Additional Challenges and Solutions: The Post-Censal Denominator Small-Area Population Estimates and the Massachusetts Approach

Robert Knorr
Director of Environmental Epidemiology
Bureau of Environmental Health
Massachusetts Department of Public Health

2018 CSTE Annual Meeting, West Palm Beach, FL
June 10, 2018

Presentation Outline

1. Why do we need sub-county population estimates and what were our options?
2. UMass Donahue Institute method
    1. What is it?
    2. Error associated with the estimates
    3. Evaluation
        1. Childhood lead poisoning screening
        2. Life Expectancy
3. Feasibility for use by other geographies

In Massachusetts, Sub-County Estimates are Essential to Stakeholders

- Disparities at larger geographies can be smoothed over and opportunities to affect improvements in health missed
- Population growth and changing population characteristics
  - Community Needs Assessment
  - Health Impact Assessment
  - Emergency Preparedness
  - Identifying Health Disparities and Vulnerable Populations
- Interest in census-tract level life expectancy estimates only adds to the need for reliable population data

Post-Censal Population Estimation Options

- Simple linear extrapolation of U.S. Census data
  - Inherently biased as population change is not constant
  - Assumes a constant rate of growth or decline, producing larger error as the decade progresses
- Commercial products (e.g. Geolytics)
  - MA found large unexplained flaws with previously purchased estimates
  - Methods descriptions are vague at best
- The American Community Survey
  - 5-year running sample
  - Rate calculation using ACS small areas is not recommended due to large margins of error
  - Does not offer single-year age groups

Factors to Consider in Choosing an Option

- Accuracy of the estimate
- How uncertainty is accounted for
- Necessary detail
- Costs of production
- Ease of application and explanation

And the Method Massachusetts has gone with …

- The Modified Hamilton-Perry method
- Developed by the University of Massachusetts Donahue Institute
- specifically account for the three components of population change, births, deaths, and migration.
- the population at the beginning of the time period (commonly the most recent census), to which the number of births and in-migrants that occurred over the time period are added, and from which the number of deaths and out-migrants that occurred over the time period are subtracted.

UMDI – Strengths of the Methodology

- Commonly employed projection method (Hamilton-Perry)
- Requires Census data products only
- Annually adjusted for improved performance
- Capable of producing race, ethnicity, sex, and single-year age stratum

UMDI Methodology - Assigning Error

- Projected 2010 estimates from 10 and 20-year Cohort Change Ratios compared to 2010 Census
- Population estimate error was generally best explained by age and cohort size
- Population estimate errors were assigned to estimates based on base population cohort size and age
- As with estimates, errors were projected to 2020 and applied across intervening years. Error increases as you move forward in time.

Comparing Error
UMDI method vs. ACS method

Comparing the coefficient of variation (CV) for UMDI and ACS population estimates for 5-year age groups, stratified by community-based cohort size.

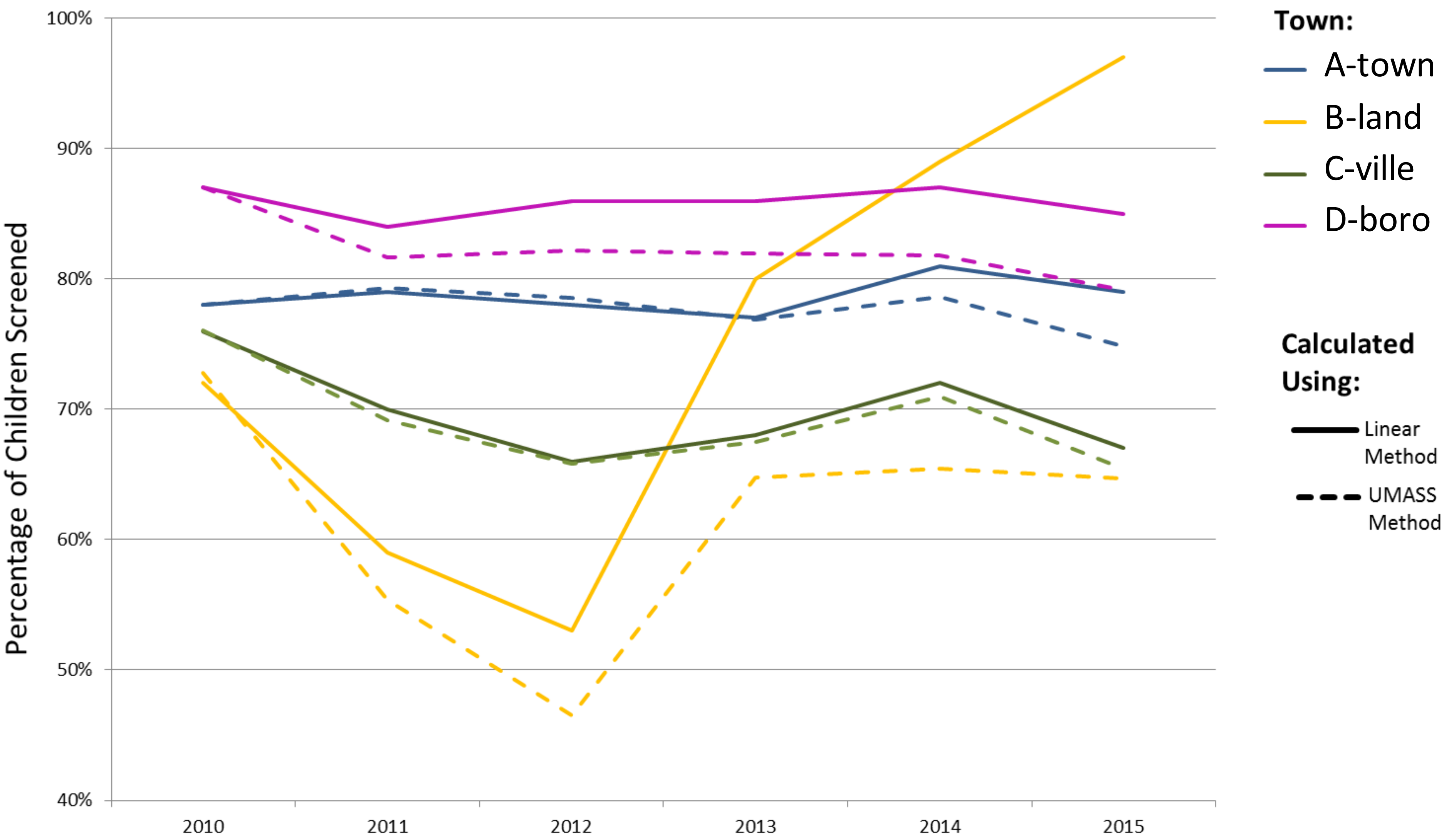
	Cohort Size	n	Under 5	5 to 9	10 to 14	15 to 19	20 to 24	25 to 29	30 to 34	35 to 39	40 to 44	45 to 49	50 to 54	55 to 59	60 to 64	65 to 69	70 to 74	75 to 79	80 to 84	85 and up
UMDI	0 to 99	3556	17	20	16	15	19	23	17	18	13	11	12	10	8	10	13	15	12	21
	100 to 199	1810	11	8	8	8	9	11	13	11	11	8	6	8	7	6	7	6	7	11
	200 to 499	3468	8	7	7	5	10	9	13	10	8	6	6	6	8	7	5	5	5	9
	500 to 999	2183	8	6	4	5	8	7	12	13	6	4	4	4	6	6	4	3	4	7
	999 to 1999	1130	7	7	5	5	7	5	8	6	4	3	3	4	5	5	4	3	2	4
	2000 to 3999	367	4	5	4	4	5	5	6	4	4	3	2	2	2	3	3	2	3	4
	4000 to 9999	92	3	5	3	4	4	5	5	4	4	3	2	2	1	NA	3	2	3	4
	10000 to 19999	19	1	2	1	NA	4	NA	NA	NA	2	4	1	1	1	1	NA	NA	NA	NA
	20000 and up	12	NA	NA	NA	4	0	1	0	1	2	NA	NA	NA	NA	NA	NA	NA	NA	NA
ACS	0 to 99	3720	51	42	43	77	117	50	51	43	39	35	31	29	42	47	40	48	54	56
	100 to 199	1825	34	31	28	41	53	35	33	31	27	26	26	25	37	40	29	30	30	32
	200 to 499	3285	23	22	22	30	38	24	23	21	20	19	19	20	28	28	20	20	21	22
	500 to 999	2222	16	15	15	21	27	17	16	14	14	14	13	13	20	19	14	14	15	15
	999 to 1999	1112	12	11	11	16	20	12	11	11	10	10	10	10	15	14	10	10	10	10
	2000 to 3999	355	8	7	8	11	15	8	7	8	8	7	7	7	11	10	7	7	6	6
	4000 to 9999	90	5	5	5	7	10	5	5	5	5	4	4	5	8	5	4	4	4	4
	10000 to 19999	18	2	3	3	NA	NA	NA	NA	3	3	1	1	2	4	5	NA	NA	NA	NA
	20000 and up	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

MA Evaluation and Implementation Work

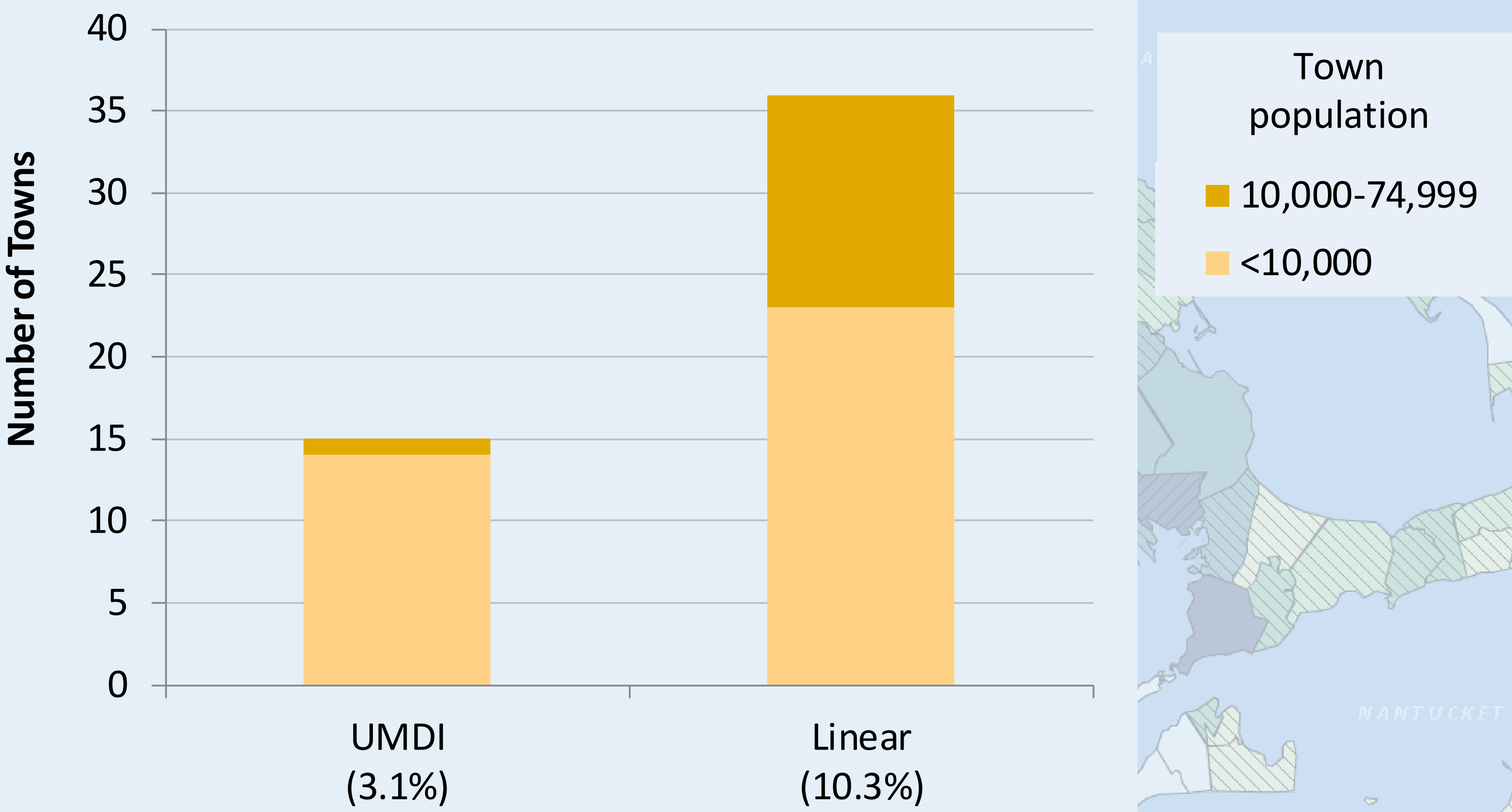
1. Review of impacts on childhood lead screening rates – UMDI vs linear extrapolation estimates

1. Review of Life Expectance estimates – UMDI vs ACS estimates.

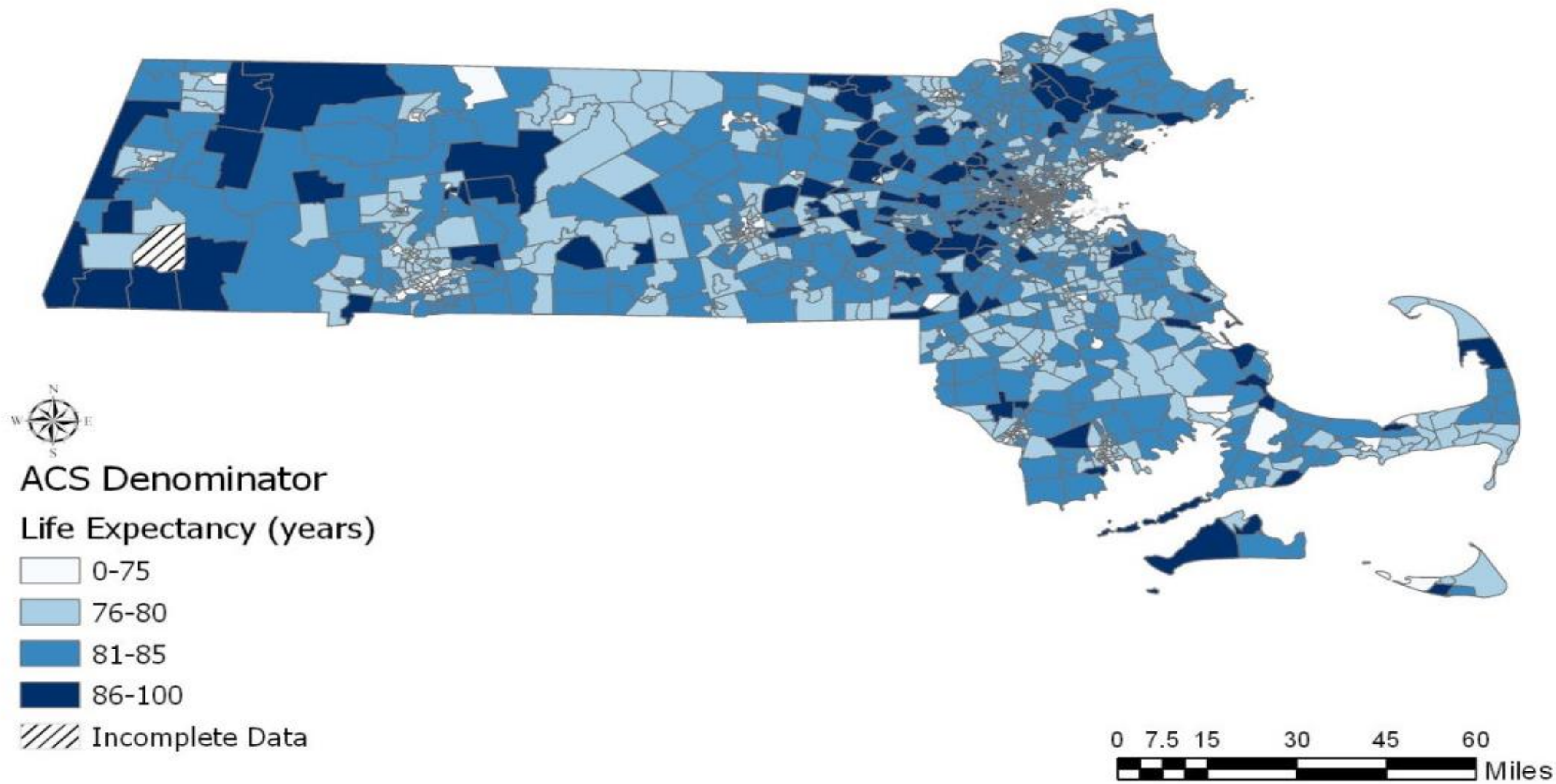
Lead Screening Rates (age 9-47 months) UMDI method vs. Linear method



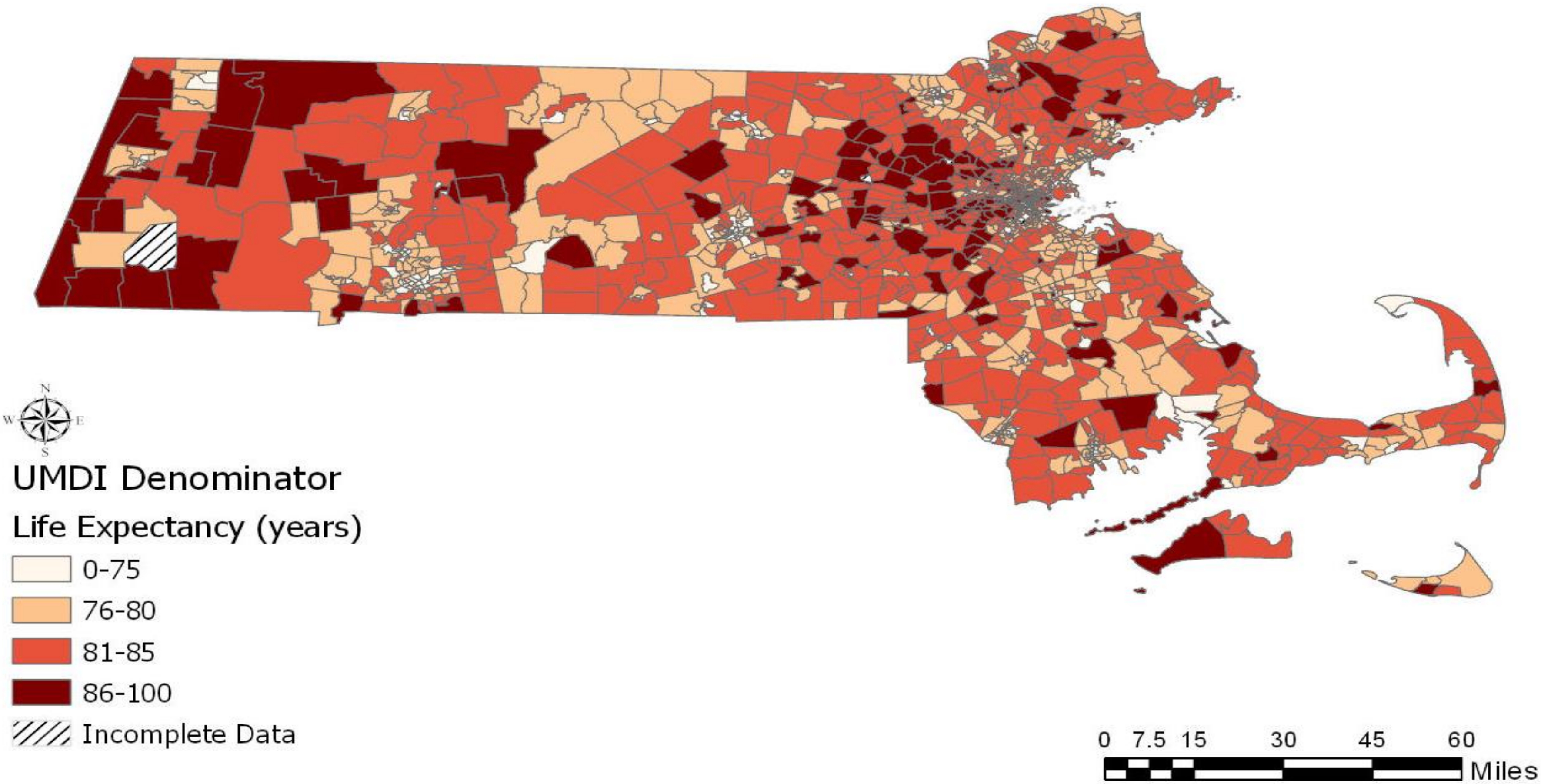
Lead Screening Rates >100%
UMDI method vs. Linear method



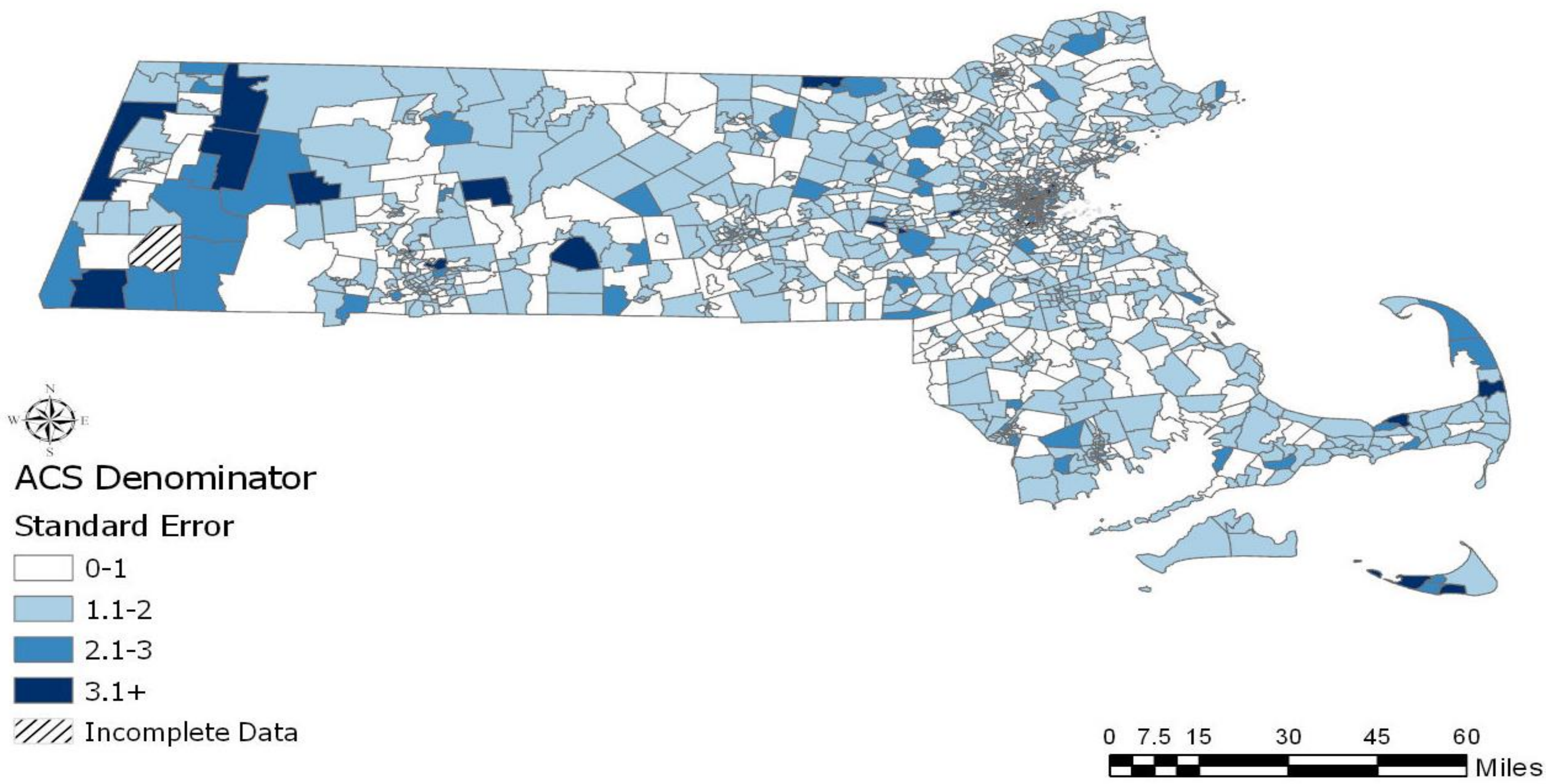
Point Estimates of Life Expectancy Using ACS Population Data as the Denominator



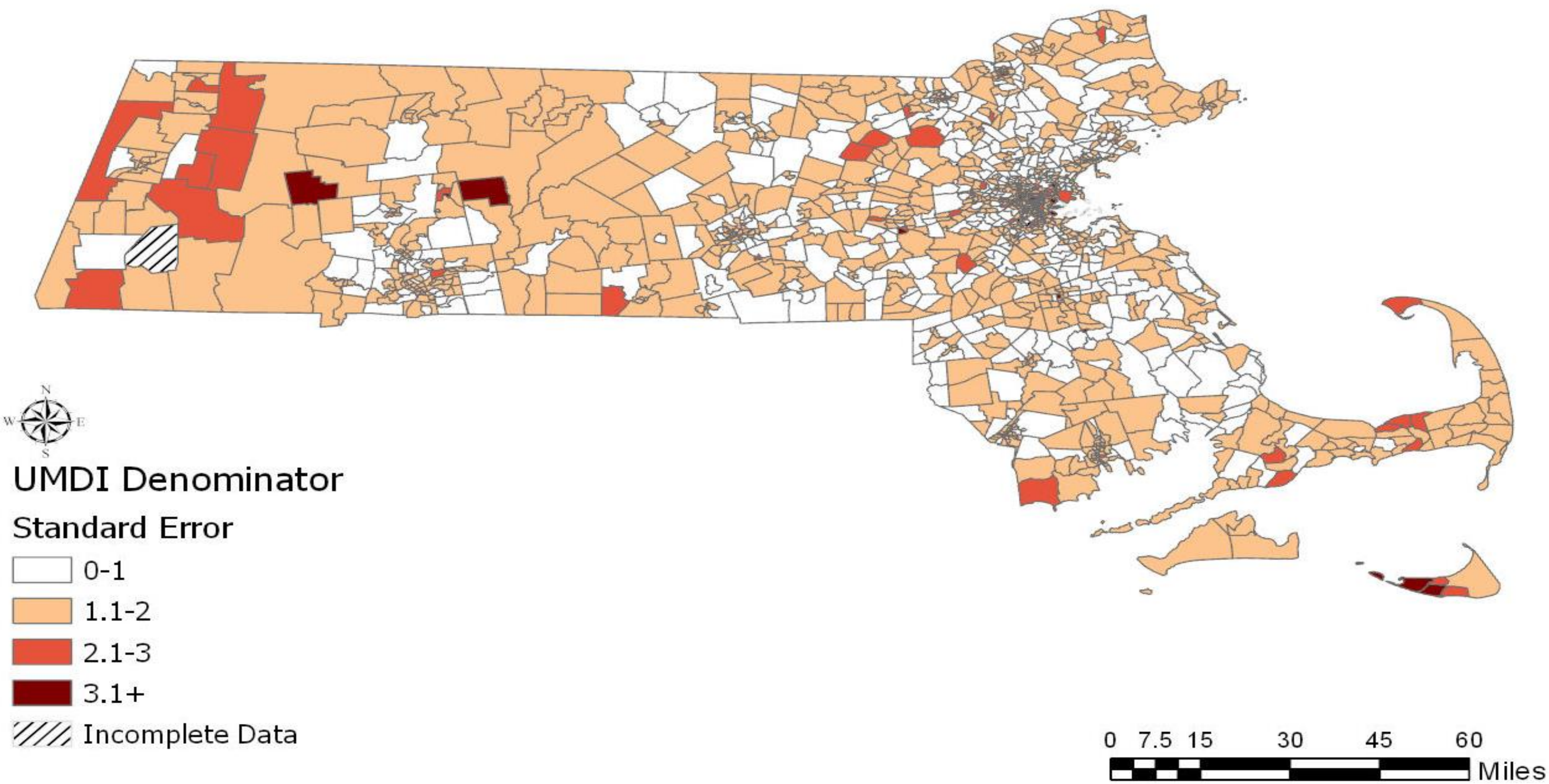
Point Estimates of Life Expectancy Using UMDI Population Data as the Denominator



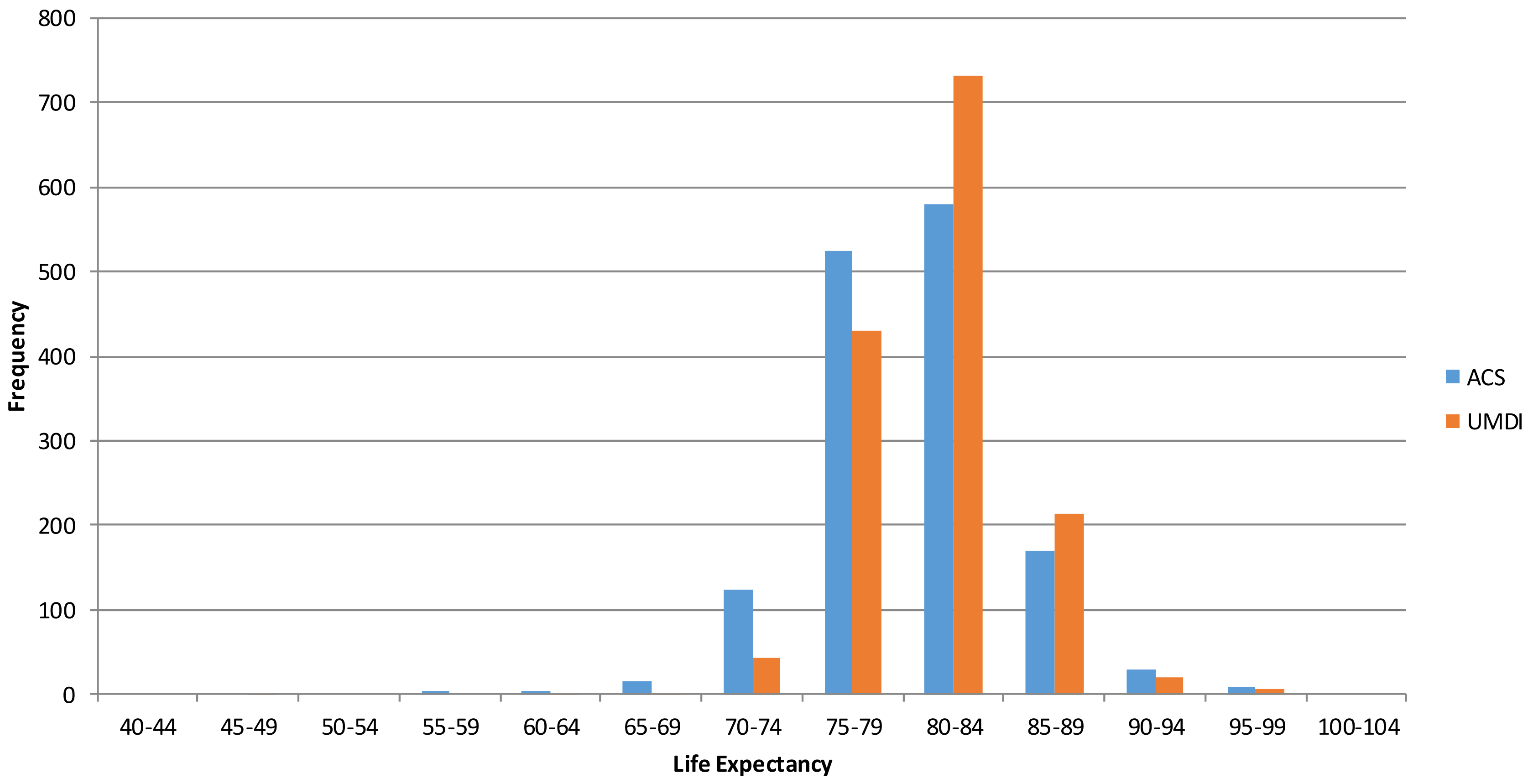
Standard Error Using ACS Population Data as the Denominator



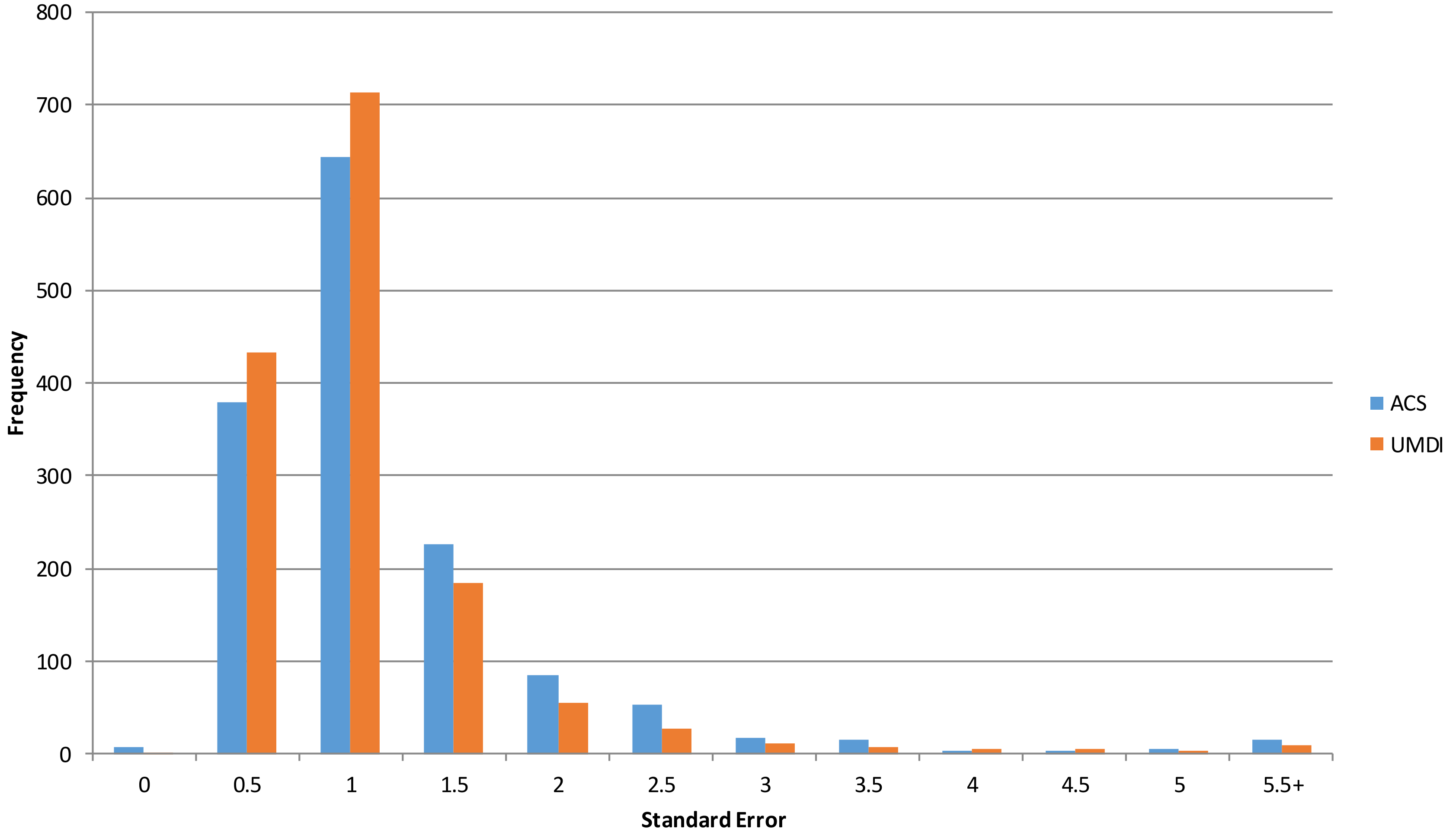
Standard Error Using UMDI Population Data as the Denominator

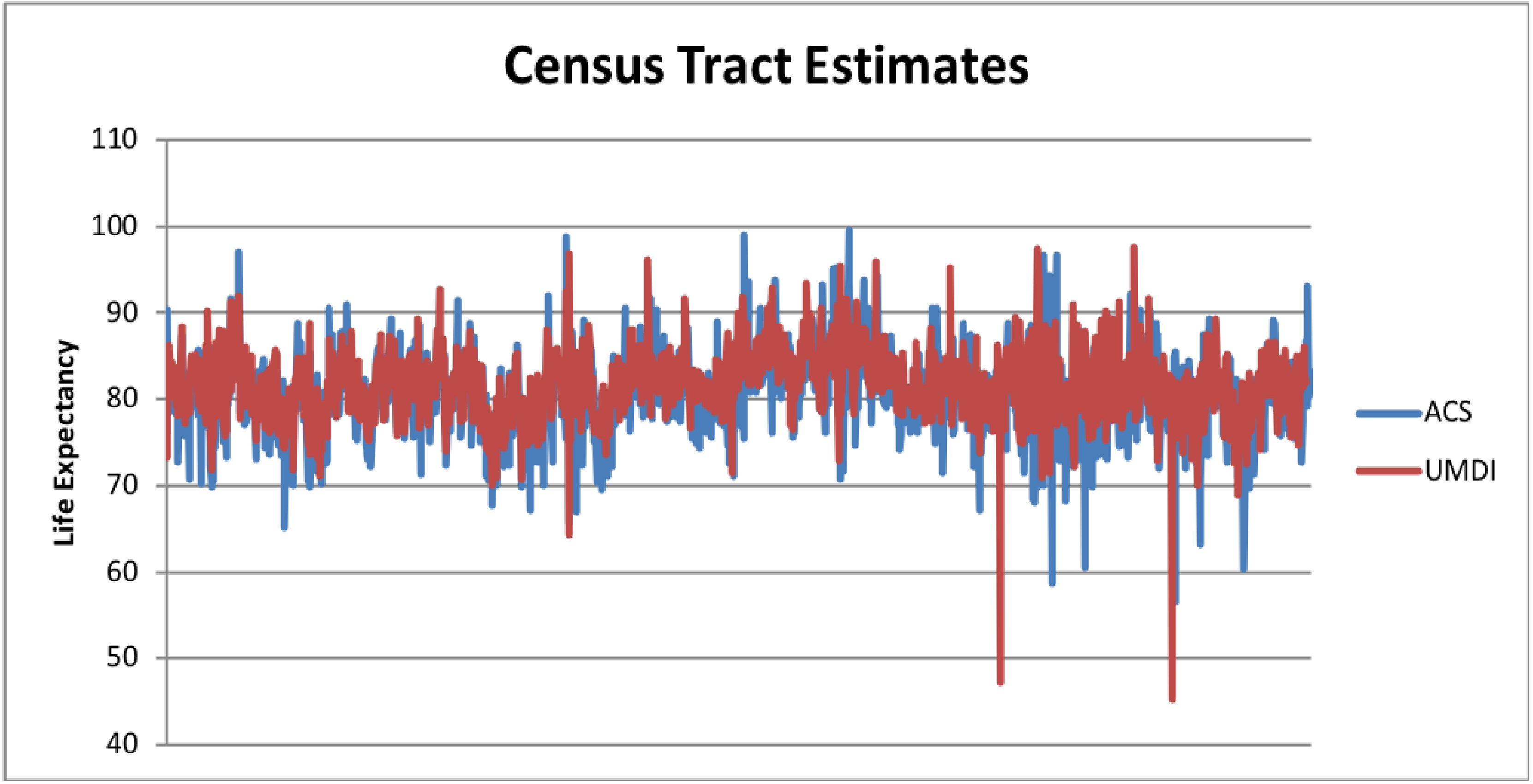


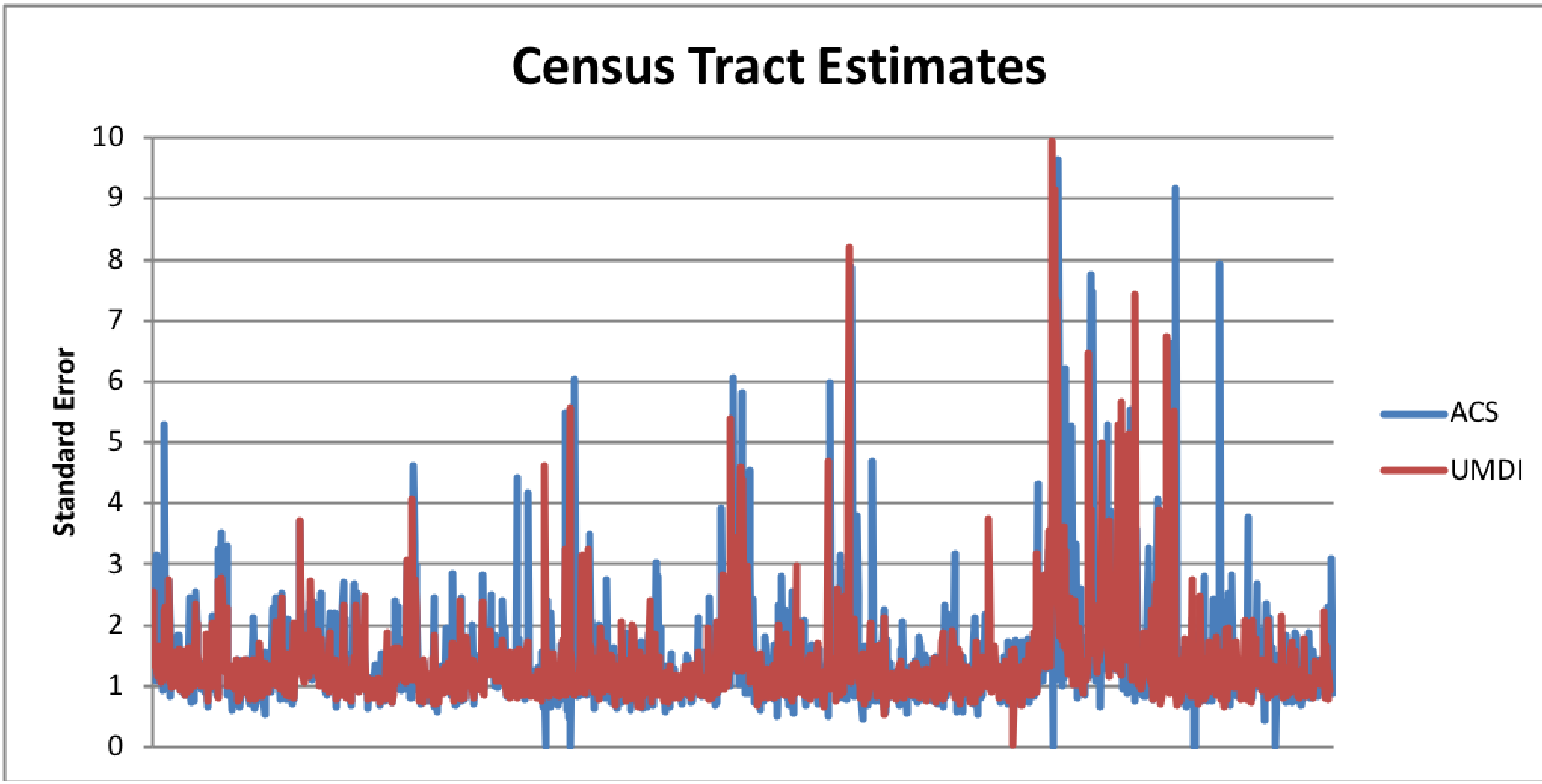
Denominator Point Estimates of Life Expectancy



Standard Error of Denominator




Census Tract Estimates
110
100
90
80
70
60
50
40
Life Expectancy
ACS
UMDI


Census Tract Estimates
Standard Error
10
9
8
7
6
5
4
3
2
1
0
ACS
UMDI

Feasibility for Application to other States?

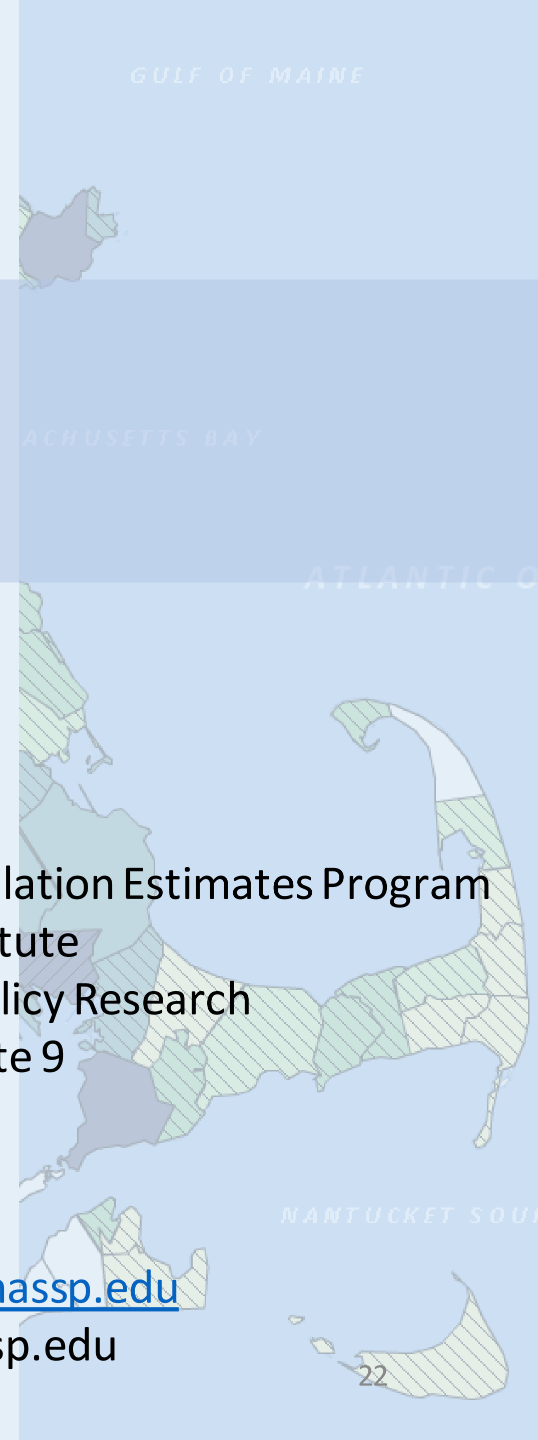
Most easily applied where:

- Geographies remain fairly stable from Census to Census and/or
- Have excellent correspondence files across time 1990-2010.

Difficulty presented when:

- Boundaries have changed or annexations have taken place from Census to Census
  - Would require efforts to document changes
  - Helpful if records available (e.g. correspondence files)
- Sub-county population sizes are small
  - A higher geographic level may be needed (i.e. tract aggregation)
  - More uncertainty in the estimates may occur

QUESTIONS?



Contact Information

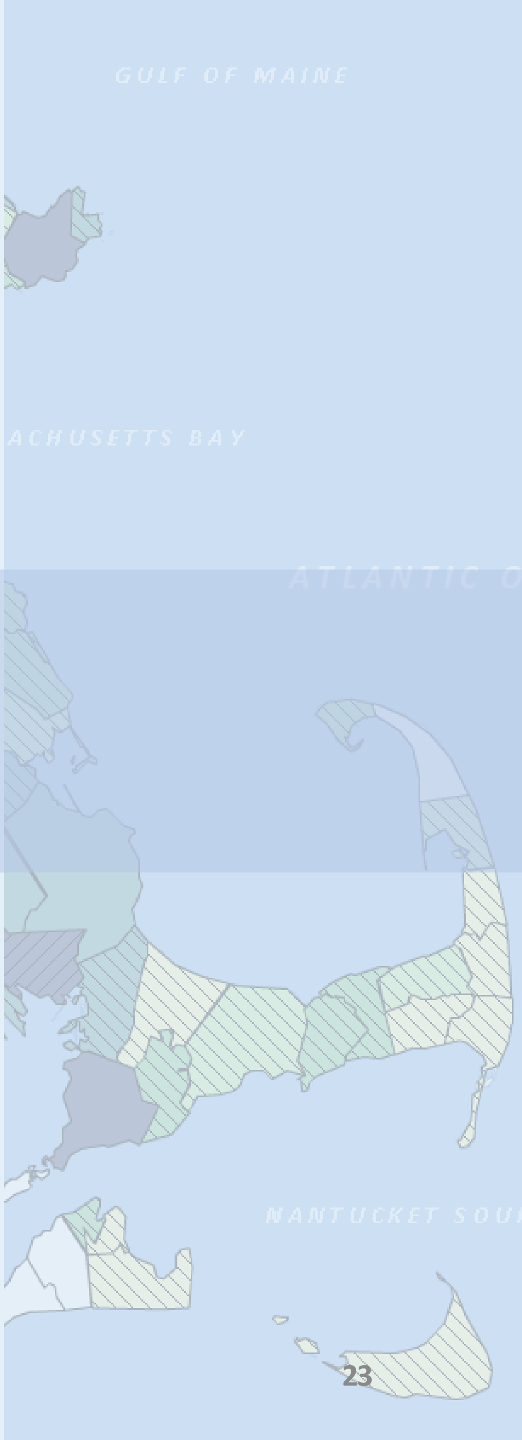
Robert Knorr
Director
Environmental Epidemiology Program
Bureau of Environmental Health
MA Department of Public Health
250 Washington Street, 7th floor
Boston, MA 02108

(617) 624-5757
robert.knorr@state.ma.us

Susan Strate
Senior Manager Population Estimates Program
UMass Donahue Institute
Economic & Public Policy Research
100 Venture Way, Suite 9
Hadley, MA 01035

(413) 577-0753
sstrate@donahue.umassp.edu
www.donahue.umassp.edu

Supplemental Slides



UMDI Methodology – Aging Forward

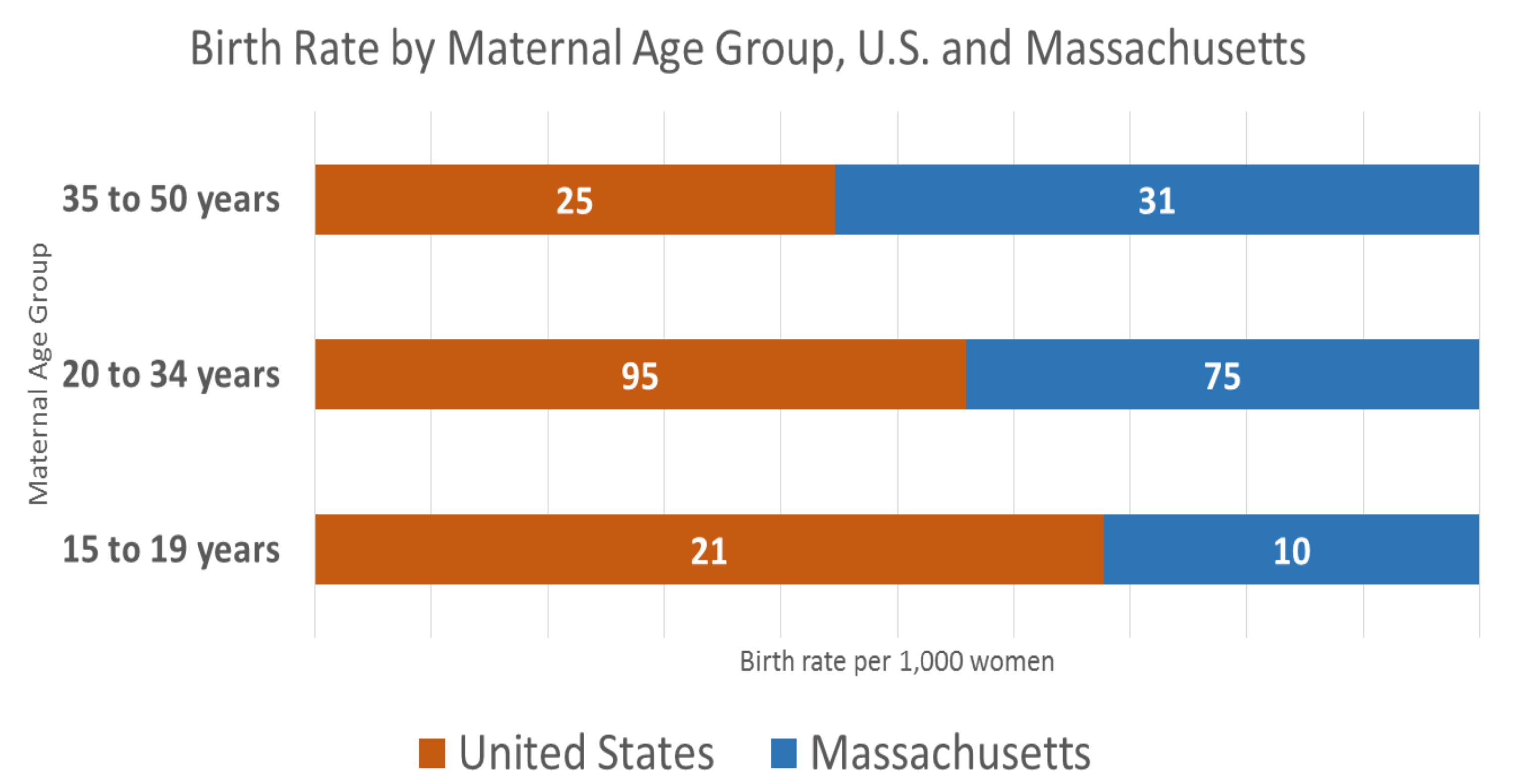
- Standard Hamilton-Perry Methodology
  - Cohort Change Ratio – "Ages" age-sex cohorts forward from one census to the next
  - Ratio is applied to base population to create future estimate.

Town	AGE	SEX	1990 Census	2000 Census	10-year CCR
ABINGTON	0-4	F	482	479	0.15
ABINGTON	5-9	F	495	500	0.15
ABINGTON	10-14	F	**451**	530	1.10
ABINGTON	15-19	F	450	462	0.93
ABINGTON	20-24	F	558	**364**	**0.81**
ABINGTON	25-29	F	608	458	1.02
ABINGTON	30-34	F	633	577	1.03
ABINGTON	35-39	F	569	689	1.13
ABINGTON	40-44	F	516	631	1.00

$$= \frac{364}{451}$$

UMASS DONAHUE INSTITUTE
ECONOMIC & PUBLIC POLICY RESEARCH

UMDI Methodology – Reflects local Shift in CWRs



U.S. Census Bureau American Community Survey 5-year dataset 2010-2014

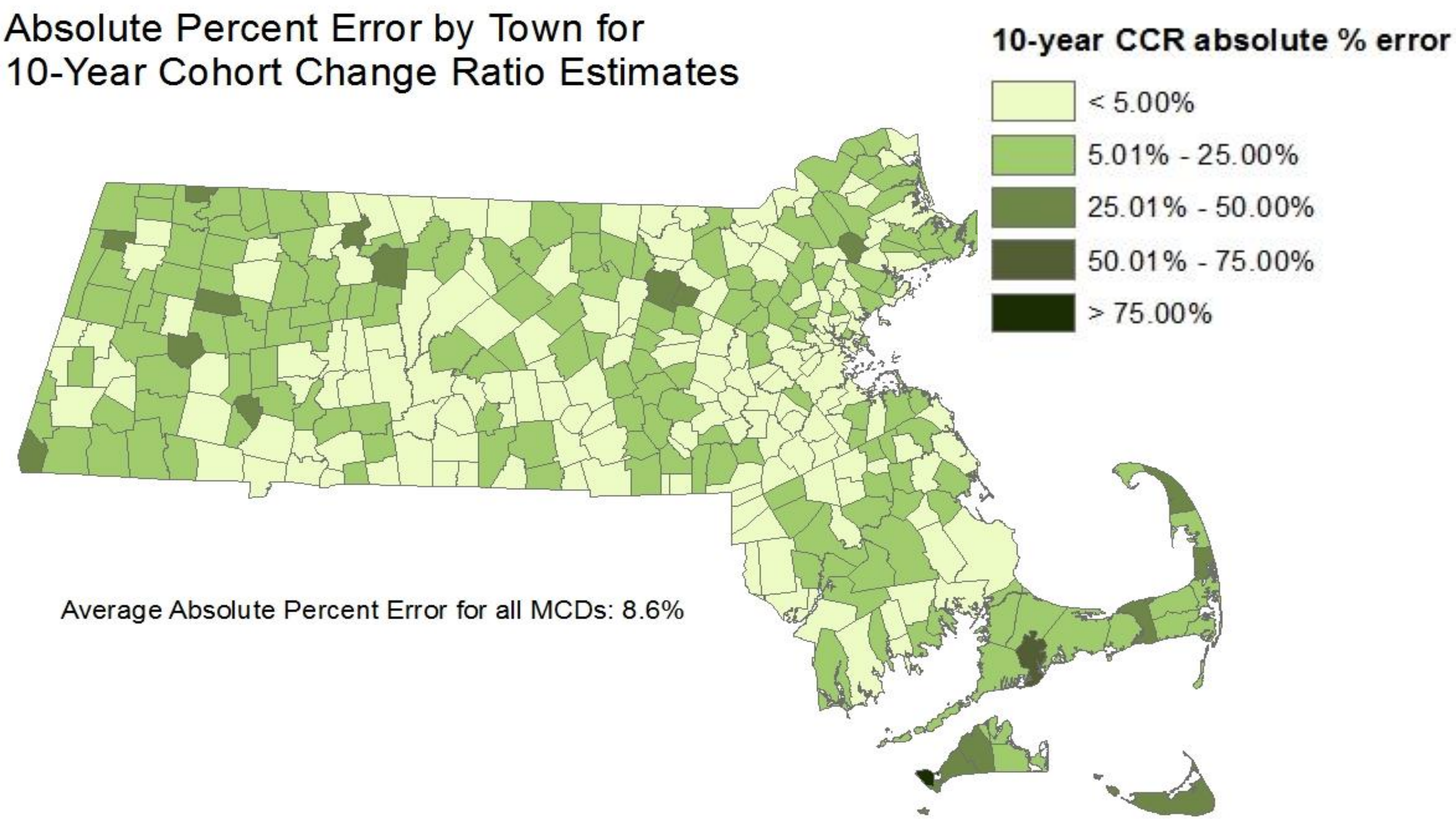
UMASS DONAHUE INSTITUTE
ECONOMIC & PUBLIC POLICY RESEARCH

UMDI Projected Population Estimate Error by Cohort Size and Age

MAPE by age and cohort-size: 10-year CCR estimates

SIZE (2000)	n	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+
0-99	3874	35%	59%	38%	31%	32%	51%	48%	45%	27%	22%	25%	22%	20%	22%	24%	26%	24%	33%
100-199	1915	25%	19%	14%	15%	14%	22%	27%	20%	16%	14%	11%	13%	11%	10%	11%	11%	12%	18%
200-499	3228	19%	14%	13%	10%	15%	19%	22%	17%	14%	11%	11%	9%	11%	11%	9%	9%	10%	17%
500-999	2115	17%	13%	10%	9%	14%	13%	21%	17%	9%	6%	6%	7%	7%	8%	8%	6%	7%	12%
1,000-1,999	1067	14%	13%	9%	9%	14%	10%	16%	12%	8%	6%	6%	6%	7%	8%	6%	6%	3%	7%
2,000-3,999	328	7%	15%	7%	9%	10%	11%	13%	9%	7%	6%	4%	3%	2%	2%	4%	3%	5%	6%
4,000-9,999	85	3%	15%	5%	5%	8%	8%	10%	9%	9%	4%	1%		1%	2%	7%	8%	7%	5%
10,000-19,999	14	5%	33%	7%						4%	3%	1%	1%						
20,000+	10				3%	3%	2%	2%	7%										
	12636	22%	24%	17%	15%	20%	28%	27%	21%	14%	11%	12%	12%	13%	14%	15%	16%	18%	27%

UMDI Projected Population Estimate Error, by Geography



Challenges with Implementing Estimates

1. Race categories

- Excludes “Other” category
- “Other” re-distributed proportionally to other race categories based on census expertise and local information.
  - “Other” captures different group of people at each census due to changes in wording of question through decades.
  - “Other” is also inconsistently applied across datasets, so these numbers are difficult to interpret regardless of denominator definition

Challenges with Implementing Estimates

2. Totals

- Estimates were modeled at the smallest cell (though "all race" and "all ethnicities" were modeled as well)
- Race re-distribution and rounding create inconsistencies depending on aggregation parameters
  - Geographies: County aggregation inconsistent between census tract and Municipality
    - Creates some problems for "Special Geographies" used by the state
  - Age: single year vs. 5-year
  - Individual races/ethnicity vs. all races/ethnicity categories

Challenges with Implementing Estimates

3. Moving Estimates

- With each vintage of County Population census data, estimates are updated for all years 2010 – most current vintage.
- Creates changes in denominators over time, and altered rates as a result.
  - Most datasets lag enough that this shouldn't impact resulting rates
  - Not the case for Lead in MA