Total number of iterations: 15000

Burn-in period: 5000

Thinning: 10

Kmax = 50

Dirichlet prior with gamma = 1/Kmax

Estimated posterior distribution of the number of clusters for each survey wave.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ACS Survey Year** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** |
| 2006-2010 | 0 | 0 | 0 | **0.397** | 0.383 | 0.158 | 0.055 | 0.007 |
| 2011-2015 | 0 | 0.001 | 0.279 | **0.387** | 0.206 | 0.074 | 0.039 | 0.014 |
| 2015-2019 | 0.124 | 0.132 | **0.269** | 0.258 | 0.137 | 0.063 | 0.014 | 0.003 |

Estimated number of observations per cluster conditionally on **mapK** (3 label switching algorithms); however, the model outputs the posterior mean of the probability of success per feature and cluster using the ECR algorithm.

|  |  |  |  |
| --- | --- | --- | --- |
| **Cluster** | **2006-2010\*** | **2011-2015\*\*** | **2015-2019** |
| 1 | 109 | 23 | 432 |
| 2 | 343 | - | 40 |
| 3 | 110 | 73 | 174 |
| 4 | 137 | 416 | 371 |
| 5 | 140 | 158 | 11 |
| 6 | 274 | 355 | 158 |
| 7 | 85 | 129 | 176 |
| 8 | 201 | 180 | 116 |
| 9 | 79 | 144 | -- |

\*Same frequencies using either label switching algorithm

\*\* Varies across label switching algorithms- cluster 5 vs. 2

* Moved 2 census tracts from cluster 2 in 2011-2015 to cluster 5 based on second highest assignment probability

## ACS 2006-2010

**Mixing weights**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 2006-2010 | | | |
|  | **Mean** | **SD** | **2.5%** | **97.5%** |
| p.1 | 0.07 | 0.02 | 0.04 | 0.12 |
| p.2 | 0.21 | 0.03 | 0.16 | 0.26 |
| p.3 | 0.08 | 0.02 | 0.04 | 0.12 |
| p.4 | 0.11 | 0.03 | 0.07 | 0.16 |
| p.5 | 0.08 | 0.01 | 0.06 | 0.12 |
| p.6 | 0.17 | 0.02 | 0.12 | 0.21 |
| p.7 | 0.07 | 0.02 | 0.04 | 0.11 |
| p.8 | 0.14 | 0.01 | 0.11 | 0.16 |
| p.9 | 0.06 | 0.02 | 0.03 | 0.09 |

**Distribution of assignment probabilities**

A graph with different colored rectangular shapes

Description automatically generated with medium confidence

|  | **cluster1 (N=109)** | **cluster2 (N=343)** | **cluster3 (N=110)** | **cluster4 (N=137)** | **cluster5 (N=140)** | **cluster6 (N=274)** | **cluster7 (N=85)** | **cluster8 (N=201)** | **cluster9 (N=79)** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Average Assignment Probability (ECR Algorithm)** |  |  |  |  |  |  |  |  |  |
| Mean (SD) | 0.682 (0.162) | 0.806 (0.173) | 0.746 (0.176) | 0.755 (0.133) | 0.711 (0.194) | 0.798 (0.181) | 0.773 (0.180) | 0.916 (0.139) | 0.695 (0.176) |
| Median [Min, Max] | 0.705 [0.336, 0.958] | 0.873 [0.175, 0.978] | 0.773 [0.161, 0.993] | 0.777 [0.472, 0.986] | 0.660 [0.343, 0.987] | 0.856 [0.167, 0.994] | 0.828 [0.372, 0.996] | 0.976 [0.302, 0.999] | 0.724 [0.381, 0.929] |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Cluster | Cluster descriptions and distribution of age across census tracts  For this, we are also going to look at the distribution of age and race/ethnicity across census tracts within the clusters | | |  |
|  | **Description** | **Age distribution** | **Race/ethnicity composition of CTs in the cluster** | **Conclusion/Summary**  **\*Also looked at the distribution of language**  **\*all at the census tract level** |
| 1 | High success probabilities for:   * Renter * White collar occupation * < HS * >= Bachelor’s * TPPR   Low house ownership, no vehicle, poverty line | Avg median age is 33.2.  Higher average proportions in the age groups:   * 20-24 * 25-34 * 35-44 * 45-54 | * Lowest avg proportion of NHW compared to other clusters.   Highest avg proportion of Hispanic or Latinos compared to other clusters.   * Second highest avg proportion of NHB | Educated  Low income (at least compared to the state median)  Multi-lingual/multi-cultural (based on distribution of language spoken at home other than EN or speak EN not “very well”  Household crowding (Crowding among housing units) |
| 2 | High success probabilities for:   * Renter * TPPR * Lack of complete plumbing + < HS   Low probabilities for all other variables ranging from 0.3-0.04. | Avg median age is 34.4.  Higher average proportions in the age groups:   * 25-34 * 35-44 * 45-54 | * Second Highest avg proportion of Hispanic or Latinos compared to other clusters. * Highest avg proportion of NHB | Low education  Low income  Multi-lingual/multi-cultural (based on distribution of language spoken at home other than EN or speak EN not “very well”  Household crowding |
| 3 | High success probabilities for:   * Owner   Medium probabilities for:   * < HS * SNAP benefits * Female household | Avg median age is 41.4.  Higher average proportions in the age groups:   * 35-44 * 45-54 | Not very diverse, on average the proportion of NHW is 85.4 | Low income  Female households  Older than above |
| 4 | High success probabilities for:   * Owner * >= Bachelors * >= HS * Median income * White collar occupation | Avg median age is 43.0.  Higher average proportions in the age groups:   * 35-44 * 45-54 | Similar distribution to the above, majority NHW (89.6) | Older in comparison to previous cluster, but seem to be more affluent or higher SES level based on the avg proportion of house ownership. |
| 5 | High probabilities for:   * Owner * >= HS * Median Income   Low for:   * >= Bachelor’s degree * White collar occupation * Renter | Avg median age is 42.6  Higher average proportions in the age groups:   * 35-44 * 45-54 | Similar distribution to the above, majority NHW (93.1) | High income  At least HS educated |
| 6 | High probabilities for:   * Owner * >= Bachelor's degree * >= HS * white collar occupation | Avg median age is 42.9  Higher average proportions in the age groups:   * 35-44 * 45-54 | Similar distribution to the above, majority NHW | Highly educated.  Medium-high income  House ownership |
| 7 | High probabilities for:   * Renter * < HS * Female household * SNAP * Below poverty | Avg median age is 37.9.  Higher average proportions in the age groups:   * 25-34 * 35-44 * 45-54 | Lower avg proportion of NHW compared to previous clusters except cluster 1  Decent avg proportion of NHB and Hisp | High unemployment (proportions compared to the median for the state)  Low ed  Female household  Higher avg of gov assistantship (SNAP)  Decent multi-language/ multi-culture  BIG renters |
| 8 | High probabilities for:   * Renter * >= Bachelor’s degree * White collar occupation * >= HS * Median income * TPPR | Avg median age is 35.9.  Higher average proportions in the age groups:   * 25-34 * 35-44 * 45-54 | Higher avg proportion of NHA than any other minority group | 2nd highest renters  Highly educated and occupation  Medium-high income  Multi-lingual (based on avg proportion of language spoken at home other than EN) |
| 9 | High probabilities for:   * Owner * < HS * Median income * unemployment | Avg median age is 40.8.  Higher average proportions in the age groups:   * 25-34 * 35-44 * 45-54 | majority NHW (84.9) | Low ed  House ownership  High unemployment proportions |

For variables: lack of plumbing and two or more rooms, 0 means there is no lack and no households with two or more ppr, respectively.

A graph of progress on a graph

Description automatically generated with medium confidence

Re-order heatmap in descending theta\_kj. This way we can see factors that are considered “more” important across clusters.

A colorful squares with black text

Description automatically generated

## ACS 2011-2015

About the survey:

* population grew 3.5% from 2006-2010 to 2011-2015.
* The two tracts in cluster 2: 25013812902 (Census Tract 8129.02, Hampden County, Massachusetts) & 25021415102 (Census Tract 4151.02, Norfolk County, Massachusetts)
* Looking at these in previous ACS, they were classified as cluster 4 and cluster 8, respectively.
* <https://www.census.gov/programs-surveys/acs/guidance/comparing-acs-data/2015/5-year-comparison.html>
* **Change in geographic boundaries** - The 2006-2010 ACS 5-year estimates used legal boundaries as of January 1, 2010. The 2011-2015 ACS 5-year estimates use legal boundaries as of January 1, 2015.—I think this has been adjusted from what I’ve seen so far
* **Also changes in the questionnaire or coding—this is a strength of our study because we are looking at each survey separately!!**
* ACTION: Moved 2 census tracts from cluster 2 in 2011-2015 to cluster 5 based on second highest assignment probability

**Mixing weights**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Mean** | **SD** | **2.5%** | **97.5%** |
| p.1 | 0.0428821 | 0.0267901 | 0.0114452 | 0.1134180 |
| p.2 | 0.0367371 | 0.0457542 | 0.0000000 | 0.1258357 |
| p.3 | 0.0520965 | 0.0147140 | 0.0307959 | 0.0869562 |
| p.4 | 0.2540061 | 0.0223038 | 0.2074207 | 0.2917182 |
| p.5 | 0.0935849 | 0.0186718 | 0.0596165 | 0.1319348 |
| p.6 | 0.2103892 | 0.0364540 | 0.1439822 | 0.2743809 |
| p.7 | 0.0921025 | 0.0181061 | 0.0553134 | 0.1272803 |
| p.8 | 0.1230779 | 0.0133807 | 0.1001930 | 0.1497132 |
| p.9 | 0.0951237 | 0.0160407 | 0.0669887 | 0.1298315 |

Note that cluster 2 only contains 2 census tracts

**Distribution of assignment probabilities—we moved the 2 cts in cluster 2 to cluster 5 based on second highest assignment probability and relabeled the clusters**

A graph of different colored rectangular shapes

Description automatically generated with medium confidenceA graph with different colored rectangular shapes

Description automatically generated

|  | **cluster1 (N=23)** | **cluster2 (N=2)** | **cluster3 (N=73)** | **cluster4 (N=416)** | **cluster5 (N=156)** | **cluster6 (N=355)** | **cluster7 (N=129)** | **cluster8 (N=180)** | **cluster9 (N=144)** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Average Assignment Probability (ECR Algorithm)** |  |  |  |  |  |  |  |  |  |
| Mean (SD) | 0.588 (0.105) | 0.323 (0.00498) | 0.713 (0.189) | 0.865 (0.170) | 0.665 (0.202) | 0.817 (0.190) | 0.763 (0.148) | 0.914 (0.130) | 0.834 (0.179) |
| Median [Min, Max] | 0.570 [0.425, 0.817] | 0.323 [0.319, 0.326] | 0.689 [0.333, 0.982] | 0.960 [0.242, 0.991] | 0.741 [0.349, 0.898] | 0.896 [0.218, 0.980] | 0.782 [0.297, 0.951] | 0.975 [0.400, 0.992] | 0.899 [0.288, 0.998] |

|  | **cluster1 (N=23)** | **cluster3 (N=73)** | **cluster4 (N=416)** | **cluster5 (N=158)** | **cluster6 (N=355)** | **cluster7 (N=129)** | **cluster8 (N=180)** | **cluster9 (N=144)** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Average Assignment Probability (ECR Algorithm)** |  |  |  |  |  |  |  |  |
| Mean (SD) | 0.588 (0.105) | 0.713 (0.189) | 0.865 (0.170) | 0.660 (0.205) | 0.817 (0.190) | 0.763 (0.148) | 0.914 (0.130) | 0.834 (0.179) |
| Median [Min, Max] | 0.570 [0.425, 0.817] | 0.689 [0.333, 0.982] | 0.960 [0.242, 0.991] | 0.737 [0.247, 0.898] | 0.896 [0.218, 0.980] | 0.782 [0.297, 0.951] | 0.975 [0.400, 0.992] | 0.899 [0.288, 0.998] |

Relabeled

|  | **cluster1 (N=23)** | **cluster2 (N=73)** | **cluster3 (N=416)** | **cluster4 (N=158)** | **cluster5 (N=355)** | **cluster6 (N=129)** | **cluster7 (N=180)** | **cluster8 (N=144)** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Average Assignment Probability (ECR Algorithm)** |  |  |  |  |  |  |  |  |
| Mean (SD) | 0.588 (0.105) | 0.713 (0.189) | 0.865 (0.170) | 0.660 (0.205) | 0.817 (0.190) | 0.763 (0.148) | 0.914 (0.130) | 0.834 (0.179) |
| Median [Min, Max] | 0.570 [0.425, 0.817] | 0.689 [0.333, 0.982] | 0.960 [0.242, 0.991] | 0.737 [0.247, 0.898] | 0.896 [0.218, 0.980] | 0.782 [0.297, 0.951] | 0.975 [0.400, 0.992] | 0.899 [0.288, 0.998] |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Cluster | Cluster descriptions and distribution of age across census tracts  For this, we are also going to look at the distribution of age and race/ethnicity across census tracts within the clusters | | |  |
|  | **Description** | **Age distribution** | **Race/ethnicity composition of CTs in the cluster** | **Conclusion** |
| 1 | High success probabilities for:   * Renter * Female household   Medium probabilities for:   * TPPR * < HS * Unemployment * Below poverty line | Avg median age is 40.4.  Higher average proportions in the age groups:   * 25-34 * 35-44 * 45-54 | majority NHW, and not very diverse | Renters  Female households  Low education  Crowded housing |
| ~~2~~ | ~~Medium probabilities for all variables but highest for~~   * ~~>= HS~~ * ~~Owner~~ * ~~>=Bachelors~~ * ~~Median income~~ | ~~Avg median age is 41.6.~~  ~~Higher average proportions in the age groups:~~   * ~~25-34~~ * ~~35-44~~ * ~~45-54~~ | ~~majority NHW, and not very diverse~~ | ~~Higher Ed~~  ~~high Income~~  ~~Note that there are only 2 census tracts in this cluster, and based on distribution of main variables I think they can be combined with cluster 6~~ |
| 3  2 | High success probabilities for:   * Owner * <HS * >= Bachelor’s degree * < HS * Median income | Avg median age is 43.2.  Same majority age groups as above | majority NHW, and not very diverse | House ownership  High ed  High income |
| 4  3 | High success probabilities for:   * Renter * Unemployment * < HS * TPPR | Avg median age is 35.4.  Same majority age groups as above | * Highest avg proportion of Hispanic or Latinos compared to other clusters. * Highest avg proportion of NHB | Low ed  Crowded housing  Low income  Multi-language/multi-cultural (EN not very well and other language other than EN)  Similar to clusters 1 &2 in 2006-2010 survey—further exploration through cross-tabulations |
| 5  4 | High probabilities for:   * Owner * >= HS | Avg median age is 44.6.  Same majority age groups as above | Similar distribution to the above, majority NHW (91.3) | At least HS educated.  House ownership |
| 6  5 | High probabilities for:   * Owner * >=bachelor’s degree * White collar * >= HS | Avg median age is 44.4.  Same majority age groups as above | Similar distribution to the above, majority NHW | Higher ed  House ownership |
| 7  6 | High probabilities for:   * >=Bachelor’s degree * Renter * White collar occupation * < HS * TPPR | Avg median age is 35.5.  Same majority age groups as above | 2nd highest avg proportion of Hispanic or Latinos and NHB compared to other clusters. | Higher ed  Renters  Multi-lingual  Crowding among housing units |
| 8  7 | High probabilities for:   * Renter * >= Bachelor’s degree * White collar occupation * >= HS * TPPR | Avg median age is 35.7.  Same majority age groups as above | Higher avg proportion of NHA than any other minority group | Multi-lingual  Crowding  High education  Med-high income |
| 9  8 | High probabilities for:   * Owner * < HS | Avg median age is 42.6.  Same majority age groups as above | majority NHW (82.5) | House ownership  Low ed  Low income |

A graph of progress bar

Description automatically generated with medium confidence

A colorful squares with numbers

Description automatically generated with medium confidence

A colorful squares with numbers

Description automatically generated with medium confidence

## ACS 2015-2019

**Mixing weights**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Mean** | **SD** | **2.5%** | **97.5%** |
| p.1 | 0.2806157 | 0.0139065 | 0.2518628 | 0.3068842 |
| p.2 | 0.0354650 | 0.0174421 | 0.0000000 | 0.0633357 |
| p.3 | 0.1124320 | 0.0139981 | 0.0836987 | 0.1391214 |
| p.4 | 0.2482094 | 0.0167872 | 0.2143773 | 0.2815114 |
| p.5 | 0.0117289 | 0.0105344 | 0.0000000 | 0.0356347 |
| p.6 | 0.1094264 | 0.0155569 | 0.0824268 | 0.1451468 |
| p.7 | 0.1154900 | 0.0135443 | 0.0892492 | 0.1433546 |
| p.8 | 0.0866326 | 0.0134639 | 0.0623886 | 0.1155193 |

**Distribution of assignment probabilities**

A graph with different colored squares

Description automatically generated with medium confidenceA graph with different colored squares

Description automatically generated

|  | **cluster1 (N=432)** | **cluster2 (N=40)** | **cluster3 (N=174)** | **cluster4 (N=371)** | **cluster5 (N=11)** | **cluster6 (N=158)** | **cluster7 (N=176)** | **cluster8 (N=116)** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Average Assignment Probability (ECR Algorithm)** |  |  |  |  |  |  |  |  |
| Mean (SD) | 0.937 (0.136) | 0.560 (0.140) | 0.862 (0.156) | 0.926 (0.118) | 0.573 (0.0601) | 0.850 (0.157) | 0.884 (0.148) | 0.853 (0.170) |
| Median [Min, Max] | 0.994 [0.223, 0.999] | 0.545 [0.310, 0.785] | 0.934 [0.289, 0.996] | 0.968 [0.276, 0.986] | 0.589 [0.415, 0.635] | 0.925 [0.308, 0.990] | 0.962 [0.351, 0.989] | 0.935 [0.329, 0.994] |

Relabed clusters

|  | **cluster1 (N=432)** | **cluster2 (N=40)** | **cluster3 (N=180)** | **cluster4 (N=375)** | **cluster5 (N=159)** | **cluster6 (N=176)** | **cluster7 (N=116)** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Average Assignment Probability (ECR Algorithm)** |  |  |  |  |  |  |  |
| Mean (SD) | 0.937 (0.136) | 0.560 (0.140) | 0.845 (0.180) | 0.920 (0.134) | 0.846 (0.163) | 0.884 (0.148) | 0.853 (0.170) |
| Median [Min, Max] | 0.994 [0.223, 0.999] | 0.545 [0.310, 0.785] | 0.927 [0.289, 0.996] | 0.968 [0.219, 0.986] | 0.923 [0.277, 0.990] | 0.962 [0.351, 0.989] | 0.935 [0.329, 0.994] |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Cluster | Cluster descriptions and distribution of age across census tracts.  For this, we are also going to look at the distribution of age and race/ethnicity across census tracts within the clusters | | |  |
|  | **Description** | **Age distribution** | **Race/ethnicity composition of CTs in the cluster** | **Conclusion** |
| 1 | High success probabilities for:   * Renter   Medium probs for:   * TPPR * <HS * Unemployment | Avg median age is 36.1  Higher average proportions in the age groups:   * 25-34 * 35-44 * 45-54 | * Highest avg proportion of Hispanic or Latinos compared to other clusters. * Highest avg proportion of NHB | Multi-cultural/ multi-lingual  Renters  Low education  Low income |
| 2 | High success probabilities for:   * Renter * TPPR * Lack of complete plumbing + < HS   Low probabilities for all other variables ranging from 0.3-0.04. | Avg median age is 38.6  Higher average proportions in the age groups:   * 20-24 * 25-34 * 35-44 * 45-54 | * Second Highest avg proportion of Hispanic or Latinos and NHB compared to other clusters. | Multi-cultural  Crowding among housing units  Low education |
| 3 | High success probabilities for:   * Owner * < HS   Medium probabilities for:   * Female household * Below poverty * unemployment | Avg median age is 43.7  Higher average proportions in the age groups:   * 25-34 * 35-44 * 45-54 | Clusters 3-7 have similar distributions the one thing to notice | House ownership  Low education  Low income |
| 4 | High success probabilities for:   * Owner * >= Bachelors * >= HS * White collar occupation   More affluent census tracts? | Avg median age is 44.6  Higher average proportions in the age groups:   * 35-44 * 45-54 | Majority NHW not much diversity | House ownership  High education  High income  (similar to cluster 4 in 2006-2010)  Affluent CTs |
| ~~5~~ | ~~High probabilities for:~~   * ~~Owner~~ * ~~>= bachelor’s~~ * ~~Female household~~ * ~~White collar occupation~~ * ~~Below poverty~~ | ~~Avg median age is 40.1~~  ~~Higher average proportions in the age groups:~~   * ~~25-34~~ * ~~35-44~~ * ~~45-54~~ | ~~Decent proportion of NHB~~ | ~~Higher education~~  ~~Female households~~  ~~Only 11 Cts here and all have avg proportion below poverty line >= median value for the state.~~ |
| 6  5 | High probabilities for:   * Owner * >= HS   Low for everything else | Avg median age is 46.7  Higher average proportions in the age groups:   * 35-44 * 45-54 | Majority NHW not much diversity | Not multi-lingual (based on distribution of language)  House ownership  At least HS education |
| 7  6 | High probabilities for:   * >= Bachelors * White collar occupation * >= HS * Renter | Avg median age is 35.6  Higher average proportions in the age groups:   * 20-24 * 25-34 * 35-44 * 45-54 | Decent proportion of Hispanics | Renters  High education  Multi-lingual |
| 8  7 | High probabilities for:   * Renter * >= Bachelor’s degree * White collar occupation * < HS * Median income * TPPR | Avg median age is 35.8  Higher average proportions in the age groups:   * 25-34 * 35-44 * 45-54 | * Second Highest avg proportion of Hispanic or Latinos and NHB compared to other clusters. | Multi-lingual (similar distribution of language as cluster 1)  Combination of high ed and < HS  Crowding |