



Methodological notes for the Los Angeles Indices of Neighborhood Change

Contents

- 1) Overview
- 2) Los Angeles Index of Neighborhood Change Methodology
- 3) Los Angeles Displacement Pressure Index Methodology

Overview

The Los Angeles Indices of Neighborhood Change is a project of the Los Angeles Innovation Team, as part of its wider efforts to develop City programs to reduce displacement in revitalizing areas.

Like many cities around the United States and world, Los Angeles has struggled with reconciling the advantages of revitalization with the displacement that often accompanies investment. Even more so than other major US Cities, Los Angeles has experienced a rapid rise in rents without a proportional increase in income. Since 2000, median gross rent in Los Angeles has gone up nearly 40% when adjusting for inflation. Meanwhile, median household income has only gone up 6% over the same time period. Nearly 60% of rental households (over 510,000 households) are now considered "rent-burdened." Nearly 1/3 (31.8% or over 281,000 households) are severely rent burdened, paying over 50% of their household incomes on rent. (Sources: 2000 Decennial Census; 2016 American Community Survey, 1-Year Sample)

As post-recession investment returns and spreads to new neighborhoods throughout the region, once neglected areas have become desirable to residents/businesses new to LA and those priced out of more affluent areas. In neighborhoods with little to no new housing/commercial construction over the past 30 years, the lack of supply leads to higher market prices for available units. Although the Innovation Team's goal was to introduce programs to mitigate the effects of these market surges and ensure existing residents also enjoyed the fruits of investment, we initially operated with very little visibility into where change was occurring, beyond public opinion. To address this, we developed two measures that allowed us to compare current and anticipated demographic change in Los Angeles' neighborhoods:

- The Los Angeles Index of Neighborhood change, which compares six metrics across Los Angeles ZIP Codes.

- The Los Angeles Displacement Pressure Index, which predicts where change is likely to occur in the near future at the Census Tract level, with additional emphasis on residents most likely to face high displacement pressures as a result.

Both indices were developed with a careful eye towards the academic literature on the subject, and underwent several revisions based on new data releases and incremental methodological improvements.

Los Angeles Index of Neighborhood Change Methodology

Features

The Los Angeles Index of Neighborhood Change (LAINC) compares change among six factors across Los Angeles city ZIP Codes between 2000 and 2014. These factors are in line with past research on the topics of gentrification and displacement and include:

1. % change in Ratio of Low Income to High Income tax filers (from 2005 to 2013).
We define low and high income as follows:

Low Income: $\leq \$25K$ Adjusted Gross Income tax filers who also received an earned income tax credit. This amount represents approximately 50% of the Los Angeles Household Median Income.

High Income: $\geq \$75K$ Adjusted Gross Income. This amount represents approximately 50% above the Los Angeles Household Median Income.

2. % change in Median Household Income (from 2000 to 2014)
3. % change in Average Household Size (from 2000 to 2013)
4. Change in % of White Non-Hispanic residents (from 2000 to 2014)
5. Change in % of residents ≥ 25 with Bachelor's Degrees (from 2000 to 2014)
6. % change in Median Gross rent (from 2000 to 2014)

Sources: 2000 Decennial Census | 2013&2014 American Community Survey (5-Year Estimate) | 2005 and 2013 IRS Tax Return Data

Normalization and Weighting

The above features were normalized using the following formula

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Samples were further weighted based on the pre-normalized p-value of the data points. The normalized sample was multiplied by the maximum confidence level at which the comparison was statistical significant. For instance, if a ZIP Code’s household change indicator was confident at 99%, it received a .99 weighting. All samples that were not significant at even an 80% confidence threshold received a weighting of .5.

Following the weighting of individual measures, we adjusted each feature based on the percent of data points that were statistically significant at a 90% confidence level as follows:

- % Change Low to High Income: **.197** (100% significant ZIP Codes)
- % Change Median HH Income: **.179** (91% significant ZIP Codes)
- % Change Median HH Size: **.125** (63% significant ZIP Codes)
- Change in % White Non-Hispanic: **.153** (77% significant ZIP Codes)
- Change in % >=25 with Bachelor’s Degree: **.167** (85% significant ZIP Codes)
- % Change in Median Gross Rent: **.179** (91% significant ZIP Codes)

Results

Since the raw index has little independent meaning, we classified each score according to a change category, according to each ZIP Code’s index Z-Score as follows:

Z- Score	Classification
>= 2	Very High Change
1-2	High Change
.5-1	Medium Change
0-.5	Low Change
<0	No/Minimal Change

Results of the Top 30 ZIP Codes

Rank	ZIP Code	Neighborhood	Raw Score	Classification
1	90014	Downtown	0.814	Very High Change
2	90013	Downtown/Arts District	0.680	Very High Change
3	90017	Westlake/West Downtown	0.661	Very High Change
4	90015	Downtown/Pico Union	0.649	Very High Change
5	90026	Silver Lake/Echo Park	0.587	High Change
6	90028	Hollywood	0.570	High Change
7	90012	Chinatown	0.548	High Change
8	90038	Hollywood	0.545	High Change
9	90291	Venice	0.539	High Change
10	90029	East Hollywood/Silver Lake	0.537	High Change
11	90039	Silver Lake/Atwater Village/Frog Town	0.528	High Change
12	91601	North Hollywood	0.525	High Change
13	90027	Los Feliz	0.520	High Change
14	90042	Highland Park/Montecito Heights	0.496	High Change
15	90004	East Hollywood/Larchmont	0.483	High Change

Rank	ZIP Code	Neighborhood	Raw Score	Classification
16	90020	Koreatown/Windsor Square	0.472	Medium Change
17	90031	Frog Town/Lincoln Heights	0.469	Medium Change
18	90006	Pico Union	0.466	Medium Change
19	90065	Glassell Park/Cypress Park/Mt Washington	0.462	Medium Change
20	90005	Koreatown	0.460	Medium Change
21	90019	Mid-City/Arlington Heights	0.441	Medium Change
22	90010	Koreatown	0.437	Medium Change
23	90034	Palms/Mid-City	0.434	Medium Change
24	91605	Sun Valley	0.433	Medium Change
25	90032	El Sereno	0.429	Medium Change
26	90023	Boyle Heights	0.428	Medium Change
27	90048	Beverly Grove	0.424	Low Change
28	91606	North Hollywood	0.422	Low Change
29	90046	Hollywood/Hollywood Hills	0.414	Low Change
30	90016	West Adams/Mid-City/Baldwin Hills	0.410	Low Change

Los Angeles Displacement Pressure Index Methodology

Features

In order to identify Census Tracts where there are populations most vulnerable to future displacement pressures, we developed a list of close to twenty possible features, based on past research findings, as well as data availability. We subsequently trimmed the list to exclude collinear features, and features with a less definitive connection to our objectives. Sources and descriptions of the final seven features are listed below:

1. % of renter occupied households that pay $\geq 50\%$ of household income in rent

Source: American Community Survey, Five Year-Estimate (2014)

Description/Justification: This feature looks at rent burden among Los Angeles Census Tracts. Past researchers have found that rent burden has a positive relationship with in-migration to an area (Chapple and Loukaitou-Sideris, 2015). While rent burden is traditionally defined as households that pay $>30\%$ of household income on rent, area rent hikes have left nearly 60% of Los Angeles rental households burdened by the 30% definition. Thus, measuring rent burden at 30% reveals less variation among Census Tracts. To better capture the extremes of rent vs income, we use the highest ACS rent threshold of $\geq 50\%$ (or “severely rent burdened”).

2. Percent of Occupied Housing Units that are Renter Occupied

Source: American Community Survey, Five Year-Estimate (2014)

Description/Justification: As an area grows more desirable, rents tend to increase at accelerated rates. The impact of a neighborhood’s rent increases is proportional to the percent of households that rent. Indeed, past studies (Kennedy and Leonard 2001; Chapple 2009) have established that the % of renters in an area is positively correlated with how likely the tract is to gentrify.

3. Affordable Housing Units due to expire by 2023 (weighted by year of expiration)

Source: The City of Los Angeles Housing Element

Description/Justification: A vital way to ensure that lower income communities have housing is through a supply of affordable housing. While affordable housing is necessary regardless of whether a neighborhood has gentrified or is likely to gentrify in the future, the gap between market-rate and affordable housing is higher in more affluent neighborhoods, including those experiencing rapid change. A drop of affordable housing through expiring affordability covenants adds potential pressures for both traditional displacement and exclusionary displacement.

4. Proximity to Highly Changed ZIP Code

Source: The Los Angeles Index of Neighborhood Change (LAINC) – developed by the i-team

Description/Justification: Gentrification is rarely self-contained. As an area becomes more attractive to affluent residents, it often reaches a saturation point, whereby those who could once afford to move to a neighborhood no longer can due rising prices. The areas on the outskirts of said neighborhoods will often also change as a result. This concept has been implicitly explored through a study that looks at housing appreciation in lower income neighborhoods bordering higher income neighborhoods (Guerrieri, Hartley, Hurst 2010).

5./6. Proximity to current rail stations/Proximity to planned rail stations

Source: Los Angeles City Data

Description/Justification: Several past researchers have noted the link between Transit Oriented Development and subsequent gentrification (Chapple and Loukaitou-Sideris, 2015; Chapple 2009; MAPC 2014). In addition to upcoming transit, we also wanted to consider existing transit stations. As Los Angeles' transit system grows wider and more interconnected, we anticipate that the value/desirability of current stations will also increase.

7. Change in housing price projections (2015-2020)

Source: ESRI Community Analyst

Description/Justification: A spike in housing prices is often used to measure neighborhood change (Ding, Hwang, and Divringi 2015). Recent sales data trends reveal a number of neighborhoods that are currently below the City's median home value but are expected to appreciate more than the City as a whole. Using ESRI's housing projections, we considered tracts with housing prices that were <80% of the LA Median value in 2015 and are projected to increase at a higher rate than LA as a whole between 2015 and 2020.

Inclusion thresholds

In order to help ensure that we are focusing on tracts where there is still a high proportion of low-income residents vulnerable to displacement (as opposed to, for instance, Bel Air or the Silver Lake hills), an eligibility threshold was set for Census Tracts. The Census Tracts that do not meet the eligibility threshold are thus not included in the Los Angeles Displacement Pressure Index. For this analysis,

Census Tracts that had <40% of households earning below the Los Angeles City median income were excluded from consideration.

Weighting

To calculate a final displacement pressure score, we used a weighted average of feature scores. Each feature was normalized and multiplied by the individual samples' coefficients of variation (the value divided by its standard error).

Here, normalization also used the following equation:

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Proximity measures (those measuring a tract's distance to another feature), affordable housing, and housing price projections underwent additional as described below:

Current/Future Rail

For the current/future rail scores, we set a threshold distance of .5 miles. All tracts that were further than .5 miles received a score of 0. We then normalized the remaining tracts within .5 miles for a base score. Future rail scores were further scaled, receiving a full score if the station was set to open in 2016, 90% of the score if the line was set to open between 2017 and 2020 and 60% of the score for all stations set to open beyond 2020.

Proximity to high change zip codes

Tracts with centroids within a mile of one of the Top 15 changing Zip Codes (as defined by the LAINC) were considered. All other tracts received a score of 0 for this feature. The remaining tracts received a new weighted score comprised of two normalized measures:

- 1) The distance to the closest high-change tract
- 2) The index change score of the closest high-change tract

Affordable Housing

For affordable housing, we considered only the tracts that contained affordable housing. All other tracts received a score of 0. The remaining tracts received a new weighted score comprised of two normalized measures:

- 1) The average years until the affordable housing expires (1/3 weight)
- 2) The actual number of affordable units (2/3 weight)

Housing Price Projections

For the housing price projection variable, we isolated tracts with comparatively low current home values, defined here as tracts with an average value below the 40th percentile (tracts with an average home value below \$377,553). In addition, we only considered tracts where the projected home value was expected to increase more than Los Angeles as a whole (or >27.4%).

Adjusted features were then weighted according to a variety of factors including:

- How well the feature aligns with findings of past analyses
- The sampling error of the raw data
- Alignment with known on-ground conditions
- Isolation of features (absence of feature does not mean that displacement pressure is not present)
- Applicability to the broader project goals

Relative weights were assigned as such:

x3

- % of renter occupied households that pay >50% of household income in rent

x2

- Proximity to Highly Changed ZIP Code (scaled by distance to the high change ZIP code and the relative intensity of the past change)
- Proximity to planned rail stations (scaled by distance and how soon the station is expected to open)
- Change in housing price projections (2015-2020)
- Percent renter occupied housing units

x1

- Affordable Housing Units due to expire by 2023 (scaled by year of expiration)
- Proximity to current rail stations (scaled by distance)

Two initial variables (change in multi-family unit sales and proximity to parks) were included during our initial index build but were eliminated following dimensionality reduction as described in the next section.

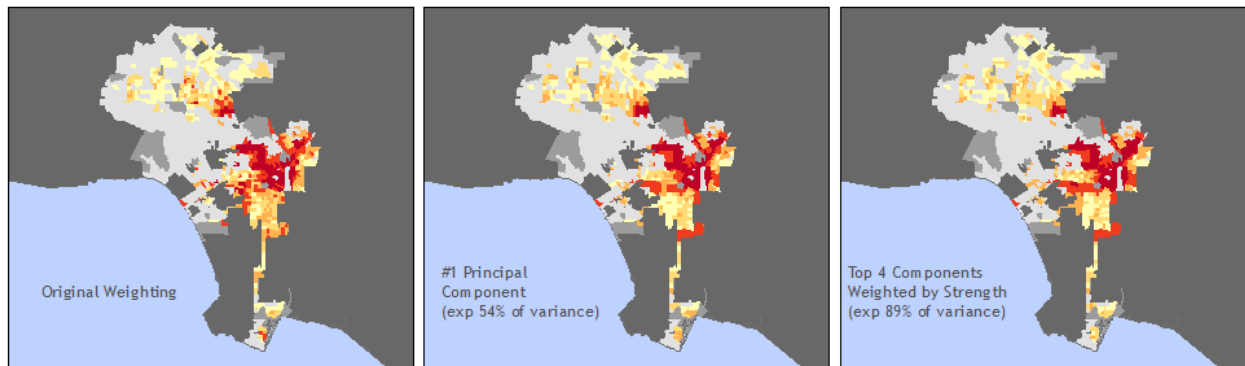
Additional Weighting Considerations

After our initial weighting, we considered alternative weighting using various machine learning techniques. First, we looked into the feasibility of regression, comparing 2000 to 2014 to determine effects on our outcome variable (LAINC score for linear regression, presence of gentrification for classification models). Given concerns around unavailability of some measures in 2000, differing geographies for the LADPI vs the LAINC and the non-absolute (or necessarily consistent) nature of our outcome variable, we decided not to pursue regression. As an alternative, we implemented Principal Component Analysis (PCA) to reclassify our nine features into uncorrelated components of varying explanatory value.

A subsequent analysis revealed that two of our features had virtually no correlation to the most explanatory components. These features (Park Acreage per Resident and Change in Multi Family Unit Sales) had no more than a .03 correlation with any of the first components. Park Acreage had a high correlation with the least variable component and Multi Family Unit Sales had a high correlation with the second least variable component. The two least explanatory components accounted for a total of <1% of the overall variability of our features. Presented with an opportunity to simplify our features, we

decided to remove these two and only look at the seven that have been described in detail earlier in this document.

For our alternate weighting schemes, we looked at just the top principal component (explaining 54% of variance) and the top four components (89% overall variance), aggregated according to each component's load factor. We then compared the results of the original weight to the two PCA weighted measures. This comparison revealed very similar overall vulnerability distribution. While the PCA methods place the index on a more statistically robust foundation, for the sake of interpretability, we continued with our initial weights, minus the two variables that contributed little to the overall variability.



For further clarification of the methodologies described in this document or additional questions about the Los Angeles Indices of Neighborhood Change, please contact Alex Pudlin at Alex.Pudlin@lacity.org.