as well as in the north shore communities of Metropolitan Boston such as Lynn and Revere.

Figure 9
Schools with growing shares of recent Hispanic Immigrant student enrollments



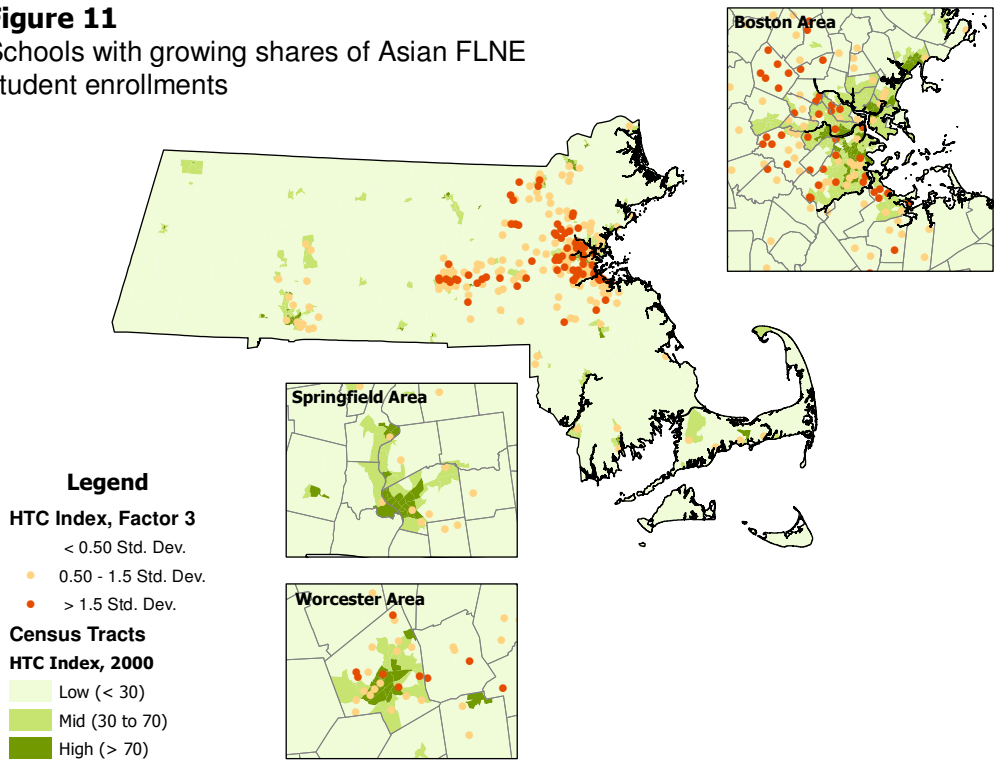
The second factor primarily identifies changes in the relative concentration of Low-Income African American households. The Hispanic population has a negative score on the second factor meaning that, by and larger, schools with growing shares of low-income African Americans are not necessarily the same places where there are growing shares of Hispanics – although (as show in Figure 10) there is still some overlap. The two language characteristics (FLNE and LEP) also have fairly low scores on this factor, suggesting that this measure is likely capturing demographic shifts among native-born African Americans rather than pockets of recent international refugees. These shifts are most notably occurring to the northwest of Boston in communities such as Everett and Malden, and in the southeastern communities of Randolph and Brockton. Worcester, and to lesser extent other gateway cities such as Lowell, have also seen growth according to this index.

Figure 10
Schools with growing shares of Low-Income African-American student enrollments



The third factor identifies changing location patterns among Asians, for whom English is not their first language. These may or may not be recent immigrants – as many first and even second generation Asian households do not speak English at home. The low income variable also scores negatively on this factor – further suggesting that this factor does not reflect recent immigrants but rather intergenerational communities of rising affluence and spatial mobility. Coinciding with this interpretation, Figure 11 shows a somewhat typical suburbanization pattern, with the share of Asian populations spreading throughout the inner suburbs of the Boston metro area particularly Lexington, Newton, Winchester, Somerville, and the western portions of the City of Boston. Growing shares of Asian students also appear in communities along the Interstate 495 corridor, with particular concentrations in somewhat more urban places such as Worcester and Acton.

Figure 11
Schools with growing shares of Asian FLNE student enrollments



Summary

This report describes the key findings from our analysis of demographic changes in school enrollments and how these changes might help to identify areas where the residents may be undercounted in the upcoming decennial census. We found that much of this change is occurring in places that are not designated by the Census Bureau as historically "Hard-to-Count" areas, and provide maps to show the general location where these demographic changes are occurring.

This paper provides merely a cursory overview of the potential uses and value of the databases developed during the course of this study. Along with this report, we offer an electronic spreadsheet compiled in a Microsoft excel workbook format that lists, by school, the numbers of and shares of students with different demographic characteristics for all the available years from 2000 onward. This information is also provided in a Geographic Information System database, which can be queried and mapped. In this report, we only highlight schools with large changes based upon a single indicator (percentage point change). With the spreadsheet or the GIS the user can develop their own measures of demographic change and investigate, in much greater detail, changes occurring in particular regions, neighborhoods or even schools. The appendix provides a description for each of these companion work products.

Appendix: Description of Companion Works Products

- Companion Guide to Using the Report: Using School Age Pops to Identify Hard-To-Count Populations
    - A users guide on how to analyze the data
- A full set of state-wide maps (paper and electronic) as well as a GIS database. This database includes specifics on the mapped schools for easy reference or for use in other projects (i.e. generating mailings or completing further analysis)
- A Breakdown of the Findings from the Report: Using School Age Pops to Identify Hard-To-Count Populations
    - A text document briefing the results and findings of the final report
- An Excel table composed of the raw data used for the study also broken down by county. It includes annual enrollment counts and shares for students distinguished by race, English proficiency, and low income status for all years from 2000 to 2009.