Memo: Population change methods for HIA based on gentrification and displacement typology

For Bullitt foundation grant

From: Lisa K. Bates, PhD; Seyoung Sung, GRA. Portland State University

Date: 09-02-16

Goal:

Estimate GHG changes caused by population movements associated with gentrification and displacement of low-income households/in-migration of high-income households.

Steps:

1. Assign each neighborhood a position in a typology of neighborhood change.
2. Use regression analysis of the 2000-2015 time period to estimate probabilities of a neighborhood experiencing displacement/replacement
3. Predict probabilities of neighborhood change for future time period based on 2015 conditions
4. Use past period population change and predicted probabilities to estimate expected values for population migration
5. [somebody else will use these to estimate GHG]
6. Assign each neighborhood a position in a typology of neighborhood change:

The Updated Gentrification + Displacement Typology v.2016 builds upon the Typology v.2013 and subsequent work by UC Berkeley’s Urban Displacement Lab to develop a typology that:

* tracks to the v.2013 typology definitions of vulnerability and demographic changes;
* utilizes more housing and real estate data to update for changing market conditions;
* Attempts to assign a type to all tracts in the city including those that are not undergoing substantial change in their character.

The Typology v.2016 is (like the v.2013) based on how conditions in a census tract compare to the citywide median, in an analysis of aggregate status shifts. The Typology v.2016 captures residents’ vulnerability to displacement, changes in the overall demographics of the tract, housing market changes, and the rates of in- and out-migration of households of different income levels.

*Simplified categorization for displacement modeling*

This categorization looks at tract income levels and demographic change as the primary indicators for typing.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Type | Tract Income in Year 1 | Tract  Income in Year 2 | Vulnerable Population in Year 1 | Demographic Change between Year 1 and 2 | Vulnerable Population in Year 2 |
| A | Low | Low | Yes | No | Yes |
| B | Low | Low | Yes | Yes | Yes |
| C | Low | - | Yes | Yes |  |
| D | - | MH |  | Yes  (Either 2000-2010 or 2010-2015) |  |
| E | MH | MH |  |  | No |
| F | MH | MH | No |  | No |

Type A

Persistently low-income areas with vulnerable populations.

Type B

At risk of gentrification due to demographic changes, but remains low income with vulnerable populations.

Type C

Includes multiple dimensions of gentrification and displacement. Starting incomes are low and demographic change is significant.

Type D-E-F

Maintains the moderate-high income level during the time period, with limited or no vulnerable populations. Type D continues to have demographic change with respect to vulnerable populations; types E and F do not have significant vulnerable groups present.

2. Use regression analysis of the 2000-2015 time period to estimate probabilities of a neighborhood experiencing displacement/replacement

The regression analysis is designed to be parsimonious to the predictors of a change in neighborhood status and to align with the interests of this project, which is to consider how transit investments affect household moves due to neighborhood market changes (and ultimately how those moves affect GHG emission).

We tested several models for predicting the probability of displacement, using different dependent variables, using a logit link function suitable for dichotomous dependent variables.

Predictor variables are:

* The 2010 typology category. This is expected to soak up most of the variance in the model, as the initial type should be the most direct predictor of the 2015 type.
* Transit stations. As this model is related to new transit investments, directly analyzing the impacts of being in a transit station area/ new TOD on gentrification and displacement is useful.
* New market units > median. New housing units may reduce displacement as in-migrants can utilize this supply rather than displacing existing residents through out-bidding on housing.
* New subsidized units = yes. New subsidized units are not frequently brought online, but with income restrictions on occupancy they will reduce displacement. It is helpful to understand what the impact of income restricted housing is separately from market rate housing.
* Price appreciation > median
* Rent appreciation > median

The final model is based on a prediction of the loss of low-income households in the tract.

Logit models result in coefficients that are interpreted as “odds ratios” --these can be somewhat confusing to interpret but should be understood as the increased/decreased likelihood of an event, given a predictor condition. In this case we are predicting the odds or likelihood of the loss of low-income households in a tract. The likelihood of loss of low-income households given the typology status is understood as the likelihood compared to a tract in type D-E-F (a higher income tract).

Model results in words:

* Tracts in Type A, B, C are more likely to have loss of low-income households than high-income tracts.
* Tracts with a light rail transit station are slightly less likely to have loss of low-income households.(10% less likely)
* Tracts with substantial new market rate construction are 15% less likely to have loss of low-income households.
* Tracts with new subsidized units are much less likely to have loss of low-income households.
* Tracts where housing prices are appreciating greater than average are twice as likely to have loss of low-income households.
* Tracts where rents are being raised greater than average are slightly less likely to have loss of low-income households.

Potentially counter-intuitive results around displacement of low-income households from low-income tracts and areas with increasing rents are probably understood by considering the low-cost housing market. One, households are leaving the City of Portland for cheaper housing elsewhere, even from predominantly low income tracts--they are leaving the analysis space altogether. This combines with a second factor. Second, rents at both the high and low end of the market are increasing substantially right now. Households with low incomes will increase their cost burden for housing to maintain rental housing even when rents go up and become unaffordable, as we can see in the very high rates of cost burden for low income renters.

3. Predict probabilities of neighborhood change for future time period based on 2015 conditions

The logit model results in beta coefficients that are interpreted as odds-ratios, which can be challenging to understand as predictors. Therefore we use the coefficients to calculate predicted probabilities for each tract.

Predicted probability = e^(b0+b1x1+b2x2…+bnxn) / 1 + [e^(b0+b1x1+b2x2…+bnxn)]

We use the model to PROJECT a predicted probability for change in the coming time period, based on change in the previou time period. Each tract has a projected predicted probability for losing low-income households that is based on the 2015 typology category.

We have also provided a projected predicted probability for each tract *if it were to have a light rail station* by estimating the probability based on transit=1.

4. Use past period population change and predicted probabilities to estimate expected values for population migration

We find the expected value of population change for each tract as the product of the projected probability of population change for each tract with the characteristics of population change for tracts \*of that type\* for 2010-2015.

The average population change for 2010-2015 by household size, income, and age (the HIA) is calculated by typology category (ABCDEF). That same change happens going forward, with the probability estimated by the model, is the expected value of change.

We assign this proportionally to TAZ based on centroid location within a tract.

5. Creation of new HIA categories

Expected value of projected population change is available for ONLY SOME of the HIA categories. The HIA has 64 categories of household size, income, and age. However, the data we are using to understand low income housing and household matching has different categories for household size, income, and age. The reconciliation of categories results in 14 distinct household types that align with the HIA.

* Household size: CHAS data have Households with 2,3,4; or households of 5 and more. We include the former under household size 3 and the latter labeled under household size 5 and applied to H4

|  |  |
| --- | --- |
| **H** | **household size** |
| 1 | 1 |
| 3 | 2 to 4 |
| 5 | 5 or more |

* Household income: CHAS data adjusts household income as a percentage of median by household size. We have reallocated households of different sizes into the income categories of the HIA based on the median point.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Income Limits by HUD (for CHAS data)** | | | | **Corresponding table with the income range by Metro HIA 2015** | | | |
| CHAS 2009-2013 | HH size | | |  | HH size | | |
| 1 | 3 | 5 | 1 | 3 | 5 |
| below 30% | 15150 | 19450 | 23350 | below 30% | i1 | i1 | i1 |
| below 50% | 25200 | 32400 | 38900 | below 50% | i1 | i2 | i3 |
| below 80% | 40350 | 51850 | 62250 | below 80% | i3 | i3 | i3 |
| median income |  | 66240 | 77760 | above 80% | i3 | i4 | i4 |

* Household age: CHAS data addresses elderly and non-elderly households only. We have labeled these non-elderly households as a0 in our dataset and would defer to the modeling team to assign them to a category.

|  |  |
| --- | --- |
| **A** | **Household type** |
| 0 | Elderly (age 62 or over) |
| 4 | Non-elderly |

Projected expected value of population change by HIA(14 types) are below. These figures add to the 2015 HIA to create final population numbers for the projection.

|  |
| --- |
| **Newly created HIA categories** |
| h1i1a0  h1i1a4  h1i3a0  h1i3a4  h3i1a0  h3i1a4  h3i2a0  h3i2a4  h3i3a0  h3i3a4  h5i1a0  h5i3a0  h5i4a0 |

6. Reconciliation of HIA categories

HIA categories originally from Metro was adjusted to 14 new HIA categories by following the order.

1) Check whether the category has senior person (62+)

2) Match categories by size

3) Match categories by income level

Due to using different source of the data for the projection, re-categorization was performed as closely as possible to the original HIA categorization. The table shows rearranged HIA categories.

|  |  |
| --- | --- |
| **HIA by Metro** | **HIA for the Projection** |
| h1i1a1 | h1i1a0 |
| h1i1a2 | h1i1a0 |
| h1i1a3 | h1i1a0 |
| h1i1a4 | h1i1a4 |
| h1i2a1 | h1i3a0 |
| h1i2a2 | h1i3a0 |
| h1i2a3 | h1i3a0 |
| h1i2a4 | h1i3a4 |
| h1i3a1 | h1i3a0 |
| h1i3a2 | h1i3a0 |
| h1i3a3 | h1i3a0 |
| h1i3a4 | h1i3a4 |
| h1i4a1 | h1i3a0 |
| h1i4a2 | h1i3a0 |
| h1i4a3 | h1i3a0 |
| h1i4a4 | h1i3a4 |
| h2i1a1 | h3i1a0 |
| h2i1a2 | h3i1a0 |
| h2i1a3 | h3i1a0 |
| h2i1a4 | h3i1a4 |
| h2i2a1 | h3i2a0 |
| h2i2a2 | h3i2a0 |
| h2i2a3 | h3i2a0 |
| h2i2a4 | h3i2a4 |
| h2i3a1 | h3i3a0 |
| h2i3a2 | h3i3a0 |
| h2i3a3 | h3i3a0 |
| h2i3a4 | h3i3a4 |
| h2i4a1 | h3i4a0 |
| h2i4a2 | h3i4a0 |
| h2i4a3 | h3i4a0 |
| h2i4a4 | h3i3a4 |
| h3i1a1 | h3i1a0 |
| h3i1a2 | h3i1a0 |
| h3i1a3 | h3i1a0 |
| h3i1a4 | h3i1a4 |
| h3i2a1 | h3i2a0 |
| h3i2a2 | h3i2a0 |
| h3i2a3 | h3i2a0 |
| h3i2a4 | h3i2a4 |
| h3i3a1 | h3i3a0 |
| h3i3a2 | h3i3a0 |
| h3i3a3 | h3i3a0 |
| h3i3a4 | h3i3a4 |
| h3i4a1 | h3i4a0 |
| h3i4a2 | h3i4a0 |
| h3i4a3 | h3i4a0 |
| h3i4a4 | h3i3a4 |
| h4i1a1 | h5i1a0 |
| h4i1a2 | h5i1a0 |
| h4i1a3 | h5i1a0 |
| h4i1a4 | h3i1a4 |
| h4i2a1 | h5i1a0 |
| h4i2a2 | h5i1a0 |
| h4i2a3 | h5i1a0 |
| h4i2a4 | h3i2a4 |
| h4i3a1 | h5i3a0 |
| h4i3a2 | h5i3a0 |
| h4i3a3 | h5i3a0 |
| h4i3a4 | h3i3a4 |
| h4i4a1 | h5i4a0 |
| h4i4a2 | h5i4a0 |
| h4i4a3 | h5i4a0 |
| h4i4a4 | h3i3a4 |

7. Calculation of Projection

1) Aggregated the HIA 2015 data prepared by Metro from TAZ to Tract

2) Combined tract level HIA 2015 data followed by the 14 categories

e.g. h1i1a0 =h1i1a1+ h1i1a2+ h1i1a3

3) had proportions for each category (64 categories) as a matrix by category and tract

e.g. The proportion of original population number of H1i1a1 to h1i1a0

= h1i1a1/ (h1i1a1+ h1i1a2+ h1i1a3)

4) In order to calculate the expected value, multiplication of the projected probability of displacement by the population change for tracts of that type was performed.

5) Projection, the number of population change, was added up the base.

6) The final population number for HIA in 14 categories was disaggregated to original 64 categories through multiplying the final number by the proportions (step 3)

7) The final projection number was disaggregated at the TAZ level.

8) Zero out all negative numbers was done for the final projection.

8. [somebody else will use these to estimate GHG]

DATASETS:

Excel workbook is RE\_HIA\_2015\_fullset

Individual sheets:

PRprobsDISPLACE is projected probability that a tract will experience low income household displacement

AVE.CHANGE(%)by types calculates the population change (%) actually observed in tracts of Type (ABCDEF) during 2010-2015

Expected value(change pop) is the projected probability of displacement \* the population change for tracts of that type, by tracts (each TAZ is assigned the expected value --which is a percent change in population --by the tract in which its centroid is found)

[RE\_HIA2015\_BASE is the data prepared by Metro for 2015 TAZ populations]

RE\_HIA\_HH\_CH\_PROJECTION is the number (+ or -) of population \*change\* from the base

RE\_HIA\_PROJECTION\_FINAL is the final population number for 14 HIA categories in each TAZ (base data + projected change data)

RE\_HIA 2015\_RESULTS\_64COLUMNS is the final population number for 64 HIA categories in each TAZ.

RE\_HIA 2015\_RESULTS\_ZERO OUT is the final result that all negative numbers were zero out for 64 categories in each TAZ.